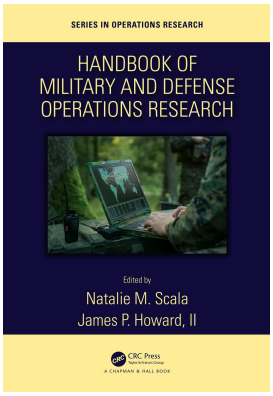

Handbook of Military and Defense Operations Research

Natalie M. Scala, James P. Howard

Modern Methods for Characterization of Social Networks through Network Models

Publication details
https://test.routledgehandbooks.com/doi/10.1201/9780429467219-8
Christine M. Schubert Kabban, Fairul Mohd-Zaid, Richard F. Deckro
Published online on: 02 Mar 2020

Chapter 8

Modern Methods for Characterization of Social Networks through Network Models

Christine M. Schubert Kabban, Fairul Mohd-Zaid and Richard F. Deckro

8.1 Introduction 171
8.1.1 Social Network Analysis Background: A Historical Perspective 172
8.1.2 Emergence of Quantitative Approaches to Social Network Analysis 173
8.2 Characterization of a Network 174
8.2.1 Graph Properties 174
8.2.2 Random Graph Generating Models 176
8.2.2.1 Mixtures of Random Graph Generating Models 180
8.2.3 Network Model Summary 180
8.3 Methods to Assess the Accuracy of a Network Model 181
8.3.1 Graph Matching 181
8.3.2 Graph Classification 182
8.3.3 Summary Graph Matching and Graph Classification 184
8.4 Statistical Testing Methods 184
8.4.1 Summary of Statistical Testing Methods 188
8.5 Future Directions 188
References 189

8.1 Introduction

> *The measure of strategic success is ultimately the change in relevant actor behavior, in addition to physical results on the battlefield.*
>
> (Office of the Joint Chiefs of Staff, 2016, p. 1)

This chapter 1) provides a background for the usefulness of network models as a means to enable social network analysis (SNA), 2) discusses the importance of network characterization via a network model and 3) provides detail on the methods useful for such characterization, with an emphasis on newly emerging statistical methods. These methods are important to the operational analyst as a means to characterize and monitor networks of interest through the network model. The application of these methods extends beyond social networks and can be applied to any network of interest. Both students and early practitioners interested in SNA will benefit from the rich background provided within this chapter and may use the references and resources herein to learn more detail on any of the methods discussed. Further, data is used to demonstrate emerging statistical methods which may be used to monitor a network of interest.

An emphasis is placed on the statistical methods due to their flexibility, sensitivity, and ease of use. Students and practitioners should be able to follow, apply, and use these methods to both characterize and monitor the networks that are built to represent the structures of interest to them. Important features for the operations researcher to identify when characterizing a network of interest are whether or not the network has changed, whether the network behaves as expected, or whether the network does not exhibit characteristics as expected. The latter feature is of particular interest in that it represents emergent operational tasks intent on identifying fabricated, synthetic network structures which mask the true relations and intent within the network of interest. Identification of these features facilitates planning and response, both offensively and defensively. For instance, understanding and unmasking the real actors and relationships in a synthetic network may enable responders to pre-empt negative world events through network characterization and monitoring tasks.

8.1.1 Social Network Analysis Background: A Historical Perspective

The study of social, political, and military structures have been key intelligence considerations throughout the ages. Historically, the keys to understanding a foe's or an ally's strengths (and weaknesses) consisted of the lord/vassal, clans, and tribe structures, reinforced by familial, marriage, and financial links. Social network analysis (SNA) is a tool for the operations researcher to understand, monitor, and exploit these structures. Indeed, analysts have charted the alliances, coalitions, and military order of battles for centuries in order to derive factors providing operational advantages. Today, our interconnected world has created numerous additional considerations supplementing the traditional historical factors into which an analyst must be tuned in order to monitor and preserve his or her own network, while at the same time understanding and exploiting the weaknesses in a foe's network. SNA, with its associated mathematical tools enables the analyst to monitor, understand, and exploit network structures. These structures of our world abound today in complex ways.

> *The initial decade of the twenty-first century has been characterized by governments attempting to mitigate the effects of terrorist organizations, insurgent groups, organized criminal enterprises, drug cartels, human trafficking, piracy, and cybercrime. These entities utilize support networks composed of money laundering, weapons smuggling, illegal technology proliferation, and other illicit activities. Dealing with this myriad of interconnected organizations and activities has led to the development of nontraditional analytic techniques in support of strategies to cope with these threats to national security. One such analytic technique brought to bear on this problem set is social network analysis (SNA), not necessarily a new technique, but novel in its relatively recent application to the national security arena. As such, governments' initial unfamiliarity with SNA has now transitioned to various instantiations in levels of application and expertise in numerous organizations.*
>
> (Morris & Deckro, 2013, p. 70)

In addition, the nature of warfare has evolved over time. Von Clausewitz stated, "War is merely the continuation of policy by other means," implying wars were fought for political purposes to implement a far-reaching and/or overall strategy. While the classic image of war termination includes creating the conditions where a foe will acquiesce to

unconditional surrender (whether it actually occurred or not), it has not generally been the outcome of modern conflicts. In addition, even when surrender is attained, winning the peace is problematic; the harsh terms of the Versailles Treaty ultimately lead to the Second World War. The Second World War termination evolved into the Cold War. The United States still maintains troops in Europe and throughout the world. While the US clearly demonstrated its military prowess in the Gulf Wars, ultimately toppling the Baathist regime, forces remain in Iraq attempting to stabilize the nation.

