

ORIGINAL ARTICLE

# On the impact of network size and average degree on the robustness of centrality measures

Christoph Martin\*  and Peter Niemeyer

Institute of Information Systems, Leuphana University of Lüneburg 21335 Lüneburg, Germany  
(e-mail: [niemeyer@uni.leuphana.de](mailto:niemeyer@uni.leuphana.de))

\*Corresponding author. Email: [cmartin@uni.leuphana.de](mailto:cmartin@uni.leuphana.de)

Special Issue Editor: Hocine Cherifi

## Abstract

Measurement errors are omnipresent in network data. Most studies observe an erroneous network instead of the desired error-free network. It is well known that such errors can have a severe impact on network metrics, especially on centrality measures: a central node in the observed network might be less central in the underlying, error-free network. The robustness is a common concept to measure these effects. Studies have shown that the robustness primarily depends on the centrality measure, the type of error (e.g., missing edges or missing nodes), and the network topology (e.g., tree-like, core-periphery). Previous findings regarding the influence of network size on the robustness are, however, inconclusive. We present empirical evidence and analytical arguments indicating that there exist arbitrary large robust and non-robust networks and that the average degree is well suited to explain the robustness. We demonstrate that networks with a higher average degree are often more robust. For the degree centrality and Erdős–Rényi (ER) graphs, we present explicit formulas for the computation of the robustness, mainly based on the joint distribution of node degrees and degree changes which allow us to analyze the robustness for ER graphs with a constant average degree or increasing average degree.

**Keywords:** centrality measures, measurement error, missing data, robustness

## 1. Introduction

Networks are used to model various real-world phenomena. Typical use cases include (online) social networks, web graphs, protein–protein interaction networks, infrastructure networks, and many more (Newman, 2003). Networks are, however, sensitive to errors in the data underlying the network. The reasons for such errors are manifold. When collecting data for a social network, for example, actors may be missing on the day of the survey or the number for the nomination of possible friends may be limited by the survey questionnaire (Wang et al., 2016). The collection of protein–protein interaction data is, depending on the method used, inevitably associated with uncertainty, which is consequently also part of the network constructed from this data (De Las Rivas & Fontanillo, 2010). When creating co-authorship or citation networks, authors or papers can be included multiple times or not at all, for example, due to incorrect spelling (Erman & Todorovski, 2015; Schulz, 2016). All these errors affect the outcome of network analysis methods and thus the conclusions that depend on these methods (Marsden, 1990; Kossinets, 2006).

In the field of network analysis, centrality measures are commonly used to analyze the position of nodes in a network. These measures map a real number to every node in the network which

can be used to rank the nodes. It is well known that errors in the network data can have a severe impact on centrality measures. For example, the best ranked actor might actually not be the best in the erroneous network. We measure this impact using the concept of robustness of centrality measures, which is the correlation between the centrality values in the error-free and the erroneous network.<sup>1</sup> Previous studies have used the Pearson correlation to measure the robustness (Bolland, 1988; Costenbader & Valente, 2003; Borgatti et al., 2006). Like most recent studies, we use a rank correlation (Kim & Jeong, 2007; Wang et al., 2012; Holzmann et al., 2019; Martin & Niemeyer, 2019). The effects of errors on the robustness of centrality measures depend on several variables, for example, the type of centrality measure, the type and extent of the error, the network topology (e.g., tree-like, core-periphery), and how we measure the robustness (Frantz et al., 2009; Smith & Moody, 2013). Few studies have addressed the issue of robustness of centrality measures from an analytical perspective. Ghoshal & Barabási (2011) investigated the existence of super-stable nodes w.r.t. degree and PageRank. Platig et al. (2013) investigated the joint occurrence of missing and false links. Tsugawa & Ohsaki (2015) adapted this approach to measuring the robustness focusing on most central nodes.

In this article, we investigate the robustness of empirical networks that vary in size and structure. Existing studies are inconclusive about the relationship between network size and robustness. No relationship between size and robustness is noticeable in the empirical part of Niu et al. (2015). In Costenbader & Valente (2003) and Borgatti et al. (2006), the authors observed that larger network size could be related to both, higher and lower robustness, depending on the network structure. In Wang et al. (2012), the smaller network is usually more robust than the larger one. In contrast, Smith & Moody (2013) noticed that larger networks are frequently more robust. For a comprehensive review of the existing work on the robustness of centrality measures, we refer to Smith et al. (2017). To the best of our knowledge, however, there exist no studies that explicitly analyze the relationship between average degree and robustness. Moreover, previous studies have mostly been concerned with smaller networks (approx. less than 1000 nodes). This raises the question whether the concept of robustness of centrality measures is at all relevant in the context of larger networks. In contrast to existing studies, we specifically investigate the relationship between the size as well as the average degree and the robustness of centrality measures. In addition, we provide analytical results for this relationship based on the interpretation of the robustness as a probability.

This article is an extensively extended version of the conference paper (Martin & Niemeyer, 2020). In particular, Section 3 contains more detailed results, Section 4 now includes additional experiments, especially the section concerning the configuration model is new, and Section 5 is also novel and contains essential results for the contribution of this article.

To examine these contrary observations regarding the network size and the robustness in greater detail, we proceed as follows: first, we investigate the robustness of the degree, the eigenvector centrality, and the PageRank in 24 empirical networks coming from diverse domains, focusing on the relationship between global measures and the robustness (Section 3). We hardly observe any association between network size and robustness but a high correlation between average degree and robustness. This observation holds for all considered centrality measures and error types that involve removing nodes or edges. We further investigate the effect of network size on the robustness using the Erdős-Rényi (ER) and the Barabási-Albert (BA) random graph model (Section 4). For both models, we observe that robustness is independent of network size if the average degree remains constant. If the average degree increases, then centrality measures in BA graphs become more robust, in contrast to ER graphs. We also make these observations in our experiments with the configuration model where random graphs are generated based on the degree distributions of the empirical networks.

In Section 5, we introduce an analytical approach for the robustness. We derive explicit expressions for the robustness of the degree centrality in ER graphs. We use these expressions to prove that for ER networks of different size but with the same average degree, the robustness of the

degree centrality remains stable. As a consequence, there exist robust and non-robust networks of varying sizes, at least w.r.t. the degree centrality. We also provide arguments, based on the variance of the degree and the variance of the degree change, as to why the robustness increases or decreases with increasing average degree, depending on the type of network.

## 2. Methods

A graph  $G(V, E)$  consists of a node set  $V$  and an edge set  $E, E \subseteq \binom{V}{2}$ . We denote the number of nodes in  $G$  by  $N$  and the number of edges by  $M$ . All graphs considered in this paper are undirected, unweighted, and simple, that is, they neither contain loops nor multiple edges. The adjacency matrix of a graph is denoted by  $A$ , where  $A_{i,j} = 1$  if there is an edge between node  $v_i$  and  $v_j$  (i.e.,  $\{v_i, v_j\} \in E(G)$ ) and 0 otherwise. The neighborhood of a node  $u$  is  $N(u) = \{v: \{u, v\} \in E(G)\}$ . It is the set of nodes that are connected to  $u$ . The degree is the number of connections that a node has,  $\text{degree}(u) = |N(u)|$ . The degree of an edge is the sum of the degree values of the end nodes,  $\text{degree}(\{u, v\}) = \text{degree}(u) + \text{degree}(v)$ . We denote the degree sequence of a graph  $G$  by  $ds(G)$ .

### 2.1 Centrality measures

Centrality measures map a real number to every node in the graph and thus imply a ranking on the nodes. These measures solely depend on the structure of the graph and not on, for example, additional information about the nodes (Koschützki et al., 2005). By  $c_G(u)$ , we denote the centrality value for a specific node  $u$  in a graph  $G$  w.r.t. a centrality measure  $c$ . If the context permits, we do not explicitly mention the graph. The vector of centrality values for all nodes in  $G$  is defined as  $c(G) = (c_G(v_1), \dots, c_G(v_N))$ .

The most straightforward centrality measure is the degree centrality which was already discussed above,  $\text{degree}(u) = |N(u)|$ . The eigenvector centrality and the PageRank are both feedback measures. They are defined recursively, the centrality value of a node depends on the centrality values of its neighbors. If  $G$  is connected, then the eigenvector centrality of a node  $u$  is defined by the unique solution to  $\text{evc}(u) = \frac{1}{\lambda} \sum_{v \in N(u)} \text{evc}(v)$ , where  $\lambda$  is the largest eigenvalue of  $A$  (Bonacich, 1987). The unique solution to  $\text{PageRank}(u) = d \sum_{v \in N(u)} \frac{\text{PageRank}(v)}{\text{degree}(v)} + (1-d)$  defines the PageRank, with  $d$  as damping factor (in our case 0.85) (Brin & Page, 1998). Originally introduced for directed graphs, this concept is also applicable for undirected graphs. One of the main differences between these two measures is that, in case of the eigenvector centrality, all neighbors of a node receive the total centrality value of this node. In contrast, in case of the PageRank, neighbors of a node only receive a fraction of the nodes centrality value, which depends on the total amount of neighbors of this node.

### 2.2 Error mechanisms

When collecting data, external factors and the selection of the sampling method can lead to inaccurate network data. We use four procedures to model the impact of errors on information about nodes and edges. We call these procedures error mechanisms. They model an error that affects the nodes or edges of a graph. Their inputs are a graph  $G$  and a parameter  $\alpha$  which controls the intensity of the error. The procedure returns one graph from the set of all possible erroneous versions of the graph  $G$ . In this study, we use the following error mechanisms:

- add edges (e+):**  $\alpha N$  edges are added to the graph. If  $\alpha N$  is not an integer, then we use the smallest integer that is greater than or equal to  $\alpha N$  ( $\lceil \alpha N \rceil$ ). The new edges are chosen uniformly at random from the  $\binom{N}{2} - M$  possible edges.
- remove edges unif. (e-):**  $\lceil \alpha N \rceil$  edges are removed from the graph. The edges are chosen uniformly at random from  $E(G)$ .

**remove edges degree (e-(p))** also removes  $\lceil \alpha N \rceil$  edges. The edges are, however, chosen with probability proportional to the edge degree (i.e.,  $P(\{u, v\}) = \frac{\text{degree}(\{u, v\})}{\sum_{e \in E(G)} \text{degree}(e)}$ ).

**remove nodes (n-):**  $\lceil \alpha N \rceil$  nodes are removed from the graph. The nodes are chosen uniformly at random from  $V(G)$ .

