Properties of Erdős-Rényi Graphs

Hang Chen¹, Vahan Huroyan¹, and Stephen Kobourov¹

University of Arizona, Tucson, AZ, USA
{hangchen,vahanhuroyan,kobourov}@email.arizona.edu

Abstract. We study graph properties, such as diameter, density and clustering coefficients, and the relationships between these graph properties. Whereas for graphs on small number of vertices we can exactly compute the average values and range for each property of interest, this becomes infeasible for larger graphs. Thus, we analyse graph properties and the relationships between the properties using Erdős-Rényi graphs, which model well the underlying graph space. Specifically, we focus on thirteen properties of interest in graph theory and graph mining. We numerically study the behavior of these properties, as the number of vertices increase, for the Erdős-Rényi model and compare the results with known theoretical bounds, where applicable. We also use linear and non-linear models to predict a specific property based on the other properties.

Keywords: Graph Mining · Graph Properties · Erdős-Rényi graphs

1 Introduction

Understanding the most descriptive graph properties and the relations between them is important in theory as well as in applications such as graph mining. The recent development of the fields such Graph Neural Networks (GNN)⁶⁷⁴⁴ and graph anonymization [1] requires a careful analysis of various graph properties. GNNs are a natural extension of the deep learning algorithms over the graphs, that is, the input of the algorithm is a graph instead of a vector. In GNNs it is important to have descriptive and unique properties for graphs’ vertices, edges. Since these algorithms are usually heavy, it is important to find the most descriptive set of parameters. Graph anonymization aims to provide privacy protection when working with graph data. Graphs that arise in social media contain sensitive information. Thus, publishing results and conducting research on these graph might be problematic. Summarizing these graphs by their properties and creating graphs with similar properties reduces these concerns.

A graph is a pair \( G = (V, E) \), where \( V \) is a set of vertices, and \( E \) is a set of edges that connect pairs of vertices. One of the standard ways of generating random graphs is the Erdős-Rényi model, which takes two parameters: the number of vertices \( |V| \) and the probability of selecting any edge \( 0 \leq p \leq 1 \). The main focus of this work is to study the behaviour of properties of graphs generated by the Erdős-Rényi models with \( p = 1/2 \) and \( p = \log(|V|)/|V| \). Some natural questions that arise include: How are these properties correlated? Can
we use a subset of properties to predict other properties? What are the most important properties for such predictions? However, summarizing a graph using a fixed number of properties can be misleading, as it is possible to have multiple different graphs with exactly the same values for each of the properties under consideration (even discounting isomorphic copies); see Chen et al. [13,14]. Thus, it is very important to consider properties that are as descriptive as possible. With this in mind, we added more graph properties to original list of Chen et al. and removed some that we have seen are high correlated with the rest.

The main motivation behind learning the relationships between graph properties is to find connections between them and hopefully come up with models to estimate some of the graph properties based on the others. Since some of the properties can be computationally expensive to compute exactly, it makes sense to approximate such properties based on other properties that can be efficiently computed. We discuss two basic models for property prediction which can be useful for larger graphs.

1.1 Related Work

There is a great deal of related work on graph mining, exploration of graph properties and graph generators. Applications of graph mining range from bioinformatics and chemistry, to software engineering and social science. The efficient calculation of the various graph properties is crucial in graph mining. These properties range from basic, e.g., vertex count and edge count, to complex, e.g., clustering coefficients [22,25,31] and average path length [11,12,31]. These are widely used in graph mining applications and each captures and represents some important information about graphs. The node and edge connectivity may be used to describe the resilience of graphs [10,27]. Another commonly used graph property is the degree distribution. Many real-world graphs, including communication, citation, biological and social graphs follow a power-law shaped degree distribution [5,11,135]. Other real-world graphs have been found to follow an exponential degree distribution [21,38,42]. Degree assortativity is of a particular interest in the study of social graphs and is calculated based on the Pearson correlation between the vertex degrees of connected pairs [33].

Graph anonymization [1,39] is another motivation for studying graph properties. Examples of such algorithms include $k$-neighborhood anonymity, edge randomization and cluster based generalization; see survey by Wu et al. [13]. Another example includes Ying et al. [45], where the goal is to preserve the spectral information of the underlying graph.

