Trace complexity

Flavio Chierichetti

Dipartimento di Informatica
Sapienza University of Rome

Abstract. Prediction tasks in machine learning usually require deducing a latent variable, or structure, from observed traces of activity — sometimes, these tasks can be carried out with a significant precision and statistical significance, while sometimes getting any useful prediction requires an unrealistically large number of traces.

In this talk, we will study the trace complexity of (that is, *the number of traces needed for carrying out*) two prediction tasks in social networks: the network inference problem and the number of signers problem.

The first problem [1] consists of reconstructing the edge set of a network given traces representing the chronology of infection times as epidemics spread through the network. The second problem's [2] goal is to guess the unknown number of signers of some email-based petitions, when only a small subset of the emails that circulated is available.

These two examples will allow us to make some general remarks about social networks' prediction tasks.

References

1. B. D. Abrahao, F. Chierichetti, R. Kleinberg, and A. Panconesi. Trace complexity of network inference. In I. S. Dhillon, Y. Koren, R. Ghani, T. E. Senator, P. Bradley, R. Parekh, J. He, R. L. Grossman, and R. Uthurusamy, editors, *KDD*, pages 491–499. ACM, 2013.
2. F. Chierichetti, J. M. Kleinberg, and D. Liben-Nowell. Reconstructing patterns of information diffusion from incomplete observations. In J. Shawe-Taylor, R. S. Zemel, P. L. Bartlett, F. C. N. Pereira, and K. Q. Weinberger, editors, *NIPS*, pages 792–800, 2011.