

SEMANTIC ANALYSIS OF FINE-SCALE STRUCTURES IN THE WIKIPEDIA NETWORK

MILES KENYON

ABSTRACT. The analysis of the geometries of complex networks has been an exceptionally fruitful area of research in recent years. The online encyclopedia Wikipedia is a particular example, where an understanding of the underlying structure may aid us in areas as far-ranging as search algorithms to epistemology. This paper presents a method of automated cluster detection in the Wikipedia network using spectral techniques. In particular, we sample the site's math-related articles to generate a graph where the individual articles are considered nodes, and the hyperlinks between them are the edges. We are able to embed this graph in Euclidean space and cluster based on inter-node distances. The very high semantic relatedness among articles within any cluster begins to reveal the existence of a high-dimensional "semantic space."