

# Density decompositions of networks

Glencora Borradaile<sup>1</sup>, Theresa Migler<sup>2</sup>, and Gordon Wilfong<sup>3</sup>

<sup>1</sup> School of EECS Oregon State University

<sup>2</sup> Cal Poly, San Luis Obispo

<sup>3</sup> Bell Labs

**Abstract.** We introduce a new topological descriptor of a network called the density decomposition which is a partition of the nodes of a network into regions of uniform density. The decomposition we define is unique in the sense that a given network has exactly one density decomposition. The number of nodes in each partition defines a density distribution which we find is measurably similar to the degree distribution of given *real* networks (social, internet, etc.) and measurably dissimilar in synthetic networks (preferential attachment, small world, etc.).

## 1 Introduction

A better understanding of the topological properties of real networks can be advantageous for two major reasons. First, knowing that a network has certain properties, e.g., bounded degree or planarity, can sometimes allow for the design of more efficient algorithms for extracting information about the network or for the design of more efficient distributed protocols to run on the network. Second, it can lead to methods for synthesizing artificial networks that more correctly match the properties of real networks thus allowing for more accurate predictions of future growth of the network and more accurate simulations of distributed protocols running on such a network.

We show that networks decompose naturally into regions of uniform density, a *density decomposition*. The decomposition we define is unique in the sense that a given network has exactly one density decomposition. The number of nodes in each region defines a distribution of the nodes according to the density of the region to which they belong, that is, a *density distribution* (Section 2). Although density is closely related to degree, we find that the density distribution of a particular network is not necessarily similar to the degree distribution of that network. For example, in many synthetic networks, such as those generated by popular network models (e.g. preferential attachment and small worlds), the density distribution is very different from the degree distribution (Section 3.1). On the other hand, for *all* of the real networks (social, internet, etc.) in our data set, the density and degree distributions are measurably similar (Section 3). Similar conclusions can be drawn using the notion of *k-cores* [30], but this suffers from some drawbacks which we discuss in Section 2.3.

### 1.1 Related work

We obtain the density decomposition of a given undirected network by first orienting the edges of this network in an *egalitarian*<sup>4</sup> manner. Then we partition the nodes based on their indegree and connectivity in this orientation.

*Fair* orientations have been studied frequently in the past. These orientations are motivated by many problems. One such motivating problem is the following telecommunications network problem: Source-sink pairs  $(s_i, t_i)$  are linked by a directed  $s_i$ -to- $t_i$  path  $c_i$  (called a *circuit*). When an edge of the network fails, all circuits using that edge fail and must be rerouted. For each failed circuit, the responsibility for finding an alternate path is assigned to either the source or sink corresponding to that circuit. To limit the rerouting load of any vertex, it is desirable to minimize the maximum number of circuits for which any vertex is responsible. Venkateswaran models this problem with an undirected graph whose vertices are the sources and sinks and whose edges are the circuits. He assigns the responsibility of a circuit's potential failure by orienting the edge to either the source or the sink of this circuit. Minimizing the maximum number of circuits for which any vertex is responsible can thus be achieved by finding an orientation that minimizes the maximum indegree of any vertex. Venkateswaran shows how to find such an orientation [32]. Asahiro, Miyano, Ono, and Zenmyo consider the edge-weighted version of this problem [3]. They give a combinatorial  $\{\frac{w_{max}}{w_{min}}, (2 - \epsilon)\}$ -approximation algorithm where  $w_{max}$  and  $w_{min}$  are the maximum and minimum weights of edges respectively, and  $\epsilon$  is a constant which depends on the input [3]. Klostermeyer considers the problem of reorienting edges (rather than whole paths) so as to create graphs with given properties, such as strongly connected graphs and acyclic graphs [19]. De Fraysseix and de Mendez show that they can find an indegree assignment of the vertices given a particular properties [13]. Biedl, Chan, Ganjali, Hajiaghayi, and Wood give a  $\frac{13}{8}$ -approximation algorithm for finding an ordering of the vertices such that for each vertex  $v$ , the neighbors of  $v$  are as evenly distributed to the right and left of  $v$  as possible [8]. For the purpose of deadlock prevention [34], Wittorff describes a heuristic for finding an acyclic orientation that minimizes the sum over all vertices of the function  $\delta(v)$  choose 2, where  $\delta(v)$  is the indegree of vertex  $v$ . This objective function is motivated by a problem concerned with resolving deadlocks in communications networks [35].