British General Sir Rupert Anthony Smith, in his text *The Utility of Force: The Art of War in the Modern World* (Smith, 2007), makes the point that warfare has shifted from "War between the people" (nations fighting to a distinct victory) to "War amongst the people" (non-nation states waging indefinite warfare) (Smith, 2007). The recent *Joint Concept for Human Aspects of Military Operations* (JC-HAMO) (Office of the Joint Chiefs of Staff, 2016) recognizes that these wars "among the people" coupled with the need to "win the peace" requires considering the effects on humans in planning *and* conducting military operations. It states that

> *Joint Force will enhance operations by impacting the will and influencing the decision making of relevant actors in the environment, shaping their behavior, both active and passive, in a manner that is consistent with U.S. objectives*
>
> (Office of the Joint Chiefs of Staff, 2016, p. 1).

JC-HAMO goes on to point out that in planning conflict and building the peace, one must consider influence operations which include the social, cultural, physical, informational, and psychological elements that create a desired effect on behavior. Building on the use of SNA in counter-terrorism studies, it is the application of an analysis approach that can aid future joint planning. An understanding of SNA approaches will continue to be of value in the intelligence preparation of the battlespace, but with the adoption of the joint concept on human effects in planning operations, it will also be a planning analysis tool.

8.1.2 Emergence of Quantitative Approaches to Social Network Analysis

Humans remain the driving force at the center of technological advances and societal relations. As such, the study and analysis of human relations has expanded beyond the original interests of early sociologists. Over the last century, relations among individuals became represented by graphs in which individuals (nodes) were connected to each other through edges which represent the particular relations of interest; some of the earliest such graphs are attributed to the works of Jacob Moreno in the mid-1930s (Moreno, 1934). Dubbed sociograms by early sociologists, these graphical depictions of the social interactions among individuals are now a primary tool from which information and characteristics of the underlying social network are derived. Such information may be used to glean understanding of the potential implications that individual and collective group behavior may have on society or assets of interest.

Although initial analysis of social networks was conducted through qualitative assessments, the use of graphs enabled quantitative assessments through the use of mathematically based tools. Now these tools and assessments are paramount due to the growth of both the size and the application of social networks.

Today, the intelligence and operational communities widely use SNA in order to analyze relationships between individuals within groups of interest (Havig et al., 2012), and in many cases, to also monitor and possibly influence a network of interest. From a security and defense standpoint, individuals within the network may represent comrades or potential threats; or as applied to non-humans, could represent relations among assets or critical resources such as a network of computers, the power grid, or water sources. Irrespective of the network itself, network analysis, as a tool to understand, monitor, and potentially act upon a network of interest, is not possible without first characterizing the underlying structure of the network.

8.2 Characterization of a Network

Characterizing a network includes the compilation of operations and evaluations that comprehensively describes a specific network and facilitates the reproduction of the nodes and connections within a network by graphical or mathematical representation. These representations allow the analyst to modify, examine, and understand the relationships within the network in order to anticipate how the network may be affected by particular actions and to determine if the network is changing, e.g., if individuals are leaving or somehow now hidden from a network or if the network of individuals is growing.

There is not just one tool that can be used to accomplish network characterization. When possible, however, a mathematical model of the network should be constructed. Such network models may be specific to a network of interest or may be generated based upon common properties inherent within the network. As such, network models are efficient mathematical tools for the operations analyst. Not only can the network model be used to gain knowledge of network behavior, but it can be used to visualize the network. Network visualization methods are specific to the network structure and purpose for studying the network (Blaha, Arendt & Mohd-Zaid, 2014). However, through the use of the network model and, potentially, its visualization, an analyst may be able to study and respond to specific inquiries of interest regarding a real world network without the need to physically interact within the network. Network models, then, are the means by which planning and response is enabled in operational environments.

The next section describes graph properties and common network models that have formed the basis of many of the models that exist today. For further reading on the basic definitions and principles covered in the following sections, see Borgatti et al. (2009), Brass (1995), Newman (2010), and Wasserman and Faust (2009), among others.

8.2.1 Graph Properties

Many network models are based upon graphs which represent the relationships within the network of interest. Table 8.1 provides basic definitions and notation typically used for graphs and networks.

The notation is based upon network representation as graphs and is constructed from fundamentals in graph theory (West, 2001) and (Wasserman & Faust, 2009). A graph is defined by its nodes, V, and edges, E, and as the triple $G = (V, E)$. Simple

TABLE 8.1: Definitions for Common Graph Terms

Terms	Definitions
Graph, Network	A graph G is defined as a triple, $G = (V, E)$, with node set, V, and edge set, E.
Nodes, Vertices	A node set, $\{1,\ldots,N\}$, is a set of points of size N that makes up a graph.
Edges, Links	An edge set, $E \subset \{V \times V\}$, is a set of ordered pairs, $e_{i,j} = (i, j)$, that connects two nodes in V.
Directed, Undirected	A graph is directed if $e_{i,j} \neq e_{j,i}$ and is undirected if $e_{i,j} = e_{j,i}$
Simple Graph, Multigraph	A simple graph is one such that there is only one instance of any particular edge, $e_{i,j}$, if it exists. A multigraph is a graph with multiple edges connecting the same pair of nodes.
K Regular Ring Lattice	A graph with N nodes where each node is connected to K neighbors on each side.
Degree	The number of edges connected to a given node.
Closeness	Inverse of the sum of pairwise distances between a node and other nodes in a graph.
Betweenness	The frequency that a particular node lies in the shortest path between all other paired nodes.
Clustering Coefficient	The proportions of local relationships amongst neighbors compared to the potential that all of the neighbors are connected.
Group Clustering Coefficient	The proportion of existing triangular relationships or triads in a graph over the number of all potential triangles.

graphs contain only a single edge between nodes and no edge which connects a node to itself. A graph is undirected if edges exist without associated direction between nodes. Special forms exists if nodes connect to neighbors in specific ways, such as being connected to exactly four neighbors as demonstrated in Figure 8.1 which is an example of what is called a 4-regular ring lattice on six nodes. Figure 8.2 gives an illustration of an undirected simple graph, $G = (V, E)$ with node set $V = \{1,2,3,4,5,6\}$ and edge set $E = \{(1,3), (1,5), (1,6), (2,4), (2,5), (2,6), (3,5), (3,6), (5,6)\}$.