These error types often occur during the data collection and are commonly used in existing studies; hence, we deploy them to provide comparability. The error mechanisms that remove edges degree (e-(p)) is an example of a nontrivial error mechanism, since the probability of an edge being affected depends on the position of that edge. For a more detailed discussion of the error mechanisms as random graphs, see Martin & Niemeyer (2019).

### 2.3 Robustness of centrality measures

To quantify the impact of errors in data collection on centrality measures, we use the concept of robustness, which measures how the ranking of nodes, induced by the centrality measure, changes. For two graphs,  $G$  and  $H$ , a centrality measure  $c$ , and a correlation  $corr$ , we denote the robustness by  $r_{corr,c}(G, H)$ , where  $H$  is the erroneous graph (a “modified” version of  $G$ , i.e.,  $H$  is on the same node set as  $G$  or on a subset of that node set). If  $G$  and  $H$  are not on the same node set then, similar to Wang et al. (2012), we only consider nodes that exist in both graphs. Since the robustness is defined as a correlation, the values for the robustness of a centrality measure are in  $[-1, 1]$ .

In the same way as Kim & Jeong (2007) and Holzmann et al. (2019), we use Kendall’s  $\tau$  (“tau-b”) rank correlation coefficient (Kendall, 1945) to measure the robustness of centrality measures. In this case, the robustness is defined as follows:

$$r_{\tau,c}(G, H) = \frac{n_c - n_d}{\sqrt{(n_c + n_d + n_t)(n_c + n_d + n_t')}} \tag{1}$$

The number of concordant pairs and discordant pairs w.r.t.  $c(G)$  and  $c(H)$  are  $n_c$  and  $n_d$ , respectively. A pair of nodes  $u, v$  is concordant if  $(c_G(v) - c_G(u)) \cdot (c_H(v) - c_H(u)) > 0$  and discordant if  $(c_G(v) - c_G(u)) \cdot (c_H(v) - c_H(u)) < 0$ . Ties in  $c(G)$  (i.e.,  $c_G(v) - c_G(u) = 0$ ) are denoted by  $n_t$  and ties in  $c(H)$  (i.e.,  $c_H(v) - c_H(u) = 0$ ) are denoted by  $n_t'$ .

Goodman and Kruskal’s rank correlation coefficient  $\gamma$  (Goodman & Kruskal, 1954) is closely related to Kendall’s  $\tau$  and the robustness using this measure is defined as follows:

$$r_{\gamma,c}(G, H) = \frac{n_c - n_d}{n_c + n_d} \tag{2}$$

If all pairs are either concordant or discordant w.r.t. the centrality  $c$  (i.e., there are no ties), then both measures are equal. Since  $(r_{\gamma,c}(G, H) + 1)/2$  can be interpreted as the probability that two randomly chosen nodes have the same order in  $c_G$  and  $c_H$  w.r.t.  $c$ , this measure is more accessible for an analytical perspective and we use it in Section 5. For the empirical part, we use Kendall’s  $\tau$  to provide comparability to existing studies.

### 2.4 Random graph models

In this paper, we use the ER model, the BA model, and the configuration model.

The ER random graph model, introduced by Erdős & Rényi (1959), has two parameters: the number of nodes  $n$  and the edge probability  $p$ . Since all node pairs are connected with the same probability ( $p$ ), the degree distribution of the nodes in this model follows a binomial distribution.

In contrast, the BA model is based on the idea of preferential attachment. Consequently, the probability that a new node will connect to an existing node is proportional to the degree of the existing node. This model also has two parameters. In addition to the number of nodes  $n$ ,

the parameter  $m$  specifies the number of connections that a new node makes to existing nodes. Due to this generation process, the degree distribution of the nodes in graphs generated by this model follows a power-law distribution (Barabási & Albert, 1999).

The configuration model is a method to create random graphs based on existing degree sequences (Newman et al., 2001). In this model, there are no other parameters apart from the degree sequence. First, an empty graph on  $n$  nodes is created ( $n$  is given by the degree sequence). Next, every node  $u$  receives degree( $u$ ) stubs (here, degree( $u$ ) is the desired degree of node  $u$ ). Finally, pairs of stubs are chosen and connected with equal probability. This procedure might result in graphs with multiple edges and loops. For our study, however, we ignore those (i.e., we work with the simple versions of these graphs).<sup>2</sup>

### 3. Experiments with empirical networks

In this section, we investigate, for empirical networks, the relationship between the robustness of centrality measures on the one side and the corresponding network size and average degree on the other. We consider the following centrality measures: degree, eigenvector centrality, and PageRank. Both the eigenvector centrality and the PageRank are feedback measures and fast to calculate (Koschützki et al., 2005). However, they have rarely been considered simultaneously in previous studies. Since PageRank can be very stable in scale-free networks (Ghoshal & Barabási, 2011), a comparison with the eigenvector centrality is therefore interesting, and we will see that both measures behave differently with regard to their robustness. For the calculation of the betweenness and the closeness for all nodes in a graph, the all-pairs shortest path problem has to be solved. The running time for that is at least quadratic (Brandes, 2001). Since the centrality values in the simulation part of this study (described in the following section) have to be recalculated numerous times, these measures are not considered.

#### 3.1 Experimental setup and data

For our empirical study, we use all the undirected and unweighted networks available through the Koblenz Network Collection (Kunegis, 2013) at the beginning of 2019. Hence, the networks used in this study can be seen as a random sample of networks that stem from different domains and therefore differ from each other in structure and size. The 24 real-world networks and descriptive statistics for them are listed in Table 1. As part of the data preprocessing, we have removed any existing loops. If a network consists of several components, we only consider the largest connected component and hence all networks are connected.

To analyze the effects of different errors on the robustness of centrality measures in the empirical networks, we use a simulation-based experimental procedure. An iteration of the experiment is performed as follows: starting from a network  $G$  (one of the 24 networks listed in Table 1), we apply the error mechanism with the intensity  $\alpha$ . The resulting modified network is called  $H$ . Finally, we calculate the robustness of the centrality measure  $c$ :  $r_{\tau,c}(G, H)$  (as defined in Section 2.3). We repeat this procedure 100 times and compute the mean and the standard deviation of the robustness for each network for all combinations of centrality measure (degree, eigenvector centrality, and PageRank), error mechanism (add edges, remove edges uniform, remove edges proportional to the edge degree, and remove nodes), and error level ( $\alpha \in \{0.1, 0.2, \dots, 0.5\}$ ).

#### 3.2 Observations for empirical networks

We start with the results aggregated across all networks. Similar to previous studies in this area (as discussed in Section 1), we observe that the robustness declines with an increasing level of error.

**Table 1.** Descriptive statistics about the largest connected components of the networks used in our study. The average degree is abbreviated as  $\langle d \rangle$ 

Name	Nodes	Edges	$\langle d \rangle$	Density	Transitivity	Source
zachary	34	78	4.6	1.4e-01	2.6e-01	(Zachary, 1977)
dolphins	62	159	5.1	8.4e-02	3.1e-01	(Lusseau et al., 2003)
pdzbase	161	209	2.6	1.6e-02	2.9e-03	(Beuming et al., 2005)
jazz	198	2,742	27.7	1.4e-01	5.2e-01	(Gleiser & Danon, 2003)
vidal	2,783	6,007	4.3	1.6e-03	3.5e-02	(Rual et al., 2005)
facebook	4,039	88,234	43.7	1.1e-02	5.2e-01	(Leskovec & Mcauley, 2012)
CA-GrQc	4,158	13,422	6.5	1.6e-03	6.3e-01	(Leskovec et al., 2007)
powergrid	4,941	6,594	2.7	5.4e-04	1.0e-01	(Watts & Strogatz, 1998)
reactome	5,973	145,778	48.8	8.2e-03	6.1e-01	(Joshi-Tope et al., 2005)
CA-HepTh	8,638	24,806	5.7	6.6e-04	2.8e-01	(Leskovec et al., 2007)
pgp	10,680	24,316	4.6	4.3e-04	3.8e-01	(Boguñá et al., 2004)
CA-HepPh	11,204	117,619	21.0	1.9e-03	6.6e-01	(Leskovec et al., 2007)
CA-AstroPh	17,903	196,972	22.0	1.2e-03	3.2e-01	(Leskovec et al., 2007)
CA-CondMat	21,363	91,286	8.5	4.0e-04	2.6e-01	(Leskovec et al., 2007)
deezer-RO	41,773	125,826	6.0	1.4e-04	7.5e-02	(Rozemberczki et al., 2019)
deezer-HU	47,538	222,887	9.4	2.0e-04	9.3e-02	(Rozemberczki et al., 2019)
deezer-HR	54,573	498,202	18.3	3.3e-04	1.1e-01	(Rozemberczki et al., 2019)
brightkite	56,739	212,945	7.5	1.3e-04	1.1e-01	(Cho et al., 2011)
livemocha	104,103	2,193,083	42.1	4.0e-04	1.4e-02	(Zafarani & Liu, 2009)
petster-cat	148,826	5,447,464	73.2	4.9e-04	1.1e-02	(Dünker & Kunegis, 2015)
douban	154,908	327,162	4.2	2.7e-05	1.0e-02	(Zafarani & Liu, 2009)
gowalla	196,591	950,327	9.7	4.9e-05	2.3e-02	(Cho et al., 2011)
dblp	317,080	1,049,866	6.6	2.1e-05	3.1e-01	(Yang & Leskovec, 2012)
petster-dog	426,485	8,543,321	40.1	9.4e-05	1.4e-02	(Dünker & Kunegis, 2015)

Therefore, we will subsequently focus on an error level of  $\alpha = 0.2$ , since the results for the other error levels yield the same conclusions and the impact of the error level is not our main objective.

When looking at the average across all networks (Table 2), degree centrality is always the most robust. For the removal error mechanisms, the PageRank is more robust than the eigenvector centrality. In the case of additional edges, the opposite effect can be observed. Regarding the standard deviation, the ranking is constant across all error types, degree centrality varies least, followed by PageRank. The robustness of the eigenvector centrality fluctuates the most, sometimes the standard deviation is two to three times as large as for the first mentioned measures. With regard to the effect of the type of measurement error on robustness, degree centrality and PageRank behave similarly. The absence of edges proportional to the edge degree has the weakest effect and spurious edges has the strongest. For eigenvector centrality, on the other hand, the first error type has the strongest influence on the robustness.

