Degree Landscapes in Scale-free Networks

Jacob Bock Axelsen,1,2 Sebastian Bernhardsson,2 Martin Rosvall,2 Kim Sneppen,1 and Ala Trusina1

1NBI, Blegdamsvej 17, Dk 2100, Copenhagen, Denmark
2Department of Theoretical Physics, Umeå University, 901 87 Umeå, Sweden

(Dated: July 12, 2018)

PACS numbers: 89.75.-k,89.75.Fb,87.16.Yc,87.16.Xa

The broad degree-distribution in many real-world networks makes it meaningful to investigate the topological organization of nodes in terms of their degree. It has been found that many social networks are assortative, with correlated degrees of adjacent nodes, but that technological and biological networks often are disassortative, with anticorrelated degrees of adjacent nodes. The degree correlation-profile, generated by comparison between the network and its randomized counterparts without degree correlations, uncovers in the Internet an overrepresentation of links between intermediate- and low-degree nodes and a slight overrepresentation of links between the nodes of highest degrees. Contrary, most biological networks have an underrepresentation of links between the hubs. To characterize the organization beyond correlations between adjacent nodes, Trusina et al. introduced the hierarchy measure \( F \), \( F \in [0,1] \), as the fraction of shortest paths between all pairs of nodes that are degree hierarchical. The degrees of the nodes along a degree-hierarchical shortest path are organized in strictly ascending, strictly descending, or first strictly ascending and then strictly descending order. It was found that biological networks with decentralized hubs stand out from other networks with a very low value of \( F \).

Here we generalize the presented findings in a landscape analogue, with mountains (high degree nodes) and valleys (low degree nodes). With this interpretation, social networks form smooth landscapes, the Internet one single mountain with first ascending and then descending hierarchical paths, whereas biological networks form rough landscapes with several mountains and broken hierarchical paths. To quantify the topology and make it possible to compare different networks, we in this paper measure the typical width of individual mountains and the separation between different mountains (Fig. 1).

To complement the methods to generate random networks (random one-mountain landscapes) and completely hierarchical networks (peaked one-mountain landscapes) with preserved degree-sequences, we here suggest a method to generate ridge landscapes (Fig. 2). In its simplest implementation, we assign a random rank to every node in a network, and organize the nodes hierarchically based on their rank. This method creates non-random networks, distinguished by a separation of hubs (disassortative with low \( F \)). We argue that the ridge landscapes can represent networks organized under different spatial constraints put on real-world networks during their evolution.

We start by reviewing the method presented by Trusina et al. to generate degree-hierarchical networks, here denoted degree-rank hierarchies. In the same time, we in detail present the suggested method to generate ridge landscapes (Fig. 2). In its simplest implementation, we assign a random rank to every node in a network, and organize the nodes hierarchically based on their rank. This method creates non-random networks, distinguished by a separation of hubs (disassortative with low \( F \)). We argue that the ridge landscapes can represent networks organized under different spatial constraints put on real-world networks during their evolution.
this way the degree of every node is kept constant and the nodes are globally organized in decreasing rank-order. To be able to investigate networks between the random-rank hierarchical networks respectively the degree-rank hierarchical networks, and random networks, we allow for random link-swaps without the constraints set by the rank of the nodes. A probability \( \varepsilon \) to make a random link-swap corresponds in this way to an error rate in the creation of the extreme networks. When \( \varepsilon \to 1 \) the methods become equivalent to the randomization of networks with remained degree-sequence suggested in [5], see Fig. 2(c).