Characteristics of a graph can be described through the relationships of the nodes and their edges. As defined in Table 8.1, five common measures of the characteristics of a graph can be summarized through nodal degree, closeness, betweenness, clustering coefficient and group clustering coefficient, in addition to many more. Table 8.2 provides the calculated values of these measurements for the graph example in Figure 8.2. The focus on these five, rather than others, is based upon the work of Guzman et al. (2014) in which an exploration of 24 different measures across many graphs could be divided into just five groups of highly correlated measures; the measures mentioned in Table 8.1 are single representations from each of these groups. Whereas the general definitions are provided in Table 8.1, there are various ways to compute these measures (Wasserman & Faust, 2009). For instance, closeness and betweenness may be standardized by dividing by the number of nodes or node pairs. When standardized, values for closeness or betweenness may be used to compare graphs of different sizes.

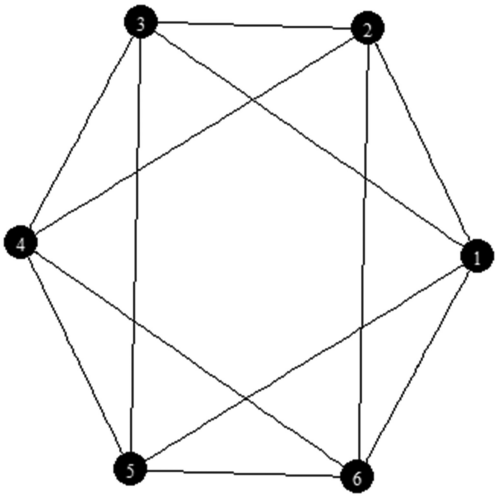

FIGURE 8.1: A 4-regular Ring Lattice on Six Nodes.

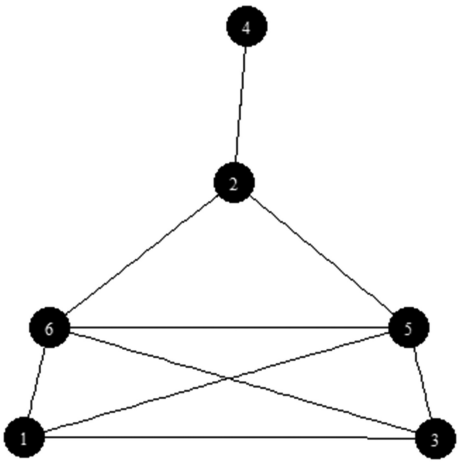

FIGURE 8.2: Undirected Simple Graph of Size n = 6.

8.2.2 Random Graph Generating Models

Random graph generating models are designed to model specific network properties of interest. As such, these models contain both fixed and random parameters so that properties, such as the size or connectivity within a network, could vary while other properties of the network could remain fixed. Therefore, families of possible network representations are generated by varying the random parameters while keeping constant those that are fixed. There are three basic random graph generating properties that form the foundation for many real-world network models: random behavior, scale-free behavior, and small-world behavior. Each of these are now described in turn.

TABLE 8.2: Nodal Measures for Graph in Figure 8.2

Node	Degree	Closeness	Betweenness	Clustering Coef.
1	3	0.1250	0	1.0000
2	3	0.1429	4	0.3333
3	3	0.1250	0	1.0000
4	1	0.0909	0	0.0000
5	4	0.1667	2	0.6667
6	4	0.1667	2	0.6667

Clustering Coef. – Clustering Coefficient

Further, algorithms that generate graphs from each of the models can be found in many software packages today.

Erdös and Rényi proposed the first commonly used random graph generating model (1959) in which a graph is generated by connecting any pair of nodes with an edge with probability p, and in which each edge is independent from every other edge. This results in a graph of N nodes and m edges having an equal probability of

$$p^m \left(1-p\right)^{\binom{N}{2}-m} \tag{8.1}$$

for all possible undirected simple graphs of N nodes and m edges. Although not mentioned earlier, the graph example in Figure 8.2 is an Erdös–Rényi graph with parameter p = 0.5.

Consequently, the graph has a group clustering coefficient of $C = k/N$ where k is the mean degree, and the statistical distribution of the degrees for all possible realizations of the Erdös–Rényi networks follows the Poisson probability distribution as the number of nodes grows large. One downside to the Erdös–Rényi graph generating algorithm is that it is not scale-free (Barabási & Albert, 1999), a property possessed by many real-world social networks such as the World Wide Web (Albert, Jeong & Barabási, 1999). A scale-free network is defined as one that has a power law degree distribution between nodes, which typically means that many nodes have few connections and only a smaller proportion of nodes have many connections. The Erdös–Rényi generated graph contains nodes that are randomly connected. However, given its history, the Erdös–Rényi algorithm is widely used as a baseline for comparisons of network metrics and classification.