As we look at the relationship between robustness and global network measures, we notice that there are both large networks that are very sensitive to errors (e.g., douban) and small networks that are very robust (e.g., Jazz). The mean values of the robustness for every network are listed in Table 3.



**Table 2.** Mean and standard deviation of the robustness of centrality measures in empirical networks, aggregated over all networks in Table 1

Error mechanism	e+		e−		e−(p)		n−	
	Mean	Sd	Mean	Sd	Mean	Sd	Mean	Sd
Degree	0.85	0.06	0.89	0.05	0.91	0.05	0.90	0.05
Eigenvector	0.75	0.18	0.81	0.11	0.73	0.15	0.79	0.15
PageRank	0.73	0.07	0.82	0.07	0.86	0.07	0.83	0.07

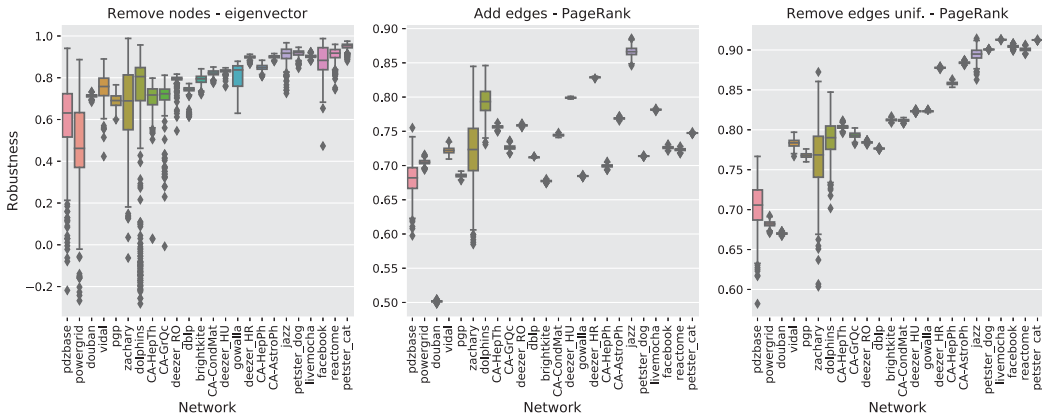
**Table 3.** Robustness of the centrality measures in empirical networks (mean values)

Centrality	Degree				Eigenvector				PageRank			
	e+	e−	e−(p)	n−	e+	e−	e−(p)	n−	e+	e−	e−(p)	n−
Network												
CA-AstroPh	0.89	0.94	0.97	0.94	0.83	0.92	0.89	0.90	0.77	0.88	0.94	0.89
CA-CondMat	0.86	0.90	0.93	0.90	0.79	0.86	0.81	0.82	0.74	0.81	0.88	0.83
CA-GrQc	0.84	0.88	0.94	0.89	0.60	0.77	0.62	0.71	0.73	0.79	0.88	0.81
CA-HepPh	0.82	0.93	0.98	0.93	0.63	0.87	0.92	0.85	0.70	0.86	0.96	0.86
CA-HepTh	0.85	0.88	0.92	0.88	0.63	0.76	0.71	0.70	0.76	0.80	0.86	0.81
brightkite	0.80	0.89	0.95	0.89	0.84	0.80	0.71	0.79	0.68	0.81	0.91	0.80
dblp	0.84	0.88	0.93	0.88	0.71	0.79	0.54	0.74	0.71	0.78	0.84	0.80
deezer_HR	0.91	0.93	0.95	0.93	0.84	0.92	0.89	0.90	0.83	0.88	0.92	0.88
deezer_HU	0.89	0.90	0.91	0.90	0.81	0.85	0.80	0.83	0.80	0.82	0.85	0.83
deezer_RO	0.87	0.88	0.90	0.88	0.77	0.81	0.65	0.79	0.76	0.78	0.84	0.78
dolphins	0.86	0.85	0.85	0.86	0.69	0.79	0.74	0.68	0.79	0.79	0.79	0.80
douban	0.66	0.78	0.82	0.78	0.93	0.71	0.71	0.71	0.50	0.67	0.76	0.71
facebook	0.93	0.96	0.97	0.96	0.47	0.91	0.85	0.88	0.73	0.90	0.92	0.91
gowalla	0.82	0.91	0.94	0.91	0.86	0.81	0.58	0.81	0.68	0.82	0.91	0.83
jazz	0.92	0.93	0.92	0.94	0.91	0.92	0.91	0.91	0.87	0.89	0.90	0.92
livemocha	0.88	0.95	0.96	0.95	0.95	0.90	0.86	0.90	0.78	0.91	0.92	0.92
pdzbase	0.81	0.82	0.84	0.83	0.76	0.61	0.54	0.60	0.68	0.70	0.73	0.71
petster_cat	0.88	0.94	0.85	0.96	0.98	0.87	0.60	0.95	0.75	0.91	0.80	0.91
petster_dog	0.87	0.94	0.90	0.94	0.97	0.92	0.64	0.92	0.71	0.90	0.90	0.91
pgp	0.80	0.87	0.94	0.87	0.67	0.73	0.69	0.69	0.69	0.77	0.88	0.78
powergrid	0.80	0.79	0.80	0.79	0.13	0.59	0.58	0.50	0.70	0.68	0.70	0.69
reactome	0.89	0.96	0.97	0.96	0.59	0.93	0.91	0.91	0.72	0.90	0.94	0.91
vidal	0.83	0.87	0.90	0.87	0.87	0.76	0.63	0.75	0.72	0.78	0.84	0.78
zachary	0.80	0.83	0.85	0.85	0.76	0.67	0.64	0.67	0.72	0.76	0.79	0.80

In the following, we discuss the relationship between global network measures and robustness in more detail. Table 4 lists the rank correlation between the average robustness and the respective values for the global network measures. For all removal error types, the robustness tends to be higher with increasing average degree. We observe almost perfect correlation for cases where edges or nodes are missing uniformly at random and still high correlation values when edges are

**Table 4.** Empirical networks: rank correlation between global measures and the average robustness

Centrality	Degree				Eigenvector				PageRank			
	e+	e−	e−(p)	n−	e+	e−	e−(p)	n−	e+	e−	e−(p)	n−
Avg. degree	0.77	0.97	0.63	0.98	0.27	0.92	0.52	0.93	0.43	0.96	0.72	0.95
Density	0.18	0.01	0.01	0.03	−0.38	0.02	0.30	−0.11	0.26	0.07	0.00	0.15
Network size	0.02	0.31	0.16	0.29	0.52	0.26	−0.14	0.43	−0.18	0.26	0.22	0.18
Transitivity	0.26	0.23	0.58	0.23	−0.63	0.27	0.50	0.04	0.22	0.16	0.49	0.23



**Figure 1.** Illustrative examples for the three different behaviors of the robustness in empirical networks. The networks are sorted their average degree (ascending). The median robustness is indicated in each box, and whiskers are 1.5 times the interquartile range.

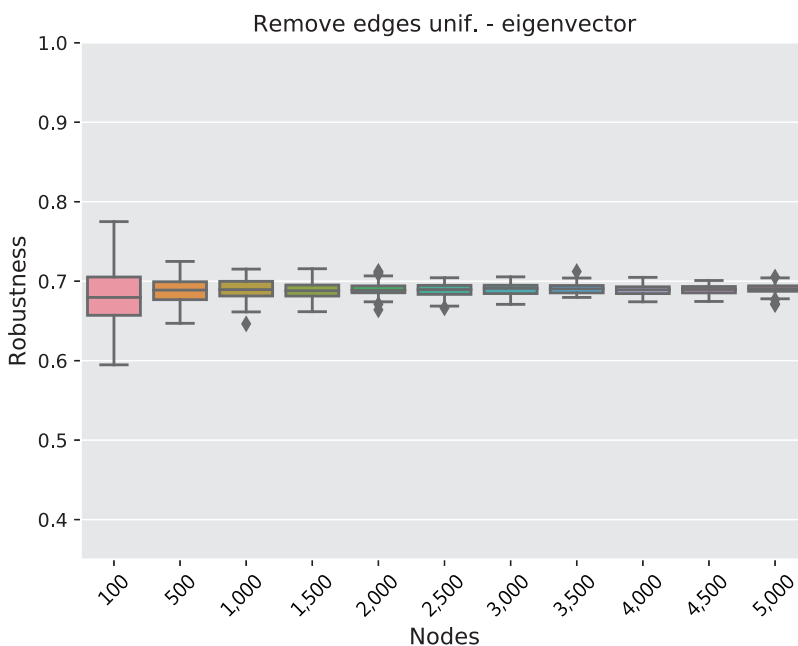
missing proportional. For the degree centrality, the correlation is also high for the case of spurious edges. For PageRank and eigenvector centrality, this is, however, not the case. While for the transitivity, a moderate correlation with the robustness can still be observed, the number of nodes as well as the density are, in most cases, basically uncorrelated with the robustness. This observation may come rather unexpected since growing networks often show “densification,” which means the average degree grows with the number of nodes (Leskovec et al., 2007).

Figure 1 shows the behavior of robustness for three groups in each panel in an exemplary fashion (results for all combinations are shown in Figure A1 in the appendix). The robustness of the eigenvector centrality in case of missing nodes is depicted in the first panel. There is a recognizable association, but in this case the variance is higher than in most other cases. The same effect can be observed with the eigenvector centrality also in connection with missing edges (uniformly). The middle panel shows the observation for PageRank and add edges. This behavior is typical for all centrality measures under the influence of additional edges, there is no obvious pattern. The last panel shows the combination of PageRank and missing edges uniform. In this case, the relationship between average degree and robustness is most prominent. Robustness is higher when the average degree is higher. The variance of robustness is low. This behavior occurs for PageRank and degree for all cases of missing edges (uniform and proportional) and missing nodes.

