Benet2017 Abstract: Dynamics Based Features for Graph Classification

Leonardo Gutiérrez Gomez, Benjamin Chiêm, Jean-Charles Delvenne[1]

1. Université catholique de Louvain, Louvain-la-Neuve, Belgium

Numerous social, medical, engineering and biological challenges can be framed as graph-based learning tasks. Here, we propose a new feature based approach to network classification. We are interested in predicting the most likely label for a whole network among a set of classes. For example, in chemoinformatics, bioinformatics, one is interested in predicting the toxicity or anti-cancer activity of molecules, seen as graphs linked by chemical bounds. Social networks classification is suitable for many social, marketing and targeting proposes.

To do so, understanding the structural decomposition of networks is crucial. Indeed, community detection algorithms on networks aim to disentangle meaningful patterns that are shared for groups of nodes along the network. In particular, dynamics-based approaches [1] as a general community detection framework, play a key role in our work. Certainly, when a dynamics takes place on a network, it is constrained by the network structure and then could potentially reveal interesting features of the organization of the network. Given this interdependence between dynamics and structure, we are able to extract meaningful features of the network across time scales useful for prediction purposes.

Having a random walk dynamic on a network, we aim to compute assortativity patterns of nodes attributes. This attributes can be structural (i.e pagerank, clustering coefficient) or external metadata. In this dynamical approach, associativity translates in computing covariances between attributes of two extremities of a randomly chosen edge. However, we extend to generalized assortativities by computing this covariances at the two end points of a randomly selected path or t-hop of length $t = 1, 2, 3....$ We aggregate those assortativity patterns for different nodes attributes in a single feature vector that is used as a global fingerprinting of the network.

This method is evaluated empirically on established benchmark network datasets. Our feature vectors are used to learn support vector machines and random forest classifiers for binary and multiclass network classification problem. Results reveal that our dynamic based features are competitive and often outperform state of the art accuracies coming from neural networks [2] and kernel machines [3] approaches.

[1] Jean-Charles Delvenne, Michael T. Schaub, Sophia N. Yaliraki, and Mauricio Barahona. The stability of a graph partition: A dynamics-based framework for community detection. *Modeling and Simulation in Science, Engineering and Technology,* page 221–242, 2013.

[2] Mathias Niepert, Mohamed Ahmed, and Konstantin Kutzkov. Learning convolutional neural networks for graphs. May 2016.

[3] Nino Shervashidze, Pascal Schweitzer, Erik Jan van Leeuwen, Kurt Mehlhorn, and Karsten M. Borgwardt. Weisfeiler-lehman graph kernels. *J. Mach. Learn.* Res., 12:2539– 2561, November 2011.