Cambridge Elements

The Structure and Dynamics of Complex Networks

# Modularity and Dynamics on Complex Networks

Renaud Lambiotte and Michael T. Schaub

ISSN 2516-5763 (online)
ISSN 2516-5755 (print)

Cambridge Elements

Elements in the Structure and Dynamics of Complex Networks
edited by
Guido Caldarelli
*Ca' Foscari University of Venice*

# MODULARITY AND DYNAMICS ON COMPLEX NETWORKS

Renaud Lambiotte
*University of Oxford*

Michael T. Schaub
*RWTH Aachen University*

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, 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 the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning, and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9781108733533
DOI: 10.1017/9781108774116



First published 2021

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

ISBN 978-1-108-73353-3 Paperback
ISSN 2516-5763 (online)
ISSN 2516-5755 (print)

# Modularity and Dynamics on Complex Networks

Elements in the Structure and Dynamics of Complex Networks

DOI: 10.1017/9781108774116
First published online: December 2021

Renaud Lambiotte
*University of Oxford*

Michael T. Schaub
*RWTH Aachen University*

**Authors for correspondence:** Renaud Lambiotte,
renaud.lambiotte@some.ox.ac.uk;
Michael T. Schaub, schaub@cs.rwth-aachen.de

**Abstract:** Complex networks are typically not homogeneous, as they tend to display an array of structures at different scales. A feature that has attracted a lot of research is their modular organisation (i.e., networks may often be considered as being composed of certain building blocks, or modules). In this Element, the authors discuss a number of ways in which this idea of modularity can be conceptualised, focusing specifically on the interplay between modular network structure and dynamics taking place on a network. They discuss, in particular, how modular structure and symmetries may impact on network dynamics and, vice versa, how observations of such dynamics may be used to infer the modular structure. They also revisit several other notions of modularity that have been proposed for complex networks and show how these can be related to and interpreted from the point of view of dynamical processes on networks.

**Keywords:** modularity, networks, time scale, dynamics, block models



ISBNs: 9781108733533 (PB), 9781108774116 (OC)
ISSNs: 2516-5763 (online), 2516-5755 (print)

# Contents

1 Introduction 1

2 Background Material 6

3 Modularity, Community Detection, and Clustering in Networks 22

4 Timescale Separation and Dynamics on Modular Networks 33

5 Symmetries and Dynamics on Modular Networks 44

6 Dynamical Methods for Assortative Communities 54

7 Dynamical Methods for Disassortative Communities and General Block Structures 66

8 Perspectives 77

References 80

## 1 Introduction

Over the last 20 years, networks and graphs have become a near-ubiquitous modelling framework for complex systems. By representing the entities of a system as nodes in a graph and encoding relationships between these nodes as edges, we can abstract systems from a variety of domains with the same mathematical language, including biological, social, and technical systems (Newman, 2018a). Network abstractions are often used with one of the two following perspectives in mind. First, graphs and networks provide a natural way to describe relational data (i.e., datasets corresponding to 'interactions' or correlations between pairs of entities). A prototypical example here is online social networks, in which we measure interactions between actors and can derive a network representation of the social system based on these measurements. We may then try to explain certain properties of the social system by modelling and analysing the network (e.g., by searching for interesting connectivity patterns between the nodes). Second, networks are often used to describe distributed dynamical systems, including prototypical examples such as power grids, traffic networks, or various other kinds of supply or distribution networks. The edges of the network are in this context not the primary object of our modelling. Rather, we are interested in understanding a dynamical process that takes place on this network. More specifically, we often aim to comprehend how the network structure shapes this dynamics (e.g., in terms of its long-term behaviour). In reality, of course, both these perspectives are simplifications in that for many real systems, there are typically uncertainty and dynamics associated to both node and edge variables which make up the network: think, for instance, of a rumour spreading on a social network, where both node variables (the infection state) and the network edges (who is in contact with whom) will be highly dynamic and uncertain. We may not know the exact status of each individual; moreover, edges will change dynamically, and their presence or absence may not be determined accurately.

