RESEARCH   **Open Access**

CrossMark

The many facets of community detection in complex networks

Michael T. Schaub[1,2,3]*, Jean-Charles Delvenne[2,4], Martin Rosvall[5] and Renaud Lambiotte[3]

*Correspondence:
mschaub@mit.edu
[1]Institute for Data, Systems, and Society, Massachusetts Institute of Technology, MA, 02139 Cambridge, USA
[2]ICTEAM, Université catholique de Louvain, B-1348 Louvain-la-Neuve, Belgium
Full list of author information is available at the end of the article

Abstract

Community detection, the decomposition of a graph into essential building blocks, has been a core research topic in network science over the past years. Since a precise notion of what constitutes a community has remained evasive, community detection algorithms have often been compared on benchmark graphs with a particular form of assortative community structure and classified based on the mathematical techniques they employ. However, this comparison can be misleading because apparent similarities in their mathematical machinery can disguise different goals and reasons for why we want to employ community detection in the first place. Here we provide a focused review of these different motivations that underpin community detection. This problem-driven classification is useful in applied network science, where it is important to select an appropriate algorithm for the given purpose. Moreover, highlighting the different facets of community detection also delineates the many lines of research and points out open directions and avenues for future research.

Keywords: Community detection, Graph partitioning, Modularity, Block model

Introduction

Sparked by the work of Newman and Girvan (Newman 2006b; Newman and Girvan 2004) on Modularity in Complex Systems, the area of community detection has become one of the main pillars of network science research. The promise that we can gain a deeper understanding of a system by discerning important structural patterns within a network has spurred a huge number of studies in this area. However, as has become abundantly clear by now, this problem has no canonical solution. In fact, even a general definition of what constitutes a community is still lacking. The reasons for this are not only grounded in the computational difficulties of tackling community detection. Furthermore, various research areas view community detection from different perspectives, which the lack of a consistent terminology illustrates: 'network clustering', 'graph partitioning', 'community', 'block' or 'module detection' all carry slightly different connotations. This jargon barrier creates confusion as soon as readers and authors have different preconceptions and intuitive notions are not made explicit.
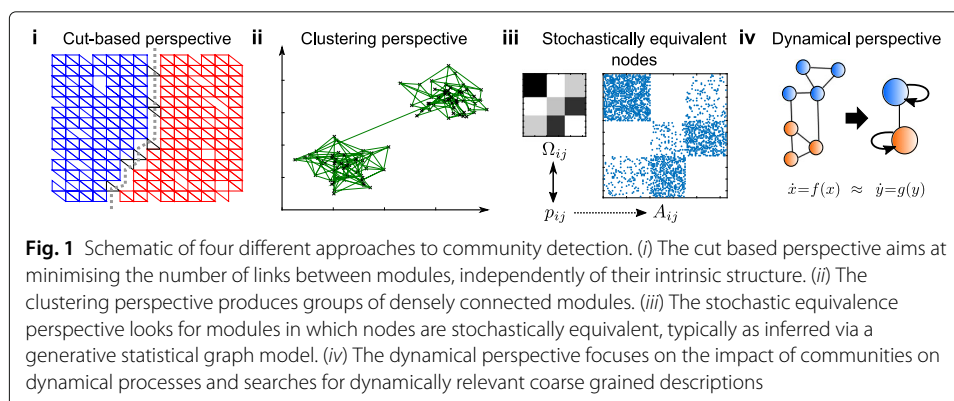
We argue that community detection should not be considered as a well-defined problem, but rather as an umbrella term with many facets. These facets emerge from different goals and motivations of what it is about the network that we want to understand or achieve, which lead to different perspectives on how to formulate the problem of

community detection. Therefore, it is critical to be aware of these underlying motivations when selecting and comparing community detection methods. Thus, rather than an in-depth discussion of the technical details of different algorithmic implementations (Schaeffer 2007; Fortunato 2010; Coscia et al. 2011; Parthasarathy et al. 2011; Newman 2012; Malliaros and Vazirgiannis 2013; Xie et al. 2013; Fortunato and Hric 2016), here we focus on the conceptual differences between different perspectives on community detection.

By providing a problem-driven classification, however, we do *not* argue that the different perspectives are unrelated. In fact, in some situations, different mathematical problem formulations can lead to similar algorithms and methods, and the different perspectives can offer valuable insights. For example, for undirected networks, optimizing the objective function Modularity (Newman and Girvan 2004), initially proposed from a clustering perspective, can be interpreted both as optimizing a particular stochastic block model (Newman 2016) and a particular diffusion process on the networks (Delvenne et al. 2013), In other situations, however, such relations are not apparent.

Neither do we argue that there is a particular perspective that is a priori better suited for any given network. In fact, no method can consistently perform best on all kinds of networks (Peel et al. 2016). Community detection is an unsupervised learning task and we cannot know what are the quantities of interest for the analysis. Instead, to understand how useful a particular method is, we must take into account the context of why the researcher is interested in the communities (Von Luxburg et al. 2012).

In the following, we unfold different aims underpinning community detection and discuss how the resulting problem perspectives relate to various applications. We focus on four broad perspectives that have served as motivation for community detection in the literature: (*i*) community detection as minimization of some form of constraint violation; (*ii*) community detection framed as a discretised analogue of data clustering, in which densely knit groups of nodes are to be found; (*iii*) community detection aiming to identify structurally equivalent nodes in a network, leading to notions such as stochastic block models; and (*iv*) community detection looking for simplified descriptions of the dynamical flows occurring on the network, that is, some form of dynamical model reduction (see Fig. 1). While this categorization is not unique, we believe that it can help clarifying concepts about community detection and be a guide to using an appropriate method for a particular purpose.

