

Cambridge Elements

Elements in the Structure and Dynamics of Complex Networks

edited by

Guido Caldarelli

Ca' Foscari University of Venice

DESCRIPTIVE VS. INFERENCE COMMUNITY DETECTION IN NETWORKS

Pitfalls, Myths, and Half-Truths

Tiago P. Peixoto

Central European University, Vienna



CAMBRIDGE
UNIVERSITY PRESS

Cambridge University Press & Assessment
978-1-009-11300-7 — Descriptive vs. Inferential Community Detection in Networks
Pitfalls, Myths and Half-Truths
Tiago P. Peixoto
Frontmatter
[More Information](#)



Shaftesbury Road, Cambridge CB2 8EA, United Kingdom
One Liberty Plaza, 20th Floor, New York, NY 10006, USA
477 Williamstown Road, Port Melbourne, VIC 3207, Australia
314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre,
New Delhi – 110025, India
103 Penang Road, #05–06/07, Visioncrest Commercial, Singapore 238467

Cambridge University Press is part of Cambridge University Press & Assessment,
a department of the University of Cambridge.

We share the University's mission to contribute to society through the pursuit of
education, learning and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9781009113007
DOI: 10.1017/9781009118897

© Tiago de Paula Peixoto 2023

This work is in copyright. It is subject to statutory exceptions and to the provisions
of relevant licensing agreements; with the exception of the Creative Commons version
the link for which is provided below, no reproduction of any part of this work may
take place without the written permission of Cambridge University Press.

An online version of this work is published at <http://dx.doi.org/10.1017/9781009118897>
under a Creative Commons Open Access license CC-BY-NC-ND 4.0 which permits
re-use, distribution and reproduction in any medium for non-commercial purposes
providing appropriate credit to the original work is given. You may not
distribute derivative works without permission. To view a copy of this license, visit
<https://creativecommons.org/licenses/by-nc-nd/4.0>

All versions of this work may contain content reproduced under license from third
parties. Permission to reproduce this third-party content must be obtained from these
third-parties directly. When citing this work, please include a reference to the
DOI 10.1017/9781009118897

First published 2023

A catalogue record for this publication is available from the British Library.

ISBN 978-1-009-11300-7 Paperback
ISSN 2516-5763 (online)
ISSN 2516-5755 (print)

Cambridge University Press & Assessment has no responsibility for the persistence
or accuracy of URLs for external or third-party internet websites referred to in this
publication and does not guarantee that any content on such websites is, or will
remain, accurate or appropriate.

Descriptive vs. Inferential Community Detection in Networks

Pitfalls, Myths, and Half-Truths

Elements in the Structure and Dynamics of Complex Networks

DOI: 10.1017/9781009118897
First published online: May 2023

Tiago P. Peixoto
Central European University, Vienna

Author for correspondence: Tiago P. Peixoto, peixotot@ceu.edu

Abstract: Community detection is one of the most important methodological fields of network science, and one which has attracted a significant amount of attention over the past decades. This area deals with the automated division of a network into fundamental building blocks, with the objective of providing a summary of its large-scale structure. Despite the importance and widespread adoption of community detection there is a noticeable gap between what is arguably the state-of-the-art and the methods that are actually used in practice in a variety of fields. This Element attempts to address this discrepancy by dividing existing methods according to whether they have a “descriptive” or an “inferential” goal. While descriptive methods find patterns in networks based on context-dependent notions of community structure, inferential methods articulate a precise generative model, and attempt to fit it to data. In this way, they are able to provide insights into the mechanisms of network formation, and separate structure from randomness in a manner supported by statistical evidence. We review how employing descriptive methods with inferential aims is riddled with pitfalls and misleading answers, and thus should be in general avoided. We argue that inferential methods are more typically aligned with clearer scientific questions, yield more robust results, and should be in many cases preferred. We attempt to dispel some myths and half-truths often believed when community detection is employed in practice, in an effort to improve both the use of such methods as well as the interpretation of their results. This title is also available as Open Access on Cambridge Core.

Keywords: community detection, stochastic block models, network clustering, statistical inference, Bayesian inference

© Tiago de Paula Peixoto 2023

ISBNs: 9781009113007 (PB), 9781009118897 (OC)
ISSNs: 2516-5763 (online), 2516-5755 (print)

Contents

1	Introduction	1
2	Descriptive vs. inferential community detection	2
3	Modularity maximization considered harmful	22
4	Myths, pitfalls, and half-truths	29
5	Conclusion	65
	References	66