**4. Experiments with random graphs**

In Section 3, we examined the robustness of 24 empirical networks from different domains. We observed that there exist small and robust as well as large and nonrobust networks with respect to centrality measures. In addition, we have analyzed the relationship between the robustness of





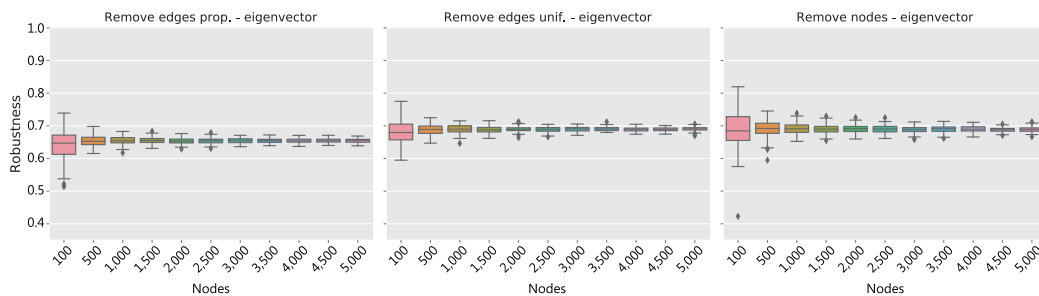
**Figure 2.** The behavior of the robustness in ER graphs with increasing network size (abscissa) for fixed average degree is shown in the figure above. The error level is 0.2.

centrality measures in these networks with different global network measures. We observed that there is little association between network size and robustness. We found, however, that in many cases the higher the average degree of the network, the higher the robustness. To study this effect in more detail, we conduct further experiments in this chapter. We use different random graph models to control the average degree and to measure the effects of its change on robustness. For this purpose, we choose two different perspectives. In Section 4.1, we keep the average degree constant and increase the size of the network. In Section 4.2, we control the average degree while keeping the network size constant.

#### 4.1 Experiments with constant average degree

In this section, we use the ER model and the BA model to investigate the behavior of the robustness when the average degree is fixed while the network size increases. The experimental setup is similar to that of Section 3.1. Instead of using empirical networks, however, we generate ER and BA graphs with an average degree of 10 and a network size  $n \in (100, 500, 1,000, 1,500, \dots, 10,000)$ , which we call  $G$ . Then, we apply the error mechanism with the intensity  $\alpha$  to  $G$  which results in the erroneous network  $H$  and calculate the robustness of the centrality measure  $c$ :  $r_{\tau,c}(G, H)$ . We repeat this procedure 100 times for the two random graph models and the varying values for the network size for all combinations of centrality measure (degree, eigenvector centrality, and PageRank), error mechanism (add edges, remove edges uniform, remove edges proportional to the edge degree, and remove nodes), and error level ( $\alpha \in \{0.1, 0.2, \dots, 0.5\}$ ).

The results for the ER graphs are very homogeneous, and for all centrality measures and error mechanisms we observe the same behavior: robustness does not change with increasing network size. However, the variance decreases with increasing network size. We limit the discussion to an error level of  $\alpha = 0.2$  since the results for the other error levels are conclusive with these results. Figure 2 is symptomatic for all other cases. It shows the robustness (ordinate) of the eigenvector centrality when nodes are missing, the network size is shown on the abscissa. It is noticeable here



**Figure 3.** Results for the robustness of centrality measures in BA graphs. Here, the network size increases while the average degree remains constant, and the error level is 0.2. For the network size values up to 5,000 are shown for better readability; for larger values, hardly any changes occur.

that the variance decreases sharply with the first increases in network size. Above a size of approx. 2,000, the change is hardly visible.

For the BA graphs, we observe, with two exceptions, the same behavior as for the ER graphs. Figure 3 shows the three different characteristics (all results in this figure are for the eigenvector centrality). The middle panel in Figure 3 represents the robustness behavior in BA graphs in almost all cases. The robustness is independent of the network size, only the variance decreases with increasing size, whereas the variance is relatively small already. The outer panels show the two exceptional cases. The absence of edges proportional to the degree of edge (left panel) reduces the robustness of the eigenvector centrality with increasing network size. If nodes are missing (right panel), the robustness is, as in most other cases, independent of the network size, but the variance is much larger and declines hardly with increasing size.

#### 4.2 Experiments with increasing average degree

In this part, we analyze the impact of changes to the average degree on the robustness of centrality measures. We conduct two types of experiments, based on ER and BA graphs and based on the configuration model. The procedure for the first experiments is similar to the experimental setup described in Section 4.1. The difference between these experiments is the generation of the random graphs. For the experiments in this section, we fix the network size at  $n = 1,000$  and select the parameters  $p$  and  $m$  in such a way that we obtain networks with an average degree between 4 and 100.

With the second type of experiments, we examine the effects of changes in the average degree on the robustness of centrality measures in a more realistic setting. We use the degree sequences of the empirical networks from Section 3. We “scale” the degree sequences and generate more dense versions of the underlying networks using the configuration model. We use these generated networks as input graph  $G$  to analyze the robustness of centrality measures in these types of networks, and the remaining procedure is analogous to the experiments already described in this section.

For the results concerning the ER graphs, the pattern consists of two parts, independent of the type of error. The robustness of the eigenvector centrality and the PageRank is, except for the smallest initial increases, constant and thus independent of the increase in the average degree. With degree centrality, on the other hand, robustness decreases with increasing average degree. The decreases occur especially during the initial increases in the average degree (approx. the range between 4 and 25), here the robustness decreases by 0.1. These observations can also be found in the rank correlation between average degree and the robustness (Table 5). While degree centrality here always shows strongly negative correlation, in most other cases no or weakly negative correlation can be observed.

**Table 5.** The rank correlations between the average degree and the robustness are listed for the cases of BA and ER graphs under the influence of different error mechanisms with an error level of 0.2

Error mechanism	BA graphs				ER graphs			
	e+	e−	e−(p)	n−	e+	e−	e−(p)	n−
Centrality								
Degree	0.91	0.92	0.90	0.87	−0.66	−0.64	−0.64	−0.51
Eigenvector	0.82	0.93	0.95	0.69	−0.07	0.02	0.04	0.15
PageRank	0.93	0.94	0.93	0.91	−0.33	−0.25	−0.41	−0.02

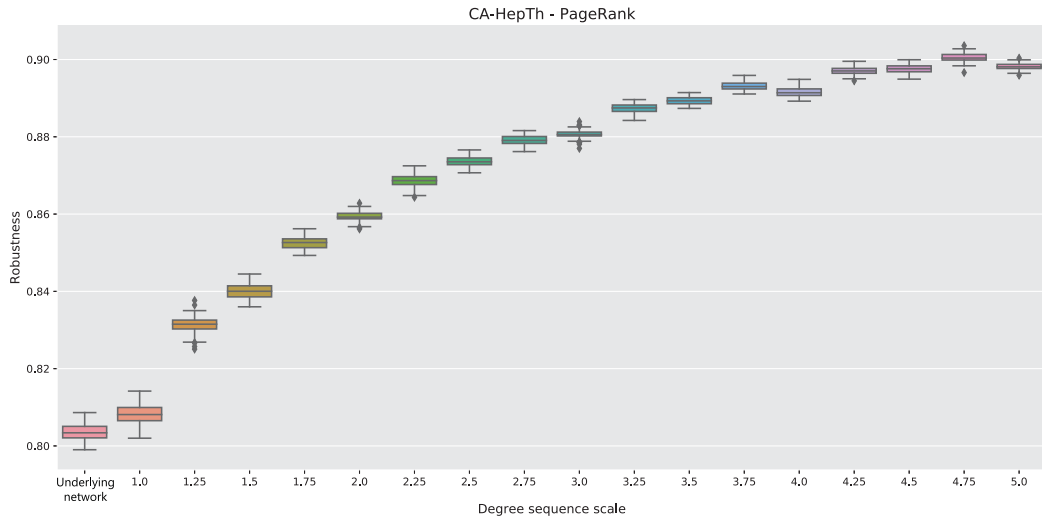
The results for the BA graphs show a consistent pattern. In all cases, regardless of centrality measure and error type, a higher average degree is accompanied by a higher robustness. There is a very high, positive rank correlation between the average degree and the associated robustness (see Table 5). The increases in robustness associated with the increase in the average degree are particularly strong for initial increases, and further increases still have a positive effect on the robustness, but this effect diminishes. The only exception to this is the eigenvector centrality, which resembles a linear relationship. For networks with a skewed degree distribution, like BA graphs, increasing the average degree results in the existence of strong hubs. This effect explains, at least partially, the high correlation between average degree and robustness (Ghoshal & Barabási, 2011). The variance is slightly higher for the error type missing nodes than for the other error types. In the case of eigenvector centrality, the variance is much higher in this case.

In the previous section, we have observed that for BA graphs, a higher average degree is associated with higher robustness. Although BA graphs have a skewed degree distribution, a property that many empirical networks also have, these networks are nevertheless otherwise rather artificial. Therefore, we now make the previous experiment a little more realistic. For this purpose, we use the degree sequences of empirical networks and manipulate them to increase the average degree and generate networks in order to analyze the robustness of centrality measures in these networks.

The experiment's basic design is similar to that of the experiments in the previous section. The difference lies in the way the networks are created. We take the degree sequences of the empirical networks (Table 1) and create several “scaled” versions of them and generate networks based on these degree sequences using the configuration model. To scale a degree sequence of a network  $G$ , we take a factor  $s$  and multiply each entry  $ds(G)_i$  of the degree sequence  $ds(G)$  by this factor. If  $ds(G)_i \cdot s$  is not natural number, we take the integer part of it and add 1 with the probability of the fractional part. For example, if the original degree is 9 and  $s = 1.25$ , then the scaled degree is  $11 + \text{Bern}(0.25)$  where  $\text{Bern}$  is the Bernoulli distribution. We scale the degree with factors between 1 and 5 ( $s \in (1.0, 1.25, 1.5, \dots, 5)$ ) and repeat the whole procedure 100 times for every combination. We also calculate the robustness of the underlying networks (the networks from which we obtain the degree sequences).

First, we compare the robustness of centrality measures in “unscaled” random graphs (i.e.,  $s = 1$ ) with the robustness of the corresponding empirical network. In the case of the degree centrality, the robustness values of both networks are, in the vast majority of cases and regardless of the type of error, similar. This is, however, not the case the eigenvector centrality and the PageRank. For these measures, the robustness of the random graph and the underlying network are only similar in about 50% of the cases. We observe no clear pattern which could explain this behavior.

The characteristic behavior regarding the influence of the scaling factor  $s$  on the robustness of centrality measures in the random graphs is shown in Figure 4. The robustness increases with increasing scaling factor (i.e., for networks with higher average degree). The initial increases of the scaling factor have a larger influence on the robustness than the subsequent increases. The rank correlations between the scaling factor of the degree sequence and the robustness of centrality



**Figure 4.** Configuration model results for one network (CA-HepTh). The scaling factor is listed on the abscissa. Additionally, the first entry is the robustness of the underlying network.

measures w.r.t. random graphs with these degree sequences for all networks, error mechanisms, and centralities are listed in Table 6. For most cases, we observe very high correlations. Multiple realizations of the same experiment yield similar results (i.e., the variance is low). This observation strongly suggests that the robustness is a property of the degree sequence. Notable exceptions are the results for smaller networks (zachary, dolphins, jazz, and pdzbase). In these cases, the correlations are lower and the variances are higher compared to the other networks. We suspect that this instability is related to the size of these networks, and they all contain less than 200 nodes.