Figure 2 shows topologies generated with the different models. They all originate from a random scale-free network (shown in Fig. 2(e)) with degree distribution, \( P(k) \propto k^{-2.5} \) and system size \( N = 400 \), generated with the method suggested in [5]. The extreme networks, the perfect random-rank hierarchy in Fig. 2(c) and the perfect degree-rank hierarchy Fig. 2(g) (\( \varepsilon = 0 \)), surround the networks with increasing error-rate towards the random scale-free network with \( \varepsilon = 1 \) in the middle (Fig. 2(e)). The perfect degree-rank hierarchy consists of a tightly connected core of large degree nodes, that forms a very peaked mountain. The landscape is not too far from the, although flatter, random case. The random-rank hierarchy, on the other hand, forms a very stringy and non-random structure — a ridge landscape. The length of the string is of the order \( D \propto N \), with very long pathways that break the small-world property found in most real-world networks. However, as for the original small-world scenario proposed by [14], the large diameter of the stringy scale-free networks collapses if small-per-

turbations exist in the hierarchical organization. If we generate the network with a small error rate \( \varepsilon \), the diameter of the network collapses as seen in Fig. 2(d). Note that the color gradient indicates that the random-rank hierarchy is still intact at this stage, and that the hubs (mountains) are separated. The degree-rank hierarchy in Fig. 2(f) is rewired with a higher error rate \( \varepsilon = 0.5 \), while still maintaining a high level of hierarchical organization.

In both cases, the two organizing principles leads to higher clustering [14], more triangles, than in the random counterparts (not shown). This is expected, as organization along any coordinate tends to make friends of friends more alike. The effect is stronger in the degree-rank hierarchy, since the clustering automatically increases further when the hubs with their many links are connected.

In Fig. 2 we also quantify the degree-hierarchical organizations of the scale-free networks organized by, respectively, degree- and number rank. For the random scale-free network with degree distribution \( P(k) \propto k^{-2.5} \) and \( N = 1000 \) nodes, \( \mathcal{F} = 0.83 \pm 0.05 \). The networks organized hierarchically according to degree-rank (as in Fig. 2(b)) have \( \mathcal{F} = 1 \) as expected. Further, when introducing a finite error rate \( \varepsilon \) for link rewirings toward the degree-hierarchy we find that its topology is robust in the sense that both diameter (not shown) and \( \mathcal{F} \) remain unchanged for even quite large errors. The perfect random-rank hierarchy has a much lower degree hierarchical organization, \( \mathcal{F} = 0.13 \pm 0.05 \). Because of the collapsing diameter, the random-rank hierarchy is not as robust as the degree-rank hierarchy to errors in the rewiring.

Figure 3 shows, in increasing degree-hierarchical order,
a number of real-world networks as degree landscapes: Yeast in (a) is the protein-protein interaction network in *Saccharomyces Cerevisia* [16], Manhattan in (b) is the dual map of Manhattan with streets as nodes and intersections as links [16], and the Internet in (c) is the network of autonomous systems [17]. The topological maps are not based on the real space the networks are embedded in, but the Kamada-Kawai algorithm in Pajek [18].

To quantify the degree landscapes independently of the layout dimension, we introduce two measures. First, in- 
spired by the information horizons in networks [18, 19, 20], we present a revised hierarchy-measure (Degree-Rank), the random-rank hierarchy (Random-Rank) [3].

*FIG. 3* Real-world networks as degree landscapes. The coloring of the altitudes are relative to the summit altitude. Yeast in (a) is the protein-protein interaction network in *Saccharomyces Cerevisia* [16], Manhattan in (b) is the dual map of Manhattan with streets as nodes and intersections as links [16], and the Internet in (c) is the network of autonomous systems [17]. The topological maps are not based on the real space the networks are embedded in, but the Kamada-Kawai algorithm in Pajek [18].

Wide mountains. However, the randomized counterparts of the two latter networks, with more peaked mountain landscapes, are both more degree hierarchical than the real networks.

We define the width of a mountain as the length where 50 percent of the paths are hierarchical. Figure 4(a) shows that the average width of the mountains in the random-rank hierarchy and yeast is about 4. In Manhattan and the Internet it is larger, about 6, and in the degree-rank hierarchy it is by definition the network diameter.

