



HAL
open science

Algorithmic Amplification of Politics and Engagement Maximization on Social Media

Paul Bouchaud

► **To cite this version:**

Paul Bouchaud. Algorithmic Amplification of Politics and Engagement Maximization on Social Media. 2023. <hal-04189507>

HAL Id: hal-04189507

<https://hal.science/hal-04189507v1>

Preprint submitted on 28 Aug 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Algorithmic Amplification of Politics and Engagement Maximization on Social Media

Paul Bouchaud¹²

¹ École des Hautes Études en Sciences Sociales, France

² Institut des Systèmes Complexes Paris Île-de-France, France
`paul.bouchaud@iscpif.fr`

Abstract. This study examines how engagement-maximizing recommender systems influence the visibility of Members of Parliament’s tweets in timelines. Leveraging engagement predictive models and Twitter data, we evaluate various recommender systems. Our analysis reveals that prioritizing engagement decreases the ideological diversity of the audiences reached by Members of Parliament and increases the reach disparities between political groups. When evaluating the algorithmic amplification within the general population, engagement-based timelines confer greater advantages to mainstream right-wing parties compared to their left-wing counterparts. However, when considering users’ individual political leanings, engagement-based timelines amplify ideologically aligned content. We stress the need for audits accounting for user characteristics when assessing the distortions introduced by personalization algorithms and advocate addressing online platform regulations by evaluating the metrics platforms aim to optimize, beyond the mere algorithmic implementation.

Keywords: recommender systems, algorithmic audits, social media regulation, ideology

1 Introduction

In recent years, social media platforms have become critical arenas for political discourse, providing a platform for politicians, political organizations, and news outlets to engage with vast audiences. To optimize user engagement and advertising revenue, these platforms employ sophisticated deep-learning algorithms to curate and prioritize content [5]. The implications of such engagement-maximizing recommender systems have sparked intense scholarly debate and public scrutiny. Previous research has pointed out the potential influence of algorithmic ranking on political discourse by amplifying emotionally-charged content [12, 21] and by contributing to polarization [8, 13, 24]. Nevertheless, limited access to large-scale data has posed challenges in fully investigating the effects of these algorithms. Independent audits of social networks have predominantly relied on enlisting volunteers to provide their data [12, 13, 20] and on “sock-puppet audits” [22, 23], in which one creates artificial users. While providing valuable

insights, "sock-puppet audits" audits are hindered by the limitations in replicating human digital behavior [9] and the number of fake accounts researchers can create. On the contrary, enlisting volunteers may be susceptible to potential selection bias [2] and may offer less experimental control, yet they do showcase real-life effects of social media algorithms. Alternatively, studies carried out by, or in collaboration with, corporations offer valuable insights, benefiting from access to non-public data and the means to conduct large-scale experiments. For instance, by leveraging proprietary user information and conducting a multi-year controlled experiment involving almost two million users, Huszár et al [1] unveiled how Twitter's recommender unevenly amplified tweets from politicians based on their ideological leaning. Recently, through a collaboration with Meta, Guess et al. [25] found that switching users from algorithmic feeds to reverse-chronological feeds significantly reduced their platform usage, increased their exposure to political content, to content from moderate friends and ideologically mixed sources on Facebook.

Amid the growing regulatory scrutiny, such as with the Digital Services Act in the EU or the proposed Online Safety Bill in the UK, there emerges a pressing need to enhance methodologies for auditing social media platforms. This article aims to showcase the potential of empirically grounded social media simulations to both strengthen audit methodologies and assess prospective policies. Specifically, by leveraging engagement predictive models trained on behavioral data, we investigate different recommender systems on timelines, focusing on the visibility of tweets authored by Members of Parliament (MPs). Our study seeks to i) illustrate how the optimization of social media timelines for engagement leads to a reduction in ideological diversity among the audience of MPs, and ii) underscore the significance of accounting for individual user characteristics when auditing social media platforms.

2 Methods

In this section, we outline the development of engagement predictive models and their use to simulate timelines. Additionally, we introduce the amplification metrics employed to analyze the recommender systems.