Among the remaining (larger) networks, it is noticeable that the correlations for the douban network are considerably lower. This might be due to the structure of the douban network has a low average degree (4.2) and approx.  $2/3$  of the nodes have a degree of one. Hence, there is little structure implied by its degree sequence.

## 5. Analysis and discussion of the degree centrality

In our experiments in Sections 3 and 4, we observed that, in many cases, not the network size but the average degree is correlated with the robustness of centrality measures. An increasing average degree often leads to a higher robustness, which was observed in the experiments with the empirical networks and also in the experiments with graphs generated by the BA and configuration model. However, the opposite effect was also observed. In the case of ER graphs, the robustness of the degree centrality decreases with an increasing average degree. In this section, we take a more detailed look at the two scenarios involving ER and BA graphs.

In the following, we focus on the degree centrality, as it is the most accessible for an analytical perspective (Platig et al., 2013; Tsugawa & Ohsaki, 2015; Murai & Yoshida, 2019). First (Section 5.1), we derive an expression for the robustness based on the interpretation of the robustness as a probability. Then, we argue that this interpretation is closely related to the robustness measured with Kendall's  $\tau$  but more accessible for an analytical perspective. The exact expression for the robustness depends on the type of error. In Section 5.2, we derive these expressions for the case of ER graphs and use them in Section 5.3 to show that, for ER graphs and sufficiently large network size, the robustness is independent of the network size, as long as the average degree is

**Table 6.** Configuration model: rank correlations between the scaling factor of the degree sequence and the average robustness of centrality measures w.r.t. random graphs with these degree sequences

Centrality Error Network	Degree				Eigenvector				PageRank			
	e+	e−	e−(p)	n−	e+	e−	e−(p)	n−	e+	e−	e−(p)	n−
zachary	0.60	0.75	0.46	0.81	0.81	0.90	0.96	0.76	0.84	0.87	0.84	0.91
dolphins	0.88	0.79	0.82	0.94	0.96	0.97	0.96	0.96	0.96	0.91	0.90	0.94
jazz	0.66	0.38	0.26	0.96	0.72	0.66	0.50	0.96	0.82	0.78	0.56	0.96
pdzbase	0.63	0.34	−0.31	0.32	0.62	0.96	0.96	0.75	0.66	0.63	0.41	0.65
vidal	0.85	0.79	0.18	0.76	1.00	1.00	1.00	0.97	0.91	0.87	0.78	0.84
facebook	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
powergrid	0.81	0.82	0.85	0.82	1.00	1.00	1.00	1.00	0.78	0.85	0.85	0.85
CA-GrQc	0.87	0.87	0.84	0.85	1.00	1.00	1.00	1.00	0.99	0.97	0.97	0.97
reactome	0.96	0.91	0.91	0.91	1.00	1.00	1.00	1.00	1.00	1.00	0.99	1.00
CA-HepTh	0.87	0.87	0.84	0.87	1.00	1.00	1.00	1.00	0.99	0.97	0.96	0.97
pgp	0.84	0.69	−0.15	0.74	1.00	1.00	1.00	0.97	0.91	0.82	0.56	0.82
CA-HepPh	0.94	0.88	0.82	0.88	1.00	1.00	1.00	0.99	1.00	1.00	0.99	1.00
CA-AstroPh	0.97	0.93	0.93	0.93	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
CA-CondMat	0.91	0.91	0.90	0.91	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
deezer_RO	0.88	0.90	0.91	0.91	1.00	1.00	1.00	1.00	1.00	0.99	0.99	0.99
deezer_HU	0.93	0.93	0.93	0.93	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
deezer_HR	0.97	0.96	0.96	0.96	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
brightkite	0.85	0.72	−0.25	0.72	1.00	1.00	1.00	0.96	1.00	0.87	0.84	0.88
livemocha	0.97	0.93	0.91	0.91	1.00	1.00	1.00	1.00	1.00	0.99	0.96	0.99
petster_cat	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.97	1.00	1.00	1.00	1.00
douban	0.34	−0.57	−0.69	−0.57	−0.51	0.87	0.62	0.82	0.57	−0.10	−0.69	−0.12
gowalla	0.90	0.82	0.34	0.84	1.00	1.00	0.99	0.81	1.00	0.96	0.90	0.94
dblp	0.87	0.87	0.85	0.87	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
petster_dog	0.97	0.96	0.93	0.96	1.00	1.00	1.00	0.99	1.00	1.00	1.00	1.00

constant. In Section 5.4, we analyze the behavior of the robustness of the degree centrality when the average degree is increasing in ER and BA graphs in more detail.

**5.1 Analytical approach for the robustness**

To analyze the degree centrality in more detail, we use the following terms:  $G$  is the unmodified graph and  $H$  is the erroneous graph (a “modified” version of  $G$ , i.e.,  $H$  is on the same node set as  $G$  or on a subset of that node set).

We used Kendall’s  $\tau$  to measure the robustness in our experiments to provide comparability to existing studies. Goodman and Kruskal’s rank correlation coefficient  $\gamma$  (Goodman & Kruskal, 1954), as explained in Section 2.3, allows us to develop an analytical approach. Therefore, we will use it in the remainder of this section. Note that both measures differ in their definitions only

when there are ties. We can rewrite the robustness of the degree centrality with respect to  $G$  and  $H$  (as stated in Equation (2)) in terms of the probability of concordant ( $P_c$ ) and discordant ( $P_d$ ) pairs:

$$\begin{aligned}
 r_{\gamma, \text{degree}}(G, H) &= \frac{n_c - n_d}{n_c + n_d} = \frac{2n_c - (n_c + n_d)}{n_c + n_d} \\
 &= 2 \frac{n_c}{n_c + n_d} - 1 = 2 \frac{\frac{n_c}{n}}{\frac{n_c}{n} + \frac{n_d}{n}} - 1 \\
 &= 2 \frac{P_c}{P_c + P_d} - 1.
 \end{aligned}
 \tag{3}$$

Now we will derive how these probabilities and thus also the robustness can be calculated.

The error level is denoted by  $\alpha$ ; from the context, it becomes apparent whether this refers, for example, to the level of deleted nodes or edges. Additionally, consider two nodes  $v_1$  and  $v_2$  drawn randomly from  $V(H)$ . Now, let  $D_i$  denote the random variable for the degree of node  $v_i$  in  $G$  and  $X_i$  denote the random variable for the degree change of node  $v_i$  (i.e., the difference of the degree of node  $v_i$  in  $G$  and in  $H$ ),  $D_1, X_1$  and  $D_2, X_2$  are independent and identically distributed (i.i.d). On this basis, let  $P(D_1 = d_1, X_1 = x_1, D_2 = d_2, X_2 = x_2)$  be the joint probability that specific values for  $d_1, x_1, d_2, x_2$  occur together. We abbreviate this by  $P(d_1, x_1, d_2, x_2)$ .

Summing  $P(d_1, x_1, d_2, x_2)$  over the quadruples that correspond to concordant (discordant) pairs of nodes, we can calculate the probability for  $(v_1, v_2)$  to be concordant (discordant). For example, for the case of missing edges, the probability for  $(v_1, v_2)$  to be concordant is

$$P_c = \sum_{\substack{d_1 < d_2; d_1 - x_1 < d_2 - x_2 \\ d_1 > d_2; d_1 - x_1 > d_2 - x_2}} P(d_1, x_1, d_2, x_2),
 \tag{4}$$

and the probability for  $(v_1, v_2)$  to be discordant is

$$P_d = \sum_{\substack{d_1 < d_2; d_1 - x_1 > d_2 - x_2 \\ d_1 > d_2; d_1 - x_1 < d_2 - x_2}} P(d_1, x_1, d_2, x_2).
 \tag{5}$$

The robustness as defined in Equation (3) is thus a function of the probabilities defined in Equations (4) and (5).

### 5.2 Expressions for the robustness and error types

In the previous section, we showed how the robustness can be expressed in terms of the probability of pairs to be concordant or discordant. These probabilities depend on the type of error and the degree distribution of the graph. In this section, we derive explicit expressions for  $P(d_1, x_1, d_2, x_2)$  for the case of missing edges, missing nodes, and additional edges in ER graphs with  $n$  nodes and edge probability  $p$ .

Since  $D_1, D_2$  and  $X_1, X_2$  are i.i.d., we can express  $P(d_1, x_1, d_2, x_2) = P(d_1, x_1) \cdot P(d_2, x_2)$ . To derive the actual expression for  $P(d_i, x_i)$ , we use the fact that  $P(d_i, x_i) = P(x_i | d_i) \cdot P(d_i)$ , where  $P(d_i)$  is the probability that node  $v_i$  has a degree of  $d_i$ . For ER graphs, this is the binomial distribution  $D_i \sim \text{Bin}(n, p)$ . Independent of the specific error mechanism,  $P(x_i | d_i)$  describes the effects of that error on node  $v_i$ . Hence, to create explicit expressions to that allow us to calculate the actual robustness, we solely need to define  $P(x_i | d_i)$  for the corresponding type of graph and error. In the following, we provide these for the case of missing edges, missing nodes, and additional edges in ER graphs.

**Missing edges.** For the case of missing edges,  $P(x_i | d_i)$  is the probability that  $x_i$  edges are removed from a node with degree  $d_i$ . This is binomially distributed with  $P(X_i = x_i | D_i = d_i) \sim \text{Bin}(d_i, \alpha)$ . The fraction of missing edges is denoted by  $\alpha$ . The restrictions for the quadruples which are used for the calculation of robustness are the same as in Equations (4) and (5).



- The marginal distribution is as follows:

$$P(D_i = d_i, X_i = x_i) = :P(d_i, x_i) = \binom{d_i}{x_i} \alpha^{x_i} (1 - \alpha)^{d_i - x_i} \binom{n}{d_i} p^{d_i} (1 - p)^{n - d_i}, \tag{6}$$

for  $i \in \{1, 2\}$ .

