Graph embedding procedure via dissimilarity mapping for social network comparison

Domenico De Stefano

A key issue in social network analysis is related to the comparison between several observed networks in $\mathscr{G} = \{g_1, \ldots, g_K\}$. Let $g_k = (V_k, E_k, \upsilon_k, \nu_k)$ be the k-th network where V_k is the nodeset, E_k is the edgeset (i.e. a set composed of pairs of nodes, $E_k \subseteq V_k \times V_k$), υ_k is the set of labels of the nodes, ν_k is the set of labels of the edges. Assuming that the process of comparing social networks is equivalent to evaluating the structural similarity between their underlying graphs, to this end we propose a special graph embedding procedure based on the so-called *dissimilarity mapping*.

In general, a graph embedding procedure can be viewed as the extraction of descriptive numerical features from graphs. In this work, for graph embedding we mean a definition of an explicit mapping from the graph (network) domain to a vector space $\Re^K$. The proposed procedure is built on the general dissimilarity mapping proposed in Duin & Pekalska (2005). To this end we need to define: *i*) a general dissimilarity function in the graph domain $d(g_i, g_j)$; *ii*) a mapping $\phi : \mathscr{G} \to \Re^K$.

Embedding of graphs into vector spaces establishes access to a number of exploratory statistical tools. In particular it allows to define a distance among the networks in $\mathscr{G}$, projecting them as points in a multivariate space. We will show how the proposed graph embedding procedure represents a flexible tool in exploratory social network analysis as well as in social network modeling framework.

Reference

E. Pekalska and R.P.W. Duin (2005). *The Dissimilarity Representation for Pattern Recognition, Foundations and Applications*, World Scientific, Singapore.

Domenico De Stefano
DEAMS Dep. of Economics, Business, Mathematics and Statistics
University of Trieste, P.le Europa 1, 34127 Trieste, Italy
E–mail: domenico.destefano@econ.units.it