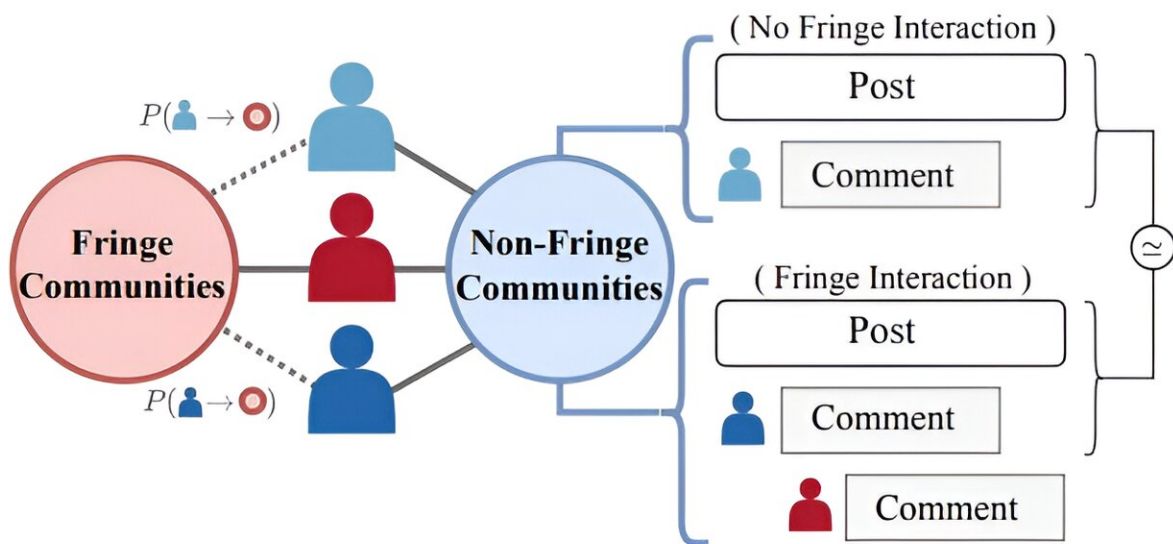


Cross-community interactions with fringe users increase the growth of fringe communities on Reddit

August 5 2024, by Tanya Petersen



We study whether interacting with fringe users (red) causes non-fringe users (blue) to join fringe communities. We match users that involuntarily received comments from fringe users (blue; “treatment”) with users who did not, but had similar chances to do so (light blue; “control;” light blue \approx blue). Then, we estimate the causal effect of fringe interactions by comparing the fraction of users that went on to join the fringe community in the treatment ($P(\text{blue} \rightarrow \text{red}\bullet)$) and in the control group ($P(\text{light blue} \rightarrow \text{red}\bullet)$). Credit: *arXiv* (2023). DOI: 10.48550/arxiv.2310.12186

New EPFL research has found that the exchange of comments between members and non-members of fringe communities (fringe interactions) on mainstream online platforms attracts new members to these extremist groups. It has also suggested potential ways to curtail this growth.

Fringe communities promoting [conspiracy theories](#) and extremist ideologies have thrived on mainstream online platforms, consistently raising the question of how this growth is fueled. Now, researchers in EPFL's School of Computer and Communication Sciences (IC) studying this phenomenon have found what they believe to be a possible mechanism supporting this growth.

In their [paper](#) posted to the preprint server *arXiv*, "Stranger Danger! Cross-Community Interactions with Fringe Users Increase the Growth of Fringe Communities on Reddit," scientists from the Data Science Laboratory (DLAB) have outlined fringe interactions—the exchange of comments between members and non-members of fringe communities.

Applying text-based causal inference techniques to study the impact of fringe interactions on the growth of fringe communities, the researchers focused on three prominent groups on Reddit: r/Incel, r/GenderCritical, and r/The Donald. The data for this research was collected between 2016 and 2020. All these communities have been banned and are no longer active.

"It's well known that there are extremist communities on the big online social platforms and equally known that these platforms try to moderate them, however even with moderation, these communities keep growing. So, we wondered what else could be driving this growth?" explained Giuseppe Russo, a postdoctoral researcher at DLAB.

"We found that fringe interactions—basically entailing an exchange of comments between a 'vulnerable user' and a user already active in an

extremist community—were bringing people on the edge towards these groups. And the comments aren't obvious, like 'join my community'—they are topical discussions that attract the attention of other people and a clever way to skip moderation policies," he continued.

Users who received these interactions were up to 4.2 percentage points more likely to join fringe communities than similar, matched users who did not. This effect was influenced by the characteristics of communities where the interaction happens (e.g., left vs. right-leaning communities) and the language used in the interactions. Toxic language had a 5 percentage point higher chance of attracting newcomers to fringe communities than non-toxic interactions.

Once a vulnerable user had engaged in an exchange of comments, finding an extremist community online was just a click away, as every user profile immediately shows the groups in which someone has been active. The researchers also found that this recruitment method is unique to extremist communities—they didn't find it occurring in climate, gaming, sporting or other communities—supporting the hypothesis that this could be an intended strategy.

"The recruitment mechanism is very simple, it's literally just one exchange of comments. One small discussion is enough to drive vulnerable users towards these bad communities and once they are there, many previous studies have shown that they radicalize and become active members of these extremist communities," said Russo.

The research results raise important questions—is this recruitment mechanism relevant enough to warrant the attention of online platforms' moderation policies and, if yes, how can we stop it from helping fuel the growth of these extremist communities?

Based on the observational period considered, the researchers estimate approximately 7.2%, 3.1% and 2.3% of users that joined r/Incels, r/GenderCritical, and r/The Donald respectively, did so after interacting with members of those groups.

They say that this suggests that community-level moderation policies could be combined with sanctions applied to individual users, for example reducing the visibility of their posts or limiting the number of comments they can make in more susceptible communities. The researchers believe that these measures may diminish the impact of fringe interactions and slow down the growth of fringe communities on mainstream platforms.

"From a societal perspective, I think this is extremely relevant for many reasons, one of which is that we have seen that the impact of these fringe communities is not exclusively online, there is an offline impact too. We have seen riots and terrorism, we have seen women targeted and killed by members of these communities, so by reducing access to these communities we definitely reduce the risk offline, which is what we really care about in the end," concluded Russo.

More information: Giuseppe Russo et al, Stranger Danger! Cross-Community Interactions with Fringe Users Increase the Growth of Fringe Communities on Reddit, *arXiv* (2023). [DOI: 10.48550/arxiv.2310.12186](https://doi.org/10.48550/arxiv.2310.12186)

Provided by Ecole Polytechnique Federale de Lausanne

Citation: Cross-community interactions with fringe users increase the growth of fringe communities on Reddit (2024, August 5) retrieved 5 August 2024 from <https://phys.org/news/2024-08-community-interactions-fringe-users-growth.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.