Aggregating Discrete Information from Inconsistent Sources

Nir Ailon

School of Mathematics, Institute for Advanced Study, Princeton NJ, USA
nailon@math.ias.edu
www.math.ias.edu/~nailon

Abstract

An output of a meta-search engine is to be computed from outputs of several search engines. Microarray data is to be clustered based on outputs of several clustering heuristics. A ranking of hotels is to be compiled from qualitative information ("number of stars" rating) collected from several hotel critics.

More generally: How do we combine possibly contradicting discrete information from different sources into one consistent output? These questions, lying in the intersection of combinatorial optimization and social choice theory, have enjoyed a recent flurry of activity in theoretical and experimental computer science research. I will survey recent work and present new and improved approximation algorithms.
Based on joint work with Moses Charikar and Alantha Newman.

1   Introduction

The talk will be mostly based on the papers [1, 2]. Other related papers are [3, 4, 5, 6, 7, 8].

References

[1] N. Ailon. Aggregation of partial rankings, p-ratings and top-m lists. In *Proceedings of the 18th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2007.

[2] N. Ailon, M. Charikar, and A. Newman. Aggregating inconsistent information: ranking and clustering. In *Proceedings of the 37th Annual ACM Symposium on Theory of Computing (STOC)*, pages 684–693, 2005.

[3] J. Aslam and M. Montague. Condorcet fusion for improved retrieval. In *Proceedings of the 11th International Conference on Information and Knowledge Management*, pages 538–548, 2002.

[4] D. Coppersmith, L. Fleischer, and A. Rudra. Ordering by weighted number of wins gives a good ranking for weighted tournamnets. In *Proceedings of the 17th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2006.

[5] C. Dwork, R. Kumar, M. Naor, and D. Sivakumar. Rank aggregation revisited. *Manuscript*, 2001.

[6] R. Fagin, R. Kumar, M. Mahdian, D. Sivakumar, and E. Vee. Comparing and aggregating rankings with ties. In *Proceedings of the ACM Symposium on Principles of Database Systems (PODS)*, pages 47–58, 2004.

[7] V. Filkov and S. Skiena. Integrating microarray data by consensus clustering. In *Proceedings of International Conference on Tools with Artificial Intelligence (ICTAI)*, pages 418–425, Sacramento, 2003.

[8] A. Gionis, H. Mannila, and P. Tsaparas. Clustering aggregation. In *Proceedings of the 21st International Conference on Data Engineering (ICDE)*, Tokyo, 2005.