**Intermediacy of publications in citation networks**

**Abstract**

Citation networks of scientific publications offer fundamental insights into the structure and development of scientific knowledge. We propose a new indicator, referred to as intermediacy, for tracing the evolution of scientific knowledge over time. Given two publications, an older and a more recent one, intermediacy identifies publications that appear to play a major role in the intellectual development from the older to the more recent publication. The identified publications are important in connecting the older and the more recent publication in the citation network. After providing a formal definition of intermediacy, we study its mathematical properties. We then present two empirical analyses, one in which intermediacy is used to trace intellectual developments at the interface between the community detection and the scientometric literature and one in which intermediacy is applied to the literature on peer review. We show both mathematically and empirically how intermediacy differs from main path analysis, which is the most popular approach for analyzing intellectual developments in citation networks. Main path analysis tends to favor longer paths over shorter ones, whereas intermediacy has the opposite tendency. Compared to main path analysis, our conclusion is that intermediacy offers a more principled approach for tracing the evolution of scientific knowledge.

**Discussion**

Citation networks provide valuable information for analyzing how scientific knowledge evolves over time. The most commonly used approach for tracing the evolution of scientific knowledge in citation networks is main path analysis (Hummon & Doreian, 1989). However, the idea of a main path is relatively poorly understood. The algorithmic definition of a main path is clear, but the underlying conceptual motivation remains somewhat obscure. As we have shown in this paper, main path analysis has the tendency to favor longer paths over shorter ones. We consider this to be a counterintuitive property that lacks a convincing justification.

The notion of intermediacy, introduced in this paper, offers an alternative to main path analysis. It provides a principled approach for identifying publications that appear to play a major role in the intellectual development from an older to a more recent publication. The older publication is referred to as the target, while the more recent publication is called the source. Publications with a high intermediacy are important in connecting the source and the target publication in a citation network. As we have shown, intermediacy has two intuitively reasonable properties, referred to as path addition and path contraction. Because of the path contraction property, intermediacy tends to favor shorter paths over longer ones. This is a fundamental difference with main path analysis. Intermediacy also has a free parameter that can be used to fine-tune its behavior. The parameter enables interpolation between two extremes. In one extreme, intermediacy identifies publications located on a shortest path between the source and the target publication. In the other extreme, it identifies publications located on a large number of edge independent source-target paths.

We have also tested intermediacy empirically. Two empirical analyses have been presented, one in which intermediacy was used to trace intellectual developments at the interface between the community detection and the scientometric literature and one in which intermediacy was applied to the literature on peer review. The first analysis has shown that intermediacy yields results that appear sensible from the point of view of a domain expert. The second analysis has provided empirical confirmation of the mathematical differences between intermediacy and main path analysis.

There are various options for further research. First of all, a more extensive mathematical analysis of intermediacy could be carried out, possibly resulting in an axiomatic foundation for intermediacy. The ideas underlying intermediacy may also be used to develop other types of network indicators, such as an indicator of the connectedness of two nodes in a network. In empirical analyses, intermediacy could be applied not only in citation networks of scientific publications, but for instance also in patent citation networks or even in completely different types of networks, such as ???.