Fig. 1 Schematic of four different approaches to community detection. (*i*) The cut based perspective aims at minimising the number of links between modules, independently of their intrinsic structure. (*ii*) The clustering perspective produces groups of densely connected modules. (*iii*) The stochastic equivalence perspective looks for modules in which nodes are stochastically equivalent, typically as inferred via a generative statistical graph model. (*iv*) The dynamical perspective focuses on the impact of communities on dynamical processes and searches for dynamically relevant coarse grained descriptions

Schaub *et al. Applied Network Science* (2017) 2:4

Page 3 of 13

Minimizing constraint violations: a cut based perspective

One of the earliest graphs partitioning applications was in the area of circuit layout and design (Alpert and Kahng 1995; Fortunato 2010). This spurred the development of the now classical Kernighan-Lin algorithm (Kernighan and Lin 1970) and the work by Donath and Hoffman (1972, 1973), who were among the first to suggest the use of eigenvectors for graph partitioning. For instance, we might be confronted with a graph which describes the signal flows between different components of a circuit. To implement the circuit in an efficient way, our goal is now to partition the graph into a fixed number of approximately equally sized groups with a small number of edges between those groups. The edges that run between the groups are commonly denoted as the cut. Our aim is thus to minimize this cut while keeping some kind of *balanced groups*, which is an important ingredient in this context.

To make this more precise let us consider one specific variant of this scheme, known as *ratio cut* (Hagen and Kahng 1992). Let us denote the adjacency matrix of an undirected network with n nodes by A, where $A_{ij} = 1$ if there is a connection from node i to node j, and $A_{ij} = 0$ if the nodes are not connected. We can now write the problem of optimizing the ratio cut for a bipartition of all vertices $\mathcal{V}$ into two communities $\mathcal{V}_1$ and $\mathcal{V}_2 = \mathcal{V} \backslash \mathcal{V}_1$ as follows (Hagen and Kahng 1992; Von Luxburg 2007):

$$\min_{\mathcal{V}_i} \text{RatioCut}(\mathcal{V}_1, \mathcal{V}_2) := \min_{\mathcal{V}_i} \sum_i \frac{\text{cut}(\mathcal{V}_i, \mathcal{V} \backslash \mathcal{V}_i)}{|\mathcal{V}_i|}, \tag{1}$$

where $\text{cut}(\mathcal{V}_1, \mathcal{V}_2) := \sum_{i \in \mathcal{V}_1, j \in \mathcal{V}_2} (A_{ij} + A_{ji})/2$ is the sum of the (weighted) edges between the two vertex sets $\mathcal{V}_1, \mathcal{V}_2$. Related problem formulations also occur in the context of parallel computations and load scheduling (Spielman and Teng 1996; Pothen 1997), where approximately equally sized portions of work are to be sent to different processors, while keeping the dependencies between those tasks minimal. Further applications include scientific computing (Pothen 1997; Spielman and Teng 1996), where partitioning algorithms can be used to divide the coordinate meshes arising in the context of discretizing and solving partial differential equations. Image segmentation problems may also be phrased in terms of cut-based measures (Shi and Malik 2000; Von Luxburg 2007).

Investigating these types of problems, has led to many important contributions for partitioning graphs, in particular in relation to spectral methods. The connection of spectral algorithms to cut-based problem formulations arises naturally by considering relaxations of the original, combinatorially hard discrete optimization problems such as (2), or other related objective functions such as the average or normalized cuts. This can be best seen when rewriting the above optimzation problem as follows:

$$\min_f \quad f^T L f \tag{2}$$

$$\text{subject to} \quad f \perp \mathbf{1} \quad \|f\| = \sqrt{n} \tag{3}$$

$$\text{where} \quad f_i := \begin{cases} -\sqrt{|\mathcal{V}_2|/|\mathcal{V}_1|} & \text{if } i \in \mathcal{V}_1 \\ \sqrt{|\mathcal{V}_1|/|\mathcal{V}_2|} & \text{if } i \in \mathcal{V}_2 \end{cases} \tag{4}$$

Here the Laplacian matrix of the network has been defined as $L = D - A$, where D is the diagonal degree matrix with $D_{ii} = \sum_j A_{ij}$. We mention here the seminal work of (Fiedler 1973; 1975), who realized already in the 70s that the second smallest eigenvalue of the Laplacian is associated to the connectivity of the graph, and that the associated

Schaub *et al. Applied Network Science* (2017) 2:4

Page 4 of 13

eigenvector can thus be used to compute spectral bi-partitions. Such spectral ideas led to many influential algorithms and methods, see, for example, (Von Luxburg 2007) for a tutorial on spectral algorithms.

In this cut-based problem formulation, there is no specification as to how the found groups in the partition should be connected internally. While there is the implicit constraint that the groups must not split into groups with an even smaller cut, there is no specification that the found groups of nodes are densely connected internally. Indeed, the type of graphs considered in the context of cut-based partitions are often of a mesh or grid like form, and for these kind of graphs several guarantees can be given in terms of the quality of the partitions obtained by spectral algorithms (Spielman and Teng 1996). While such non-dense groupings emerging from 'non-clique structures' (Schaub et al. 2012) can also be dynamically relevant (see Section "Communities as dynamical building blocks"), they are likely to be missed when employing a community notion that focuses on finding dense groupings as discussed next.