2.1 Engagement predictive models

In order to develop predictive models for user engagement, the first step was to obtain a training dataset. Unfortunately, the extensive behavioral data needed for such training are proprietary assets of the companies and not publicly available. Although Twitter had collaborated with ACM RecSys to release extensive datasets for engagement prediction challenges in 2020 [3] and 2021 [4], these challenges' sole focus was to enhance prediction accuracy, neglecting to address wider ethical implications; also the datasets were no longer accessible for our research. Consequently, we compiled a training dataset using the Horus data-donation browser add-on [12], installed by 1.6k volunteers. During the first quarter of

2023, we collected over 2.5 million tweet impressions on volunteers’ timelines, recording 43k likes and 8k retweets.

In addition to the captured tweets, we enriched our training dataset of historical and account specific features, such as the previously liked and retweeted tweets or its number of followers. To assess the similarity between Twitter accounts, we leveraged a follow graph large of 260 million nodes and 700 million edges, built via snowballing starting from half a million seed accounts. After pruning poorly connected nodes (degree less than 50), we used node2vec [15] to embed the follow graph into low-dimensional representations, while preserving neighborhood relationships. The graph of follow allows to retrieve meaningful proximity between accounts and is notably leveraged by Twitter within its recommendation pipelines [11, 30].

Building upon insights from the ACM RecSys Challenges [31], we trained gradient-boosting machines, specifically LightGBM [16], over a set of hand-crafted features. During the training process, several features emerged as particularly significant to predict likes and retweets: the time delay between tweet publication and its impression, the average word length in the tweet, and the distance in the follow space between the tweet’s author and the last authors liked and retweeted by the user. For further information on the training process and the complete list of features available to the models, refer to [6].

Overall, our models achieved an average precision of 71.3% for predicting likes and 88.4% for predicting retweets, which, compared to a naive baseline, corresponds to a cross-entropy gain of 30.5% for likes and 63.5% for retweets. One should note that our aim is not to claim beating the current state-of-the-art with new models, but to train models that adequately predicts engagement for the purpose of the present demonstration.

To evaluate the models’ generalization ability, we leveraged the findings from Barbiero et al. [17] and computed the convex hull of the training dataset in the feature space. Remarkably, a substantial portion (88.9%) of the testing dataset samples lay outside the training set’s convex hull, necessitating the models to extrapolate for making predictions. We observed that the models achieved high accuracy for both testing points inside and outside the training convex hull, demonstrating successful generalization. Additionally, we found that the samples from the inference dataset (introduced later) closely resembled the training dataset, with an average cosine similarity with the closest training sample of 0.95. In summary, our models effectively generalized from the training samples, and the samples from the simulated timelines exhibited high similarity with the training data.

2.2 Timelines simulation

The models’ ability to generalize, along with our approach of relying solely on publicly available data for inference, enables us to estimate likes and retweets probability for a wide range of Twitter accounts. We opted to simulate timelines for a set U of 2690 French public Twitter accounts. These accounts were sampled such as reflecting the same distribution of political orientations as that of the

general population, within those having retweeted or been retweeted in political-related tweets at least five times in 2022. We collected the likes and retweets history of each chosen account, along with the list of accounts it followed.

To build the corpus of tweets that may appear in the simulated timelines, we collected the tweets published or retweeted by 96k of our users’ friends (covering 71% of users’ friends) over a 30-day period from February 18, 2023, to March 19, 2023. For each user, we simulated timelines by having them log in at random hours on seven randomly chosen days and reading a set of L tweets recommended by the system. Both the hour of connection and the session length L are drawn from the empirical distributions determined by our data-donation initiative (the average session length equals $L = 35$ tweets).

To gain a deeper understanding of the mechanism in play, we implemented a range of recommender systems producing timelines with a fraction ϕ of tweets that maximize engagement and a fraction $1 - \phi$ of the most recent tweets. To build the timelines, one first aggregates all tweets published or retweeted by the user’s friends (in our database) within the 18-hours period prior the user’s session. The aggregation window was chosen to balance the size of the tweet pool and the associated computational costs, see a sensitivity analysis in [6]. Using engagement predictive models, one estimates the probability of each tweet being liked or retweeted by the user if impressed in the current timeline. Following Twitter’s design [28], one then sums the two probabilities to obtain an engagement score. The user is then served a timeline made of the ϕL tweets with the highest engagement score and of $(1 - \phi)L$ most recent tweets.

