

**Automatic Segmentation of Dynamic Network Sequences with Node Labels**

**Abstract:**

Given a sequence of snapshots of flu propagating over a population network, can we find a segmentation when the patterns of the disease spread change, possibly due to interventions? In this paper, we study the problem of segmenting graph sequences with labeled nodes. Memes on the Twitter network, diseases over a contact network, movie-cascades over a social network, etc. are all graph sequences with labeled nodes. Most related work on this subject is on plain graphs and hence ignores the label dynamics. Others require fix parameters or feature engineering. We propose SNAPNETS, to automatically find segmentations of such graph sequences, with different characteristics of nodes of each label in adjacent segments. It satisfies all the desired properties (being parameter free, comprehensive and scalable) by leveraging a principled, multi-level, flexible framework which maps the problem to a path optimization problem over a weighted DAG. Also, we develop the parallel framework of SNAPNETS which speeds up its running time. Finally, we propose an extension of SNAPNETS to handle the dynamic graph structures and use it to detect anomalies (and events) in network sequences. Extensive experiments on several diverse real datasets show that it finds cut points matching ground-truth or meaningful external signals and detects anomalies outperforming non-trivial baselines. We also show that the segmentations are easily interpretable, and that SNAPNETS scales near-linearly with the size of the input. Finally, we show how to use SNAPNETS to detect anomaly in a sequence of dynamic networks.