Maximizing internal density: the clustering perspective

A different motivation for community detection arises when considered in the context of data clustering. We use the term clustering, which itself has been a synonym for many different things, in the following sense: For a set of given data points in a possibly high-dimensional space, the goal here is to partition the points into a number of groups, such that points within a group are 'close' (or 'similar'); and points in different groups are more distant to each other. To achieve this goal one often constructs a proximity or similarity graph between the points and tries to group nodes together that are closer to each other than they are to the rest of the graph. This approach results again in a form of community detection problem where the closeness between nodes is described by the presence and weight of the edges between them.

Although minimizing the cut size and maximizing the internal number of links are closely related, there are differences here contrasting with the cut-based perspective outlined above, pertaining to the typical constraints and search space associated with these objective functions. First, when employing a clustering perspective there is normally no a priori information as to the number of groups we are looking for. Second, we do not necessarily require the groups to be balanced in any way, rather we would like to find an 'optimal' split into densely knit groups irrespective of their relative sizes.

Unsurprisingly, finding an optimal clustering is again a computationally hard problem. Further, as Kleinberg has shown (Kleinberg 2003), there are no clustering algorithms satisfying a certain set of intuitive properties we might require from a clustering algorithm in continuous spaces; and similar problems also arise in the discrete setting for clustering of graphs (Browet et al. 2016).

Nevertheless, there exists a large number of methods that follow a clustering like paradigm and separate the nodes of a graph into cohesive groups of nodes, often by optimizing a quality function. An important clustering metric in this context is the so called conductance (Andersen et al. 2006; Kannan et al. 2004; Kloster and Gleich 2014; Spielman and Teng 2013). Optimizing the (global) conductance has been initially introduced as a way to produce a global bi-partition similarly to the 2-way ratio-cut. However, more lately this quantity has been successfully employed as a local quality function to

Schaub *et al. Applied Network Science* (2017) 2:4

Page 5 of 13

find localized clusters around one or more seed nodes. Given a set of nodes $\mathcal{V}_k \subset \mathcal{V}$, a potential community, its local conductance can be written as:

$$\phi(\mathcal{V}_k) := \frac{\sum_{i \in \mathcal{V}_k, j \notin \mathcal{V}_k} A_{ij}}{\min\{\text{vol}(\mathcal{V}_k), \text{vol}(\mathcal{V} \backslash \mathcal{V}_k)\}}, \tag{5}$$

where $\text{vol}(\mathcal{V}_k) := \sum_{i \in \mathcal{V}_k} \sum_j A_{ij}$ is the total degree of the nodes in set $\mathcal{V}_k$, commonly termed its volume in analogy with (continuous) geometric objects. Interestingly, it has been shown that in specific contexts the conductance can be a good predictor of some latent group structures in several real-world applications (Yang and Leskovec 2015)[1].

Moreover, a local perspective on community detection has two appealing properties: First, the definition of a cluster does not depend on the global graph structure but only on the relative local density. Second, only a portion of a graph needs to be accessed, which is advantageous if there are computational constraints (very large graphs), or we are only interested in a particular subsystem. In such cases, we would like to avoid having to apply a method to the whole graph to find, for example, the cluster containing a particular node in the graph.

The Newman-Girvan Modularity (Newman and Girvan 2004; Newman 2006b) is arguably one of the most common clustering measures used in the literature and was originally proposed from the clustering perspective discussed here. It is a global quality function and aims to find the community structure of the network as a whole. Given a partition $\mathcal{P} = \{\mathcal{V}_1, \ldots, \mathcal{V}_k\}$ of a graph into k groups, the Modularity of $\mathcal{P}$ can be written as:

$$Q(\mathcal{P}) := \frac{1}{2m} \sum_{q=1}^k \sum_{i,j \in \mathcal{V}_q} \left[A_{ij} - \frac{d_i d_j}{2m} \right], \tag{6}$$

where $d_i = \sum_j A_{ij}$ is the degree of node i and $2m = \sum_i d_i$ is the total weight of all edges in the graph. By optimizing the Modularity measure over the space of all partitions, one aims to identify groups of nodes that are more densely connected to each other than one would expect according to a statistical null model of the graph. This statistical null model is commonly chosen to be the configuration model with preserved degree sequence.

However, a by-product of this choice of a global null-model is the tendency of Modularity to balance the size of the modules in terms of their total connectivity. While different variants of Modularity aim to account for this effect (Fortunato 2010), this makes Modularity also interpretable as a trade-off between a cut-based measures and an entropy (Delvenne et al. 2013). In fact, Modularity can be seen as a proxy for all the perspectives discussed in this article. The optimization of Modularity is usually performed by means of spectral or greedy algorithms (Fortunato 2010; Newman 2006a; Blondel et al. 2008). While there are problems with this approach, such as its resolution limit (Fortunato and Barthélemy 2007) and other spurious effects (Fortunato and Barthélemy 2007; Good et al. 2010; Guimera et al. 2004; Lancichinetti and Fortunato 2011), the general idea has triggered development of a plethora of algorithms that follow a similar strategy (Fortunato 2010). Several works have addressed some of these shortcomings, for instance by incorporating a resolution parameter, or explicitly accounting for the density inside each group (Chen et al. 2014; Chen et al. 2015).

By grouping *similar* nodes that link to similar nodes into communities, we constrain ourselves to finding *assortative* group structure (Fortunato and Hric 2016). Stated

differently, if we ordered the nodes in the network according to the underlying group structure, the adjacency matrix would be close to block diagonal. While we may also have hierarchical clusters with clusters of clusters etc., such an assortative structural organisation might be too restrictive if we want to analyse, for example, social networks or capture the organisation of bipartite networks. If we aim to define groups based on more general connectivity patterns, this leads naturally to notions such as the stochastic equivalence, which we will consider in the next section.