Barabási and Albert proposed a model based on two mechanisms that govern the scale-free power law distribution of real world networks (Barabási & Albert, 1999). They defined the two mechanisms to be: (i) networks expand continuously by the addition of new nodes, and (ii) new nodes attach preferentially to existing nodes that are already well-connected. The model operates by first starting with an initial number of nodes N each having degree m. This is followed by an iterative process of adding one node with m edges where the edges are connected to an existing node i with degree d_i based on the following preferential attachment probability, denoted as $\pi\left(d_i\right)$:

$$\pi\left(d_i\right) = \frac{d_i}{\sum_j d_j} \tag{8.2}$$

which is the probability that node i will be attached to the new node. As developed, they claimed that their model indicates that the development of large networks is governed by robust selforganizing phenomena that is not specific to the domain, be it social, biology, or the World Wide Web (Barabási & Albert, 1999; Albert, Jeong & Barabási, 1999). Mathematically, the Barabási–Albert network has a group clustering coefficient (C) that is approximately $C \sim N^{-3/4}$ and the distribution of the degrees across all nodes follows a Pareto distribution with an exponent parameter of $\beta = 2$ (Barabási & Albert, 1999). Figure 8.3 is a representation of a Barabási–Albert graph of size $n = 6$ with $m = 2$. Note that the first and second nodes have considerably higher degrees in comparison to the other nodes due to preferential attachment. For completeness, the nodal measures for the example in Figure 8.3 are provided in Table 8.3.

The empirical degree distribution of the Barabási–Albert graph is different than what was originally derived, especially for relatively small graphs (Mohd-Zaid, Schubert Kabban & Deckro, 2017). Some have also suggested that the preferential attachment model causes biases on the connection of the high degree nodes in the Barabási–Albert graph (Li, Zhang & Small, 2011; Zhang, Small & Judd, 2015). Zhang, Small, and Judd proposed an optimal scale-free network model as an alternative to the Barabási–Albert model that accounts for the degree correlation in the Barabási–Albert model.

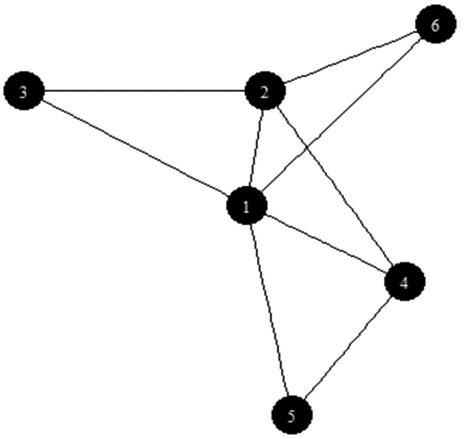

FIGURE 8.3: Barabási–Albert Graph of Size $n = 6$ with $m = 2$.

TABLE 8.3: Nodal Measures for Graph in Figure 8.3

Node	Degree	Closeness	Betweenness	Clustering Coef.
1	5	0.2000	4.0	0.4000
2	4	0.1667	1.5	0.5000
3	2	0.1250	0.0	1.0000
4	3	0.1429	0.5	0.6667
5	2	0.1250	0.0	1.0000
6	2	0.1250	0.0	1.0000

Clustering Coef. – Clustering Coefficient

In addition, to date only Leydesdorff (2007) has studied the skewness and kurtosis of degree, betweenness, and closeness of various empirical networks.

A random graph generating model that produces small-world properties was introduced by Watts and Strogatz (1998). Small-world networks are networks where 1) the shortest path, L, between most pair of nodes in the network grows proportionately to the logarithm of the network size, N, such that $L \propto \log N$, and 2) the clustering coefficient is larger than that of a random network of the same size. The algorithm for this model functions by first starting with a ring lattice of size N (see an example of a ring lattice in Figure 8.1). This is then followed by rewiring each edge in the lattice with probability γ such that duplicates and self-loops are excluded. However, the nodes remain closely connected through short paths. Many real-world networks such as the neural network of the worm *Caenorhabditis elegans*, the power grid of the western United States, and the collaboration network of film actors are shown to possess small-world characteristics (Watts & Strogatz, 1998). That is, the average of the shortest paths between every pair of nodes is relatively small, and the ratio between the network's clustering coefficient is large compared to that of a random network. Further, the degree distribution for the Watts–Strogatz graph has been shown to follow a closed form distribution as described by Barat and Weigt (2000) and has a group clustering coefficient of

$$C \simeq \frac{3(k-2)}{4(k-1)}(1-\gamma)^3 \tag{8.3}$$

where k is the mean degree. Table 8.4 lists the average shortest path and the clustering coefficient characteristic of the real world networks listed above (Watts & Strogatz, 1998). Notice that despite the large size of these networks, the shortest path between nodes is on average as low as 2–3 nodes. Even the power grid has a relatively low shortest path as the average shortest path in a random graph of that size was reported as 12.7 (Watts & Strogatz, 1998). Further, the size of the clustering coefficient compared to a purely random network is large, demonstrating the clustering of nodes in these graphs. However, one disadvantage of the Watts–Strogatz algorithm is that it produces a network that does not have a scale-free power law distribution. Hence, neither the Barabási–Albert nor the Watts–Strogatz algorithm is capable of simultaneously modeling these real world properties in networks.

Other variations on these random graph generating models exist. Seshadhri, Kolda, and Pinar (2012) proposed a modified Erdös–Rényi model where community is defined as a subgraph that is internally highly connected and must contain a dense Erdös–Rényi subgraph. The proposed Block Two-Level Erdös–Rényi (BTER) model creates

TABLE 8.4: Average Shortest Path, $L_{average,}$ and Ratio of Clustering Coefficient with Respect to a Random Network of the Same Size, $C_{network}/C_{random,}$ for Some Real World Networks

Network	**Size**	$L_{average}$	$C_{network}/C_{random}$
Film actors	225,226	3.65	2925.9
Power grid	4,941	18.70	16
Caenorhabditis elegans	282	2.65	5.6

Data from Watts, D.J. and S.H. Strogatz, "Collective dynamics of 'small-world' networks", *Nature* 393, no 6684 (1998): 440–442.

connection via the Erdös–Rényi model within communities followed by cross-communities connections between nodes with degree one. They demonstrated that it accurately captures the scale-free and clustering properties of many real-world social networks. Morris, O'Neal, and Deckro (2014) created an algorithm for prescribed node degree connected graphs (PNDCG) that allows the user to define the scale parameter as well as the clustering coefficient of the network. The clustering coefficient is related to the transitivity concept in the social network literature where transitivity implies the idea of "a friend of a friend is a friend." Comparisons of the average clustering coefficient with those from the Erdös–Rényi and Barabási–Albert generated networks show that their algorithm is able to generate networks with a wider distribution of average clustering coefficients.

