Extract Reference Co-Occurrence (Bibliographic Coupling) Network

Description
Extracts a bibliographic coupling network from a network.

Pros & Cons
Can return sensible similarity networks based on citation, but may not make sense for other kinds of networks.

Applications
Extracting similarity between nodes. Good for then feeding into force directed algorithms, especially DrL.

Implementation Details
The algorithm in Network Workbench will take in a directed, undirected, or hypergraph network. The undirected edges will be transformed into two directed edges. All edge weights are ignored. A new network is returned with the same nodes (and their attributes) with weighted undirected edges where each edge corresponds to a single similarity score (ranging from 0.0 to 1.0) between two nodes. The similarity score uses the following formula: similarity = sharedCitations / sqrt(citationCount1*citationCount2) where sharedCitations is the number of citations which they both cite, citationCount1 is the total number of citations the first node cited, and citationCount2 is the total number of citations the second node cited. This algorithm takes a single parameter, which states the maximum number of top bibliographic coupling scores to keep. If 0 is specified, then every possible bibliographic coupling score will be kept.

Usage Hints
Any network can be given to this algorithm, though it makes most sense on directed networks, especially citation networks.

Acknowledgments
Todd Holloway implemented the first version in Perl. Bruce Herr later reimplemented the algorithm in C++ for scalability and speed. The C++ version is in Network Workbench.

References
Garfield, E. "From Bibliographic Coupling to Co-Citation Analysis via Algorithmic Historio-Bibliography"

Links
- Source Code

* See Also*