In our work we show that the density decomposition can isolate the densest subgraph. The densest subgraph problem has been studied a great deal. Goldberg gives an algorithm to find the densest subgraph in polynomial time using network flow techniques [16]. There is a 2-approximation for this problem that runs in linear time [10]. As a consequence of our decomposition, we find a subgraph that has density no less than the density of the densest subgraph less one. There are algorithms to find dense subgraphs in the streaming model [4, 15]. There are algorithms that find all densest subgraphs in a graph (there could be many such subgraphs) [29].

<sup>4</sup> An egalitarian orientation is one in which the indegrees of the nodes are as balanced as possible as allowed by the topology of the network.

We consider many varied real networks in our study of the density decomposition. We find our results to be consistent across biological, technical, and social networks.

## 2 The density decomposition

In order to obtain the density decomposition of a given undirected network we first orient the edges of this network in an egalitarian manner. Then we partition the nodes based on their indegree and connectivity in this orientation.

The following procedure, the PATH-REVERSAL algorithm, finds an egalitarian orientation [9]. A *reversible path* is a directed path from a node  $v$  to a node  $u$  such that the indegree of  $v$ ,  $\delta(v)$ , is at least greater than the indegree of  $u$  plus one:  $\delta(v) > \delta(u) + 1$

Arbitrarily orient the edges of the network.  
While there is a reversible path  
reverse this path.

Since we are only reversing paths between nodes with differences in indegree of at least 2, this procedure converges; the running time of this algorithm is quadratic [9]. The orientation resulting from this procedure suggests a hierarchical decomposition of its nodes:

Let  $k$  be the maximum indegree in an egalitarian orientation.

*Ring  $k$*  ( $R_k$ ) contains all nodes of indegree  $k$  and all nodes that reach them.

Iteratively, given  $R_k, R_{k-1}, \dots$ , and  $R_{i+1}$ ,  $R_i$  contains all the remaining nodes with indegree  $i$  along with all the remaining nodes that reach them.

By the termination condition of the above procedure, only nodes of indegree  $k$  or  $k - 1$  are in  $R_k$ . Further, nodes in  $R_i$  must have indegree  $i$  or  $i - 1$ . By this definition, an edge between a node in  $R_i$  and a node in  $R_j$  is directed from  $R_i$  to  $R_j$  when  $i > j$  and all the isolated nodes are in  $R_0$ . The running time to give this decomposition is bounded by the running time to find an egalitarian orientation,  $O(|E|^2)$ .

Density can be defined in two ways: either as the ratio of number of edges to number of nodes ( $\frac{|E|}{|V|}$ ) or as the ratio of number of edges to total number of possible edges ( $\frac{2|E|}{|V|(|V|-1)}$ ). In this discussion we use the former definition. This definition of density is closely related to node degree (the number of edges adjacent to a given node): the density of a network is equal to half the average total degree.

We *identify* a set  $S$  of nodes in a graph by merging all the nodes in  $S$  into a single node  $s$  and removing any self-loops (corresponding to edges of the graph both of whose endpoints were in  $S$ ). Our partition  $R_k, R_{k-1}, \dots, R_0$  induces regions of uniform density in the following sense:

**Density Property** For any  $i = 0, \dots, k$ , identifying the nodes in  $\cup_{j>i} R_j$  and deleting the nodes in  $\cup_{j<i} R_j$  leaves a network  $G$  whose density is in the range  $(i-1, i]$  (for  $|R_i|$  sufficiently large).

In particular,  $R_k$  isolates a *densest* region in the network. Consider the network  $G_i$  formed by identifying the nodes  $\cup_{j>i} R_j$  and deleting the nodes in  $\cup_{j<i} R_j$ ; this network has one node (resulting from identifying the nodes  $\cup_{j>i} R_j$ ) of indegree 0 and  $|R_i|$  nodes of indegree  $i$  of  $i-1$ , at least one of which must have indegree  $i$ . Therefore, for any  $i$ , the density of  $G_i$  is at most  $i$  and density at least

$$\frac{(|R_i| - 1)(i - 1) + i}{|R_i| + 1} \xrightarrow{|R_i| \gg i} i - 1.$$

In Section 2.1, we observe that this relationship between density and this decomposition is much stronger.