8.2.2.1 Mixtures of Random Graph Generating Models

Each of the random graph generating models described focus on a network property (i.e., scalefree or smallworld) and attempts to model that specific property through a single unified distribution for a network measure such as degree. More recent work has focused on using multiple distributions for a single network measure or a combination of network measures. Such models are called mixture models and are motivated by the lack of fit of any one individual property across similar real-world data sets. A mixture model is a single probabilistic model that combines several distributions, each from a different subpopulation, in which the distribution for each subpopulation is governed by different parameter settings. These distributions then are combined into one model through the use of a weighting term that weights each of the distributions; the sums of the weights must equal one. Parameter estimation for the mixture model is then accomplished through statistical techniques such as Expectation-Maximization or Bayesian methods.

With respect to networks, mixture models have been applied in two ways: either as a mixture of different distributions for the same network property, such as randomness, or more recently, as a mixture of varying properties, such as randomness with scale-free and small-world properties. Examples include the works of Thomas, Stoica, and Beuscart (2010) who modeled popularity with respect to social media as a mixture of Gaussian distributions, Zanghi, Ambroise, and Miele (2008) who created an Erdös–Rényi based mixture model to examine affiliation, and Daudin, Picard, and Robin (2008) who created a mixture of Poisson distributions to describe the Erdös–Rényi-based degree distribution within a bacteria network. Each of these examples focuses on mixtures of similar random variable distributions to model a network property. More recently, Durante, Dunson, and Vogelstein (2017) developed a nonparametric, Bayesian based method to model a network using a mixture of network properties. Applied to human brain data, their method demonstrates the feasibility and usefulness in creating a model that mixes several network properties in order to accurately model real-world data.

8.2.3 Network Model Summary

Network models are only useful if such models are able to characterize a network accurately enough for the intended analysis. This section discussed various network models and means by which network models are generated and combined (via mixtures) in order to reproduce network properties of interest. However, when we use specific network properties, such as small-world properties, in order to build a network

model, how do we determine that the resulting model is accurate with respect to the real-world network of interest? In order to assess the accuracy of a network model, alternate techniques are used based upon mathematical or statistical methods. These methods for assessing the accuracy of a network model are discussed next.

8.3 Methods to Assess the Accuracy of a Network Model

There are several methods to determine whether or not the generated model of the network, a model based upon parameters, truly represents the real-world data of the network of interest to an acceptable level. These methods are based upon numerical and more recently, statistical techniques, and are categorized through the intent of the method, that is, via graph matching or graph classification methods.

8.3.1 Graph Matching

Graph matching methodology has developed over many decades and a comprehensive survey of methods and applications for graph matching over a span of three decades from the early 1970s to 2004 was conducted by Conte et al. (2004). In general, graph matching is categorized into two main groups: Exact Matching and Inexact Matching. Exact Matching methods consider edge preserving matches or edge mappings such as graph isomorphisms, graph homomorphisms, or maximum common subgraphs (MCS). Such methods are NP-complete with the exception of graph isomorphism which has not yet been shown to be NP-complete. Currently, MCS methods are only able to handle relatively small graphs due to their computational complexity. Exact Matching methods include other techniques such as Tree Searches which work to iteratively expand the set of paired matched nodes, group theory, subgraph decomposition, or decision trees on a library of known graphs.

Inexact Matching is an approximated matching technique that is more tolerant to slight differences between the graphs, and is achieved by finding a mapping that minimizes a stated matching cost. Here, cost is defined as either deformations or differences between the matched graphs or as a set of graph edit operations. Graph edit operations are operations where a new graph is formed from an original by adding or deleting a node or an edge. This can be performed through one of two ways such as optimal inexact matching, which finds a solution that gives the global minimum matching cost, or approximate matching that only guarantees that the solution is a locally minimized cost.

Optimal matching techniques are often more computationally expensive than Exact Matching itself, but approximate matching techniques are faster than Exact Matching, often exhibiting polynomial matching time. Like Exact Matching, Inexact Matching can also be broken into subcategories; specifically, Tree Searches, Continuous Optimization, Spectral Methods, and other techniques that do not fall into the former groups such as decomposition methods, neural networks, genetic algorithms, bipartite matching, and methods based on local properties. Regardless, the appeal of a faster computation time, which is usually polynomial, makes approximate matching techniques desirable if the application requires a timely analysis.

An updated survey was recently conducted by Livi and Rizzi (2013) with a focus on Inexact Matching methods that are divided into Graph Edit Distance (GED) based,

Graph Kernel-based, and Graph Embedding based techniques. GED based techniques fall in the family of tree searches whereas kernel and embedding based techniques fall in the continuous optimization family where the discrete problem is transformed into the continuous space in order to take advantage of the vast amount of metrics and matching techniques available in that space. Kernel-based techniques transform the graphs onto an induced feature space for evaluation whereas Graph Embedding based techniques transform the graphs to obtain a general vector representation and utilizes the metric space to compute dissimilarities. Livi and Rizzi (2013) suggested that all of the types of methods considered have polynomial complexity, but the computation cost is still dependent on a set of parameters that have to be tuned specifically to the application at hand.

8.3.2 Graph Classification

Graph classification methods classify a graph to the closest matched graph from a library of known graphs. Many such methods exist, and several of the more common methods are described here.