Furthermore, we incorporated a heuristic to guarantee that each account would only appear once within the timeline. This heuristic is stronger than the one employed by Twitter [30] and aim to bolster the robustness of our conclusions by counteracting possible trivial distorting effects. One should note that, this article does not address feedback loops; successive timelines are treated independently, and one always referred to the real past engagements when predicting engagement. Moreover, while we explore a range of recommenders spanning from reverse chronological feeds to those maximizing engagement, we do not view the former as ”neutral”, chronological feeds are recency-based rankings that primarily reward individuals who post frequently. We chose such ranking as a baseline due to their simplicity and widespread adoption.

2.3 Metrics

Relative Amplification In this study, we focus on the reach of tweets published by elected officials members of the French National Assembly. Adopting the approach of Huszár et al. [1], we define the reach amplification $a_R(T, U)$ of a set T of tweets within an audience U as the ratio between the number of users in U who saw at least one tweet from T in their engagement-maximizing timelines and in their reverse-chronological timelines.

By focusing solely on reach, the measure $a_R(T, U)$ used by [1] may overlook distortions in terms of impressions, particularly given the significant differences in political group sizes, see Table 1. To address this limitation, we introduce the

impression amplification, $a_I(T, U)$, as the ratio between the number of impressions of tweets from the set T in the engagement-based or reverse-chronological timelines of users in U .

$$a_R(T, U, d) = \frac{\sum_{u \in U} \mathbb{I}(feed^{eng}(u, d) \cap T \neq \emptyset)}{\sum_{u \in U} \mathbb{I}(feed^{chrono}(u, d) \cap T \neq \emptyset)} - 1 \quad \& \quad a_I(T, U, d) = \frac{\sum_{u \in U} |feed^{eng}(u, d) \cap T|}{\sum_{u \in U} |feed^{chrono}(u, d) \cap T|} - 1$$

with $feed^{eng, chrono}(u, d)$ the set of tweets figuring in the timelines, either engagement-based or reverse-chronological, of user u at day d .

As in [1], we refer as group amplification the amplification of the set T of all tweets published by members of a given political party, and normalize it such that an amplification ratio of 0% corresponds to an equal reach in engagement-based and reverse-chronological timelines. Higher amplifications indicate that engagement predictive models assigns higher scores to the set of tweets, leading them to appear more frequently than they would in a reverse-chronological timelines. We emphasize that both measures of amplification are relative to our choice of considering the reverse-chronological timelines as a baseline. To ensure the robustness of our analysis, we performed bootstrapping over the users and days of impressions for all the amplification measures, subsequently reported with 95% bootstrap confidence interval. To explore more thoroughly the consequences of engagement maximization, we form subgroups $\tilde{U} \subset U$ made of users with a particular political leaning.

Group	Communist	LFI	Ecologist	Socialist	LIOT	Democrat	Renaissance	Horizons	Republican	National Rally
Abbreviation	PCF	LFI	EELV	PS	LIOT	MODEM	RE	HOR	LR	RN
Members	22	75	23	27	21	51	162	26	59	88

Table 1: Political group membership at the French National Assembly (XVIIth legislature), from left to right. The abbreviation corresponds to the main political party composing the group at the National Assembly. "LFI" stands for "La France Insoumise" (far-left) and "LIOT" for "Libertés, indépendants, outre-mer & territoires" (catch-all group). We excluded from our analysis 5 non-attached MPs

Political Leaning To determine the political leanings of Twitter accounts, we leveraged the retweet graph associated with all tweets from French political figures or containing French political keywords published in 2022. Such graph, collected within the Politoscope project, have been shown to be reliable indicators of ideological alignment [19]. While a clustering analysis of the retweet graph do provide categorical political leanings of each user, we seek a more granular analysis and adopted the methodology used in [6, 12] to obtain a continuous opinion scale.