## 2.1 Density and the Density Decomposition

In this section we discuss the following three properties:

**Property D1** The density of a densest subnetwork is at most  $k$ . That is, there is no denser region  $R_j$  for  $j > k$ .

**Property D2** The density decomposition of a network is unique and does not depend on the starting orientation.

**Property D3** Every densest subnetwork contains only nodes of  $R_k$ .

These properties allow us to unequivocally describe the density structure of a network. We summarize the density decomposition by the *density distribution*:  $(|R_0|, |R_1|, \dots, |R_{k-1}|, |R_k|)$ , i.e. the number of nodes in each region of uniform density. We will refer to a node in  $R_i$  as having *density rank*  $i$ .

The subnetwork of a network  $G$  *induced* by a subset  $S$  of the nodes of  $G$  is defined as the set of nodes  $S$  and the subset of edges of  $G$  whose endpoints are both in  $S$ ; we denote this by  $G[S]$ . First we will note that both the densest subnetwork and the subnetwork induced by the nodes of highest rank have density between  $k-1$  and  $k$ . Recall that  $k$  is the maximum indegree of a node in an egalitarian orientation of  $G$  and that  $R_i$  is the set of nodes in the  $i^{\text{th}}$  ring of the density decomposition. We will refer to  $R_k$  as the densest ring.

We use the following two lemmas to prove Property D1.

**Lemma 1.** *The density of the subnetwork induced by the nodes in  $R_k$  is in the range  $(k-1, k]$ .*

We could prove Lemma 1 directly with a simple counting argument on the indegrees of nodes in  $R_k$  or by using a network flow construction similar to Goldberg's and the max flow-min cut theorem [16].

**Lemma 2.** *The density of a densest subnetwork is in the range  $(k-1, k]$ .*

The upper bound given in Lemma 2 may be proven directly by using a counting argument for the indegrees of vertices in an egalitarian orientation of the densest subnetwork or by using the relationship between the density of the maximum density subgraph and the psuedoarboricity [20].

This upper bound proves Property D1 of the density decomposition. Property D1 has been proven in another context. It follows from a theorem of Frank and Gyarfas [12] that if  $\ell$  is the maximum outdegree in an orientation that minimizes the maximum outdegree then the density of the network,  $d$ , is such that  $\lceil d \rceil \leq \ell$ .

**Corollary 1.** *The subgraph induced by the nodes of  $R_k$  is at least as dense as the density of the densest subgraph less one.*

Note that the partition of the rings does not rely on the initial orientation, or, more strongly, nodes are uniquely partitioned into rings, giving Property D2.

**Theorem 1.** *The density decomposition is unique.*

We can prove this by noting that the maximum indegree of two egalitarian orientations for a given network is the same [9, 3, 32]. For a contradiction, we consider two different egalitarian orientations of the same graph that yield two distinct density decompositions. We then compare corresponding rings in each orientation and find that they are in fact the same.

The following theorem relies on the fact that the density decomposition is unique and proves Property D3.

**Theorem 2.** *The densest subnetwork of a network  $G$  is induced by a subset of the nodes in the densest ring of  $G$ .*

We could prove Theorem 2 directly by comparing the density of the subgraph induced by the vertices in the densest subgraph intersected with the vertices in  $R_k$  and the density of the densest subgraph. Or we could use integer parameterized max flow techniques [3, 14].

Note that there are indeed cases where the densest subgraph is induced by a strict subset of nodes in the top ring. For example, consider the graph,  $G$ , consisting of  $K_3$  and  $K_4$  with a single edge connecting the two cliques.  $K_4$  is the densest subgraph in  $G$ , however all of  $G$  is contained in the top ring ( $R_2$ ).