Nodes with similar structural roles and stochastic block models

Within social network analysis, a common goal is to identify nodes within a network that serve a similar structural role in terms of their connectivity profile. Accordingly, nodes are similar if they share the same kind of connection patterns to other nodes. This idea is captured in notions such as *regular equivalence*, which states that nodes are regularly equivalent if they are equally related to equivalent others (Everett and Borgatti 1994; Hanneman and Riddle 2005). A relaxation of this idea is *stochastic equivalence* (Holland et al. 1983), which means that nodes are equivalent if they connect to equivalent nodes with equal probability.

One of the most popular techniques to model and detect such kind of relationship in network data is the use of stochastic block models (SBMs) (Holland et al. 1983; Nowicki and Snijders 2001) and associated inference techniques. These models have their roots in the social networks literature (Anderson et al. 1992; Holland et al. 1983), and provide a flexible framework for modelling block structures within a network. When considering block models, we are interested in identifying node groups such that nodes within a community connect to nodes in other communities in an 'equivalent way' (Fortunato and Hric 2016).

Consider a network composed of n nodes divided into k classes. The standard SBM is defined by the set of node class labels and the affinity matrix Ω. More precisely, the link probability between two nodes i, j belonging to class c_i and c_j is given by:

$$p_{ij} := \mathbb{P}(A_{ij}) = \Omega_{c_i c_j}.$$

Under an SBM, nodes within the same class thus have exactly the same probabilities to connect to nodes of another class. This is the mathematical formulation of having stochastically equivalent nodes within each class. Finding the latent groups of nodes in a network now amounts to inferring the model parameters that provide the best fit for the observed network. That is, find the SBM with the highest likelihood.

The standard SBM assumes that the expected degree of each node is a Poisson binomial random variable (a Binomial random variable with possibly non-identical success probabilities in each trial). Because inferring the most likely SBM typically results in grouping nodes based on their degree in empirical networks with broad degree distributions, it can be advantageous to include a degree correction into the model. In the degree corrected SBM (Karrer and Newman 2011), the probability p_{ij} for a link to appear between two nodes i, j depends both on their class labels c_i, c_j and their respective degree parameters d_i, d_j (each entry A_{ij} might be a Bernoulli or a Poisson random variable such as in (Karrer and Newman 2011)):

$$p_{ij} \sim d_i d_j \Omega_{c_i c_j}.$$

Schaub *et al. Applied Network Science* (2017) 2:4

Page 7 of 13

Thus, while edges in real-world networks tend to be correlated from effects such as triadic closure (Fortunato 2010), by construction edges are conditionally independent random variables in SBMs. Moreover, most common SBMs are defined for unweighted networks or networks with integer weights by modelling the network as a multi-graph. Though there are generalizations (Aicher et al. 2014; Peixoto 2015), this is still an area comparably less studied.

In contrast to the notions of community considered above, with stochastic equivalence we are no longer interested in maximising some internal density or minimising a cut. To see this, consider a bipartite graph that from a cut- or density-based perspective contains no communities (one may even see bipartite structure as 'anti-communities'). From the stochastic equivalence perspective, however, we would say that this graph contains two groups because nodes in each set only connect to nodes in the other set.

When adopting an SBM to detect such structural organisation of the links, we explicitly adopt a statistical model for the networks. The network is essentially an instance of an ensemble of possible networks generated from such a model[2]. This model based approach comes with several advantages: First, by defining the model we effectively declare what is signal and what is noise in the data under the SBM. We can thus provide a statistical assessment of the observed data with, for example, p-values under the SBM. In other words, we can identify patterns that cannot be reasonably explained from density fluctuations of edges inherent to any realisation of this model. Second, we are, for example, able to generate new networks from our model with a similar group structure, or predict missing edges and impute data. Third, we can make strong statements about the detectability of groups within a graph. For example, precise criteria specify when any algorithm can recover the planted group structure for a graph created by an SBM (Decelle et al. 2011; Mossel et al. 2013). By fitting an SBM to an observed adjacency matrix it is possible to recover such a planted group structure down to its theoretical limit (Mossel et al. 2013; Massoulié 2014). While these criteria apply to networks generated with SBMs and not real networks in general, in which case we do not know what kind of process created the network (Peel et al. 2016), it is nevertheless a remarkable result since it highlights that there are networks with undetectable block patterns.

Many benchmark graphs proposed in the literature, such as the commonly used LFR benchmarks (Lancichinetti et al. 2008), can be seen as specific types of SBMs. Results on these benchmarks graphs should therefore be interpreted with the SBM perspective in mind, especially with respect to the detectability limit. Finally, this model based approach also offers ways to estimate the number of communities from the data by some form of model selection, including hypothesis testing (Bickel and Sarkar 2016), spectral techniques (Krzakala et al. 2013; Saade et al. 2014), the minimum description length principle (Peixoto 2013), or Bayesian inference (Yan 2016).

Communities as dynamical building blocks

Let us now consider a fourth alternative motivation for community detection, focusing on the processes that take place on the network. All notions of community outlined above are effectively structural in the sense that they are mainly concerned with the composition of the graph itself or its representation as an adjacency matrix, respectively. However, in many cases one of the main goals of applying tools from network science is to understand the *behavior* of a system. While the topology of a system puts constraints on the dynamics

that can take place on the network, the network topology alone cannot explain the system behavior. Whence, instead of finding a coarse grained description of the adjacency matrix, we might be interested in finding a coarse grained description of the dynamics acting on top of the network.

