Abstract

The Web consists of numerous Web communities, news sources, and services, which are often exploited by various entities for the dissemination of false or otherwise malevolent information. Yet, we lack tools and techniques to effectively track the propagation of information across the multiple diverse communities, and to capture and model the interplay and influence between them. Furthermore, we lack a basic understanding of what the role and impact of some emerging communities and services on the Web information ecosystem are, and how such communities are exploited by bad actors (e.g., state-sponsored trolls) that spread false and weaponized information.

In this thesis, we shed some light on the complexity and diversity of the information ecosystem on the Web by presenting a typology that includes the various types of false information, the involved actors as well as their possible motives. Then, we follow a data-driven cross- platform quantitative approach to analyze billions of posts from Twitter, Reddit, 4chan's Politically Incorrect board (/pol/), and Gab, to shed light on: 1) how news and image-based memes travel from one Web community to another and how we can model and quantify the influence between the various Web communities; 2) characterizing the role of emerging Web communities and services on the information ecosystem, by studying Gab and two popular Web archiving services, namely the Wayback Machine and archive.is; and 3) how popular Web communities are exploited by state-sponsored actors for the purpose of spreading disinformation and sowing public discord.
In a nutshell, our analysis reveal that small fringe Web communities like 4chan's /pol/ and The Donald subreddit have a disproportionate influence on mainstream communities such as Twitter with regard to the dissemination of news and image-based memes. We find that Gab acts as the new hub for the alt-right community, while for Web archiving services we find that they are popular on fringe Web communities and that they can be misused by Reddit moderators in order to penalize ad revenue from news sources with conflicting ideology. Finally, when studying state-sponsored actors, we find that they exhibit substantial differences compared to random users, that their tactics change and evolve over time, and that they were particularly influential in spreading news on popular mainstream communities like Twitter and Reddit.