

Networked Time Series Analysis and Clustering

Mihovil Bartulovic and Tom Magelinski

In network science, a trail is defined as a time-series of categorical states, and can be used to model many interesting phenomena. Trail clustering, then, is the problem of discovering similar trails in a trail dataset. For example, a trail dataset containing trails of url shares by users on social media may be clustered to find similar users. First, we will cover the necessary background in time series and network analysis. Then, we will cover trail analysis basics and trails comparison methods. From there, we will expand on different challenges faced in trail analysis, including heterogeneous state-types, higher-order states, and highly dynamic data. Finally, we will consider the scalability of the approaches discussed. A case study of social media analysis will be used to demonstrate the discussed trail clustering methods in practice.