

A Polynomial Time Approximation Scheme for k-Consensus Clustering

Date Tuesday, January 26

Time 3 pm

Location 303 Mudd

Abstract: We introduce a polynomial time approximation scheme for the metric Correlation Clustering problem, when the number of clusters returned is bounded (by k). Consensus Clustering is a fundamental aggregation problem, with considerable application, and it is analysed here as a metric variant of the Correlation Clustering problem. The PTAS exploits a connection between Correlation Clustering and the k -cut problems. This requires the introduction of a new rebalancing technique, based on minimum cost perfect matchings, to provide clusters of the required sizes.

Both Consensus Clustering and Correlation Clustering have been the focus of considerable recent study. There is an existing dichotomy between the k -restricted Correlation Clustering problems and the unrestricted versions. The former, in general, admit a PTAS, whereas the latter are, in general, APX-hard. This result extends the dichotomy to the metric case, responding to the result that Consensus Clustering is APX-hard to approximate. Joint work with Tom Coleman.