No matter under what perspective we are interested in networks, it should be intuitively clear that networks with some kind of 'modular structure' may be of interest to us. For now, consider modular structure simply in terms of a network made of dense clusters that are loosely connected with each other. From the perspective of relational data, modular structure may be indicative of a hidden cause that binds a set of nodes together: this corresponds to the idea of homophily in social networks (McPherson, Smith-Lovin, & Cook, 2001), which can lead to the formation of communities of tightly knit actors. From the perspective of dynamics, it is often impractical to keep a full description of a dynamical process on a network when the number of dynamical units is too large. In many cases, it is unclear whether such finely detailed data is necessary to understand the global phenomena of interest, as relevant observables can often be obtained by aggregating microscopic information into macroscopic information (i.e., aggregating information over many nodes). This kind of dimensionality reduction of the dynamics is particularly successful if there exist roughly homogeneously connected blocks of nodes (i.e., a modular network structure (Simon & Ando, 1961)).

As the title indicates, this Element will primarily adopt a dynamical perspective on network analysis. Accordingly, our core objective will be to explore the relations between modular structure and dynamics on networks; but we will also explain how certain aspects of the analysis of relational data can be interpreted from this lens. However, an exhaustive exposition of methods to characterise and uncover modules (also called blocks, clusters, or communities) in networks will not be the main focus of our exposition. We refer the interested reader to the extensive literature on this topic for more detailed treatments; see, for example, Doreian, Batagelj, and Ferligoj (2020); Fortunato and Hric (2016); Newman (2018a); Schaeffer (2007).

## Network Dynamics and Network Structure

It is well-known that there exists a two-way relationship between dynamics on graphs and the underlying graph structure (Schaub et al., 2019b). On the one hand, the structure of a network affects dynamical processes taking place on it (Porter & Gleeson, 2016). In the simplest case of a linear dynamical system, this relationship derives from the spectral properties of a matrix encoding the graph, most often the adjacency matrix or the graph Laplacian. On the other hand, dynamics can help reveal salient features of a graph. This includes the identification of central nodes or the detection of modules in large-scale networks.

To illustrate this two-fold relation between network structure and network dynamics on an intuitive level, let us consider random walks on networks. Random walks are often used as a model for diffusion, and there is much research on the impact of network structure on different properties of random-walk dynamics (Masuda, Porter, & Lambiotte, 2017). In particular, degree heterogeneity, finite size effects and modular structure can all make diffusion processes on networks quantitatively and even qualitatively different from diffusion on regular or infinite lattices. At the same time, random walks are key to many algorithms that uncover various types of structural properties of networks. For example, the classical PageRank method for identifying important nodes may be interpreted in terms of a random walk(Gleich, 2015). Indeed, as we will discuss, several algorithms use trajectories of dynamical processes such as random walks to reveal mesoscale network patterns.

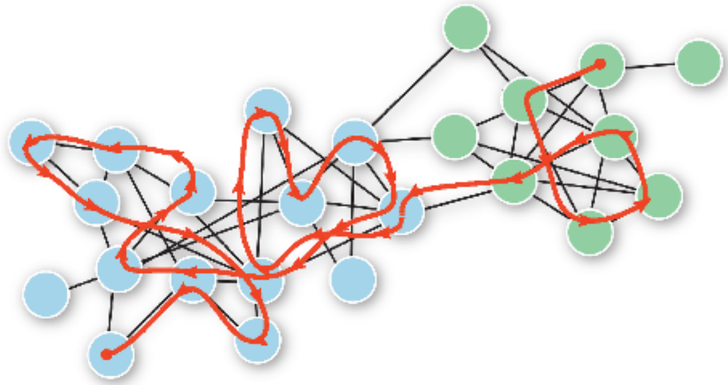

**Figure 1 Modularity and Dynamics on Networks.** Our main ambition is to understand relationships between modular structure of a network, here highlighted in different node colours, and a dynamics taking place on it, here illustrated with the red trajectory on the network. The two complementary questions at the core of this Element are: (1) How does the modular structure of a network affect dynamics? (2) How can dynamics help us characterise and uncover the modular structure of a network?

## Outline of This Element

In this Element, we try to provide a basic overview of the topic of modularity and dynamics on complex networks. Our exposition is structured as follows. In Section 2, we first discuss some background material in Network Science and then review classical notions of modular structure in networks in Section 3. In Sections 4 and 5, we discuss the interplay between dynamics and network structure in terms of timescale separation and symmetries, and how these aspects can be used to reduce the complexity of the description of network dynamics. In Section 6, we then explain how we can detect so-called assortative community structure, primarily based on the notion of timescale separation. Section 7 then discusses the definition and detection of more general (dynamical) block structure, leveraging ideas from linear systems theory and symmetry reduction. In Section 8, we conclude with a short discussion on avenues for future work and additional perspectives.