Frequent Subgraph Mining (FSG) is a method that takes a graph and a minimum support threshold ε% to generate all connected subgraphs that occur in at least ε% of the graph (Inokuchi, Washio & Motoda, 2000). Other methods, such as that proposed by Moonesinghe et al. (2007), uses the subgraphs generated by FSG to construct a binary feature vector which is then used to compute the maximum entropy of the graph in an iterative fashion until convergence occurs. This method was shown to perform comparatively to an AdaBoost and a support vector machine (SVM) classifier.

Ketkar, Holder, and Cook (2009) presented an empirical comparison of the major approaches for graph classification, namely SubdueCL (Gonzalez, Holder & Cook, 2002), FSG with SVM (FSG = flexible segmentation graph), a walk-based (direct product) kernel, FSG with AdaBoost, and DT-CLGBI which is a combination of FSG and decision trees (Nguyen et al., 2006). SubdueCL is the pioneering algorithm for graph classification, and functions by creating a decision list from subgraphs and performing an isomorphism test with a new graph for classification. FSG with SVM works by using FSG to create a feature vector that is then used as input for SVM classification. The walk-based kernel is created by taking the direct product of two graphs as a similarity measure. FSG with AdaBoost works by using FSG to create a list of subgraphs and AdaBoost to create a list of positive and negative examples from the subgraphs that result in the upper bound of the gain that is associated with the supergraph. DT-CLGBI combines aspects of frequent subgraph mining and decision trees. The algorithms were compared using a chemical compound dataset as well as artificial network data generated using an in-house data generating technique (Ketkar, Holder & Cook, 2009). The result showed that walk-based kernel performed poorly when the average degree is high and SubdueCL performed poorly when the graph is disconnected. Other methods performed similarly to one another.

Jin, Young, and Wang (2009) also proposed a graph classification method by deriving classification rules based on pattern co-occurrence from the subgraphs. The method only performs pattern mining once and as such, this method results in faster computation time than other methods that require multiple such iterations. The method can be integrated into any subgraph mining algorithm by organizing patterns into groups of co-occurrence rules to form a rule set. That is, whenever a pattern is generated, the discrimination score of every rule is calculated with the pattern's inclusion and then

the pattern is inserted into the rule that yields the greatest increase in discrimination score. The algorithm then finds a co-occurrence rule set that maximizes the number of graphs that can be classified correctly. The authors compared their method against LEAP (Yan et al., 2008) with SVM (LEAP+SVM) and gPLS (partial least squares), showing that their technique performed comparably to the other techniques with faster computation time by magnitudes.

A technique for comparing graphs using subgraphs structures was introduced by Macindoe and Richards (Macindoe & Richards, 2010). This is performed by computing three summarizing features from the subgraphs, and then making a comparison using the earth mover's distance between the distributions of summarizing features to that of subgraphs from other graphs. The summarizing features used by the proposed method are Leadership (also known as Group Closeness) (Freeman, 1978), Bonding (also known as Group Clustering Coefficient) (Wasserman & Faust, 2009), and Diversity (Richards & Wormald, 2009) measures as previously defined. The results of Macindoe and Richards' analysis suggested that graphs can be shown to be similar based on the full graph structure but dissimilar by their local structures.

Employing the statistical knowledge obtained from nodal attributes, Gilbert, Valveny, and Bunke (2012) proposed four fuzzy graph embedding methods that utilize known statistical techniques, namely fuzzy k-means and Gaussian Mixture Models (GMM). These techniques were then applied to SVM for performing graph classification. These methods were compared against the k-Nearest Neighbor classifier as well as another GED-based embedding method on select datasets with the labels removed in order to illustrate the generalizability of their methods on unlabeled graphs. The results from their experiments showed that the proposed methods performed generally no better than the two reference methods; however, they claim that their methods were more computationally efficient.

By using already available network measures, Li et al. (2012) presented a graph classification technique that utilizes a feature vector of twenty graph measures which is then applied to SVM for classification. This approach did not have an overall better accuracy, but was consistently faster in comparison to other kernel-based graph classification techniques such as Random Walk, Shortest Path, Cyclic Pattern, Subtree, and Graphlet and Subgraph kernels. Further, this approach using the full graph measures was better than that using subgraph features and the most important features included: average clustering coefficient, number of nodes, number of eigenvalues, number of edges, energy (which is the squared sum of the eigenvalues of the adjacency matrix), and average degree. The smaller feature set was found to be sufficient to capture most of the important structural properties of the graph.

Ugander, Backstrom, and Kleinberg (2013) proposed a coordinate system based on triadic structure within subgraphs for characterizing possible sub-networks within a social network. This method begins by using a Markov Chain to model the frequency space of triadic evolution for a size $k = 3, 4$ subgraphs within a graph. They demonstrated their method on a Facebook® dataset by performing classification of sub-networks of various sizes into neighborhood, groups, and events through logistic regression. This was compared to the performance of using only global graph features such as size of k largest components, size of k-core, number of components in k-core, number of compositions in k-core, degeneracy, size of k-brace, and number of components in k-brace. The results showed that the proposed method performed much better than that which just uses global measures.

Lagraa et al. (2014) proposed a new distance measure for comparing graphs using modular decomposition for obtaining prime graphs. This is then used to compare with

other network's prime graphs using probe distance, which measures the number of edit operations needed to transform one graph into a second graph. Modular decomposition is first used to obtain prime graphs which are graphs that have only trivial modules. The authors then used select datasets to perform comparisons and classifications. The resulting distance and computation time were then compared against those obtained using regular edit distance and star distance, which was proposed by Zeng et al. (2009). The results showed that the prime distance is only comparable to the star distance in terms of runtime and acts as an upper bound for the star distance.

