Social networking analysis helps identify cancer biomarkers

June 4 2015

The advent of online social networks has led to the rapid development of tools for understanding the interactions between members of the network, their activity, the connections, the hubs and nodes. But, any relationships between lots of entities, whether users of Facebook and Twitter, bees in a colony, birds in a flock, or the genes and proteins in our bodies can be analyzed with the same tools.

Now, research published in *International Journal of Data Mining and Bioinformatics* shows how social network analysis can be used to understand and identify the biomarkers in our bodies for diseases, including different types of cancer.
Tansel Özyer, Serkan Ucer and Taylan Iyidogan of the Department of Computer Engineering, at TOBB University, in Ankara, Turkey, explain how the detection of disease biomarkers in general and cancer biomarkers in particular has become an important task in medical research and diagnostics. They have now used the tools of Social Network Analysis (SNA) to help them unravel the connections and identify the biomarkers present in patient genomic microarray data.

By analogy with a social network, the team views genes as actors or members of the social network and similarities between different genes are considered to be the connections between these actors. Genomic databases can be vast, given that the human genome comprises some 20000 genes, and so such an approach can, they suggest, dramatically decrease the number of features that must be analyzed to find useful biomarkers. Once identified and understood, such biomarkers can then be tested for in screening programs for people at risk of a given disease or for diagnosis should the present with particular symptoms.

The team has demonstrated proof of principle with three types of cancer: lymphoma, colon cancer and leukemia. "We showed how our approach is capable of effectively detecting cancer biomarkers out of high-dimensional genomic data," the team reports. "We combined clustering and classification into the developed framework to help in detecting the links between the various genes within the model and to validate the outcome, respectively." The next step will be to optimize the approach and to extend it to protein-protein interactions, protein-gene interactions, disease-protein interactions, disease-drug interactions all with a view to improving diagnostics and tailoring therapy for the individual patient based on the outcomes of their personal biological network analysis.