Using the node2vec algorithm [15], we generated embeddings for each Twitter account within the retweet graph, then capturing its underlying structure. We anchored the positions, on our opinion scale, of the far-left leader, Jean-Luc Mélenchon, and the far-right leader, Marine Le Pen, to -0.75 and $+0.75$, respectively. For determining the opinion of the centrist leader, Emmanuel Macron, we

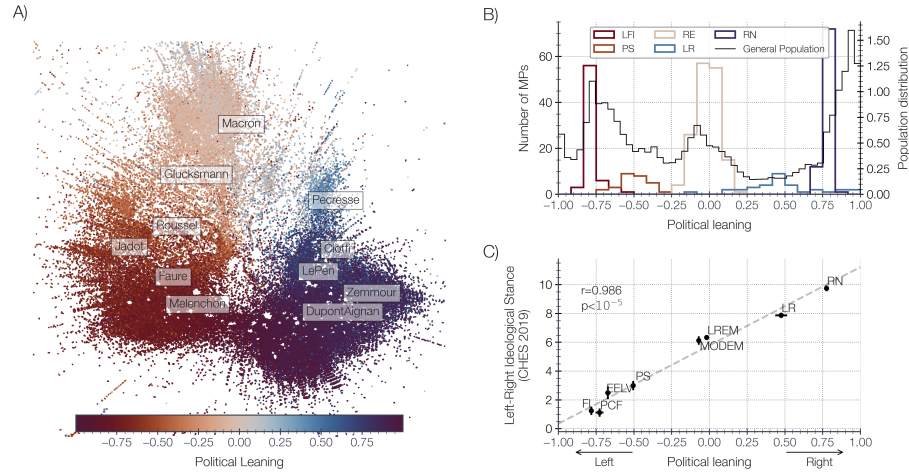


Fig. 1: A) The graph of retweets associated with political tweets published in 2022, spatialized using ForceAtlas2 [26]. Nodes are colored based on the assigned numerical opinion. B) The distribution of assigned political leanings for Members of Parliament belonging to the far-left (LFI), left (PS), centrist (RE), right (LR), and far-right (RN) groups are presented. The distributions of opinion in the general population, i.e. all Twitter accounts in the political retweet graph, is displayed in black. C) The average position on the opinion scale of MPs in each political group is compared to the estimation made by political experts within the 2019 Chapel Hill Expert Survey [7]; error bars represent bootstrap standard errors on the average position.

averaged the opinion of the two opposite leaders, weighted by the angular similarities between their embeddings and Emmanuel Macron’s, resulting in an opinion of -0.02 .

For each Twitter account, we then computed the angular similarity between its embedding and the embeddings of the three leaders. The account’s political leaning was subsequently determined as the average opinion of the two closest leaders, weighted by their angular similarities. To account for the circular nature of the French political landscape, where anti-system activists bridge far-left and far-right militants [10] (as shown in Figure 1.A), we implemented a periodic boundary condition at ± 1 in our numerical opinion estimates. This approach ensures interpretability, with negative values indicating left-leaning accounts, positive values indicating right-leaning accounts, and supporters of the current French President Emmanuel Macron being around zero. The numerical scale aligns with the political group of Members of Parliament, see Figure 1.B, the assessment of political experts (2019 Chapel Hill Expert Survey [7]) see Figure 1.C, and a clustering analysis [19]. A visual representation of the political land-

scape is shown in Figure 1, where nodes are color-coded based on their assigned numerical opinion, providing a clear interpretation of the political landscape.

Audience Diversity Having estimated the political leaning of each user, we examine the ideological composition of Members of Parliament’s audiences across various recommender systems i.e. different values of ϕ . Specifically, we compute the average absolute difference of opinion, taking account the periodicity at ± 1 , between a Member of Parliament and the users who have encountered their tweets, weighted by the number of impressions.

3 Results

3.1 Relative Amplification

We display on Figure 2 the reach and impression amplifications, within the general population, of tweets published by members of each political groups. The tweets from the governing centrist party Renaissance appear $a_I(T_{RE}, U) = 128.4\%_{[123.3, 134.6]}$ more in the engagement-based timelines than in the recency-based ones, the tweets from the Republican party are amplified by $a_I(T_{LR}, U) = 29.8\%_{[27.1, 33.1]}$ while those of Socialist MPs by $a_I(T_{PS}, U) = 3.3\%_{[0.1, 6.4]}$. By performing a permutation test over the MPs of the Socialist and Republican groups, one conclude, with statistical significance $p < 0.05$, that the mainstream right-wing party benefit more from engagement-based timelines than left-wing party, both under reach and impression amplification metrics.

