Journal of Theoretical Biology 348 (2014) 21–32

Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/yjtbi

Higher-order structure and epidemic dynamics in clustered networks

Martin Ritchie [a], Luc Berthouze [b,c], Thomas House [d], Istvan Z. Kiss [a,*]

[a] School of Mathematical and Physical Sciences, Department of Mathematics, University of Sussex, Falmer, Brighton BN1 9QH, UK
[b] Centre for Computational Neuroscience and Robotics, University of Sussex, Falmer, Brighton BN1 9QH, UK
[c] Institute of Child Health, University College London, London WC1E 6BT, UK
[d] Warwick Mathematics Institute, University of Warwick, Gibbet Hill Road, Coventry CV4 7AL, UK

HIGHLIGHTS

- Networks of equal clustering may show significantly different higher-order structures.
- We present an efficient motif counting algorithm.
- Novel order-four transitive-type metrics permit more accurate network description.
- We conjecture on the correct motif counting cardinality for use in transitive ratios.

ARTICLE INFO

Article history:
Received 17 October 2013
Received in revised form
13 January 2014
Accepted 21 January 2014
Available online 30 January 2014

Keywords:
Clustering
Motif counting
Loops
Networks
Epidemic

ABSTRACT

Clustering is typically measured by the ratio of triangles to all triples regardless of whether open or closed. Generating clustered networks, and how clustering affects dynamics on networks, is reasonably well understood for certain classes of networks (Volz et al., 2011; Karrer and Newman, 2010), e.g. networks composed of lines and non-overlapping triangles. In this paper we show that it is possible to generate networks which, despite having the same degree distribution and equal clustering, exhibit different higher-order structure, specifically, overlapping triangles and other order-four (a closed network motif composed of four nodes) structures. To distinguish and quantify these additional structural features, we develop a new network metric capable of measuring order-four structure which, when used alongside traditional network metrics, allows us to more accurately describe a network's topology. Three network generation algorithms are considered: a modified configuration model and two rewiring algorithms. By generating homogeneous networks with equal clustering we study and quantify their structural differences, and using SIS (Susceptible-Infected-Susceptible) and SIR (Susceptible-Infected-Recovered) dynamics we investigate computationally how differences in higher-order structure impact on epidemic threshold, final epidemic or prevalence levels and time evolution of epidemics. Our results suggest that characterising and measuring higher-order network structure is needed to advance our understanding of the impact of network topology on dynamics unfolding on the networks.

1. Introduction

Network modelling is an essential tool in characterising a wide range of phenomena: infectious diseases, brain activity, chemical reactions, social interactions, the internet, etc. Any system that involves interactions of its constituent components may be modelled as a network. The versatility of networks as a modelling paradigm may be further augmented by running dynamical processes on the network such as epidemics or neuronal activity. A network's structure can have a dramatic effect on the processes that run on the network which is currently parameterised by low-order structure alongside the degree distribution. As we shall see, with epidemiological processes the presence of higher-order structure affects how a disease spreads through a network, and the effect of such structures on neuronal dynamics is known to be significant (Sporns et al., 2005; Honey et al., 2009; Gallos et al., 2012; Lynall et al., 2010; Kaiser and Hilgetag, 2010). In this paper we aim to go beyond open and closed triples and give a more comprehensive description of networks in terms of higher-order structure frequency (specifically order-four structures) and their distribution around nodes. In particular, we will examine existing clustered network generating algorithms with respect to their ability, or otherwise, to control higher-order network structure which

---

[*] Corresponding author.
E-mail address: i.z.kiss@sussex.ac.uk (I.Z. Kiss).

http://dx.doi.org/10.1016/j.jtbi.2014.01.025

sometimes may be regarded as a by-product of generating low-order structure that can preclude a correct interpretation of the impact of clustering. The paper is structured as follows. We first introduce and describe a set of clustered network generating algorithms. We follow with a presentation of the network metrics (including a description of the motif identifying/counting algorithm) that we propose to quantify similarities and differences between the generated networks. We then analyse and discuss the impact of higher-order structural differences, at identical degree distribution and equal clustering, on SIS and SIR epidemics. Finally, we discuss how our motif-counting results and newly proposed measure for higher-order structures could be used to parameterise pairwise-like models with closure at the level of quadruples.

