Journal of Informetrics 11 (2017) 669–682

Contents lists available at ScienceDirect

Journal of Informetrics

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

Heterogeneity in an undirected network: Definition and measurement☆

Xiaojun Hu[a], Loet Leydesdorff[b], Ronald Rousseau[c,d,*]

[a] Medical Information Center, Zhejiang University School of Medicine, Hangzhou 310058, China
[b] University of Amsterdam, Amsterdam School of Communication Research (ASCoR), PO Box 15793, 1001 NG Amsterdam, The Netherlands
[c] KU Leuven, Dept. of Mathematics, Celestijnenlaan 200B, B-3001 Leuven, Belgium
[d] University of Antwerp (UA), Faculty of Social Sciences, B-2000 Antwerp, Belgium

ARTICLE INFO

Article history:
Received 25 March 2017
Received in revised form 7 May 2017
Accepted 8 May 2017

Keywords:
Heterogeneity
Layered networks
External links
Interdisciplinarity
Protein-disease network

ABSTRACT

Although heterogeneous networks are ubiquitous, a precise definition is lacking, in our opinion. We submit a definition of network heterogeneity and elaborate on a resulting approach to its measurement. This measure is denoted as HE. As an illustration HE is applied to examples from the fields of informetrics and neurosciences. Yet, it is pointed out that our definition has universal applicability.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

Networks may be homogeneous in the sense that all nodes are of the same type or they may contain different types of nodes. In computer architecture studies, the terminology "heterogeneous network" often refers to a network of computers and other devices of different origins or made by different manufacturers. On the internet the term heterogeneous network may refer to a constellation using different data link protocols. A bipartite network such as that of scientists and papers, which links a scientist to each paper of which he or she is an author and vice-versa links a paper to its authors, is a simple example of a heterogeneous network of authors and papers. Similarly, the nodes in a network of bibliographically coupled or co-cited articles can be characterized by the journal in which they are published or, at a next level, the journal category which is assigned to the journal.

We recall that according to the Merriam-Webster dictionary the term "heterogeneity" refers to the quality of consisting of dissimilar or diverse ingredients or constituents. In this contribution the term network "heterogeneity" refers to a property of a network in which links connect different types of nodes. Of course, whether two nodes are considered to be of different types depends on the application one has in mind. An interesting example is the case that being of different types means belonging to different communities as the result of a community detecting algorithm. In the first part of this article each

☆ This article is an extended version of a conference paper presented at the 21st STI conference in València (Spain) (Hu, Leydesdorff, & Rousseau, 2016).

* Corresponding author at: KU Leuven, Dept. of Mathematics, Celestijnenlaan 200B, B-3001 Leuven, Belgium.
E-mail addresses: xjhu@zju.edu.cn (X. Hu), loet@leydesdorff.net (L. Leydesdorff), ronald.rousseau@kuleuven.be (R. Rousseau).

http://dx.doi.org/10.1016/j.joi.2017.05.003
1751-1577/© 2017 Elsevier Ltd. All rights reserved.

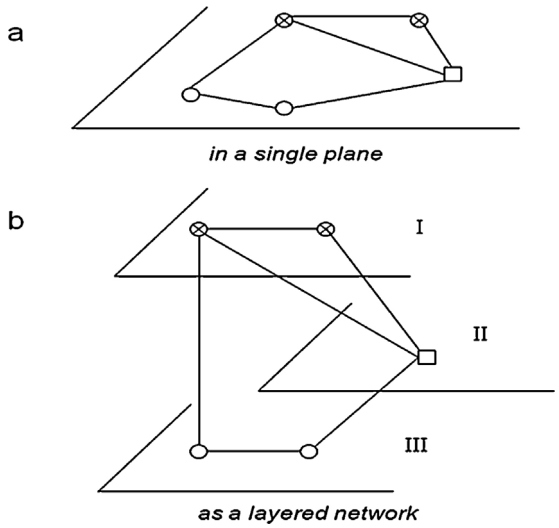

Fig. 1. a. Representation of a heterogeneous network in a single plane. b. Representation of this heterogeneous network as a layered network.

node is assumed to be of a single type only. Further on, however, this condition will be relaxed. Once a choice about the types of nodes to be distinguished is made, it is still an open question how to define a measure of heterogeneity.

As heterogeneous networks are ubiquitous in the real world we are convinced that a precise definition of heterogeneity is required (Han, 2009). A more precise definition than that used in (Han, 2009) with a resulting measure for network heterogeneity is proposed in this contribution.

In this study networks are assumed to be undirected. Links between nodes of the same type are called internal links, while links between nodes of different types will be called external links. From our perspective, nodes which are not related, i.e., not linked, never add to the heterogeneity of the network. The focus is on the heterogeneity in terms of links and not of nodes.

The term heterogeneity in a network, as used in this contribution, implies that external links are gauged positively, while internal links do not contribute to a network's heterogeneity. Moreover, we prefer evenly distributed, i.e. more balanced, external links above unevenly distributed ones (Hill, 1973). This is a choice we made, admitting that other choices, depending on the application one has in mind, can be made. In this contribution a high heterogeneity value refers to a tightly woven net between dissimilar things. We value this property so high that networks without external links receive the same heterogeneity value as homogeneous networks. Indeed, when different types of nodes are never connected the network is actually a disjoint union of homogeneous networks. For all intents and purposes such a network cannot be considered heterogeneous. In summary, units that determine the value of a heterogeneity measure are links between nodes of different types. We stress the point that these units are not nodes but external links.

2. Heterogeneous networks and an interpretation as layered networks

If there are N types of nodes these can also be considered as belonging to N layers, leading to a layered network. Links connecting different types of nodes, i.e. external links, then become links between different layers. In this way the theory of layered networks (Boccaletti et al., 2014) can be applied to heterogeneous networks.

The following example (Fig. 1) illustrates how a heterogeneous network in a single plane can be reconfigured as a layered one.

3. Literature review

We subdivide this literature review into two parts: a part on the measurement of heterogeneous networks and a part on recent progress on the measurement of diversity. As many colleagues before us measured heterogeneity through a diversity measure the two parts are closely related. Furthermore, we will propose an adaptation of a diversity measure to characterize the heterogeneity of networks.

3.1. Heterogeneity of networks

Often the term heterogeneous network is used without any concrete definition or method to measure a network's heterogeneity. Older examples of studies without a concrete definition include: Law (1986), Leydesdorff (2010a,b), Tijssen (1998), and White (2001). Tijssen studied relations in the academic-industry network in the case of process engineering. White

Download English Version:

https://daneshyari.com/en/article/4968061

Download Persian Version:

https://daneshyari.com/article/4968061

Daneshyari.com