## 2.2 Interpretation of density rank

We can interpret orientations as assigning responsibility: if an edge is oriented from node  $a$  to node  $b$ , we can view node  $b$  as being *responsible* for that connection. Indeed several allocation problems are modelled this way [9, 2, 32, 3, 17]. Put another way, we can view a node as wishing to shirk as many of its duties (modelled by incident edges) by assigning these duties to its neighbors (by orienting the linking edge away from itself). Of course, every node wishes to shirk as many of its duties as possible. However, the topology of the network

may prevent a node from shirking too many of its duties. In fact, the egalitarian orientation is the assignment in which every node is allowed to simultaneously shirk as many duties as allowed by the topology of the network. For example, consider two situations in which a node has degree 7, in the first situation,  $a$  is the center of the star network with 8 nodes, in the second situation,  $b$  is a node in the clique on 8 nodes. Although nodes  $a$  and  $b$  both have degree 7, in the star network  $a$  can shirk all of its duties, but in the clique network  $b$  can only shirk half of its duties. There is a clear difference between these two cases that is captured by the density rank of  $a$  and  $b$  that is not captured by the degree of  $a$  and  $b$ . For example, if these were co-authorship networks, the star network may represent a network in which author  $a$  only co-authors papers with authors who never work with anyone else whereas the clique network shows that author  $b$  co-authors with authors who also collaborate with others. One may surmise that the work of author  $b$  is more reliable or respected than the work of author  $a$ .

**Theorem 3.** *For a clique on  $n$  nodes, there is an orientation where each node has indegree either  $\lfloor n/2 \rfloor$  or  $\lfloor n/2 \rfloor - 1$ .*

A proof for Theorem 3 can be given by construction of such an egalitarian orientation or by using a non-linear programming approach [26].

### 2.3 Relationship to $k$ -cores

A  $k$ -core of a network is the maximal subnetwork whose nodes all have degree at least  $k$  [30]. A  $k$ -core is found by repeatedly deleting nodes of degree less than  $k$  while possible. For increasing values of  $k$ , the  $k$ -cores form a nesting hierarchy (akin to our density decomposition) of subnetworks  $H_0, H_1, \dots, H_p$  where  $H_i$  is an  $i$ -core and  $p$  is the smallest integer such that  $G$  has an empty  $(p + 1)$ -core. For networks generated by the  $G_{n,p}$  model, most nodes are in the  $p$ -core [23, 27] For the preferential attachment model, all nodes except the initial nodes belong to the  $c$ -core, where  $c$  is the number of edges connecting to each new node [1].

These observations are similar to those we find for the density distribution (Section 3) and many of the observations we make regarding the similarity of the degree and density distributions of real networks also hold for  $k$ -core decompositions [24]. However, the local definition of cores (depending only on the degree of a node) provides a much looser connection to density than the density decomposition, as we make formal in Lemma 3.

The density of the top core may be less than the density of the top ring. Also, there are graphs for which the densest subgraph is not contained in the top core.

**Lemma 3.** *Given a core decomposition  $H_0, H_1, \dots, H_k$  of a network, the subnetwork formed by identifying the nodes in  $\cup_{j>i} H_j$  and deleting the nodes in  $\cup_{j<i} H_j$  has density in the range  $[\frac{i}{2}, i)$  for  $|H_i|$  sufficiently large.*

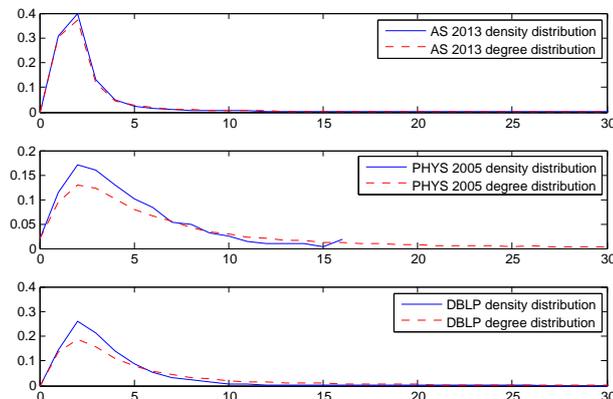


Fig. 1: In the AS network nodes represent autonomous systems and two autonomous systems are connected if there is a routing agreement between them [38] (44,729 nodes and 170,735 edges). In the PHYS network nodes represent condensed matter physicists and two physicists are connected if they have at least one co-authored paper [25] (40,421 nodes and 175,692 edges). In the DBLP network, nodes represent computer scientists and two computer scientists are connected if they have at least one co-authored paper [37] (317,080 nodes and 1,049,866 edges). The (truncated) normalized density and degree distributions are displayed. The degree distributions have long diminishing tails. AS 2013 has 67 non-empty rings, but rings 31 through 66 contain less than 1.5% of the nodes; ring 67 contains 0.75% of the nodes. DBLP has 4 non-empty rings denser than ring 30 that are disconnected; rings 32, 40, 52 and 58 contain 0.02%, 0.01%, 0.03% and 0.04% of the nodes, respectively.

