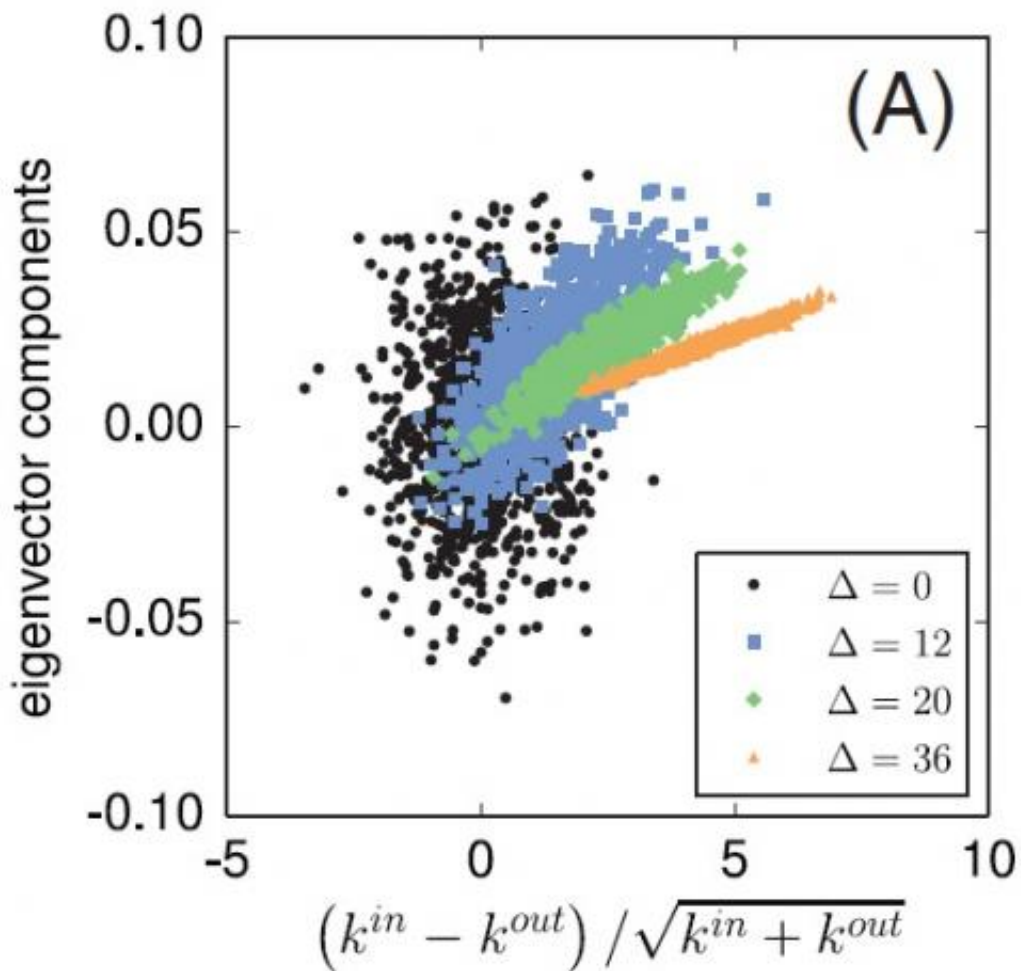


# Network paradox may help algorithms overcome 'universal limitation'

May 20 2014, by Lisa Zyga



An analysis of community detection in networks shows that detection algorithms are better at identifying ill-defined communities than well-defined communities. This paradox suggests that algorithms may have greater success if they redefine their definition of a community. Credit: Radicchi, et al. ©2014 EPLA

(Phys.org) —Sometimes paradoxes can be frustrating, but other times they can reveal something that was previously hidden. A new paradox in the field of network science, presented in a recent issue of *EPL* by Filippo Radicchi, Assistant Professor at Indiana University, seems to fall into the latter category.

While networks are used to represent a diverse range of interactions—from social to biological to physical—all networks share many of the same basic features. One common feature is their internal organization in communities. In intuitive terms, communities are well-connected subgroups (such as social cliques) within the larger network. In network terms, community structure is defined as a subgroup of nodes whose total number of internal links (links with each other) is larger than its total number of external links (links with nodes outside the subgroup).

While identifying communities defined in this way may seem fairly straightforward, community detection is actually quite complicated in large, complex networks. Nevertheless, it is crucial for understanding the structural and dynamical properties of networks, and thus using networks to their full potential.

For this reason, researchers have developed algorithms aimed at community detection. The problem is, when a community's total number of internal links is only a little bit larger than its total number of external links, the algorithms may not detect the community. That is, the algorithms can only detect communities whose average value of the difference between the numbers of internal and external links surpasses a certain detectability threshold. The detectability threshold value changes in response to different network properties, although it is of course always greater than zero. Virtually all community detection algorithms face this limitation, making the limitation appear universal.

In his new paper, Radicchi has discovered something that may call for a reconsideration of this seemingly universal limitation. His findings cut right to the heart of the definition of a community.

The traditional definition of a community, as stated above, is based on averages. That is, if a community's total number of internal links is larger than its total number of external links, the average node also has more internal than external links. However, any individual node could be very different than the average node. For example, some nodes may have more external links than internal links. Also, some nodes may have lots of links, both internal and external, while other nodes have very few links of either kind.

Looking at nodes in this way is important because these properties affect the value of the detectability threshold—but not in the way that one would expect. As Radicchi explains, some communities are more high quality and better defined than others. High-quality, well-defined communities are those in which many individual nodes have more internal than external links. On the other hand, low-quality, ill-defined communities are those in which many nodes have more external than internal links, even though the community's total number of internal links is still larger than the total number of external links (so it still falls under the definition of a "community").

It would seem that high-quality communities would be easier to detect than low-quality ones. But this is exactly the opposite of what Radicchi found.

Instead, his main results counterintuitively show that the value of the detectability threshold is inversely proportional to the quality of the community. He thinks that the reason why high-quality communities are more difficult to detect for algorithms is that the algorithms are actually detecting subgroups that do not fully fit the intuitive definition of

communities.

This finding suggests that, if current algorithms could be modified to better detect the intuitive properties of communities, then they would likewise be better at detecting high-quality communities than low-quality communities. By the same token, these better algorithms might no longer be restricted by the "universal limitation" detectability threshold.

As Radicchi explained, improving community detection algorithms could have wide-reaching implications for many different types of networks.

"The detection of communities in graphs represents a way to 'simplify' the system under observation," Radicchi told *Phys.org*. "Identifying the way in which [nodes](#) are organized in communities and how communities are organized in macro communities (i.e., hierarchical structure) is, in a certain sense, similar to drawing a map of the network. Depending on their level in the hierarchical structure, [communities](#) can play the same role as those played by cities, counties, states, countries and continents in the organization of locations in geographical maps. Concrete applications of community detection algorithms range from the design of efficient navigation protocols in the Internet to the creation of efficient systems of recommendation of commercial products to customers."

**More information:** Filippo Radicchi. "A paradox in community detection." *EPL*, 106 (2014) 38001. DOI: [10.1209/0295-5075/106/38001](https://doi.org/10.1209/0295-5075/106/38001)

© 2014 Phys.org

Citation: Network paradox may help algorithms overcome 'universal limitation' (2014, May 20) retrieved 27 April 2024 from

<https://phys.org/news/2014-05-network-paradox-algorithms-universal-limitation.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.