To gain a deeper understanding of the underlying mechanisms, we segment users on their political leaning and investigate the reach and impression amplification of MPs’ tweets in the timelines of left-leaning users (Figure 2.A), center-leaning users (Figure 2.B), and right-leaning users (Figure 2.C). Across all ideological subgroups, we observed that tweets from MPs who align with the users’ political views were amplified. This led to their tweets reaching a larger set of users (reach amplification) and appearing more frequently (impression amplification) in the engagement-based timelines compared to the recency-based timelines. Conversely, tweets from MPs with opposing political leanings were algorithmically reduced in visibility. For instance, tweets from ecologist MPs appeared, at least once, in the timelines of $a_R(T_{ecologist}, U^{right}) = -84.2\%_{[-88.5, -79.4]}$ less right-leaning users in engagement-based timelines than in reverse-chronological ones. By conducting permutation tests over MPs either ideologically aligned or disaligned with the users, we conclude with statistical significance that ideologically aligned MPs benefit more from engagement-based ranking than dissonant MPs.

In all examined instances, the reach amplification was found to be smaller than the impression amplification, and even overly conservative when considering MPs of the same leaning as the users. For example, when considering left-leaning users, the number of users encountering at least one tweet published by LFI MPs was equivalent in both engagement-based and recency-based timelines,

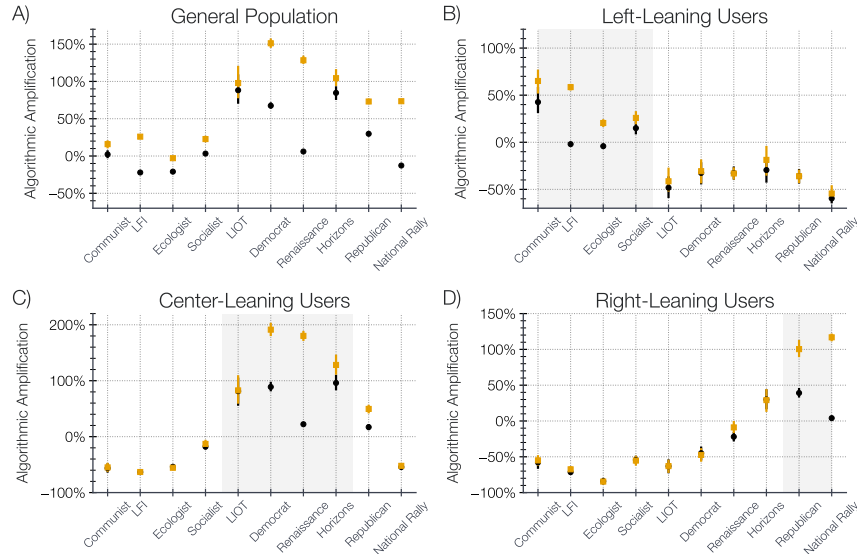


Fig. 2: Reach (black circles) and Impression (orange squares) group amplifications. Amplification computed by considering the general population (A), left-leaning users (B), center-leaning users (C), and right-leaning users (D). Error-bars correspond to 95% bootstrap confidence interval. The political groups that aligns with the users’ ideology is shaded in grey for clarity.

$a_R(T_{LFI}, U^{left}) = -1.98\%_{[-3.24, -0.58]}$. However, the impression amplification showed a substantial increase, reaching $a_I(T_{LFI}, U^{left}) = 58.6\%_{[55.0, 61.9]}$. This means that left-leaning users saw over half more tweets authored by LFI MPs in their engagement-based timelines compared to what they would have seen in recency-based timelines. Similar patterns were observed for center-leaning users with the major centrist party Renaissance (as shown in Figure 2.B) and for right-leaning users with the right-leaning Republican group (as shown in Figure 2.C).