Recently Chen et al. [13,14] consider different graph generators and the question of whether graph generators can represent and cover the space of non-isomorphic graphs. Experimental results show that no graph generator can model the underlying space of non-isomorphic graphs well. However, as we show here, if isomorphism is allowed, then the Erdős-Rényi graph generator does model the space of all graphs well.
2 Graph Properties

Table 1. The set of graph properties considered in this paper. The first column includes the name of the property. The second column shows the formula by which these properties can be computed for given graphs. The third column presents the time complexity of calculating the property and the last column includes relevant references.

| Name                  | Formula                                                                 | time       | Reference |
|-----------------------|-------------------------------------------------------------------------|------------|-----------|
| Global Clustering Coefficient | \( GCC(G) = \frac{3 \times \text{[triangle]} \text{ in the graph}}{|V|^3} \) | \( O(|V|^3) \) | [11]223   |
| Average Square Clustering | \( ASCC(G) = \frac{1}{n(n-1)} \sum_{u,v \in V} d(u,v) \) | \( O(|V|^3) \) | [26]      |
| Degree Assortativity | \( r(G) = \frac{\sum_{x \in V} xy(xy-a_{xy}b_{xy})}{\sigma_u \sigma_v} \) | \( O(|V| + |E|) \) | [33]131   |
| Density | \( \text{den} = \frac{2|E|}{|V||V|-1} \) | \( O(1) \) |           |
| Diameter | \( diam(G) = \max\{\text{dist}(v, w), v, w \in V\} \) | \( O(|V||E|) \) | [12]29223131 |
| Edge Connectivity | \( \text{Ce}: \text{the minimum number of edges to remove to disconnect the graph} \) | \( O(|V||E|) \) | [18]      |
| Degree Centrality | \( C_D(v) = \text{degree}(v) \) | \( O(1) \) | [25]32    |
| Closeness Centrality | \( C_C(v) = \frac{\sum_{u \neq v} \text{d}(u,v)}{|V|-1} \) | \( O(|V|^2 \log(|V|)) \) | [25]923132 |
| Betweenness Centrality | \( C_B(v) = \frac{\sum_{s,t \in V} \sigma_{s,t}(v)}{\sigma_{s,t}} \) | \( O(|V|^3) \) | [16]172532 |
| Eigenvector Centrality | \( C_{ei}(v) = \sum_{x \in V} A_{ei}(x) C_{ei}(u) \), \( A \) is adjacency matrix | \( O(|V|^3) \) | [14]      |
| Freeman’s Centralization | \( C_f(G) = \max_{x \in G} \{C_{ei}(x) - C_{ei}(v)\} \) | \( O(|V||E|) \) | [19]      |
| Effective Graph Resistance | \( R_G = |V| \sum_{k=1}^{n-1} \frac{1}{\mu_k} \) | \( O(|V|^3) \) | [25]114   |
| Spectral Radius | \( \rho(G) = |\lambda_1| \) | \( O(|V|^3) \) | [25]8120 |

In this section we discuss different graph properties that are widely used in the literature of graph mining, bioinformatics, social science, chemistry. The goal is to come up with a collection of descriptive properties of graphs so that each graph can be uniquely (or almost uniquely) represented as a vector of its properties. Since different fields use different graph properties, we include the ones that are frequently used in practice; the complete list can be seen in Table 1. Most of the properties appear in the earlier studies of Chen et al. [13]14, but we also add new ones (e.g., centrality measures) to make the embedded graph space more distinctive. We also exclude some of the graph properties discussed in Chen et al. because of their similarity to the others. We remark, that the
properties in Table 1 vary from simple graph measures such as density, diameter and edge connectivity to more complex ones such as degree associativity and centralization.

Note that some of the graph properties discussed below are only defined for connected graphs. With this in mind, we disregard the disconnected case in our analysis. It is known that for fixed $p$ and increasing values of $|V|$, the Erdős-Rényi model almost surely produces connected graphs [32]. We experimentally confirm this by generating 10,000 graphs for $|V| = 5, 6, \ldots, 15$ and examine the percentage of connected graphs; see Table 2. While for $p = 1/2$ we get 99% connected graphs for 13 or more vertices, that is not the case for $p = \log(|V|)/|V|$. In general every time the generator would generate a disconnected graph, we would run the generator with the same parameters until we get a connected graph.

