Turning the Tables on Facebook
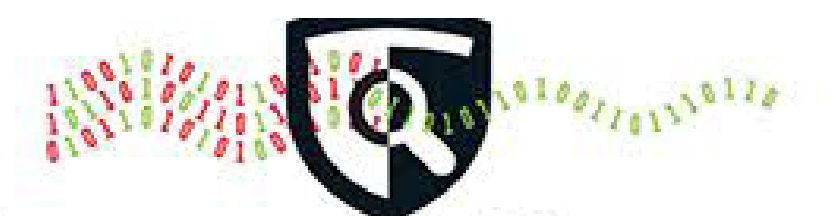
Levi Kaplan^, Muhammad Ali*, Piotr Sapiezynski^*, Alan Mislove^*

Background

Advertisers use Facebook to reach consumers with offers of not only products and services but also life opportunities, political messages, and outright scams. Each ad impression is an effect of a complex interplay between the advertiser's targeting parameters and the platform's ad delivery algorithms. We turn the tables on Facebook by using their own tools to bring much needed transparency and accountability to this process. Our work uncovered instances of discrimination, differential pricing, and other worrying phenomena, and led to real-world accountability. The papers below are the most recent publications in our Facebook series.

Takeaways

Overall, our work brings up interesting legal and ethical considerations for Facebook to take into account when designing such systems. We illustrate that there are potential harms arising from who a Facebook user is or who Facebook interprets them to be. In our first study, we show this to be harmful from the perspective of an advertiser, whose diverse ads featuring young women may not be shown to diverse audiences but rather to older men. In our second study, we show this to be harmful from the perspective of a Facebook user, who may be disproportionately delivered harmful ads like clickbait or sensitive ads like those for healthcare.

Study 1^

How does the demographic of the face in a Facebook ad affect that ad's delivery?

Study 2*

How are harmful ads on Facebook delivered, and to which subset of users?

Approach

We take the role of an advertiser and run a series of controlled advertisements containing pictures of people to understand how the demographics of a person in an advertisement influences how Facebook delivers that ad. To control for potentially confounding differences, we also create a series of computer-generated faces using StyleGAN, changing only the demographic of interest (c). We measure race using voter records from North Carolina and Florida divided by the voters' self-reported race. By running a Custom Audience using these voters, we can use delivery by state as a proxy for race.

Approach

We recruited 41 diverse participants to donate all the Facebook ads they see along with the associated targeting information. We manually coded over 3,000 of these ads and investigated how the different ad types are distributed among the participants.

Results

Lorenz curves (a) and the Gini coefficient (b) emphasize the same message: Potentially harmful ad types such such as Sensitive, Clickbait and Potentially Prohibited have a significantly higher level of inequality in contributions compared to the Neutral ads. Such ads are disproportionately shown to a small subset of users, while most users see very few of them. The long tail of online harms means that a fraction of users face problems that the majority might not even be aware of.

Results

Generally, ads with a certain demographic are delivered more to users of that same demographic. Ads containing images of Black people are delivered more to Black Facebook users. This breaks down when considering ads containing teenage and young adult women—for these ads, they aren't delivered more to women, but instead are delivered more to older men (d). This could run counter to an advertiser's intent for using images of young women for diversity campaigns. Defaulting to pictures of white men may be more harmful than previously thought, since minority individuals may not be delivered them.

(c)

(d)

(a)

(b)