

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

M. Catalano, A. Fasano, M. Giordano, and G. Rebaudo's contribution to the Discussion of 'Root and community inference on the latent growth process of a network' by Crane and Xu

This is a pre print version of the following article:

Original Citation:

Availability:

This version is available <http://hdl.handle.net/2318/1982451> since 2024-06-06T13:45:25Z

Published version:

DOI:10.1093/jrsssb/qkae051

Terms of use:

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)

M. Catalano, A. Fasano, M. Giordano, and G. Rebaudo's contribution to the Discussion of 'Root and community inference on the latent growth process of a network' by Crane and Xu

Marta Catalano

Luiss University, Roma, Italy. E-Mail: matalano@luiss.it

Augusto Fasano

Catholic University, Milano, Italy.

Matteo Giordano

University of Torino, Torino, Italy.

Giovanni Rebaudo

University of Torino, Torino, Italy.

We congratulate the authors for their methodological and theoretical contribution to the statistical literature on networks.

A natural extension of the proposed PAPER model is included, with K communities growing simultaneously and where new nodes are either assigned to an existing community or elected as a new root. The employed assignment rule is of Pólya-urn type, which leads to a logarithmic growth of the number of communities (Korwar and Hollander, 1972) and is known to coincide with the predictive scheme of exchangeable sequences associated to the Dirichlet process. The probability of creating a new community is then independent of the number of past ones, which is a distinctive feature of the Dirichlet process within the class of Gibbs-type priors (De Blasi et al., 2015). An interesting direction would be to allow for more flexible predictive schemes that ensure alternative asymptotics, ranging from power-law behaviours (via the Pitman-Yor process (Pitman, 2006) or normalized generalized gamma completely random measures (Lijoi et al., 2007)) to slower than logarithmic growth (via the single-group hierarchical Dirichlet process (Camerlenghi et al., 2018)).

An important theoretical aspect is the number of communities. While the authors provide an empirical investigation, future research could tackle the question of posterior consistency for the number of communities, along the lines of the existing results for stochastic block models (Geng et al., 2019) and in the growing literature in Bayesian nonparametric mixture models (Nobile, 1994; Miller and Harrison, 2013; Ascolani et al., 2023).

Turning to applications, the proposed model lends itself to some natural generalizations suggested by popular epidemiological models, like SIR dynamics, where

at each instant the infectious nodes can transmit the disease to their susceptible neighbours with some probability, resulting in multiple new infectious individuals at the next time. Equating new infections in SIR dynamics to added nodes in the PAPER model, a useful extension would be obtained by allowing the addition of multiple nodes at each step: for instance, if computationally feasible, a fraction of the existing nodes at that time, representing an average contact rate, or also a random number, e.g. driven by a non-homogeneous Poisson process. The SIR analogy further suggests extensions where nodes are active (i.e. accepting newly introduced nodes as neighbours) only for a limited time, representing the period during which an infectious agent can spread the disease.

Once more, we commend the authors for an outstanding paper.

References

- [1] Ascolani, F., Lijoi, A., Rebaudo, G., Zanella, G. (2023). Clustering consistency with Dirichlet process mixtures. *Biometrika*, 110, 551-558.
- [2] Camerlenghi, F., Lijoi, A., Prünster, I. (2018). Bayesian nonparametric inference beyond the Gibbs-type framework. *Scand J Statist.*, 45, 1062–1091.
- [3] Korwar, R. M., Hollander, M. (1973). Contributions to the theory of Dirichlet processes. *Ann. Probab.*, 1, 705–711.
- [4] Lijoi, A., Mena, R.H., Prünster, I. (2007). Bayesian nonparametric estimation of the probability of discovering a new species. *Biometrika*, 94, 769–786.
- [5] Geng, J., Bhattacharya, A., Pati, D. (2019). Probabilistic community detection with unknown number of communities. *J. Am. Stat. Assoc.*, 114, 893-905.
- [6] Miller, J. W., Harrison, M. T. (2013). A simple example of Dirichlet process mixture inconsistency for the number of components. In *Adv. Neural Inf. Process. Syst.*, 26, 199-206.
- [7] Nobile, A. (1994). Bayesian Analysis of Finite Mixture Distributions. Ph.D. thesis, Carnegie Mellon Univ.
- [8] Pitman, J. (2006). Combinatorial stochastic processes. *Lect. Notes Math.*, 1875.