Table 2. The percentage of connected graphs from a set of 10,000 generated by the ER model with $p = 1/2$ and $p = \log(|V|)/|V|$ for increasing values of $|V|$.

| $p = 1/2$ | 5  | 6  | 7  | 8  | 9  | 10 | 11 | 12 | 13 | 14 | 15 |
|-----------|----|----|----|----|----|----|----|----|----|----|----|
| 59.8%     | 71.3% | 81.6% | 89.1% | 94.7% | 96.4% | 98.1% | 98.9% | 99.4% | 99.7% | 99.8% |

The Global Clustering Coefficient (GCC) [11,22] measures the tendency of a graph’s vertices to cluster together. It computes the ratio of closed vertex triplets over all possible triplets. The Average Square Clustering Coefficient (ASCC) [26] computes the ratio of closed vertex squares over all possible squares for each vertex. The Average Path Length (APL) [12,25,11,31] measures the average of all shortest paths in the graph. The Degree assortativity ($r$) measures whether vertices with high degrees are connected to other vertices with high degree. The degree assortativity [34,31] takes values between $-1$ and 1. When the degree assortativity is close to 1 the vertices with high degrees tend to be connected to each other and vertices with low degrees tend to be connected with other low degree vertices. On the other hand, when the degree assortativity is close to $-1$, vertices with high degrees tend to be connected to vertices with low degrees.

The Density (den) [18] is the ratio between the number of edges of the graph and the maximum number of edges possible. The Diameter (diam) of a graph [12,29,22,31] measures the greatest distance between any pair of vertices. The Edge connectivity ($C_e$) measures the minimum number of edges that need to be removed to disconnect the graph. Edge connectivity captures the robustness of the graph, as in sparse graphs, edge connectivity can vary whereas in dense graphs, the variation in edge connectivity decreases.

Next, we discuss several centrality measures introduced by Newman [32]. We consider four of them: degree centrality, closeness centrality, betweenness centrality and eigenvector centrality. All these are vertex-based measures but they can be interpreted as graph-based measures via centralization. We use Freeman centralization which measures unevenness [19,36], where high/low centralization values represent high/low unevenness. The degree centrality [28,32] of a vertex
is proportional to its degree. The closeness centrality for a vertex measures the inverse of the average distance from the vertex to all others. Thus, it captures how close is the vertex to the center of the graph. Betweenness centrality measures the influence of a vertex over the flow of information between every pair of vertices, assuming information flows over shortest paths between vertices. Eigenvector centrality measures the influence of a vertex in the graph. It is a natural extension of the degree centrality and high eigenvector centrality means that a vertex is connected to many vertices with high eigenvector centrality values.

The last two properties that we study are spectral radius and effective resistance. The spectral radius \((R_G)\) of a graph is defined as the spectral radius of the corresponding adjacency matrix of the graph (the spectral radius of a graph is the largest absolute value of its eigenvalues). The effective resistance \((R_G)\) of a graph is defined as the sum of the effective resistances over all pairs of vertices, where the notion of resistance can be calculated by Ohm’s law, when treating the graph as an electrical circuit.

We would like to mention that some of the properties described above scale with the number of vertices in the graph (e.g., diameter, edge connectivity, effective graph resistance and spectral radius). For the sake of our analysis we normalize these properties so that the values lie between 0 and 1. The only property that we do not normalize is assortativity, which naturally lies between −1 and 1, independent on the size of the graph.

3 Behaviour of Graph Properties

This section discusses the behaviour of the graph properties of interest for graph generated with the Erdős-Rényi model. The goal is to see how these properties change, with respect to the number of vertices in the graph. For some of the properties discussed in this paper the asymptotic results are known and proved for various values of \(p\); see Table 3. However, our goal is study these results numerically. For this purpose we generate graphs based on Erdős-Rényi model with \(|V| = 5, 10, \ldots, 80\) using two standard values of \(p\): \(p = 1/2\) and \(p = \log(|V|)/|V|\). For each experiment we generate 1000 graphs and compute the 13 properties. For the properties which asymptotic bounds are known we show these bounds by a red curve; see Figures 1, 2. We remark that our experimental results match with the theoretical bounds, which is a good indication that the experimental results for properties without theoretical bounds are plausible. Bellow we discuss some of our findings.

