Zanghi, Hugo; Picard, Franck; Miele, Vincent; Ambroise, Christophe

Strategies for online inference of model-based clustering in large and growing networks.

(English) | Zbl 1194.62096

Summary: We adapt online estimation strategies to perform model-based clustering on large networks. Our work focuses on two algorithms, the first based on the stochastic approximation EM (SAEM) algorithm, and the second on variational methods. These two strategies are compared with existing approaches on simulated and real data. We use the method to decipher the connexion structure of the political web-sphere during the US political campaign in 2008. We show that our online EM-based algorithms offer a good trade-off between precision and speed, when estimating parameters for mixture distributions in the context of random graphs.

MSC:
- 62L20 Stochastic approximation
- 62P99 Applications of statistics
- 91F10 History, political science
- 65C60 Computational problems in statistics (MSC2010)
- 05C80 Random graphs (graph-theoretic aspects)

Keywords:
- graph clustering; EM algorithms; online strategies; web graph structure analysis

Software:
- Cytoscape; Mixnet

Full Text: DOI

References:


