

## VRG23-013 - Towards Trustworthy Recommendation Systems for Online Social Networks

### Abstract

Online social networks have become integral to our daily lives, providing a platform for social interactions and keeping us informed about the latest news. However, concerns have been raised about the filter bubbles that form on these platforms and their potential contribution to societal polarization. Some have attributed this to the use of recommender systems and timeline algorithms that prioritize content with extreme viewpoints. A significant challenge in this area is the feedback loop between these algorithms and human behavior, which only allows us to assess their impact after they have already influenced society. It is important to address this issue to ensure a positive societal impact of these technologies.

This project aims to understand the impact of recommender systems in social networks before they are deployed. This is arguably one of the most important challenges in the area of social network analysis, with the goal of minimizing negative societal effects and contributing to the effective regulation of social media algorithms. The project consists of three work packages. First, we develop novel methods that allow to mathematically understand the interplay of recommender systems and the polarization in online social networks. This will involve extending opinion formation models from sociology to scenarios that capture user interactions on online social networks. We will also study the computational aspects of these models, such as their efficient simulation and arising optimization problems. Second, we study the impact that adversaries can have on online social networks. This is motivated by the fact that recently state-actors used bots and disinformation campaigns to influence elections in opposing countries. Our goal is to understand the power of such adversaries and to characterize how recommender systems must be modified and improved to mitigate such attacks. Third, we consider applications and further extensions of the models we develop. We study how the user opinions can be exploited to aid fine-grained information campaigns, such as by governments who wish to inform their population about new policies or by advertisers who wish to sell their products. This part of the project will have impact on the industry in the mid-term, by enabling better targeted ad campaigns in online social networks.

Scientific disciplines:

Machine learning (50%) | Theoretical computer science (50%)

Keywords:

Recommendation, systems Approximation, algorithms Opinion, dynamics

---

VRG leader:	Stefan Neumann
Institution:	KTH Royal Institute of Technology
Proponent:	Thomas Gärtner
Institution:	TU Wien

---



---

Status: Ongoing (01.02.2024 - 31.01.2031)  
GrantID: 10.47379/VRG23013

---

Further links to the persons involved and to the project can be found under  
<https://wwtf.at/funding/programmes/vrg/VRG23-013/>