According to Figures some of the properties quickly converge to some values for both Erdős-Rényi with \(p = 1/2\) and \(p = \log(|V|)/|V|\). For example see the results for GCC, ASCC, APL, diameter, density, effective graph resistance. Others converge faster for one of the two models and slower for the other one. For example for Erdős-Rényi with \(p = 1/2\) the diameter is almost always 2, thus the normalized version of the diameter converges to 0 very quickly \((2/|V|)\), while for Erdős-Rényi with \(\log(|V|)/|V|\) it is not clear whether diameter converges to 0 or
Table 3. Known bounds for some of the properties Erdős-Rényi graphs. The first row shows the bounds for Erdős-Rényi graphs with \( p = 1/2 \) and the second row shows the bounds for Erdős-Rényi graphs with \( p = \log |V|/|V| \). Note that the bounds of this table are for the non-normalized properties. Later, in Figure 1, 2 we show the normalized versions of these bounds.

| GCC | APL | r | den | diam | \( R_G \) | \( p \) |
|-----|-----|---|-----|------|--------|-------|
| \( p = 1/2 \) | \( p \) | \( \ln |V| \) | \( |V| - 1 \) | \( 0 \) | \( 2 \) | \( 1 \) | \( \log |V| / |V| \) | \( 0 \) | \( 1 \) | \( 17 \) | \( 1 + o(1) \max(\sqrt{\Delta}, np) \) |
| \( p = \log |V|/|V| \) | \( p \) | \( \approx \ln |V| \) | \( |V| - 1 \) | \( 0 \) | \( \approx \ln |V| / |V| \) | \( 15 \) | \( \approx |V| / p \) | \( 10 \) | \( 1 + o(1) \max(\sqrt{\Delta}, np) \) |

Fig. 1. Behaviour of graph properties for 1000 graphs generated according to Erdős-Rényi model with \( p = 1/2 \) and values of \( |V| \) in the range \([5, 10, \ldots, 80]\). The results are illustrated with violin plots. The first row shows the results for Global Clustering Coefficient, Average Square Clustering, Average Path Length and Degree Assortativity. The second row shows Diameter, Density, Edge Connectivity and Degree Centrality. The third row shows Closeness Centrality, Betweenness Centrality, Eigenvector Centrality and Effective Graph Resistance. The last row shows Spectral Radius.

Another experimental observation is that edge connectivity for Erdős-Rényi with \( p = 1/2 \) converges slower than for Erdős-Rényi with \( p = \log |V|/|V| \); see Figure 2. This might be to the fact that graphs generated by Erdős-Rényi with \( p = 1/2 \) are denser than the graphs generated by Erdős-Rényi with \( p = \log |V|/|V| \) and density is correlated with edge connectivity. Note that there are
several graph properties (e.g., the centrality measures) that do not have known bounds and is not clear whether they converge.

4 Property Correlation and Prediction

In this section we explore the correlations between graph properties for the set of all graphs ($|V| \leq 7$) and then for graphs generated by Erdős-Rényi model. Next, we use linear and non-linear models to predict some graph properties based on other properties. At the end we apply some feature selection techniques to understand which features are the most important ones for predicting the others.

4.1 Exploring correlations between properties

We aim to study the correlations between the 13 graph properties discussed in Section 2. First, we compute the correlations between the properties for graphs
with \(|V| = 4, 5, 6, 7\). This analysis is special since we have the set of all graphs for \(|V| \leq 7\). Thus, we can see whether the computed correlations for a sample of graphs generated by the Erdős-Rényi model match with the correlations of the set of all graphs. For each \(|V| = 4, 5, 6, 7\) we generate 1000 graphs according to Erdős-Rényi model with \(p = 1/2\) and between each 2 properties we compute the correlation between them. We also compute this correlations for the set of all graphs with \(|V| = 4, 5, 6, 7\). In Figure 3 we report the results, by blue color we mark the point for the set of all graphs and by the red color we mark the result for the sample generated by the Erdős-Rényi model \(p = 1/2\). Figure 3 demonstrates that the results match.

