Review of: "A method to reduce false positives in a patent query"

Chung-Huei Kuan[1]

1 National Taiwan University of Science and Technology

Potential competing interests: No potential competing interests to declare.

1. The title's use of phrase "reduce false positives" is not quite accurate. According to the manuscript, the authors introduces a method to "identify which patents are really out of scope" (i.e., the false positives) and then "exclude" them from the query result "without risking removing true positives."

2. Keyword search indeed may often produce noisy query result, but this phenomenon depends on the fields used for keyword search. If it is the Title or Abstract field that is sought, the false positives would be fewer than when the Specification field is used.

3. The classification network is better known as a co-occurrence network in the scientometric community.

4. Some details of the classification system are not mentioned in the manuscript but may affect the validity of the proposed method. The IPCs assigned to a patent include at least one main classification (usually shown in bold face) and some optional secondary classifications (usually shown in normal face). So, when building the classification network, whether only the main IPCs or both are considered is not clear in the manuscript.

5. A very important aspect of the classification network is also not mentioned in the manuscript, which is the strength of connection between two nodes (i.e., the frequency of IPC co-occurrence). The connection strength would be helpful, if not vital, in identifying true communities of IPCs.

6. I also have doubt about that "some classifications can play the role of gatekeeper." In a social network, persons situated at the junctions of two or more people communities bridge these communities together. But IPCs are not humans. The significance of IPCs at the junction of two more IPC communities requires further exploration.