8.3.3 Summary Graph Matching and Graph Classification

Graph matching and classification methods are tools that can be used to determine the similarity between a network of interest and a hypothetical, proposed network. Although these methods have been refined through the introduction of advanced methods as discussed in the previous subsections, these methods still contain difficulties, especially as the size of the network grows. Based on computational approaches, graph matching and classification methods suffer from scalability issues in that computational resources and time requirements grow exponentially as the size of the network grows. Further, these methods may provide estimates as to the ability of matching a particular network of interest, but do not formally determine if the associated graph characteristics follow those that are being modeled. That is, the characteristics may be misspecified even though the overall fit is adequate. This notion is similar to when a straightline regression model is fitted to data that contains some curvature. Even though the overall error of the straight-line model may be within acceptable bounds, the fit of the straightline model may be mis-specified as there is curvature in the data. In short, the fitted model, although containing an acceptable amount of minimum error, lacks overall fit to the underlying data. Similarly, although there may be an acceptable amount of mis-matching, there still may be mis-specification, and therefore, alternate models may be more appropriate. Recall, network models are only useful if such models are able to characterize a network accurately enough for the intended analysis. Therefore, methods verifying that network characteristics and measures are appropriately modeled are paramount to network analysis.

8.4 Statistical Testing Methods

Statistical, inferential methods overcome many of the issues presented by graph matching and classification by formally testing whether or not a feature of a particular network follows the statistical distribution associated with the proposed network model property of interest. For instance, if the real world network is believed to follow that as described by an Erdös–Rényi network model, then the degree distribution should behave statistically as a Binomial distribution. Therefore, the degree distribution for the network of interest could be tested to determine if it is distributed as Binomial with specific parameters.

Recent advances in statistical inference for networks can be found in works such as Blasio, Seierstad, and Aalen (2011) and Mohd-Zaid (2016) both of which examined the linear preferential attachment property. Whereas Blasio, Seierstad, and Aalen considered statistical testing for the linear preferential attachment property within a network

in general, Mohd-Zaid directly developed a statistical test of hypothesis based upon the mechanisms governing the Barabási–Albert network model. The approach in this latter work provides a framework for setting up hypothesis testing based on the hypothesized distribution for network properties. This approach is outlined in detail for the degree distribution of a Barabási–Albert network model and may be used to derive statistical distributions for any network model.

Recall that under the property of linear preferential attachment governing the scale-free mechanism of the Barabási–Albert network model, Barabási and Albert developed equations that describe the probability that a new node would attach to an existing node (Barabási & Albert, 1999). Then, by fixing the minimum degree for any node (the minimum connections for any individual in the network), they derived the rate of change for the degree of each node over a fixed number of iterations (instances) at which a single node is added to the network on each iteration. It is from these equations that they determined that the asymptotic degree distribution follows a Pareto distribution with parameters m (the number of edges added on each iteration) and $\beta=2$. The distribution for a random variable X which is distributed according to the Pareto(m, 2) distribution is given as follows (Casella & Berger, 2002):

$$f(x \mid m, \beta = 2) = \beta m^{\beta} x^{-\beta-1} I_{[m,\infty)}(x) = 2m^2 x^{-3} I_{[m,\infty)}(x), \tag{8.4}$$

where $I_{[m,\infty)}(x)$ is an indicator function taking the value of 1 if x takes the value of m or larger, and 0 otherwise. In this manner, the Pareto distribution is defined for (positive) values of m (assuming the number of edges added on each iteration is at least 1). Knowing that the degree distribution for a Barabási–Albert network model would have the Pareto(m, 2) distribution, Mohd-Zaid developed statistical hypothesis tests based on the Pareto distribution in order to test the fit of the network to the degree structure of a Barabási–Albert network model. Three tests were developed: one for the parameter m, one for the parameter β, and one that simultaneously tests the parameters m and β. The assumption (null hypothesis) is that the parameter of interest is equal to the value assumed under the Barabási–Albert network model. The alternative hypothesis varies depending on the parameter.

Consider the test for m assuming that the value of β is known. To test whether or not m is larger than that hypothesized for the particular Barabási–Albert network of interest (denote this hypothesized value as m_1), we would reject the null hypothesis and conclude that the number of edges being added on each iteration is more than that hypothesized if the observed minimum degree for the network (call this value $x_{(1)}$) is larger than

$$x_{(1)} = m_1 / \alpha^{1/N\beta}, \tag{8.5}$$

where alpha is the Type 1 error rate, N is the total number of nodes, and β is the fixed parameter for the associated Pareto distribution. This is equivalent to testing the following:

$$H_o : m \leq m_1$$

$$H_1 : m > m_1$$

Similarly, we can test the value of β assuming that m is unknown. For this test, we would hypothesize that the value of β equals some value, (β_o), and we would reject this

value for the Pareto distribution if there is any evidence that the parameter β does not equal this value. These hypotheses are given by:

$$H_o : \beta = \beta_o$$

$$H_1 : \beta \neq \beta_o$$

Similar to the test for m, when the null hypothesis is rejected in this statistical test, it establishes that the degree distribution is not following the expected hypothesized Pareto distribution, and therefore is not following that of a Barabási–Albert network model. The test statistic to test the value of β is

$$T = \ln\left(\left(\prod_{i=1}^{N} x_i\right) \Big/ \left(x_{(1)}^{N}\right)\right), \tag{8.6}$$

where each x represents the degree for each of the N nodes and $x_{(1)}$ is the observed minimum degree for the network. This test will reject the null hypothesis, $H_o : \beta = \beta_o$, for

$$T \leq \frac{z_\alpha \sqrt{N-1} + (N-1)}{\beta_o} \quad \text{or} \quad T \geq \frac{z_{1-\alpha} \sqrt{N-1} + (N-1)}{\beta_o} \tag{8.7}$$

where z_α is the appropriate quantile of the standard normal (z) distribution.