3.2 Audience Diversity

To gain insights into the audience reached by MPs’ tweets, we present on Figure 3.A the average absolute difference of opinion between MPs and the users who encountered their tweets, as a function the fraction of engagement-maximizing tweets in the timelines, ϕ . Firstly, substantial differences in the ideological diversity of MPs’ audiences exist in reverse-chronological timelines, $\phi = 0$. Tweets from MPs belonging to the centrist party (opinion around 0) and the far-right party (+0.75) are primarily seen by ideologically aligned users (see Figure 3.C), with an average absolute difference of opinion of 0.183 and 0.182. Conversely, tweets from the mainstream left and right parties (± 0.45) are impressed on the

timelines of a more diverse set of users (see Figure 3.B, D), with an average absolute of difference of 0.300 and 0.355. Secondly, as the timelines transition from recent to engagement-maximizing tweets, there is a systematic decrease in the opinion diversity of the audiences reached by MPs, and an increase in the disparities between political groups. Compared to reverse-chronological timelines, at $\phi = 1$ the ideological diversity of the audiences reached by centrist and far-right MPs’ tweets decreased by 42.9% and 41.7% while those of Socialist and Republican MPs decreased by 30.5% and 34.3%.

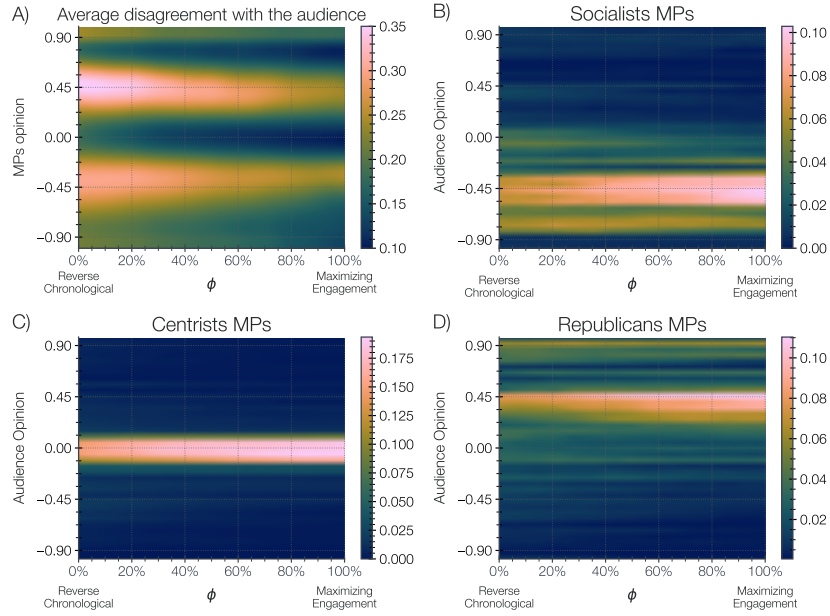


Fig. 3: A) Average absolute difference of opinion between Members of Parliament (MPs) and their respective audiences. The distributions of opinion among the audience of socialist (B), centrist (C), and republican (D) MPs are shown as a function of ϕ , representing the fraction of tweets maximizing engagement in their timelines.

4 Discussion

In this study, we employed engagement predictive models trained on behavioral data to construct timelines that users would have been exposed to under different content recommender policies. Following the work of Huszár et al. [1], we focused on how tweets from Members of Parliament are presented on users’ timelines. As [1], we find that —when aggregating over the whole population— engagement-based timelines benefit mainstream right-wing parties

more than left-wing parties. However, upon segmenting users based on their political leaning, we observed that engagement-based timelines merely favor tweets from authors ideologically aligned with the users, as discussed in [6, 12]. These findings underscore the significance of taking into account individual user characteristics, here the political leaning, during the audit process. The research and regulatory implications of an audit revealing —through aggregating over the whole population— a platform-wide preference towards right-wing political content diverge from the implications of a study which —by considering users’ political leaning— uncover an algorithmic confirmation bias, where ideologically aligned content is amplified over discordant one. Also, we showcase how crucial is the definition of algorithmic amplification, e.g. focusing on the reach of tweets may overlook an algorithmic confirmation bias revealed by inspecting the number of impressions. Finally, exploring different recommender systems, ranging from reverse-chronological to engagement-maximizing, we uncovered that: as recommender systems prioritize engagement, the ideological diversity of the audience reached by different MPs decreases, while the disparities between political groups’s reach increase.