Take air traffic as an example. An airline network, with weighted links connecting cities according to the number of flights between them, can offer some interesting insights about air traffic. For instance, in the US air traffic network, Las Vegas and Atlanta form two major hubs. However, if we instead focus on the passenger flows based on actual itineraries, the two cities show very different behavior: Las Vegas is a tourist destination and typically the final destination of itineraries, whereas Atlanta is a transfer hub onto other final destinations (Rosvall et al. 2014; Peixoto and Rosvall 2015). Thus, these airports play dynamically quite different roles in the network. Focusing on interconnection patterns alone can thus give an incomplete picture if we are interested in the dynamical behavior of a system, for which additional dynamical information should be taken into account. Conversely, a concentration of edges with high impact on the dynamics may just arise from a statistical fluctuation, if the network is seen as a realization of a particular random graph model. In this way, structural and dynamical approaches can offer complementing information.

Flow-based community detection approaches focus on specifying the modular dynamics on a given, fixed network. Consequently, depending on the dynamics of interest, the modular building blocks may look different. In general, however, they are blocks of nodes with different identities that trap the flow or channel it in specific directions. That is, they form reduced models of the dynamics where blocks of nodes are aggregated to single meta nodes with similar *dynamical function* with respect to the rest of the network. In this view, the goal of community detection is to find effective coarse-grained system descriptions of how the dynamics take place on the network structure.

This dynamical take on community detection has primarily focused on modelling the dynamics with Markovian diffusion processes (Delvenne et al. 2010; Lambiotte et al. 2014; Rosvall and Bergstrom 2008), though work of topological scales and synchronization share the same common ground (Arenas et al. 2006). Interestingly, for a simple diffusion dynamics such as a random walk on an undirected network, which is essentially determined by the spectral properties of the network Laplacian, this perspective is tightly connected with the clustering perspective discussed in Section "Maximizing internal density: the clustering perspective". This is because the presence of densely knit groups within the network can introduce a time-scale separation in the diffusion dynamics: A random walker traversing the network will initially be trapped for a significant time inside a community corresponding to the fast time-scale, before it can escape and explore the larger network corresponding to a slower time-scale. However, already for directed networks this connection between link density and dynamical behavior breaks down, even for a simple diffusion process Lambiotte et al. (2014; Rosvall and Bergstrom 2008; Schaub et al. 2012). This apparent relationship breaks down completely when focusing on longer pathways, possibly with memory effects in the dynamics (Rosvall et al. 2014; Salnikov et al. 2016).

A dynamical perspective is useful especially in applications in which the network itself is well defined, but the emergent dynamics are hard to grasp. For instance, consider the nervous system of the roundworm *C. elegans* for which there exists a distinct network.

Schaub *et al. Applied Network Science* (2017) 2:4

Page 9 of 13

A basic generative network model, such as a Barabasi-Albert graph or an SBM, might be too simple to capture the complex architecture of the network, and sampling alternative networks from such a model will not create valid alternative roundworm connectomes. Indeed, some more complicated network generative models have been proposed to model the structure of the network (Nicosia et al. 2013), and may be used to assess the significance of individual patterns compared to the background of the assumed model. However, if instead we are interested in assessing the dynamical implications of the evolutionary conserved network structure, it may be fruitful to engineer differences in the actual network and investigate how they affect the dynamical flows in the system. For instance, one can replicate experimental node ablations *in silico* and assess their dynamical impact (Bacik et al. 2016).

In the dynamical perspective we are typically interested in how short term dynamics are integrated into long term behavior of the system and seek a coarse grained description of the dynamics occurring on a given network. The network itself represents the true structure, save for empirical imperfections. This dynamical viewpoint is not tied to a particular method: for instance, it is possible to formulate generative statistical models for empirically observed pathways (Peixoto and Rosvall 2015)[3]. Compared to some the previous perspectives, the dynamical viewpoint has received somewhat less attention and has been confined mainly to diffusion dynamics. A key challenge is to extend this perspective to other types of dynamics and link it more formally to approaches of model order reduction considered in control theory. In light of the recently growing interest in control of complex systems, this could help us better understanding complex systems.

Discussion

Community detection can be viewed through a range of different lenses. Rather than looking at community detection as a generic tool that is supposed to work in a generic context, considering the application in mind is important when choosing between or comparing different methods. Each of the perspectives outlined above has its own particularities, which may or may not be suitable for the problem of interest.

We emphasize the different perspectives in the following example. Given a real-world graph generated by a possibly complex random assignment of edges, we assume that we are interested in some particular dynamics taking place on this graph such as epidemic spreading. We also assume that the graph is structured such that the dynamics exhibit a time-scale separation. If, for instance, we want to coarse grain an epidemic and identify critical links that should be controlled to confine the epidemics, then it does not matter whether or not random fluctuations generated the modules that induces the time-scale separation. In any case, these modules will be relevant for the dynamics.

Assume now that the same graph encodes interdependency of tasks in a load scheduling problem. In such a circumstance, a cut-based approach will find a relevant community structure, in that it allows an optimally balanced assignment of tasks to processors that minimises communication between processors. These communities may be very different from the ones attached to the epidemic spreading. In these two cases, we considered a single realisation of the network, and the goal was to extract useful information about its structure, independently of the possible mechanisms that generated it.

Let us now consider the same network from a stochastic equivalence perspective, and assume for simplicity that the graph is a particular realization of an Erdős-Rényi graph. In this case, an approach based on the SBM is expected to declare that there is no significant pattern to be found here at all, as the encountered structural variations can already be explained by random fluctuations rather than by hidden class labels. Thus, communities in the SBM picture are defined via the latent variables within the statistical model of the network structure, and not via their impact on the behavior of the system. In this way, different motivations for community detection can find different answers even for the very same network.