- As we assume that the edges are deleted independently of each other,

$P(d_1, x_1, d_2, x_2) = P(d_1, x_1) \cdot P(d_2, x_2)$ , and thus calculate the robustness of the degree centrality  $r_{\gamma, \text{degree}}(G, H)$  with

$$P_c = \sum_{\substack{d_1 < d_2; d_1 - x_1 < d_2 - x_2 \\ d_1 > d_2; d_1 - x_1 > d_2 - x_2}} P(d_1, x_1) \cdot P(d_2, x_2)$$

and

$$P_d = \sum_{\substack{d_1 < d_2; d_1 - x_1 > d_2 - x_2 \\ d_1 > d_2; d_1 - x_1 < d_2 - x_2}} P(d_1, x_1) \cdot P(d_2, x_2).$$

**Missing nodes.** In the case of missing nodes, the restrictions for the quadruples, which are used for the calculation of robustness and the degree distribution, are the same as above. The error level  $\alpha$  is the fraction of nodes that are missing in  $H$ . For the conditional distribution of the degree decrease  $P(x_i | d_i)$ , we note that:  $n\alpha$  nodes are deleted,  $n(1 - \alpha)$  nodes are not deleted, and  $d_i$  is the degree of node  $v_i$ —the number of neighbors “drawn” from the set of  $n$  nodes (actually  $n - 1$ , but for large  $n$ , the difference becomes negligible).

- With this, we can specify the distribution of  $P(x_i | d_i)$  as a  $HGeom(n\alpha, n(1 - \alpha), d_i)$ :

$$P(x_i | d_i) = \binom{n\alpha}{x_i} \frac{\binom{n(1-\alpha)}{d_i - x_i}}{\binom{n}{d_i}}, \text{ for } i \in \{1, 2\}. \tag{7}$$

- Hence, with  $n' = n\alpha$ :

$$P(d_i, x_i) = P(x_i | d_i) P(d_i) = \binom{n'}{x_i} \frac{\binom{n-n'}{d_i - x_i}}{\binom{n}{d_i}} \binom{n}{d_i} p^{d_i} (1 - p)^{n - d_i}, \text{ for } i \in \{1, 2\}. \tag{8}$$

- Similar to the case of missing edges, we can use the fact that  $P(d_1, x_1, d_2, x_2) = P(d_1, x_1) \cdot P(d_2, x_2)$  and thus calculate the robustness of the degree centrality  $r_{\gamma, \text{degree}}(G, H)$  with

$$P_c = \sum_{\substack{d_1 < d_2; d_1 - x_1 < d_2 - x_2 \\ d_1 > d_2; d_1 - x_1 > d_2 - x_2}} P(d_1, x_1) \cdot P(d_2, x_2)$$

and

$$P_d = \sum_{\substack{d_1 < d_2; d_1 - x_1 > d_2 - x_2 \\ d_1 > d_2; d_1 - x_1 < d_2 - x_2}} P(d_1, x_1) \cdot P(d_2, x_2).$$

**Additional edges.** While, in the case of additional edges, the degree distribution is still the same as above,  $x_i$  now refers to the degree increase in node  $v_i$  and  $\alpha$  is the fraction of edges added to the graph.

- The conditional distribution for the degree increase is  $Bin(n - d_i, \alpha \frac{p}{1-p})$  and hence,

$$P(d_i, x_i) = \binom{n}{d_i} p^i (1-p)^{(n-d_i)} \binom{n-d_i}{x_i} \left(\alpha \frac{p}{1-p}\right)^{x_i} \left(1 - \alpha \frac{p}{1-p}\right)^{(n-d_i-x_i)}, \text{ for } i \in \{1, 2\}. \tag{9}$$

- Again, we can use the fact that  $P(d_1, x_1, d_2, x_2) = P(d_1, x_1) \cdot P(d_2, x_2)$  and thus calculate the robustness of the degree centrality  $r_{\gamma, \text{degree}}(G, H)$ . It is important to note, however, that the conditions for the summations change in the case of additional edges:

$$P_c = \sum_{\substack{d_1 < d_2; d_1+x_1 < d_2+x_2 \\ d_1 > d_2; d_1+x_1 > d_2+x_2}} P(d_1, x_1) \cdot P(d_2, x_2), \tag{10}$$

and

$$P_d = \sum_{\substack{d_1 < d_2; d_1+x_1 > d_2+x_2 \\ d_1 > d_2; d_1+x_1 < d_2+x_2}} P(d_1, x_1) \cdot P(d_2, x_2). \tag{11}$$

### 5.3 The case of constant average degree

In the following, we use the expressions developed in the previous section to study the impact of increasing network size (while the average degree is constant) on the degree centrality in more detail. For the case of missing edges and additional edges, we prove that the robustness of the degree centrality is independent of the network size.

**Missing edges.** For the case of missing edges, we derived the expression for  $P(d_i, x_i)$  in Equation (6). The degree distribution can be approximated by an exponential distribution with  $\lambda = np$ , hence  $P(d_i) = \frac{e^{-\lambda} \lambda^{d_i}}{d_i!}$ . If we replace the corresponding term in Equation (6), then:

$$P(d_i, x_i) = \binom{d_i}{x_i} \alpha^{x_i} (1 - \alpha)^{d_i-x_i} \frac{e^{-\lambda} \lambda^{d_i}}{d_i!}, \text{ for } i \in \{1, 2\}. \tag{12}$$

In Equation (12), the probability and hence robustness (Equation (3)) do not directly depend on  $n$ , as long as  $\lambda = np$  is constant (which implies that the average degree stays constant), and the robustness does not change if the network size increases. This shows that the robustness of the degree centrality in ER graphs for this case does not directly depend on the network size. For arbitrary network size, there exist robust and non-robust networks. This also explains the observations from the experiments conducted in Section 4.1.

**Additional edges.** For the case of additional edges, we derived the expression for  $P(d_i, x_i)$  in Equation (9). With  $q = \alpha \frac{p}{1-p}$ , we can rewrite Equation (9) as:

$$P(d_i, x_i) = Bin(n, p; d_i) \cdot Bin(n, q; x_i) \cdot \text{correction term, for } i \in \{1, 2\}, \tag{13}$$

where the correction term is

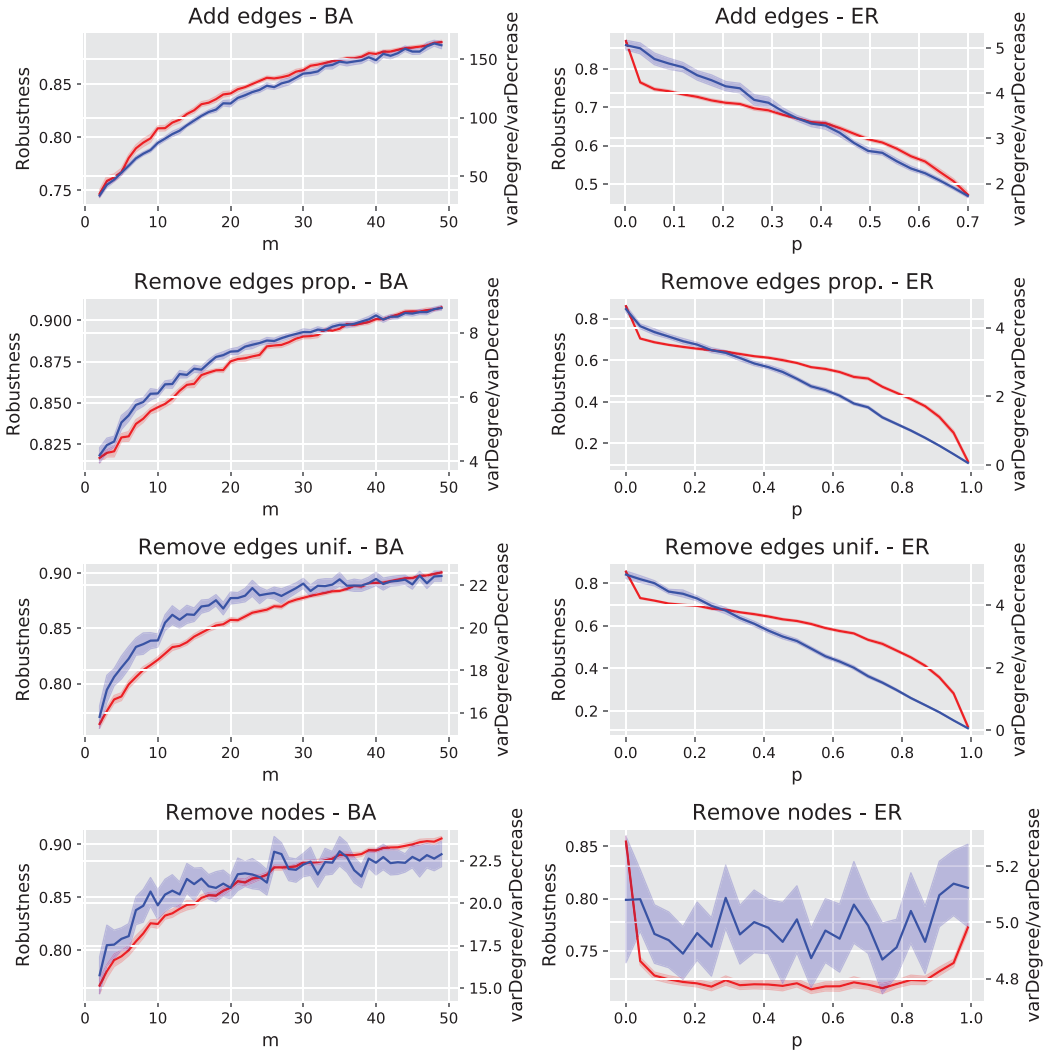
$$\frac{(n - d_i)(n - d_i - 1) \cdots (n - d_i - x_i + 1)}{n(n - 1) \cdots (n - x_i + 1)} (1 - q)^{-d_i} \tag{14}$$

and converges to 1.

Finally, we can apply the Poisson approximation ( $\lambda = np$ ) to Equation (13), so the probability no longer depends on the network size:

$$P(d_i, x_i) \approx Pois(\lambda; d_i) \cdot Pois(\alpha \lambda; x_i), \text{ for } i \in \{1, 2\}. \tag{15}$$

This result shows that, also for the case of additional edges, the robustness of the degree centrality in ER graphs does not depend on the network size for large networks.



**Figure 5.** Robustness of the degree centrality (red curve, axis labels on the left side) and ratio between the variance of the degree and the variance of the degree change (blue curve, axis labels on the right side).

