Harnessing the predictive power of virtual communities

January 30 2012

Scientists have created a new algorithm to detect virtual communities, designed to match the needs of real-life social, biological or information networks detection better than with current attempts. The results of this study by Lovro Šubelj and his colleague Marko Bajec from the University of Ljubljana, Slovenia have just been published in *European Physical Journal B*.

Communities are defined as systems of nodes interacting through links. So-called classical communities are defined by their internal level of link density. By contrast, link-pattern communities – better suited to describe real-world phenomena – are characterised by internal patterns of similar connectedness between their nodes.

The authors have created a model, referred to as a propagation-based algorithm, that can extract both link-density and link-pattern communities without any prior knowledge of the number of communities, unlike previous attempts at community detection. They first validated their algorithm on several synthetic benchmark networks and with random networks. The researchers subsequently tested it on ten real-life networks including social (members of a karate club), information (peer-to-peer file sharing) and biological (protein-protein interactions of a yeast) networks. By this, it was found that the proposed algorithm detected the real-life communities more accurately than existing state-of-the-art algorithms.

They concluded that real-life networks appear to be composed of link-

pattern communities that are interwoven and overlap with classical link-density communities. Further work could focus on creating a generic model to understand the conditions, such as the low level of clustering, for link-pattern communities to emerge, compared to link-density communities. The model could also help to explain why such link-pattern communities call the existing interpretation of small-world phenomena (six degrees of separation between nodes) into question.

Applications include the prediction of future friendships in online social networks, analysis of interactions in biological systems that are hard to observe otherwise, and detection of duplicated code in software systems.