A natural question to ask is, how big of a sample one should take to obtain comparable correlation results with that of the set of all graphs for larger values of \(|V|\). It is impossible to find an exact answer to this question, since as \(|V|\) grows the set of all graphs with \(|V|\) vertices grows too fast, exceeding the number of atoms in the universe (\(10^{78}\)) already for \(|V| = 24\). We propose to answer this question with the following stability test: For \(|V| = 100\) we sample 100, 200, 400, 800 and 1600 graphs according to Erdős-Rényi model and compute the corresponding correlations between the properties. We repeat this experiment 10 times and report the results in Figure 4. The idea is that once we start getting consistent results, that is the variation between the correlations is small, we can assume that these are the correct correlations between the properties for the total dataset. From Figure 4 we can see the violin plot range diminishes with larger sample size and the results for \(|S| = 1600\) are consistent.
We note that the correlation analysis might vary with $|V|$. Thus, we compare the correlations between the properties for graphs with $|V| = 7$ and $|V| = 100$; see Figure 5. For $|V| = 100$, the weak correlations become weaker, while strong correlations become stronger (e.g., the correlation between the four centrality measures become very strong). It is also worth noting that two properties, edge connectivity and the degree assortativity, are not correlated with the rest.

4.2 Linear and non-linear models for prediction

Studying the correlations helps to understand how these 13 properties are correlated for a given sample, although it does not fully answer our main question, which is figuring out how to use some of these properties to predict the rest.

We start with a simple experiment: for a given property, we run a basic linear regression model to see whether we can predict it by the other 12 properties. For this experiment we generate 31,000 graphs by the Erdős-Rényi model with $p = 1/2$ and $|V| = 100$. We compute the 13 properties mentioned in Table 1. Next, we randomly separate this dataset into training, dev and test datasets with 80%, 10%, 10% of the data, respectively. We use the basic linear regression model for each predictor, learn the model based on the training dataset and calculate the results for the test dataset. As a baseline we take the mean predictor, that is the mean values of the training set. As a loss we use $L_1$ error, which is the $L_1$ distance between the predictor and the true value; the results are shown in Table 4. As we can see the linear model for all of the properties significantly...
Fig. 5. The left subfigure demonstrates the correlations between the graph properties for all isomorphic graphs with \(|V| = 7\). The right subfigure demonstrates the correlations between the graph properties for 1000 graphs generated by the Erdős-Rényi model with \(p = 1/2\).

improves compared to the baseline predictor. Comparing the linear predictor to the baseline mean predictor, \(\rho\) (from .00562 to .00004) and \(R_G\) (from .00578 to .00006) improve the most.

### Table 4. Prediction results for the baseline mean estimator (the first row), the linear regression estimator (the second row) and the non-linear estimator, described in Section 4.2. The columns correspond to 13 graph properties.

|          | \(L_1\) loss | GCC | SCC | APL | \(r\)  | \(\text{den}\) | \(\text{diam}\) | \(C_e\) | \(C_d\) | \(C_c\) | \(C_b\) | \(C_e\) | \(R_G\) | \(\rho\) |
|----------|--------------|-----|-----|-----|-------|-------------|-------------|-------|-------|-------|-------|-------|-------|-------|
| mean     | .00590       | .00092 | .00017 | .01129 | .00564 | 0           | .01735      | .01634 | .01723 | .00046 | .00488 | .00578 | .00562 |
| linear   | .00299       | .00003 | .00000* | .00771 | .00000 | 0           | .01046      | .00033 | .00035 | .00013 | .00075 | .00006 | .00004 |
| non-linear | .00009       | .00002 | .00000* | .00241 | .00000* | 0           | .01004      | .00000* | .00000* | .00011 | .00067 | .00004 | .00001 |

However, the linear model has its limitations as there are likely some non-linear connections between the properties. Thus, for the next experiment we add some non-linear combinations of the properties. We use all the second order combinations, the square roots and the logarithms of the properties. To avoid complications we use the absolute values for the square roots \(\sqrt{|x|}\) and we take \(\log(1 + |x|)\) for each of the property. After adding all these features we run another linear regression for this dataset; the results are shown in Table 4. We can see some significant improvements compared to the linear regression model.
Comparing the non-linear to the linear predictor, $C_d$ (from .00033 to 0) and $C_C$ (from .00035 to 0) improve the most.