### 3 The similarity of degree and density distributions

The normalized density  $\rho$  and degree  $\delta$  distributions for three networks (AS 2013, PHYS 2005, and DBLP) are given in Figure 1, illustrating the similarity of the distributions. We quantify the similarity between the density and degree distributions of these networks using the Bhattacharyya coefficient,  $\beta$  [7]. For two normalized  $\mathbf{p}$  and  $\mathbf{q}$ , the Bhattacharyya coefficient is:

$$\beta(\mathbf{p}, \mathbf{q}) = \sum_i \sqrt{p_i \cdot q_i}.$$

$\beta(\mathbf{p}, \mathbf{q}) \in [0, 1]$  for normalized, positive distributions;  $\beta(\mathbf{p}, \mathbf{q}) = 0$  if and only if  $\mathbf{p}$  and  $\mathbf{q}$  are disjoint;  $\beta(\mathbf{p}, \mathbf{q}) = 1$  if and only if  $\mathbf{p} = \mathbf{q}$ . We denote the Bhattacharyya coefficient comparing the normalized density  $\rho$  and degree  $\delta$  distribu-

tions,  $\beta(\boldsymbol{\rho}, \boldsymbol{\delta})$  for a network  $G$  by  $\beta_{\rho\delta}(G)$ . Specifically,

$$\beta_{\rho\delta}(G) = \beta(\boldsymbol{\rho}, \boldsymbol{\delta}) = \sum_i \sqrt{\rho_i \cdot \delta_i}$$

where  $\rho_i$  is the fraction of nodes in the  $i^{\text{th}}$  ring of the density decomposition of  $G$  and  $\delta_i$  is the fraction of nodes of *total* degree  $i$  in  $G$ ; we take  $\rho_i = 0$  for  $i > k$  where  $k$  is the maximum ring index. Refer to Figure 2. For all the networks in our data set,  $\beta_{\rho\delta} > 0.78$ . Note that if we exclude the Gnutella and Amazon networks,  $\beta_{\rho\delta} > 0.9$ . We point out that the other networks are self-determining in that each relationship is determined by at least one of the parties involved. On the other hand, the Gnutella network is highly structured and designed and the Amazon network is a one-mode projection of the buyer-product network (which is in turn self-determining).

Perhaps this is not surprising, given the close relationship between density and degree; one may posit that the density distribution  $\boldsymbol{\rho}$  simply bins the degree distribution  $\boldsymbol{\delta}$ . However, note that a node's degree is its *total* degree in the undirected graph, whereas a node's rank is within one of its *indegree* in an egalitarian orientation. Since the total indegree to be shared amongst all the nodes is half the total degree of the network, we might assume that, if the density distribution is a binning of the degree distribution, the density rank of a node of degree  $d$  would be roughly  $d/2$ . That is, we may expect that the density distribution is halved in range and doubled in magnitude ( $\rho_i \approx 2\delta_{2i}$ ). If this is the case, then

$$\beta(\boldsymbol{\rho}, \boldsymbol{\delta}) \approx \sum_d \sqrt{\rho_i \delta_i} \approx \sum_d \sqrt{2\delta_d \delta_{2d}}.$$

If we additionally assume that our network has a power-law degree distribution such as  $\delta_x \propto 1/x^3$ ,

$$\beta(\boldsymbol{\rho}, \boldsymbol{\delta}) \approx \int_1^\infty \sqrt{\frac{2}{x^3} \left(2 \frac{2}{(2x)^3}\right)} dx = 0.5$$

(after normalizing the distributions and using a continuous approximation of  $\beta$ ). Even with these idealized assumptions, this does not come close to explaining  $\beta_{\rho\delta}$  being in excess of 0.78 for the networks in our data set. Further, for many synthetic networks  $\beta_{\rho\delta}$  is small, as we discuss in the next section. We note that this separation between similarities of density and degree distributions for the empirical networks and synthetic networks can be illustrated with almost any divergence or similarity measure for a pair of distributions.