In addition to the differences *between* these perspectives, there are also variations *within* each perspective. For instance, distinct plausible generative models such as the standard SBM or the degree corrected SBM will for a given graph lead to different inferred community structure. Similar variations exist in the dynamical paradigm as well: distinct natural assumptions for the dynamics, such as dynamics with memory or not, uniform across nodes or edges, etc., applied to a given graph will lead to different partitions. Also different balancing criteria, or different concepts of high internal density will be valid in different contexts.

As a matter of fact, some of the internal variations make the perspectives overlap in particular scenarios. For instance, one can compare all algorithms on simple, undirected LFR benchmark graphs (Lancichinetti et al. 2008). However, the LFR benchmark clearly imposes a density-based notion of communities. Similarly, for simple undirected networks, optimizing Modularity corresponds to the inference of a particular SBM (Newman 2016) or may be reinterpreted as a diffusion process on a graph (Delvenne et al. 2013). Nevertheless, this overlap of concepts, typically present on unweighted undirected networks, is only partial, and breaks down, for example, for directed, weighted networks, or for more complex dynamics.

Conclusions

In summary, no general purpose algorithm will ever serve all applications or data types (Peel et al. 2016), because each perspective emphasizes a particular core aspect: a cut-based method provides good separation of balanced groups, a clustering method provides strong cohesiveness of groups with high internal density, stochastic block models provide strong similarity of nodes inside a group in terms of their connectivity profiles, and methods that view communities as dynamical building blocks aim to provide node groups that influence or are influenced by some dynamics in the same way. As more and more diverse types of data are collected, leading to ever more complex network structures, including directed (Malliaros and Vazirgiannis 2013), temporal (Holme and Saramäki 2012; Sekara et al. 2016), multi-layer or multiplex networks (Boccaletti et al. 2014), the differences between the perspectives presented here will become even more striking—the same network might have multiple valid partitions depending on what question about the network we are interested in. We might moreover not only be interested in partitioning the nodes, but also in partitioning edges (Ahn et al. 2010), or even motifs (Benson et al. 2016). Rather than striving to find a 'best' community-detection algorithm for a better understanding of complex networks, we argue for a more careful treatment of what network aspects that we seek to understand when applying community detection.

Endnotes

[1] Let us emphasize here again that this fact does not mean that conductance is a more meaningful algorithm in any way, or able to reveal some generic 'ground truth'; see (Peel et al. 2016) for an extensive discussion on the relation of meta-data and structure.

[2] This ensemble assumption is also reflected in the Modularity formalism, where the observed network is compared to a null model.

[3] Note, however, that whereas the generative approach in (Peixoto and Rosvall 2015) tries to explicitly model the underlying state space of the trajectories, we may simply be interested in effectively compressing the long term behavior of the system, which is a somewhat different goal. See, for instance, the discussion in (Peixoto and Rosvall 2015; Persson et al. 2016)

Availability of data and materials
Not applicable.

Authors' contributions
All authors wrote the manuscript. All authors read and approved the final manuscript.

Competing interests
The authors declare no competing financial interests.

Author details
[1] Institute for Data, Systems, and Society, Massachusetts Institute of Technology, MA, 02139 Cambridge, USA. [2] ICTEAM, Université catholique de Louvain, B-1348 Louvain-la-Neuve, Belgium. [3] naXys and Department of Mathematics, University of Namur, B-5000 Namur, Belgium. [4] CORE, Université catholique de Louvain, B-1348 Louvain-la-Neuve, Belgium. [5] Integrated Science Lab, Department of Physics, Umeå University, SE-901 87 Umeå, Sweden.

References
Ahn YY, Bagrow JP, Lehmann S (2010) Link communities reveal multiscale complexity in networks. Nature 466(7307):761–764. doi: 10.1038/nature09182
Aicher C, Jacobs AZ, Clauset A (2014) Learning latent block structure in weighted networks. J Complex Netw 3(2):221–248. doi:10.1093/comnet/cnu026
Alpert CJ, Kahng AB (1995) Recent directions in netlist partitioning: a survey. Integration VLSI J 19(1):1–81
Andersen R, Chung F, Lang K (2006) Local graph partitioning using pagerank vectors. In: 2006 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS'06). IEEE. pp 475–486. doi:10.1109/FOCS.2006.44
Anderson CJ, Wasserman S, Faust K (1992) Building stochastic blockmodels. Soc Netw 14(1):137–161
Arenas A, Díaz-Guilera A, Pérez-Vicente CJ (2006) Synchronization reveals topological scales in complex networks. Phys Rev Lett 96(11):114–102. doi:10.1103/PhysRevLett.96.114102 http://prl.aps.org/abstract/PRL/v96/i11/e114102
Bacik KA, Schaub MT, Beguerisse-Díaz M, Billeh YN, Barahona M (2016) Flow-based network analysis of the *Caenorhabditis elegans* Connectome. PLoS Comput Biol 12(8):1–27. doi:10.1371/journal.pcbi.1005055
Benson AR, Gleich DF, Leskovec J (2016) Higher-order organization of complex networks. Science 353(6295):163–166. doi:10.1126/science.aad9029
Bickel PJ, Sarkar P (2016) Hypothesis testing for automated community detection in networks. J R Stat Soci Series B (Stat Methodol) 78(1):253–273
Blondel VD, Guillaume JL, Lambiotte R, Lefebvre E (2008) Fast unfolding of communities in large networks. J Stat Mech Theory Experiment 2008(10):P10008
Boccaletti S, Bianconi G, Criado R, Del Genio CI, Gómez-Gardeñes J, Romance M, Sendiña-Nadal I, Wang Z, Zanin M (2014) The structure and dynamics of multilayer networks. Phys Rep 544(1):1–122
Browet A, Hendrickx JM, Sarlette A (2016) Incompatibility boundaries for properties of community partitions. arXiv:160300621. https://arxiv.org/abs/1603.00621
Chen M, Kuzmin K, Szymanski BK (2014) Community detection via maximization of modularity and its variants. IEEE Trans Comput Soc Syst 1(1):46–65
Chen M, Nguyen T, Szymanski BK (2015) A new metric for quality of network community structure. arXiv:150704308
Coscia M, Giannotti F, Pedreschi D (2011) A classification for community discovery methods in complex networks. Stat Anal Data Mining 4(5):512–546