### Why Are Networks Modular?

For many years, researchers have been fascinated by the prevalence of modularity in systems as different as the World Wide Web, foodwebs, and brain networks, raising the question: are there universal mechanisms driving the evolution of networks toward a modular architecture? Among the many mechanisms that have been proposed (Meunier, Lambiotte, & Bullmore, 2010), the following profound idea of Herbert Simon (1962) stands out by its elegance. 'Nearly-decomposable' systems, as Simon calls them, allow faster adaptation or evolution of the system in response to changing environmental conditions. In Simon's view, modules represent stable building blocks that ensure the robustness of a system evolving under changing or competitive selection criteria. To illustrate this idea, Simon wrote an intuitive parable about two watchmakers, called 'Hora' and 'Tempus' (Simon, 1962):

> There once were two watchmakers, named Hora and Tempus, who manufactured very fine watches. Both of them were highly regarded, and the phones in their workshops rang frequently – new customers were constantly calling them. However, Hora prospered, while Tempus became poorer and poorer and finally lost his shop. What was the reason?
>
> The watches the men made consisted of about 1,000 parts each. Tempus had so constructed his that if he had one partly assembled and had to put it down – to answer the phone say – it immediately fell to pieces and had to be reassembled from the elements. The better the customers liked his watches, the more they phoned him, the more difficult it became for him to find enough uninterrupted time to finish a watch.
>
> The watches that Hora made were no less complex than those of Tempus. But he had designed them so that he could put together subassemblies of about ten elements each. Ten of these subassemblies, again, could be put together into a larger subassembly; and a system of ten of the latter subassemblies constituted the whole watch. Hence, when Hora had to put down a partly assembled watch in order to answer the phone, he lost only a small part of his work, and he assembled his watches in only a fraction of the man-hours it took Tempus.

This story illustrates in simple terms the potential evolutionary advantage that a modular system structure may have, and provides an argument for the ubiquity of modularity in a broad range of natural and social systems.[a] In the following, we will not dwell on why there is modular structure in the network, but rather focus on the question: how does the modularity of a network affect its behaviour and, in particular, its dynamical properties?

[a] One needs to be careful with such statements. Simon himself cautioned that many systems lack conclusive statistical evidence for being modular and may only be perceived as modular due to confirmation bias. However, the statement that many networks are modular has been validated on a large corpus of network datasets by now. See, for example, Fortunato (2010); Ghasemian, Hosseinmardi, and Clauset (2019); Leskovec et al. (2008).

**Table 1** Notation

| Symbol | Description |
|---|---|
| $n \in \mathbb{N}$ | number of nodes |
| $m \in \mathbb{R}$ | total weight of edges (number of edges for unweighted networks) |
| $\mathcal{C} \in \mathbb{N}$ | number of modules / communities |
| $\mathcal{V} = \{1, \dots, n\}$ | set of nodes / vertices |
| $i, j, \ell \in \{1, \dots, n\}$ | indices for nodes |
| $\mathcal{P} = \{\mathcal{A}_1, \dots, \mathcal{A}_\mathcal{C}\}$ | partition of the nodes into communities $\mathcal{A}_\alpha$ |
| $\mathcal{A}_\alpha$ | set of nodes within the $\alpha$th community |
| $\alpha, \beta \in \{1, \dots, \mathcal{C}\}$ | indices for communities |
| $k_i \in \mathbb{R}$ | weighted degree (strength) of node $i$ |
| $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ | weighted adjacency matrix of a network |
| $\boldsymbol{K}(\boldsymbol{A}) = \text{diag}(\boldsymbol{A}\mathbf{1}) \in \mathbb{R}^{n \times n}$ | weighted degree matrix of a network |
| $\boldsymbol{L}(\boldsymbol{A}) = \boldsymbol{K} - \boldsymbol{A}$ | combinatorial Laplacian matrix |
| $\boldsymbol{\mathcal{L}}(\boldsymbol{A}) = \boldsymbol{I} - \boldsymbol{K}^{-1/2}\boldsymbol{A}\boldsymbol{K}^{-1/2}$ | normalised Laplacian matrix |
| $\boldsymbol{L}_{\text{rw}}(\boldsymbol{A}) = \boldsymbol{I} - \boldsymbol{K}^{-1}\boldsymbol{A}$ | random-walk Laplacian matrix |
| $\boldsymbol{H} \in \{0, 1\}^{n \times \mathcal{C}}$ | partition indicator matrix with entries $H_{i\alpha} = 1$, if node $i$ is in the $\alpha$th community ($\mathcal{A}_\alpha$), and $H_{i\alpha} = 0$ otherwise |
| $\boldsymbol{h}_\alpha \in \{0, 1\}^n$ | Indicator vector of the $\alpha$th community (i.e., $\alpha$th column of $\boldsymbol{H}$) |
| $\gamma : \{1, \dots, n\} \to \{1, \dots, n\}$ | permutation function of node labels |
| $\boldsymbol{\Gamma}$ | permutation matrix associated to $\gamma$ |
| $d(x, y)$ | distance function between $x$ and $y$ |
| $\kappa(x, y)$ | kernel function of $x$ and $y$ |