### 3.1 The dissimilarity of degree and density distributions of random networks

In contrast to the measurably similar degree and density distributions of real networks, the degree and density distributions are measurably *dissimilar* for networks produced by many common random network models; including the

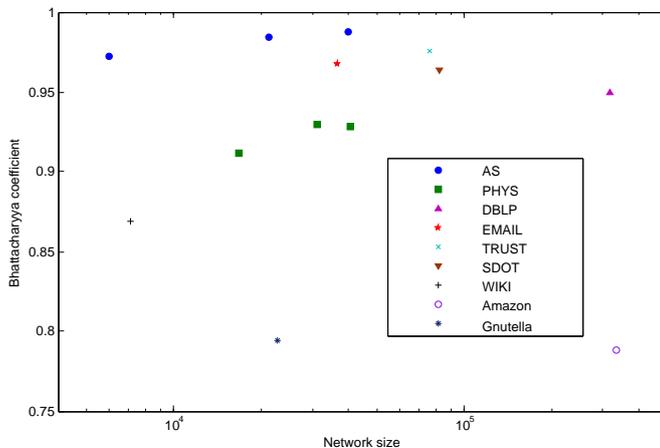


Fig. 2: Similarity ( $\beta_{\rho\delta}$ ) of density and degree distributions for 9 diverse networks. We introduced AS, PHYS, and DBLP in Figure 1. In the EMAIL network nodes represent Enron email addresses and two addresses are connected if there has been at least one email exchanged between them [18] (36,692 nodes and 183,831 edges). In the TRUST network nodes represent `epinions.com` members and two members are connected if one trusts the other [31] (75,879 nodes and 405,740 edges). In the SDOT network nodes represent `slashdot.org` members and two members are connected if they are friends or foes [22] (82,168 nodes and 504,230 edges). In the WIKI network nodes represent `wikipedia.org` users and two users are connected if one has voted for the other to be in an administrative role [21] (7,115 nodes and 103,689 edges). In the Amazon network nodes represent products and two products are connected if they are frequently purchased together [37] (334,863 nodes and 925,872 edges). In the Gnutella network nodes represent network hosts and two hosts are connected if they share files [28] (22,687 nodes and 54,704 edges). EMAIL, TRUST, and WIKI are naturally directed networks. For these networks, we ignore direction and study the underlying undirected networks. Notice that both the Amazon and Gnutella networks are highly structured. It is not surprising that these networks would have a weaker connection between the density and degree distributions.

preferential attachment (PA) model of Barabasi and Albert [6] and the small world (SW) model of Watts and Strogatz [33]. We use  $\tilde{\beta}_{\rho\delta}(M)$  to denote the Bhattacharyya coefficient comparing the expected degree and density distributions of a network generated by a model  $M$ .

*Preferential attachment networks* In the PA model, a small number,  $n_0$ , of nodes seed the network and nodes are added iteratively, each attaching to a fixed number,  $c$ , of existing nodes. Consider the orientation where each added edge is directed toward the newly added node; in the resulting orientation, all but

the  $n_0$  seed nodes have indegree  $c$  and the maximum indegree is  $c$ . At most  $cn_0$  path reversals will make this orientation egalitarian, and, since  $cn_0$  is typically very small compared to  $n$  (the total number of nodes), most of the nodes will remain in the densest ring  $R_c$ . Therefore PA networks have nearly-trivial density distributions:  $\rho_c \approx 1$ . On the other hand the expected fraction of degree  $c$  nodes is  $2/(c+2)$  [5]. Therefore  $\tilde{\beta}_{\rho\delta}(\text{PA}) \approx \sqrt{2/(c+2)}$ .

*Small-world networks* A small-world network is one generated from a  $d$ -regular network by reconnecting (uniformly at random) at least one endpoint of every edge with some probability. For probabilities close to 0, a network generated in this way is close to  $d$ -regular; for probabilities close to 1, a network generated this way approaches one generated by the random-network model  $(G_{n,p})$  of Erdős and Rényi [11]. In the first extreme,  $\tilde{\beta}_{\rho\delta}(SW) = 0$  (Lemma 4 below) because all the nodes have the same degree and the same rank. As the reconnection probability increases, nodes are not very likely to change rank while the degree distribution spreads slightly. In the second extreme, the highest rank of a node is  $\lfloor c/2 \rfloor + 1$  [36] and, using an observation of the expected size of the densest subnetwork<sup>5</sup>, with high probability nearly all the nodes have this rank. It follows that

$$\tilde{\beta}_{\rho\delta}(G_{n,p}) \approx \sqrt{\frac{c^{c/2}}{e^{-c}(c/2)!}},$$

which approaches 0 very quickly as  $c$  grows. We verified this experimentally finding that  $\tilde{\beta}_{\rho\delta}(G_{n,p}) < 0.5$  for  $c \geq 5$ .