It is important to note that our models are not trained to replicate Twitter’s recommender systems but rather to study the consequences of maximizing engagement. As a result, our simulations present noticeable differences with the findings of [1]. In our simulations, the amplification within the general population ranges from -22.0% to 88.2%. In contrast, [1] reported amplification consistently above 100%, reaching 153.1% for the Republican group. This disparity may be attributed to the convoluted interplay of factors within Twitter’s recommender systems, which blend deep-learning models with hand-crafted heuristics. For example, Twitter assigns a reputation score to accounts [29], through a variant of the PageRank algorithm considering the number of followers and age of the account [27]; owing to such heuristics one may hypothesize that Members of Parliament might experience distinct treatment. Additionally, our timelines were exclusively composed of tweets published or retweeted by users’ friends, without any injected content, unlike on Twitter. Furthermore, we acknowledge changes in the French political landscape between [1] and our study, such as the entry of a large far-right group into the assembly and shifts in alliances among other groups.

Despite these differences, the primary goals of our work remain twofold: first, to illustrate i) how optimizing for engagement unevenly amplify MPs tweets based on their ideology and leads to a reduction in ideological diversity among MPs’ audience; and ii) to underscore the importance of considering individual user characteristics while analyzing social-technical systems like social media, as it significantly impacts the conclusions of audits. Therefore, we advocate for fine-grained audits to thoroughly assess the distortions introduced by personalization algorithms. Furthermore, in conjunction with conducting comprehensive audits of the algorithms implemented by the platforms, we advocate addressing the regulation of online services through an assessment of the metrics that platforms are optimizing, such as engagement.

5 Acknowledgments

The author deeply thanks Pedro Ramaciotti Morales for his precious insights and Maziyar Panahi for enabling the collection of the large-scale retweet network. Finally, the author acknowledges the Jean-Pierre Aguilar fellowship from the CFM Foundation for Research, the support and resources provided by the Complex Systems Institute of Paris Île-de-France.

References

1. Huszár, F., Ktena, S., O'Brien, C., Belli, L., Schlaikjer, A. & Hardt, M. Algorithmic amplification of politics on Twitter. *Proc. Natl. Acad. Sci. U.S.A.* **119** (2021,12), <https://doi.org/10.1073>
2. Kmetty, Z., Stefkovics, Szamely, J., Deng, D., Anikó, K., Omodei, E., Edit, P. & Koltai, J. Determinants of willingness to donate data from social media platforms. (Center for Open Science, 2023, 3), <https://doi.org/10.31219>
3. Belli, L. et al Privacy-Aware Recommender Systems Challenge on Twitter's Home Timeline. (2020)
4. Belli, L. et al The 2021 RecSys Challenge Dataset: Fairness is not optional. *RecSysChallenge '21: Proceedings Of The Recommender Systems Challenge 2021*. (2021,10), <https://doi.org/10.1145>
5. Satuluri, V., Wu, Y., Zheng, X., Qian, Y., Wichers, B., Dai, Q., Tang, G., Jiang, J. & Lin, J. SimClusters. *Proceedings Of The 26th ACM SIGKDD International Conference On Knowledge Discovery & Data Mining*. (2020,8), <https://doi.org/10.1145>
6. Bouchaud, P. Skewed Perspectives: Examining the Influence of Engagement Maximization on Content Diversity in Social Media Feeds. (2023,6), <https://hal.science/hal-04139494>, preprint
7. Jolly, S., Bakker, R., Hooghe, L., Marks, G., Polk, J., Rovny, J., Steenbergen, M. & Vachudova, M. Chapel Hill Expert Survey trend file, 1999–2019. *Electoral Studies*. **75** pp. 102420 (2022,2), <https://doi.org/10.1016>
8. Rathje, S., Bavel, J. & Linden, S. Out-group animosity drives engagement on social media. *Proc. Natl. Acad. Sci. U.S.A.* **118** (2021,6), <https://doi.org/10.1073>
9. Ribeiro, M., Veselovsky, V. & West, R. The Amplification Paradox in Recommender Systems. (2023)
10. Chavalarias, D., Bouchaud, P. & Panahi, M. Can Few Lines of Code Change Society ? Beyond fact-checking and moderation : how recommender systems toxifies social networking sites. (2023,2), <https://hal.science/hal-04031304>, preprint
11. Rossi, W., Polderman, J. & Frasca, P. The Closed Loop Between Opinion Formation and Personalized Recommendations. *IEEE Trans. Control Netw. Syst.* *Transactions On Control Of Network Systems*. **9**, 1092-1103 (2022,9), <https://doi.org/10.1109>
12. Bouchaud, P., Chavalarias, D. & Panahi, M. Crowdsourced Audit of Twitter's Recommender Systems Impact on Information Landscapes. (2023,3), <https://hal.science/hal-04036232>, preprint
13. Milli, S., Carroll, M., Pandey, S., Wang, Y. & Dragan, A. Twitter's Algorithm: Amplifying Anger, Animosity, and Affective Polarization. (2023)
14. Bavel, J., Rathje, S., Harris, E., Robertson, C. & Sternisko, A. How social media shapes polarization. *Trends In Cognitive Sciences*. **25**, 913-916 (2021,11), <https://doi.org/10.1016>