## Notation

We use the following general mathematical notations and conventions. We denote vectors by small letters in bold such as $\boldsymbol{x}, \boldsymbol{y}$ and use $(\cdot)^\top$ to denote the transpose of a vector. Our convention is that all vectors are column vectors, and accordingly, $\boldsymbol{x}^\top$ is a row vector. We use $\mathbf{1}$ to denote the vector of all ones. Matrices are denoted by bold uppercase letters such as $\boldsymbol{A}, \boldsymbol{M}$, where $\boldsymbol{I}$ is used to denote the identity matrix. We write $\text{diag}(\boldsymbol{x})$ to denote the diagonal matrix whose diagonal entries are defined by the components of vector $\boldsymbol{x}$ and are 0 otherwise. Entries of vectors or matrices are non-bold with subscripts. For instance, vector $\boldsymbol{x}$ has entries $x_1, \dots, x_n$ and the matrix $\boldsymbol{A}$ has entries $A_{ij}$. If there is ambiguity, we may alternatively use the notation $[\boldsymbol{v}_i]_j$ to denote the $j$th entry of a vector $\boldsymbol{v}_i$ (or a matrix, accordingly). Finally, we use $\mathbb{P}(\cdot)$ and $\mathbb{E}[\cdot]$ for the probability and expectation of the statement inside the parentheses, respectively.

More specific notation regarding networks and associated objects is summarised in Table 1. These objects are explained in Section 2.

## 2 Background Material

In this section, we review some notions from algebraic and spectral graph theory as well as the theory of linear dynamical systems. These concepts will be essential for our discussions in the following sections.

### 2.1 Graph Theory

Networks provide a natural framework to represent systems composed of elements in interaction. At the core of a network representation is the inherent assumption that the system under investigation can be decomposed into nodes, representing the system elements, and edges, representing pairwise interactions between the system elements.

In the simplest setting, we assume that both the nodes and the edges of a network are all of the same type and their number is fixed. All of these assumptions can be relaxed,[1] but we will be mostly concerned with undirected (and weighted) networks in this Element. Within this setup we can represent a network mathematically by a graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$, with a set of nodes $\mathcal{V}$ of cardinality $n := |\mathcal{V}|$, and a set of edges $\mathcal{E} = \{\{i, j\} \mid i, j \in \mathcal{V}\}$. Without loss of generality, we will identify the node set $\mathcal{V}$ with the set $\{1, \ldots, n\}$. For a weighted graph, we endow the graph $\mathcal{G}$ with a weight function $w_\mathcal{G}: \mathcal{E} \to \mathbb{R}_+$, which maps each edge $\{i, j\}$ to a positive weight, $w_\mathcal{G}(\{i, j\}) = w_{ij}$.

More generally, we can consider directed graphs, meaning that node $i$ may be adjacent (connected) to node $j$ but not vice versa. This lack of symmetry leads to a number of mathematical complications that make directed networks and dynamical systems acting on directed networks far more difficult to analyse (cf. the box 'The Case of Directed Networks' in Section 2.3.3). We will thus focus on undirected networks, unless otherwise stated.

The edge set of a graph describes which nodes are adjacent (i.e., directly connected by an edge). Especially in the context of dynamical systems defined on graphs, we need to capture how a sequence of direct connections defines indirect connectivity between pairs of nodes, leading to the additional notions

[1] For simplicity, within this Element, we will concentrate on undirected graphs with positive edges weights and provide some additional discussion on directed networks. Relaxing the above modelling assumptions leads to various notions, including signed networks (Kunegis et al., 2010), multiplex networks (Kivela et al., 2014), temporal networks (Holme & Saramäki, 2019), and higher-order networks (Battiston et al., 2020; Lambiotte, Rosvall, & Scholtes, 2019; Schaub et al., 2021). These are active areas of research that we will mention further when relevant, and we invite the reader to consult the literature for further information on these topics.