**Lemma 4.** *For  $d \geq 3$ ,  $\beta_{\rho\delta}(G) = 0$  for any  $d$ -regular network  $G$  with  $d \geq 3$ .*

We can prove Lemma 4 by showing that  $\rho_d = 0$  since  $\delta_d = 1$  for regular networks.

## 4 Conclusion

We have introduced the density decomposition and summarized the decomposition with the density distribution. We found that the hierarchy of vertices within this decomposition are partitioned according to the density of the induced sub-graphs. We found that the density and degree distributions are remarkably similar in real graphs and dissimilar in synthetic networks. In further work we plan to use the density distribution to build more realistic random graph models. For more details, a full version of the paper may be found on arxiv.

## References

1. Jose Alvarez-Hamelin, Luca Dall'Asta, Alain Barrat, and Alessandro Vespignani. k-core decomposition of Internet graphs: hierarchies, self-similarity and measurement biases. *Networks and Heterogeneous Media*, 3(2):371, 2008.

<sup>5</sup> e-mail exchange between Glencora Borradaile and Abbas Mehrabian

2. Yuichi Asahiro, Jesper Jansson, Eiji Miyano, and Hirotaka Ono. Upper and lower degree bounded graph orientation with minimum penalty. In *Proceedings in Computing: The Australasian Theory Symposium*, pages 139–146, 2012.
3. Yuichi Asahiro, Eiji Miyano, Hirotaka Ono, and Kouhei Zenmyo. Graph orientation algorithms to minimize the maximum outdegree. *International Journal of Foundations of Computer Science*, 18(2):197–215, 2007.
4. Bahman Bahmani, Ravi Kumar, and Sergei Vassilvitskii. Densest subgraph in streaming and mapreduce. *Proc. VLDB Endow.*, 5(5):454–465, January 2012.
5. Albert-László Barabási. *Network Science*. Cambridge University Press, Cambridge, 2016.
6. Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.
7. Anil Kumar Bhattacharyya. On a measure of divergence between two statistical populations defined by their probability distributions. *Bull. Calcutta Math. Soc.*, 35:99–109, 1943.
8. Therese Biedl, Timothy Chan, Yashar Ganjali, Mohammad Taghi Hajiaghayi, and David R. Wood. Balanced vertex-orderings of graphs. *Discrete Appl. Math.*, 148:27–48, April 2005.
9. Glencora Borradaile, Jennifer Iglesias, Theresa Migler, Antonio Ochoa, Gordon Wilfong, and Lisa Zhang. Egalitarian graph orientations. In *Journal of Graph Algorithms and Applications*, volume 21, pages 687–708, 2017.
10. Moses Charikar. Greedy approximation algorithms for finding dense components in a graph. In *Proceedings of the Third International Workshop on Approximation Algorithms for Combinatorial Optimization*, pages 84–95, London, UK, 2000. Springer-Verlag.
11. Paul Erdős and Alfréd Rényi. On random graphs, I. *Publicationes Mathematicae (Debrecen)*, 6:290–297, 1959.
12. András Frank and A. Gyárfás. How to orient the edges of a graph? *Colloquia Mathematica Societatis János Bolyai*, 1:353–364, 1976.
13. Hubert de Fraysseix and Patrice Ossona de Mendez. Regular orientations, arboricity, and augmentation. In *Proceedings of the DIMACS International Workshop on Graph Drawing*, pages 111–118, London, UK, 1995. Springer-Verlag.
14. Giorgio Gallo, Michael Grigoriadis, and Robert Tarjan. A fast parametric maximum flow algorithm and applications. *SIAM Journal on Computing*, 18(1):30–55, February 1989.
15. David Gibson, Ravi Kumar, and Andrew Tomkins. Discovering large dense subgraphs in massive graphs. In *Proceedings of the 31st international conference on very large data bases*, pages 721–732. VLDB Endowment, 2005.
16. Andrew Goldberg. Finding a maximum density subgraph. Technical report, University of California at Berkeley, Berkeley, CA, USA, 1984.
17. Nicholas J. A. Harvey, Richard E. Ladner, László Lovász, and Tami Tamir. Semi-matchings for bipartite graphs and load balancing. *Journal of Algorithms*, 59:53–78, 2006.
18. Bryan Klimt and Yiming Yang. Introducing the Enron Corpus. In *First Conference on Email and Anti-Spam*, 2004.
19. William F. Klostermeyer. Pushing vertices and orienting edges. *Ars Combinatorial*, 51:65–75, 1999.
20. Lukasz Kowalik. Approximation scheme for lowest outdegree orientation and graph density measures. In *Proceedings of the 17th International Conference on Algorithms and Computation*, ISAAC’06, pages 557–566, Berlin, Heidelberg, 2006. Springer-Verlag.