*FIG. 4* The degree-hierarchical organization as a function of path length. $F(\ell)$ is the fraction of pair of nodes, separated by a distance $\ell$, that are connected by a degree-hierarchical path. (a) shows the two model networks: The degree-rank hierarchy (Degree-Rank), the random-rank hierarchy (Random-Rank) for $\varepsilon = 0$ and 0.1, together with the random scale-free network (Random). The real-world networks in (b-d) are the same as in Fig. 3. All networks are compared with their random counterparts (Rand) [3].

In the second landscape measure, we measure the separation between mountains to investigate how the hubs are positioned relative to each other. $d(k_{hub})$ is associated to maximum distances between nodes with degree $k$ equal or larger to the threshold value $k_{hub}$. It is defined by the distance from one hub to its most distant hub in the network, averaged over all hubs

$$d(k_{hub}) = \frac{1}{N_{k \geq k_{hub}}} \sum_{\{i\mid k_i \geq k_{hub}\}} \max_{\{j\mid k_j \geq k_{hub}\}} d_{ij},$$

with $d_{ij}$ being the length of the shortest path between $i$ and $j$, and $k_i$ the degree of node $i$. The value of $d(k_{hub})$ is highly dependent of the definition of a hub, and we therefore measure $d(k_{hub})$ for all values of $k_{hub}$. 

Counterparts (Rand) [3].
FIG. 5: Average longest distance $d(k_{hub})$ between nodes of degree $k \geq k_{hub}$ as a function of $k_{hub}$ (Eq. 4), for the same networks as in Fig. 4.

Figure 5 shows $d(k_{hub})$ for a few different networks and their random counterparts. Figure 5(a) shows that the one-mountain landscapes, the degree-rank hierarchy and the random network, both have hubs tightly connected. Contrary, the hubs in the random-rank hierarchy are extremely separated ($d(1) \approx 100$) all the way out to a very high hub-threshold value. All the real-world networks in Fig. 5(b-d) fall in between these extremes, but with a higher $d(k_{hub})$ than randomly expected for most values of $k_{hub}$. Manhattan, (Fig. 5(c)), and the Internet, (Fig. 5(d)), are close to random for really high degrees, while yeast, (Fig. 5(b)), has a separation for all sizes. The close resemblance between the random-rank hierarchy and yeast in Fig. 5 and suggests that the separation of hubs probably reflects a separation of function at all scales.

Manhattan is mainly a planned city where the largest hubs, corresponding to streets and avenues, are connected to each other in a bipartite way. This results in a $d(k_{hub})$ close to 2 for the largest hubs. The Internet is constructed with a hierarchical structure within each country, and all intermediate-degree nodes (typically connected to low degree nodes) are therefore separated from each other globally. However, the largest hubs interconnect the countries, and are therefore connected with each other. This results in a $d(k_{hub})$ close to 1 for the largest hubs.

To summarize, we have generalized the degree-organizational view of real-world networks with broad degree-distributions, in a landscape analogue with mountains (high degree nodes) and valleys (low degree nodes). To quantify the topology and to be able to compare networks, we have measured the widths of the mountains and the separation between different mountains. We found that the dual map of Manhattan consists approximately only of one mountain. This implies that typical navigation between a source and a target in the city involves first going to larger and larger streets, and then to smaller and smaller streets. The Internet shares this one-mountain landscape, but the spatial constraints are weaker and the width of the mountain is about the same despite the substantially larger network.

Finally, the topological landscape in the protein-interaction network in yeast has a topology with numerous separated hills. We suggest that this reflect functional localization, where proteins tend to be connected because of similar functions rather than because they have similar degree.
http://moat.nlanr.net/

[18] S. Valverde and R. V. Solé, Internet's critical path horizon. Eur. Phys. J. B 38, 245 (2004).
[19] A. Trusina, M. Rosvall and K. Sneppen K. Phys. Rev. Lett. 94, 238701 (2005).
[20] M. Rosvall, P. Minnhagen, and K. Sneppen. Navigating Networks with Limited Information Phys. Rev. E 71, 066111 (2005).
[21] We acknowledge the support from the Danish National Research Foundation through “Models of Life” at NBI, and the Lundbeck Foundation for funding Ph.D.-studies at NBI.