**5.4 The case of increasing average degree**

In the following, we take a closer look at the observations of Section 4.2, the behavior of the robustness of the degree centrality with increasing average degree. In Section 4.2, we observed that the robustness of the degree centrality in ER graphs decreases with increasing average degree. For BA graphs, in contrast, the robustness increased with increasing average degree. To further explore this observation, recall that a pair of nodes is discordant, if the degree change induced by the error is larger than the degree difference between those nodes in the error-free graph. We therefore suspect that the ratio of the variance of degree values to the variance of degree changes could explain the robustness.

In order to investigate this, we have repeated the experiments regarding the degree centrality from Section 4.2 and recorded the changes of the degree values. Figure 5 shows the robustness of the degree centrality dependent on the parameter that controls the average degree of the ER and BA graphs (red curve). In addition, the ratio between variance of degree and variance of degree

change (blue curve) is also shown. For both types of random graphs and all four error mechanisms, these results show a consistent pattern. If this ratio is increasing, the robustness also rises, and if it falls, the robustness also falls.

Next, we consider three cases: missing nodes, missing edges, and additional edges, for which we argue why the ratio of the variances for the case of ER graphs behaves like this. Recall,  $D \sim \text{Bin}(n, p)$  is the degree distribution for the ER graph, and  $X$  is the random variable for the degree change.

**Missing edges.** For the case of missing edges,  $\alpha$  is the probability for edge deletion and  $X \sim \text{Bin}(n, p\alpha)$ . Let  $f(p) = \text{Var}(D)/\text{Var}(X)$  be the ratio of variance of the degree to variance of the degree change, then:

$$f(p) = \frac{np(1-p)}{np\alpha(1-p\alpha)}, \text{ hence} \tag{16}$$

$$f'(p) = \frac{\alpha-1}{\alpha} \cdot \frac{1}{(1-p\alpha)^2}. \tag{17}$$

Since the first part in Equation (17) is  $< 0$  and its second part is  $> 0$ , the quotient of  $\text{Var}(D)/\text{Var}(X)$  is (strictly) monotonically decreasing.

**Additional edges.** For the case of additional edges, the total number of edges added is  $\alpha p \binom{n}{2}$  and the total number of edges that could be added to the graph is  $\binom{n}{2}(1-p)$ . Hence, the probability for a nonexistent edge to be added is  $\frac{\alpha p \binom{n}{2}}{\binom{n}{2}(1-p)} = \alpha \frac{p}{1-p}$ . Therefore,  $X$ , the degree change of a node, is distributed as follows:  $X \sim \text{Bin}(n - D, \alpha \frac{p}{1-p})$ . Now, let  $D' \sim \text{Bin}(n, 1-p)$ , then  $X \sim \text{Bin}(D', \alpha \frac{p}{1-p})$  and thus  $X \sim \text{Bin}(n, p\alpha)$ , which is the same distribution as for the case of missing edges. Consequently, the same results as derived above for the case of missing edges also hold for the case of additional edges.

**Missing nodes.** For the case of missing nodes, the degree change  $X$  is distributed as follows:  $X \sim \text{Bin}(\alpha n, p)$ . The ratio of variance of the degree to the variance of the degree change for this case is as follows:

$$\frac{\text{Var}(D)}{\text{Var}(X)} = \frac{np(1-p)}{\alpha np(1-p)} = \frac{1}{\alpha}, \tag{18}$$

which is a constant function with respect to  $p$ , in contrast to Equation (16).

Equations (17) and (18) demonstrate the difference between the edge error mechanisms (missing and additional edges) and the case of missing nodes: in the latter, the quotient  $\text{Var}(D)/\text{Var}(X)$  is a constant which explains the behavior of the degree centrality if nodes are missing in ER graphs which we observe in our experiments, see bottom right panel in Figure 5. In this case, the robustness does not change with increasing  $p$  (and thus increasing average degree). For the case of missing edges and additional edges, it is plausible that the robustness is related to the quotient  $\text{Var}(D)/\text{Var}(X)$ . If the variance of the degree change increases more strongly than the variance of the degree, the probability that concordant pairs become discordant pairs also increases. This illustrates the behavior shown in the first and third panel on the right-hand side of Figure 5.

## 6. Conclusions and final remarks

Network data are often erroneous, which compromises centrality measures and the conclusions of subsequent analyses. We investigated the robustness behavior of the degree, eigenvector

centrality, and the PageRank in empirical networks of different size and structure, as well as in random graphs, under the influence of missing and additional edges and missing nodes. We were primarily interested in the relationship between the robustness of these centrality measures and the network size, respectively, the average degree. We observed that a higher average degree was frequently associated with higher robustness for cases where nodes or edges are missing. Additionally, the degree was always the most robust measure and the variance of the eigenvector centrality was, in most cases, substantially higher than the variance of the degree centrality or the PageRank. Moreover, we observed that there exist small networks that are robust and larger networks that are not robust w.r.t. centrality measures. These results also demonstrate that the study of the robustness of centrality measures in the context of larger networks is highly relevant.

For further insight, we conducted experiments on random graphs. In the first type of experiment that we performed on ER and BA graphs, the average degree was constant, but the network size was increasing. The increasing network size did, however, not affect the robustness. In the second type of experiment, the network size was fixed, but the average degree increased. In the case of ER graphs, the robustness was either not affected by the change of the average degree (eigenvector centrality and PageRank) or even decreased (degree centrality). In the case of BA graphs, the robustness increased—for all centrality measures and errors that we considered. This was also the case for random graphs generated by the configuration model which generates more realistic networks since we used the degree distributions of the empirical networks of the first part of our study. These results suggest that centrality measures are more robust, the higher the average degree in the network, as long as the networks have a skewed degree distribution.

In the third part of our study, we introduced an analytical approach for the robustness in terms of a rank correlation. Focusing on Goodman and Kruskal's rank correlation, we derived explicit expressions for the robustness of the degree centrality in ER graphs for the case of missing nodes, missing edges, and additional edges. We showed that the robustness for these type of networks is independent of their size, as long as the average degree is constant. For arbitrary network size, there exist robust and non-robust networks w.r.t. all centrality measures used in this study. Moreover, we argued that the quotient of the variance of the degree and the variance of the degree change may explain the robustness behavior at least to some extent. In addition, we studied the behavior of this quotient analytically for ER graphs.

These findings contribute to a better understanding of the robustness of centrality measures. Researchers should, therefore, pay particular attention to error-free data collection if it is known that the particular network is sparse. When centrality measures are used on this type of network, the results may be interpreted with caution.

This study also provides a basis for further research. The findings may be incorporated into procedures for the treatment of erroneous network data. In this paper, we used a rank correlation to quantify the robustness. Future studies may examine whether similar results occur for other measures of robustness like overlap and top metrics (Borgatti et al., 2006; Murai & Yoshida, 2019).

Further investigations about sophisticated error mechanisms (e.g., extending the “missing edges proportional” error mechanism or mixtures of error mechanisms (Platig et al., 2013)) would be interesting. Another direction would be the use of random graph models in which other properties apart from the average degree can be controlled separately. Moreover, further investigation of the relationship between the robustness and the quotient of variances would be interesting.

**Acknowledgments.** We thank the anonymous reviewers and the editors of this special issue for their helpful comments.

**Conflicts of interest.** None.

## Notes

1 Although the two topics sound similar, studies on the robustness of networks have a different focus than studies about the robustness of centrality measures. The subject of studies on the robustness of networks is the question, how the functionality of a network as a whole is influenced by, for example, the removal of nodes (see Albert et al. (2000) or Callaway et al. (2000)). If the term robustness is used in this work without further specification, then the term always refers to the robustness of centrality measures.

2 We use NetworkX (version 2.2, (Hagberg et al., 2008)) to generate random graphs and calculate centrality measures.

## References

- Albert, R., Jeong, H., & Barabási, A.-L. (2000). Error and attack tolerance of complex networks. *Nature*, 406(July), 378–382.
- Barabási, A.-L., & Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(October), 509–512.
- Beuming, T., Skrabanek, L., Niv, M. Y., Mukherjee, P., & Weinstein, H. (2005). PDZBase: A protein–protein interaction database for PDZ-domains. *Bioinformatics*, 21(6), 827–828.
- Boguñá, M., Pastor-Satorras, R., Díaz-Guilera, A., & Arenas, A. (2004). Models of social networks based on social distance attachment. *Physical Review E*, 70(5), 056122.
- Bolland, J. M. (1988). Sorting out centrality: An analysis of the performance of four centrality models in real and simulated networks. *Social Networks*, 10(3), 233–253.
- Bonacich, Phillip. (1987). Power and centrality: A family of measures. *American Journal of Sociology*, 92(5), 1170–1182.
- Borgatti, S. P., Carley, K. M., & Krackhardt, D. (2006). On the robustness of centrality measures under conditions of imperfect data. *Social Networks*, 28(2), 124–136.
- Brandes, U. (2001). A faster algorithm for betweenness centrality\*. *The Journal of Mathematical Sociology*, 25(2), 163–177.
- Brin, S., & Page, L. (1998). The anatomy of a large-scale hypertextual web search engine. In *Seventh international world-wide web conference (WWW 1998)*.
- Callaway, D. S., Newman, M. E. J., Strogatz, S. H., & Watts, D. J. (2000). Network robustness and fragility: Percolation on random graphs. *Physical Review Letters*, 85(25), 5468–5471.
- Cho, E., Myers, S. A., & Leskovec, J. (2011). Friendship and mobility: User movement in location-based social networks. In *Proceedings of the international conference on knowledge discovery and data mining* (pp. 1082–1090).
- Costenbader, E., & Valente, T. W. (2003). The stability of centrality measures when networks are sampled. *Social Networks*, 25(4), 283–307.
- De Las Rivas, J., & Fontanillo, C. (2010). Protein-protein interactions essentials: Key concepts to building and analyzing interactome networks. *Plos Computational Biology*, 6(6), e1000807.
- Dünker, D., & Kunegis, J. (2015). Social networking by proxy: Analysis of Dogster, Catster and Hamsterster. In *Proceedings of international conference on world wide web companion* (pp. 361–362).
- Erdős, P., & Rényi, A. (1959). On random graphs. *Publicationes Mathematicae*, 6, 290–297.
- Erman, N., & Todorovski, L. (2015). The effects of measurement error in case of scientific network analysis. *Scientometrics*, 104(2), 453–473.
- Frantz, T. L., Cataldo, M., & Carley, K. M. (2009). Robustness of centrality measures under uncertainty: Examining the role of network topology. *Computational and Mathematical Organization Theory*, 15(4), 303–328.
- Ghoshal, G., & Barabási, A.-L. (2011). Ranking stability and super-stable nodes in complex networks. *Nature Communications*, 2, 394.
- Gleiser, P. M., & Danon, L. (2003). Community Structure in Jazz. *Advances in Complex Systems*, 6(4), 565–573.
- Goodman, L. A., & Kruskal, W. H. (1954). Measures of association for cross classifications. *Journal of the American Statistical Association*, 49(268), 732–764.
- Hagberg, A. A., Schult, D. A., & Swart, P. J. (2008). Exploring network structure, dynamics, and function using NetworkX. In *Proceedings of the 7th python in science conference (SciPy2008)* (pp. 11–15).
- Holzmann, H., Anand, A., & Khosla, M. (2019). Delusive pagerank in incomplete graphs. In L. M. Aiello, C. Cherifi, H. Cherifi, R. Lambiotte, P. Lió, & L. M. Rocha (Eds.), *Complex networks and their applications VII* (pp. 104–117). Cham: Springer International Publishing.
- Joshi-Tope, G., Gillespie, M., Vastrik, I., D'Eustachio, P., Schmidt, E., de Bono, B., Jassal, B., Gopinath, G. R., Wu, G. R., Matthews, L., et al. (2005). Reactome: A knowledgebase of biological pathways. *Nucleic Acids Research*, 33(Suppl. 1), D428–D432.



