Abstract

Understanding the collective behavior of (groups of) individuals in complex systems, even in scenarios where the individual properties of their components are known, is a challenge. From the point of view of network models, the collective actions of these individuals are often projected on a graph forming a network of *co-interactions*, which we here refer to as a *many-to-many* network. However, the volume and diversity with which these co-interactions are observed in the most varied systems, such as, for example, social media platforms, economic transactions and political behavior in voting systems, impose challenges in the extraction of patterns (structural, contextual and temporal) emerging from collective behavior and that are fundamentally related to a phenomenon under study. Specifically, the frequent presence of a large number of weak and sporadic co-interactions, which, therefore, do not necessarily reflect patterns related to the phenomenon of interest, end up introducing "noise" to the network model. The large amount of noise, in turn, may obfuscate the most fundamental behavior patterns captured by the network model, that is, the patterns that are essentially relevant to the understanding of the phenomenon under investigation. Removing such noise becomes then a key challenge.

Our goal in this dissertation is to investigate the modeling and analysis of collective behavior patterns that emerge in networks formed by co-interactions in different contexts, aiming to extract relevant and fundamental information about a target phenomenon of interest. Specifically, we tackle the extraction of structural, contextual and temporal properties associated with patterns of collective behavior that are fundamentally represented by communities extracted from the network. To this end, we propose a general strategy that addresses the aforementioned challenges. In particular, this strategy includes, as an initial step, the identification and extraction of the network *backbone*, that is, the subset of the edges that are indeed relevant to the target study. The next steps consist of the extraction of communities from this backbone as a manifestation of the existing collective behavior patterns and the characterization of the structural (topological), contextual (related to the phenomenon

of interest) and temporal (dynamic) properties of these communities. Based on this general strategy, we propose specific artifacts for some of the steps that compose it and advance the state-of-the-art, in particular with a new method for backbone extraction, a new temporal node embedding method capable of representing and extracting different temporal patterns of interest from a sequence of networks, and finally a methodology to support the selection and evaluation of backbones from a structural and contextual point of view, considering the most common scenario where there is no ground truth. Furthermore, we explore these artifacts by studying three different phenomena that require different modeling and analysis strategies. Specifically, we investigate: (i) the formation of ideological groups in the Brazilian and U.S. House of Representatives, (ii) online discussions on Instagram in Brazil and Italy, and (iii) information dissemination on WhatsApp. Overall, our results show that the proposed artifacts offer relevant contributions to the field in which this dissertation is inserted.

Keywords: Complex Networks; Collective Behavior; Backbone Extraction; Community Detection.