Decelle A, Krzakala F, Moore C, Zdeborová L (2011) Inference and phase transitions in the detection of modules in sparse networks. Phys Rev Lett 107:065701. doi:10.1103/PhysRevLett.107.065701

Delvenne JC, Yaliraki SN, Barahona M (2010) Stability of graph communities across time scales. Proc Natl Acad Sci 107(29):12755–60

Delvenne JC, Schaub MT, Yaliraki SN, Barahona M (2013) The stability of a graph partition: A dynamics-based framework for community detection. In: Dynamics On and Of Complex Networks, Volume 2. Springer. pp 221–242. doi:10.1007/978-1-4614-6729-8_11

Donath WE, Hoffman AJ (1972) Algorithms for partitioning of graphs and computer logic based on eigenvectors of connection matrices. IBM Tech Discl Bull 15(3):938–944

Donath, WE, Hoffman AJ (1973) Lower bounds for the partitioning of graphs. IBM J Res Dev 17(5):420–425

Everett MG, Borgatti SP (1994) Regular equivalence: General theory. J Math Sociol 19(1):29–52

Fiedler M (1973) Algebraic connectivity of graphs. Czechoslov Math J 23(2):298–305

Fiedler, M (1975) A property of eigenvectors of nonnegative symmetric matrices and its application to graph theory. Czechoslov Math J 25(4):619–633

Fortunato S (2010) Community detection in graphs. Phys Rep 486(3):75–174

Fortunato S, Barthélemy M (2007) Resolution limit in community detection. Proc Natl Acad Sci 104(1):36–41. doi:10.1073/pnas.0605965104 http://www.pnas.org/content/1/36.abstract

Fortunato S, Hric D (2016) Community detection in networks: a user guide. Physics Reports 659:1 – 44. arXiv:160800163. doi:10.1016/j.physrep.2016.09.002

Good BH, de Montjoye YA, Clauset A (2010) Performance of modularity maximization in practical contexts. Phys Rev E 81(4):046106. doi:10.1103/PhysRevE.81.046106 http://pre.aps.org/abstract/PRE/v81/i4/e046106

Guimera R, Sales-Pardo M, Amaral LAN (2004) Modularity from fluctuations in random graphs and complex networks. Phys Rev E 70(2):025101

Hagen L, Kahng AB (1992) New spectral methods for ratio cut partitioning and clustering. IEEE Trans Comput Aided Des Integr Circ Syst 11(9):1074–1085

Hanneman RA, Riddle M (2005) Introduction to social network methods, University of California Riverside

Holland PW, Laskey KB, Leinhardt S (1983) Stochastic blockmodels: First steps. Soc Netw 5(2):109–137

Holme P, Saramäki J (2012) Temporal networks. Phys Rep 519(3):97–125

Kannan R, Vempala S, Vetta A (2004) On clusterings: Good, bad and spectral. J ACM (JACM) 51(3):497–515

Karrer B, Newman ME (2011) Stochastic blockmodels and community structure in networks. Phys Rev E 83(1):016107

Kernighan BW, Lin S (1970) An efficient heuristic procedure for partitioning graphs. Bell Syst Tech J 49(2):291–307

Kleinberg J (2003) An impossibility theorem for clustering. In: Becker S, Thrun S, Obermayer K (eds). Advances in neural information processing systems 15. MIT Press. pp 463–470. http://papers.nips.cc/paper/2340-an-impossibility-theorem-for-clustering.pdf

Kloster K, Gleich DF (2014) Heat kernel based community detection. In: Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM. pp 1386–1395. doi:10.1145/2623330.2623706

Krzakala F, Moore C, Mossel E, Neeman J, Sly A, Zdeborová L, Zhang P (2013) Spectral redemption in clustering sparse networks. Proc Natl Acad Sci 110(52):20935–40

Lambiotte R, Delvenne JC, Barahona M (2014) Random walks, Markov processes and the multiscale modular organization of complex networks. IEEE Trans Netw Sci Eng 1(2):76–90

Lancichinetti A, Fortunato S (2011) Limits of modularity maximization in community detection. Phys Rev E 84:066122. doi:10.1103/PhysRevE.84.066122

Lancichinetti A, Fortunato S, Radicchi F (2008) Benchmark graphs for testing community detection algorithms. Phys Rev E 78(4):046110. doi:10.1103/PhysRevE.78.046110 http://pre.aps.org/abstract/PRE/v78/i4/e046110

Malliaros FD, Vazirgiannis M (2013) Clustering and community detection in directed networks: A survey. Phys Rep 533(4):95–142

Massoulié L (2014) Community detection thresholds and the weak Ramanujan property. In: Proceedings of the 46th Annual ACM Symposium on Theory of Computing. ACM. pp 694–703. doi:10.1145/2591796.2591857