15. Grover, A. & Leskovec, J. node2vec. *Proceedings Of The 22nd ACM SIGKDD International Conference On Knowledge Discovery And Data Mining*. (2016,8), <https://doi.org/10.1145>
16. Ke, G., Meng, Q., Finely, T., Wang, T., Chen, W., Ma, W., Ye, Q. & Liu, T. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. *Advances In Neural Information Processing Systems 30 (NIP 2017)*. (2017,12)
17. Barbiero, P., Squillero, G. & Tonda, A. Modeling Generalization in Machine Learning: A Methodological and Computational Study. (2020)
18. Milli, S., Pierson, E. & Garg, N. Choosing the Right Weights: Balancing Value, Strategy, and Noise in Recommender Systems. (2023)
19. Gaumont, N., Panahi, M. & Chavalarias, D. Reconstruction of the socio-semantic dynamics of political activist Twitter networks—Method and application to the 2017 French presidential election. *PLoS ONE ONE*. **13**, e0201879 (2018,9), <https://doi.org/10.1371>
20. Hargreaves, E., Agosti, C., Menasche, D., Neglia, G., Reiffers-Masson, A. & Altman, E. Biases in the Facebook News Feed: A Case Study on the Italian Elections. *2018 IEEE/ACM International Conference On Advances In Social Networks Analysis And Mining (ASONAM)*. (2018,8), <https://doi.org/10.1109>
21. Brady, W., Wills, J., Jost, J., Tucker, J. & Bavel, J. Emotion shapes the diffusion of moralized content in social networks. *Proc. Natl. Acad. Sci. U.S.A.*. **114**, 7313-7318 (2017,6), <https://doi.org/10.1073>
22. Bartley, N., Abeliuk, A., Ferrara, E. & Lerman, K. Auditing Algorithmic Bias on Twitter. *13th ACM Web Science Conference 2021*. (2021,6), <https://doi.org/10.1145>
23. Bandy, J. & Diakopoulos, N. More Accounts, Fewer Links. *Proc. ACM Hum.-Comput. Interact. On Human-Computer Interaction*. **5**, 1-28 (2021,4), <https://doi.org/10.1145>
24. Bavel, J., Rathje, S., Harris, E., Robertson, C. & Sternisko, A. How social media shapes polarization. *Trends In Cognitive Sciences*. **25**, 913-916 (2021,11), <https://doi.org/10.1016>
25. Guess, A. et al. How do social media feed algorithms affect attitudes and behavior in an election campaign?. *Science*. **381**, 398-404 (2023,7), <https://doi.org/10.1126>
26. Jacomy, M., Venturini, T., Heymann, S. & Bastian, M. ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software. *PLoS ONE ONE*. **9**, e98679 (2014,6), <https://doi.org/10.1371>
27. Twitter TweepCred. *GitHub.*, <https://github.com/twitter/the-algorithm/blob/main/src/scala/com/twitter/graph/batch/job/tweepcred>
28. Twitter Source Code for Twitter's recommendation algorithm: Heavy Ranker. *GitHub.*, <https://github.com/twitter/the-algorithm-ml/blob/main/projects/home/recap>
29. Twitter Twitter/the-Algorithm: Source Code for Twitter's recommendation algorithm. *GitHub.*, <https://github.com/twitter/the-algorithm>
30. Twitter Twitter's recommendation algorithm. *Twitter.*, https://blog.twitter.com/engineering/en_us/topics/open-source/2023/twitter-recommendation-algorithm
31. Twitter What Twitter learned from the Recsys 2020 challenge. *Twitter.*, https://blog.twitter.com/engineering/en_us/topics/insights/2020/what_twitter_learned_from_recsys2020