21. Jure Leskovec, Daniel Huttenlocher, and Jon Kleinberg. Signed networks in social media. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 1361–1370, New York, NY, USA, 2010. ACM. <http://snap.stanford.edu/data/>.
22. Jure Leskovec, Kevin Lang, Anirban Dasgupta, and Michael Mahoney. Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters. *CoRR*, abs/0810.1355, 2008. <http://snap.stanford.edu/data/>.
23. Tomasz Luczak. Size and connectivity of the k-core of a random graph. *Discrete Mathematics*, 91(1):61–68, August 1991.
24. Theresa Migler. *The Density Signature*. PhD thesis, Oregon State University, 2014.
25. Mark Newman. Fast algorithm for detecting community structure in networks. *Phys. Rev. E*, 69:066133, Jun 2004. <http://www-personal.umich.edu/~mejn/netdata/>.
26. Jean-Claude Picard and Maurice Queyranne. A network flow solution to some nonlinear 0-1 programming problems, with applications to graph theory. *Networks*, 12:141–159, 1982.
27. Boris Pittel, Joel Spencer, and Nicholas Wormald. Sudden emergence of a giant k-core in a random graph. *J. Comb. Theory Ser. B*, 67(1):111–151, May 1996.
28. Matei Ripeanu, Ian Foster, and Adriana Iamnitchi. Mapping the Gnutella network: Properties of large-scale peer-to-peer systems and implications for system design. *IEEE Internet Computing Journal*, 6:2002, 2002. <http://snap.stanford.edu/data/>.
29. Barna Saha, Allison Hoch, Samir Khuller, Louiqa Raschid, and Xiao-Ning Zhang. Dense subgraphs with restrictions and applications to gene annotation graphs. In *Proceedings of the 14th Annual international conference on Research in Computational Molecular Biology*, RECOMB'10, pages 456–472, Berlin, Heidelberg, 2010. Springer-Verlag.
30. Stephen Seidman. Network structure and minimum degree. *Social Networks*, 5(3):269–287, 1983.
31. Maryam Tahajod, Azadeh Iranmehr, and Nasim Khozooyi. Trust management for semantic web. In *Computer and Electrical Engineering, 2009. ICCEE '09. Second International Conference*, volume 2, pages 3–6, 2009. <http://snap.stanford.edu/data/>.
32. Venkat Venkateswaran. Minimizing maximum indegree. *Discrete Appl. Math.*, 143:374–378, September 2004.
33. Duncan Watts and Steven Strogatz. Collective dynamics of ‘small-world’ networks. *Nature*, 393(6684):409–10, 1998.
34. W. Wimmer. Ein Verfahren zur Verhinderung von Verklemmungen in Vermittlernetzen. <http://www.worldcat.org/title/verfahren-zur-verhinderung-von-verklemmungen-in-vermittlernetzen/>, October 1978.
35. Vaughan Wittorff. Implementation of constraints to ensure deadlock avoidance in networks, 2009. US Patent # 7,532,584 B2.
36. Pu Gao Xavier Pérez-Giménez and Cristiane Sato. Arboricity and spanning-tree packing in random graphs with an application to load balancing. In *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '14, pages 317–326. SIAM, 2014.
37. Jaewon Yang and Jure Leskovec. Defining and evaluating network communities based on ground-truth. In *Proceedings of the ACM SIGKDD Workshop on Mining Data Semantics*, MDS '12, pages 3:1–3:8, New York, NY, USA, 2012. ACM. <http://snap.stanford.edu/data/>.
38. Yu Zhang. Internet AS-level Topology Archive. <http://irl.cs.ucla.edu/topology/>.