Mossel E, Neeman J, Sly A (2013) A proof of the block model threshold conjecture. arXiv:13114115. https://arxiv.org/abs/1311.4115

Newman MEJ (2016) Equivalence between modularity optimization and maximum likelihood methods for community detection. Phys. Rev. E 94(5):052315. American Physical Society. arXiv:160602319, doi:10.1103/PhysRevE.94.052315, http://link.aps.org/doi/10.1103/PhysRevE.94.052315

Newman ME (2006a) Finding community structure in networks using the eigenvectors of matrices. Phys Rev E 74(3):036104

Newman MEJ (2006b) Modularity and community structure in networks. Proc Natl Acad Sci 103(23):8577–8582. doi:10.1073/pnas.0601602103 http://www.pnas.org/content/103/23/8577.abstract

Newman ME (2012) Communities, modules and large-scale structure in networks. Nat Phys 8(1):25–31

Newman MEJ, Girvan M (2004) Finding and evaluating community structure in networks. Phys Rev E 69(2):026,113. doi:10.1103/PhysRevE.69.026113 http://pre.aps.org/abstract/PRE/v69/i2/e026113

Nicosia V, Vértes PE, Schafer WR, Latora V, Bullmore ET (2013) Phase transition in the economically modeled growth of a cellular nervous system. Proc Natl Acad Sci 110(19):7880–7885. doi:10.1073/pnas.1300753110 http://www.pnas.org/content/110/19/7880.abstract

Nowicki K, Snijders TAB (2001) Estimation and prediction for stochastic blockstructures. J Am Stat Assoc 96(455):1077–1087

Parthasarathy S, Ruan Y, Satuluri V (2011) Community discovery in social networks: Applications, methods and emerging trends. In: Social network data analytics. Springer. pp 79–113

Peel L, Larremore DB, Clauset A (2016) The ground truth about metadata and community detection in networks. arXiv:160805878 http://arxiv.org/abs/1608.05878

Peixoto TP (2013) Parsimonious module inference in large networks. Phys Rev Lett 110(14):148701

Schaub *et al. Applied Network Science* (2017) 2:4

Page 13 of 13

Peixoto, TP (2015) Inferring the mesoscale structure of layered, edge-valued, and time-varying networks. Phys Rev E 92(4):042807

Peixoto TP, Rosvall M (2015) Modeling sequences and temporal networks with dynamic community structures. arXiv:150904740. https://arxiv.org/abs/1509.04740

Persson C, Bohlin L, Edler D, Rosvall M (2016) Maps of sparse markov chains efficiently reveal community structure in network flows with memory. arXiv preprint arXiv:160608328. https://arxiv.org/abs/1606.08328

Pothen A (1997) Graph partitioning algorithms with applications to scientific computing. In: Parallel Numerical Algorithms. Springer. pp 323–368. doi:10.1007/978-94-011-5412-3_12

Rosvall M, Bergstrom CT (2008) Maps of random walks on complex networks reveal community structure. Proc Natl Acad Sci 105(4):1118—1123

Rosvall M, Esquivel AV, Lancichinetti A, West JD, Lambiotte R (2014) Memory in network flows and its effects on spreading dynamics and community detection. Nat Commun 4630:5. doi:10.1038/ncomms5630

Saade A, Krzakala F, Zdeborová L (2014) Spectral clustering of graphs with the bethe hessian. In: Ghahramani Z, Welling M, Cortes C, Lawrence ND, Weinberger KQ (eds). Advances in Neural Information Processing Systems 27. Curran Associates, Inc. pp 406–414. http://papers.nips.cc/paper/5520-spectral-clustering-of-graphs-with-the-bethe-hessian.pdf

Salnikov V, Schaub MT, Lambiotte R (2016) Using higher-order Markov models to reveal flow-based communities in networks. Sci Rep:6:23194. doi:10.1038/srep23194

Schaeffer SE (2007) Graph clustering. Comput Sci Rev 1(1):27–64

Schaub MT, Delvenne JC, Yaliraki SN, Barahona M (2012) Markov dynamics as a zooming lens for multiscale community detection: non clique-like communities and the field-of-view limit. PloS ONE 7(2):e32210

Sekara V, Stopczynski A, Lehmann S (2016) Fundamental structures of dynamic social networks. Proc Natl Acad Sci 113(36):9977–9982. doi:10.1073/pnas.1602803113

Shi J, Malik J (2000) Normalized cuts and image segmentation. IEEE Trans Pattern Anal Mach Intell 22(8):888–905

Spielman DA, Teng SH (1996) Spectral partitioning works: Planar graphs and finite element meshes. In: Foundations of Computer Science, 1996. Proceedings., 37th Annual Symposium on. IEEE. pp 96–105. doi:10.1109/SFCS.1996.548468

Spielman DA, Teng SH (2013) A local clustering algorithm for massive graphs and its application to nearly linear time graph partitioning. SIAM J Comput 42(1):1–26

Von Luxburg U (2007) A tutorial on spectral clustering. Stat Comput 17(4):395–416

Von Luxburg U, Williamson RC, Guyon I (2012) Clustering: Science or art?, Vol. 27. http://www.jmlr.org/proceedings/papers/v27/luxburg12a/luxburg12a.pdf

Xie J, Kelley S, Szymanski BK (2013) Overlapping community detection in networks: The state-of-the-art and comparative study. ACM Comput Surv (csur) 45(4):43

Yan X (2016) Bayesian Model Selection of Stochastic Block Models. arXiv:160507057. https://arxiv.org/abs/1605.07057

Yang J, Leskovec J (2015) Defining and evaluating network communities based on ground-truth. Knowl Inf Syst 42(1):181–213