2. Material and methods

2.1. Network construction

A significant part of network research relies on networks with arbitrary degree distributions built using the configuration model. This algorithm generates networks where nodes mix at random and where the probability that two nodes are connected is simply proportional to the product of their degree. Such networks coupled with stochastic node dynamics such as SIS, SIR or neural dynamics are amenable to developing macroscopic low-dimensional ODE models that are in excellent agreement with values obtained from stochastic simulations. By construction, these networks are loop-less in the limit of large network size. Whilst such networks can be considered in many cases as realistic or plausible models of some real-world networks, there are many instances where networks have a high degree of structure that typically involves clusters of well connected nodes. Classic examples come from household models used in epidemiology (Ball and Lyne, 2001), and networks of social interactions in general. Motivated by this, there are a series of theoretical or synthetic network models that can be tuned to display increased levels of clustering (Volz et al., 2011; Karrer and Newman, 2010; Newman, 2009; Read and Keeling, 2003; Eames, 2008; Bansal et al., 2009), where clustering denotes the ratio of closed loops of length three with respect to all possible open triples, irrespective of whether they are closed or not.

The classic algorithms to generate networks with tunable clustering include (a) the spatial algorithm proposed by Read and Keeling (2003), (b) an iterative method proposed by Eames (2008), (c) a configuration model that includes clustering (Karrer and Newman, 2010) and (d) the Big-V rewiring algorithm (Bansal et al., 2009; House and Keeling). In a recent study, Green and Kiss (2010) showed that even under identical degree distributions and equal levels of clustering, networks built based on different algorithms can display a markedly different 'higher-order structure'. Whilst their analysis identified large scale structural differences amongst networks with identical degree distribution and clustering, it did not consider extending the concept of clustering involving three nodes to higher-order structures with four or more nodes. The concept of motifs is not new (Sporns et al., 2005; Karrer and Newman, 2010; Volz et al., 2011; Keeling, 1999a; House and Keeling, 2011) and understanding network structure through higher-order motifs is going to provide a level of detail which cannot be articulated by open or connected triples alone. Below we provide a brief description of the clustered network construction algorithms used in this paper.

2.1.1. Big-V rewiring

The 'Big-V' is an iterative rewiring algorithm that can introduce clustering into any given network and is commonly used by network scientists (Bansal et al., 2009; House and Keeling; Green and Kiss, 2010). At each iterative step, a chain of 5 distinct nodes (u–v–w–x–y) is selected at random and a clone network is generated where the links (u–v) and (x–y) are broken and the edges (u–y) and (v–x) are created. This leads to a single chain of 5 nodes being broken into a triangle and a disconnected pair, see Fig. 1. Local clustering for each node in the chain, as well as all of its neighbours, is computed in both the original and cloned networks and the new configuration is kept only if the level of clustering has increased.

2.1.2. Motif decomposition rewiring

MD (Motif Decomposition) is an iterative rewiring algorithm that starts with a collection of complete sub-networks that are disconnected from one another and rewires edges randomly to reduce the clustering from its maximal value of 1 to the desired level. The following steps are performed:

i. Initialise a network that is composed of m complete motifs each with n members so that $N=nm$ and $\langle k\rangle = n-1$.
ii. Categorize every edge as 'local'.
iii. For the first step only, select at random two local edges, cut them, and swap the stubs to form new edges. Mark the pair of new edges as global.
iv. Select a local and a global edge, cut them, and swap the stubs to form new edges. Mark the pair of new edges as global.
v. Check the global clustering, if the desired level has not been achieved repeat step (iv).

Fig. 2 illustrates this process being performed on a complete motif with 4 members. It should be noted that this method may work with a heterogeneous degree distribution in which case the network would need to be initialised with motifs of $k+1$ nodes for each different degree k. MD has the significant advantage that it is computationally cheap and that, in the limit of large networks, network properties can be calculated analytically (see Appendix A.1).

2.1.3. CCM (Clustered Configuration Model)

It is possible to modify the configuration model (Miller, 2009; Volz, 2008) so that it constructs networks using specified motifs. Karrer and Newman (2010) and Volz et al. (2011) have shown how to build networks using a configuration model that includes triangle motifs. This idea may be easily extended to allow for larger and more exotic

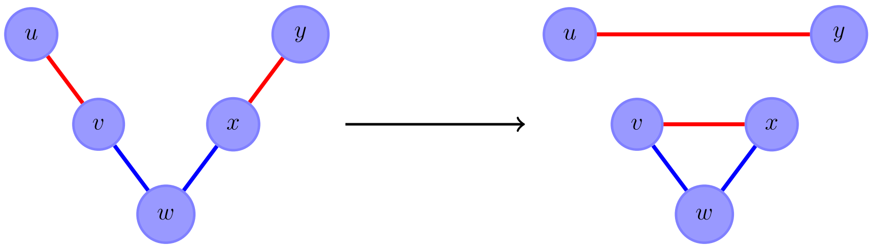

Fig. 1. A single Big-V rewiring. (a) Identify a chain of 5 nodes with 4 edges and (b) if edges (u–v) or (x–y) are already part of a triangle the cuts will not be made, otherwise rewiring is performed, and (c) independent of the outcome of (b) the algorithm will proceed to find a new chain.