- Kendall, M G. (1945). The treatment of ties in ranking problems. *Biometrika*, 33(3), 239–251.
- Kim, P. J., & Jeong, H. (2007). Reliability of rank order in sampled networks. *European Physical Journal B*, 55(1), 109–114.
- Koschützki, D., Lehmann, K. A., & Peeters, L. (2005). Centrality indices. In U. Brandes, & T. Erlebach (Eds.), *Network analysis: methodological foundations* (pp. 16–61). Springer Berlin Heidelberg.
- Kosinets, Gueorgi. (2006). Effects of missing data in social networks. *Social networks*, 28(3), 247–268.
- Kunegis, J. (2013). KONECT - The koblenz network collection. In *WWW 2013 companion - Proceedings of the 22nd international conference on World Wide Web*.
- Leskovec, J., & Mcauley, J. J. (2012). Learning to discover social circles in ego networks. In F. Pereira, C. J. C. Burges, L. Bottou, & K. Q. Weinberger (Eds.), *Advances in neural information processing systems 25* (pp. 539–547). Curran Associates, Inc.
- Leskovec, J., Kleinberg, J., & Faloutsos, C. (2007). Graph evolution: Densification and shrinking diameters. *The ACM Transactions on Knowledge Discovery from Data*, 1(1).
- Lusseau, D., Schneider, K., Boisseau, O. J., Haase, P., Sloat, E., & Dawson, S. M. (2003). The bottlenose dolphin community of doubtful sound features a large proportion of long-lasting associations. *Behavioral Ecology and Sociobiology*, 54, 396–405.
- Marsden, P. V. (1990). Network data and measurement. *Annual Review of Sociology*, 16(1), 435–463.
- Martin, C., & Niemeyer, P. (2019). Influence of measurement errors on networks: Estimating the robustness of centrality measures. *Network Science*, 7(2), 180–195.
- Martin, C., & Niemeyer, P. (2020). The role of network size for the robustness of centrality measures. In H. Cherifi, S. Gaito, J. F. Mendes, E. Moro, & L. M. Rocha (Eds.), *Complex networks and their applications VIII* (pp. 40–51). Cham: Springer International Publishing.
- Murai, S., & Yoshida, Y. (2019). Sensitivity analysis of centralities on unweighted networks. In *The world wide web conference. WWW 2019* (pp. 1332–1342). New York, NY, USA: ACM.
- Newman, M. (2003). The structure and function of complex networks. *Siam Review*, 45(2), 167–256.
- Newman, M. E., Strogatz, S. H., & Watts, D. J. (2001). Random graphs with arbitrary degree distributions and their applications. *Physical Review E, Statistical, Nonlinear, and Soft Matter Physics*, 64.
- Niu, Q., Zeng, A., Fan, Y., & Di, Z. (2015). Robustness of centrality measures against network manipulation. *Physica A: Statistical Mechanics and Its Applications*, 438, 124–131.
- Platig, J., Ott, E., & Girvan, M. (2013). Robustness of network measures to link errors. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 88(6).
- Rozemberczki, B., Davies, R., Sarkar, R., & Sutton, C. (2019). GEMSEC: Graph embedding with self clustering. *Proceedings of the 2019 IEEE/ACM international conference on advances in social networks analysis and mining 2019* (pp. 65–72). ACM.
- Rual, J.-F., Venkatesan, K., Hao, T., Hirozane-Kishikawa, T., Dricot, A., Li, N., Berriz, G. F., Gibbons, F. D., Dreze, M., & Ayivi-Guedehoussou, N. (2005). Towards a proteome-scale map of the human protein–protein interaction network. *Nature*, 1173–1178.
- Schulz, J. (2016). Using Monte Carlo simulations to assess the impact of author name disambiguation quality on different bibliometric analyses. *Scientometrics*, 107(3), 1283–1298.
- Smith, J. A., & Moody, J. (2013). Structural effects of network sampling coverage I: Nodes missing at random. *Social Networks*, 35(4).
- Smith, J. A., Moody, J., & Morgan, J. H. (2017). Network sampling coverage II: The effect of non-random missing data on network measurement. *Social Networks*, 48, 78–99.
- Tsugawa, S., & Ohsaki, H. (2015). Analysis of the robustness of degree centrality against random errors in graphs. In *Studies in computational intelligence*, vol. 597 (pp. 25–36).
- Wang, C., Butts, C. T., Hipp, J. R., Jose, R., & Lakon, C. M. (2016). Multiple imputation for missing edge data: A predictive evaluation method with application to Add Health. *Social Networks*, 45, 89–98.
- Wang, D. J., Shi, X., McFarland, D. A., & Leskovec, J. (2012). Measurement error in network data: A re-classification. *Social Networks*, 34(4), 396–409.
- Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of ‘small-world’ networks. *Nature*, 393(1), 440–442.
- Yang, J., & Leskovec, J. (2012). Defining and evaluating network communities based on ground-truth. In *Proceedings of ACM SIGKDD workshop on mining data semantics* (p. 3). ACM.
- Zachary, W. (1977). An information flow model for conflict and fission in small groups. *Journal of Anthropological Research*, 33, 452–473.
- Zafarani, R., & Liu, H. (2009). Social computing data repository at ASU.

Appendix A. Results for empirical networks

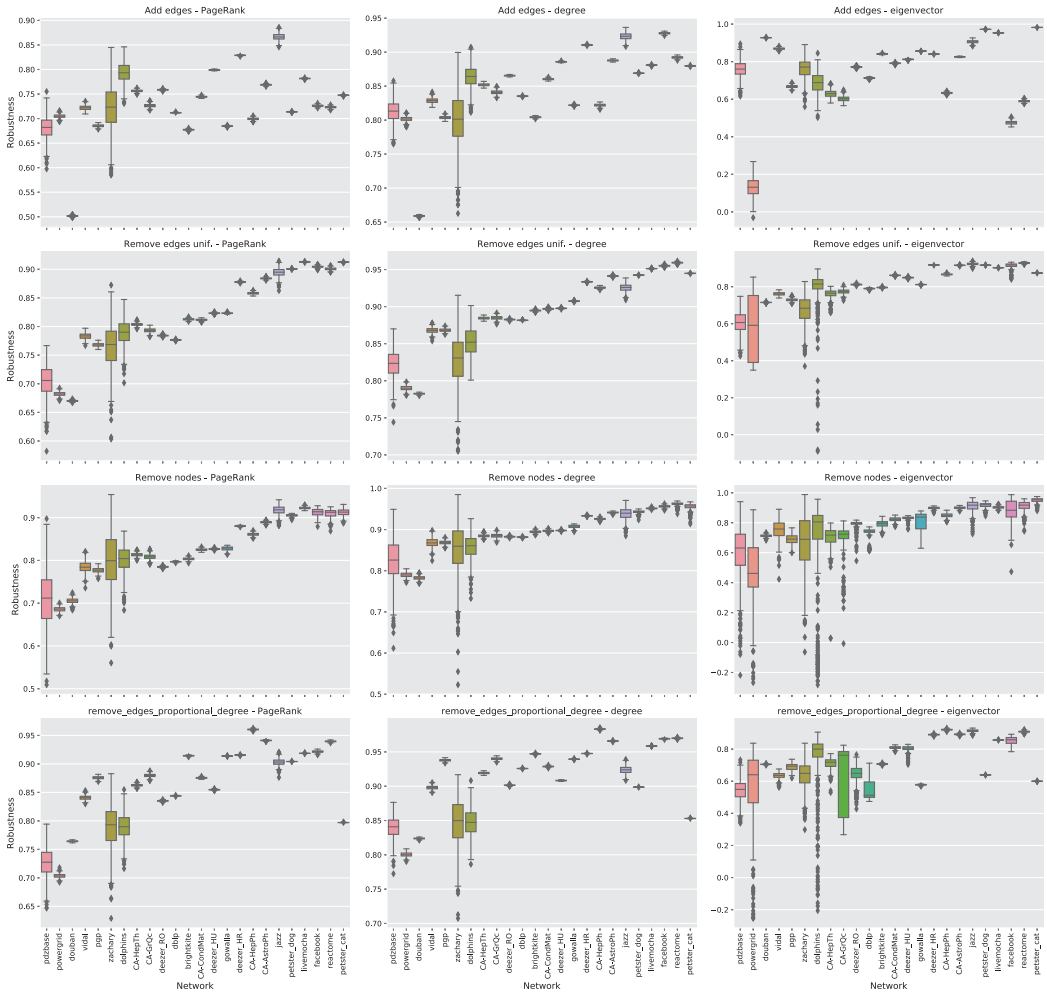


Figure A1. Results for the experiments on empirical networks as described in Section 3. The median robustness is indicated in each box, and whiskers are 1.5 times the interquartile range.

Cite this article: Martin C. and Niemeyer P. (2021). On the impact of network size and average degree on the robustness of centrality measures. *Network Science* 9, S61–S82. <https://doi.org/10.1017/nws.2020.37>