4.3 Feature selection

The analysis in Section 4.2 leads to the question: which properties are important for the prediction of other properties? One way to answer this question is to use feature selection techniques, such as LASSO. However, LASSO heavily depends on many parameters which are application dependent. Instead we propose the following feature selection technique: For each property, we fix 2 other properties that we call predictors and run a linear regression model on it. Next we pick another property and add it to the set of predictors. We run a linear regression with these 3 predictors and if the improvement on the loss is at least 20% we declare this property as an important predictor. We repeat this experiment for all possible properties and predictors and record the results in a matrix (the property importance matrix). Each cell of the property importance matrix records the number of times that the particular predictor has been important; see the left subfigure of Figure 6. For comparison, the right subfigure of Figure 6 shows the matrix of absolute values of correlations between properties. We can see that the two are similar and thus our proposed technique seems to be a plausible alternative to the correlation computation. Both figures show that density is useful predictor for GCC, SCC, effective graph resistance and spectral radius. This is useful as density is easy to compute compared to other graph properties.

Fig. 6. The left figure presents the property importance matrix described in Section 4.3. The right figure shows the absolute values of the correlation for all pairs of properties.

5 Conclusions and Future Work

This paper provides some early experimental results about the properties of Erdős-Rényi graphs. It seems that such graphs can be used to study the properties of the much larger underlying space of all graphs. Determining the natural
dimension of the space of graphs (when treating each graph as a high dimensional vector based on its properties) seems a challenging but useful research direction. Identifying a set of properties that can be efficiently computed and can be used to predict more expensive-to-compute properties would also be useful.
References

1. C. C. Aggarwal, Y. Li, and S. Y. Philip. On the hardness of graph anonymization. In 2011 IEEE 11th International Conference on Data Mining, pages 1002–1007. IEEE, 2011.
2. M. Barthelemy. Betweenness centrality in large complex networks. The European physical journal B, 38(2):163–168, 2004.
3. A. Berman and X.-D. Zhang. On the spectral radius of graphs with cut vertices. Journal of Combinatorial Theory, Series B, 83(2):233–240, 2001.
4. A. Blass and F. Harary. Properties of almost all graphs and complexes. Journal of Graph Theory, 3(3):225–240, 1979.
5. S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang. Complex networks: Structure and dynamics. Physics reports, 424(4-5):175–308, 2006.
6. P. Bonacich. Power and centrality: A family of measures. American journal of sociology, 92(5):1170–1182, 1987.
7. S. P. Borgatti. Centrality and network flow. Social networks, 27(1):55–71, 2005.
8. U. Brandes. A faster algorithm for betweenness centrality. Journal of mathematical sociology, 25(2):163–177, 2001.
9. U. Brandes. On variants of shortest-path betweenness centrality and their generic computation. Social Networks, 30(2):136–145, 2008.
10. D. Cartwright and F. Harary. Structural balance: a generalization of heider’s theory. Psychological review, 63(5):277, 1956.
11. D. Chakrabarti and C. Faloutsos. Graph mining: Laws, generators, and algorithms. ACM computing surveys (CSUR), 38(1):2, 2006.
12. D. Chakrabarti and C. Faloutsos. Graph patterns and the r-mat generator. Mining Graph Data, pages 65–95, 2007.
13. H. Chen, V. Huroyan, U. Soni, Y. Lu, R. Maciejewski, and S. G. Kobourov. Same stats, different graphs: Exploring the space of graphs in terms of graph properties. IEEE transactions on visualization and computer graphics, 2019.
14. H. Chen, U. Soni, Y. Lu, R. Maciejewski, and S. Kobourov. Same stats, different graphs. In International Symposium on Graph Drawing and Network Visualization, pages 463–477. Springer, 2018.
15. F. Chung and L. Lu. The diameter of sparse random graphs. Advances in Applied Mathematics, 26(4):257–279, 2001.
16. E. Costenbader and T. W. Valente. The stability of centrality measures when networks are sampled. Social networks, 25(4):283–307, 2003.
17. W. Ellens, F. Spieksma, P. Van Mieghem, A. Jamakovic, and R. Kooij. Effective graph resistance. Linear algebra and its applications, 435(10):2491–2506, 2011.
18. S. Even and R. E. Tarjan. Network flow and testing graph connectivity. SIAM journal on computing, 4(4):507–518, 1975.
19. L. C. Freeman. Centrality in social networks conceptual clarification. Social networks, 1(3):215–239, 1978.
20. S. Friedland. Bounds on the spectral radius of graphs with e edges. Linear Algebra and its Applications, 101:81–86, 1988.
21. R. Guimera, L. Danon, A. Diaz-Guilera, F. Giralt, and A. Arenas. Self-similar community structure in a network of human interactions. Physical review E, 68(6):065103, 2003.
22. S. Kairam, D. MacLean, M. Savva, and J. Heer. Graphprism: compact visualization of network structure. In Proceedings of the International Working Conference on Advanced Visual Interfaces, pages 498–505. ACM, 2012.
23. M. Krivelevich and B. Sudakov. The largest eigenvalue of sparse random graphs. *Combinatorics, Probability and Computing*, 12(1):61–72, 2003.

