Politecnico di Torino | ScuDo Scuola di Dottorato – Doctoral School WHAT YOU ARE, TAKES YOU FAR | UNIVERSITÀ DI TORINO

Subset Selection and Detection Problems on Opinion Dynamics Models

Roberta Raineri

PhD Thesis Summary

In the last decades, the study of Opinion Dynamics in Social Networks has gained significant attention, as it provides essential tools to understand, predict, and possibly influence the dissemination of information and collective decision-making processes. This growing relevance is not limited to academic domains, as confirmed by a McKinsey & Company report which states that "marketing-induced consumer-to-consumer word of mouth generates more than twice the sales of paid advertising", highlighting the critical economic value of understanding how opinions spread and evolve. Different examples further emphasize the tangible consequences of these dynamics. The strategic use of Twitter bots during presidential elections and during the first impeachment of U.S. President Donald Trump illustrates how automated agents can play a central role in manipulating public opinion. Similarly, for companies conducting surveys or opinion polls, it is crucial to identify which individuals should be asked in order to obtain reliable insights on the overall society with minimal cost.

This dissertation investigates mathematical models of Opinion Dynamics on Social Network, focusing on two central problems: a Subset Selection Problem and a Detection Problem. The first problem focuses on identifying the "most informative" agents, meaning the subset of individuals that, if questioned, can most accurately represent the average opinion within the entire network. We formalize a selection criterion based on variance reduction estimation and apply it across various settings built on the widely studied French-DeGroot and Friedkin-Johnsen opinion dynamics models, also accounting for phenomena such as stubborn agents and noisy interactions. Although the problem is NP-Hard, we demonstrate that, under certain reasonable assumptions, the objective function is submodular, thereby enabling the use of efficient greedy algorithms. Additionally, by working within a graph-based framework, we identify connections between the optimal selection and key network properties and topology. The second problem aims at detecting stubborn agents, such as bots or persistent opinion holders, without prior knowledge of the underlying social network. We formulate the stubborn nodes identification and their influence estimation problems as a low-rank approximation problem. We then propose a novel detection method grounded in matrix factorization techniques, specifically interpolative decomposition, that infers latent social structure and identifies stubborn behaviors directly from opinion observations. We determine sufficient conditions on the model parameters to ensure the algorithm's resilience even in the presence of noisy data or finite time observations.

Overall, this thesis bridges theoretical modeling with practical applications of opinion dynamics, offering scalable and robust solutions suited for real-world environments. The proposed methods contribute to advancing the understanding of information spread in modern interconnected societies.