Modeling Semantics and Structure of Discussion Threads

The dynamic nature of discussion threads poses interesting and challenging problems for computer scientists. Although techniques such as semantic models or structural models have been shown to be useful in a number of areas, they are inefficient in understanding discussion threads due to the temporal dependence among posts in a discussion thread. Such dependence causes that semantics and structure coupled with each other in discussion threads. In this paper, we propose a sparse coding-based model named SMSS to Simultaneously Model Semantic and Structure of discussion threads.

A discussion thread has the following four characteristics:

(a) A discussion thread has several topics;
(b) An individual post is related to a few topics;
(c) A post is related to its previous posts;
(d) The reply relations are sparse.

The characteristics (a) and (b) are related to the SEMANTIC part while (c) and (d) are related to the STRUCTURE part.