24. L. Leydesdorff. Betweenness centrality as an indicator of the interdisciplinarity of scientific journals. *Journal of the American Society for Information Science and Technology*, 58(9):1303–1319, 2007.

25. G. Li, M. Semerci, B. Yener, and M. J. Zaki. Graph classification via topological and label attributes. In *Proceedings of the 9th international workshop on mining and learning with graphs (MLG), San Diego, USA*, volume 2, 2011.

26. P. G. Lind, M. C. Gonzalez, and H. J. Herrmann. Cycles and clustering in mining networks. *Physical review E*, 72(5):056127, 2005.

27. D. Loguinov, A. Kumar, V. Rai, and S. Ganesh. Graph-theoretic analysis of structured peer-to-peer systems: routing distances and fault resilience. In *Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications*, pages 395–406. ACM, 2003.

28. G. Lohmann, D. S. Margulies, A. Horstmann, B. Pleger, J. Lepsien, D. Goldhahn, H. Schloegl, M. Stumvoll, A. Villringer, and R. Turner. Eigenvector centrality mapping for analyzing connectivity patterns in fmri data of the human brain. *PloS one*, 5(4):e10232, 2010.

29. M. McGlohon, L. Akoglu, and C. Faloutsos. Statistical properties of social networks. *Social network data analytics*, pages 17–42, 2011.

30. N. Meghanathan. Spectral radius as a measure of variation in node degree for complex network graphs. In *2014 7th International Conference on u-and e-Service, Science and Technology*, pages 30–33. IEEE, 2014.

31. A. Mislove, M. Marcon, K. P. Gummadi, P. Druschel, and B. Bhattacharjee. Measurement and analysis of online social networks. In *Proceedings of the 7th ACM SIGCOMM conference on Internet measurement*, pages 29–42. ACM, 2007.

32. M. Newman. *Networks*. Oxford university press, 2018.

33. M. E. Newman. Assortative mixing in networks. *Physical review letters*, 89(20):208701, 2002.

34. M. E. Newman. Mixing patterns in networks. *Physical Review E*, 67(2):026126, 2003.

35. M. E. Newman. The structure and function of complex networks. *SIAM review*, 45(2):167–256, 2003.

36. B. Ruhnau. Eigenvector-centrality node-centrality? *Social networks*, 22(4):357–365, 2000.

37. F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner, and G. Monfardini. The graph neural network model. *IEEE Transactions on Neural Networks*, 20(1):61–80, 2008.

38. P. Sen, S. Dasgupta, A. Chatterjee, P. Sreeram, G. Mukherjee, and S. Manna. Small-world properties of the indian railway network. *Physical Review E*, 67(3):036106, 2003.

39. C. Sun, S. Y. Philip, X. Kong, and Y. Fu. Privacy preserving social network publication against mutual friend attacks. In *Data Mining Workshops (ICDMW), 2013 IEEE 13th International Conference on*, pages 883–890. IEEE, 2013.

40. J. A. Sylvester. Random walk hitting times and effective resistance in sparsely connected Erdős-Rényi random graphs. *arXiv preprint arXiv:1612.00731*, 2016.

41. S. Wasserman, K. Faust, et al. *Social network analysis: Methods and applications*, volume 8. Cambridge university press, 1994.

42. D. Wei-Bing, G. Long, L. Wei, and C. Xu. Worldwide marine transportation network: Efficiency and container throughput. *Chinese Physics Letters*, 26(11):118901, 2009.
43. X. Wu, X. Ying, K. Liu, and L. Chen. A survey of privacy-preservation of graphs and social networks. In Managing and mining graph data, pages 421–453. Springer, 2010.
44. Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and P. S. Yu. A comprehensive survey on graph neural networks. arXiv preprint arXiv:1901.00596, 2019.
45. X. Ying and X. Wu. Randomizing social networks: a spectrum preserving approach. In proceedings of the 2008 SIAM International Conference on Data Mining, pages 739–750. SIAM, 2008.