If neither m nor β is known, then both parameters may be tested jointly such that if either one of the hypothesized values is rejected, then the conclusion would be that the network of interest does not have a degree distribution that follows the expected Pareto distribution associated with the Barabási–Albert network model and therefore, cannot be modeled as such. Such a conclusion is important in that, although the data may fit well enough (error-wise) to the network model, knowing that a particular feature of that network model is not being modeled correctly suggests that the assumed network model may be inappropriate and other models should be considered. The mathematical formulation of the hypotheses for this test is:

$$H_o : m \leq m_1 \cap \beta = 2$$

$$H_1 : m > m_1 \cup \beta \neq 2$$

To actually implement this joint test of hypothesis onto network data requires additional adjustments due to the non-independent nature of the nodes of the network (Mohd-Zaid, 2016). However, a simple version of the Pareto-based test as described above can be applied to the examples from Figures 8.2 and 8.3 to illustrate its use. First, consider the Erdös–Rényi graph in Figure 8.2. An application of the test on said graph should result in a rejection provided that the test has good power. Suppose the hypothesis that the graph is a Barabási–Albert graph is being tested on the data in Figure 8.2 with

$$H_o : m \leq 2 \cap \beta = 2$$

$$H_1 : m > 2 \cup \beta \neq 2$$

Then the test statistics for the graph based on its degree distribution given in Table 8.2 are $x_{\min} = 1$ and

$$T = \ln\left(\frac{3\times3\times3\times1\times4\times4}{1^6}\right) = 6.0684.$$

Assuming an acceptable Type-I error of $\alpha = 0.05$, the null hypothesis is rejected if

$$x_{\min} > \frac{2}{0.05^{1/(6\times2)}} = 2.5671$$

or

$$T \le \frac{-1.96\sqrt{6-1}+(6-1)}{2} = 0.3087$$

or

$$T \ge \frac{1.96\sqrt{6-1}+(6-1)}{2} = 4.6913.$$

Here, since $T \ge 4.6913$, the null hypothesis is rejected and it is unlikely that the graph follows a Barabási–Albert network model which is true in this case. Applying the same hypothesis test on the Barabási–Albert graph in Figure 8.3 gives the test statistics $x_{\min} = 2$ and $T = 2.0149$. Since $x_{\min} < 2.5671$and $0.3087 < T < 4.6913$, there is not enough evidence to reject the null hypothesis. Thus, an assumption that the graph in Figure 8.3 has a degree distribution that follows that of a Barabási–Albert network model, which is the ground truth, cannot be statistically disputed by the empirical evidence. With this knowledge, a Barabási–Albert network model may be used to represent this network, providing the analyst a means to study the characteristics of this network and to reasonably approximate what would happen to the network through the manipulation of the proposed Barabási–Albert network model. For instance, the analyst now can reasonably estimate how the network may expand, and whether or not the network is changing in structure if the network seems to be expanding in a way that is not anticipated.

Statistical hypothesis testing can be used to test the randomness property of the Erdös–Rényi network. Recall, we stated that if the network is believed to follow that described by an Erdös–Rényi network, then the degree distribution should behave statistically as a Binomial distribution. Therefore, the degree distribution for the network could be tested to determine if it is Binomial with specific parameters. Consider the same networks graphed in Figures 8.2 and 8.3. Suppose that we want to test the hypothesis that the networks in Figures 8.2 and 8.3 are random networks where each node connects independently to the other nodes with equal chance, thus $p = 0.5$. The hypothesis to be tested is then: H_o: $p = 0.5$ vs H_1: $p \neq 0.5$. For a network of size $n = 6$ where degree follows a Binomial distribution, the expected average degree is simply np, or in this case, 3. The test statistic (B) for the Binomial test is the number of instances where a node has a degree that is equal to the expected average degree. For the graph in Figure 8.2, three of the six nodes connect as expected with degree = 3 (Table 8.2). Thus, the test statistic for the Binomial test for the network graphed in Figure 8.2 is B = 3. Using a table of Binomial probabilities for $n = 6$ and a Type I error rate of 0.05, the rejection regions for this test are B > 5 or B < 1. Thus, we fail to reject the null hypothesis; there is no evidence to suggest that the graph in Figure 8.2 does not follow this property of the Erdös–Rényi network. That is, the network graphed in Figure 8.2

appears to follow the random property of the Erdös–Rényi network. The associated p-value for this test is 1.0. For Figure 8.3, B = 1 (see Table 8.3), and we once again fail to reject the null hypothesis; however, we acknowledge that the test statistic lies on the rejection region boundary. For this small graph and using the discrete Binomial distribution, the p-value for this lower bound is 0.1094 (the next possible p-value, if B = 0, is 0.0132). Given these values and the small size of these graphs, a Type I error rate of 0.05 may be too conservative for both tests.

8.4.1 Summary of Statistical Testing Methods

The use of statistical approaches to formally test network properties is a novel development in recent literature. Such approaches provide additional benefit to the network analyst in that these methods provide a means for characterization with a level of confidence that can be directly computed when network data is observed incompletely or with error. For these cases in which the error is within bounds, the real world data may still be generated and modeled using the hypothesized network model until such time that additional information is obtained. This provides the analyst with a means to observe network behavior on only a portion of the network. For example, if a portion of the network can be characterized as a Barabási–Albert network model, then the Pareto-based hypothesis test can be applied to this portion of the real world network in order to detect degradations of connections within the network (loss of nodes or edges) with good power (Mohd-Zaid, Schubert Kabban & Deckro, 2017).

8.5 Future Directions

Requirements for network characterization are rapidly evolving as SNA becomes a primary tool to passively observe and monitor the behavior within networks. As interest in network analysis and the size of the networks grows, tools that can rapidly model and generate networks, or portions of networks, are critical to time-sensitive responses. Novel methods such as the use of statistical testing to augment or supplant network classification or matching provide the analyst with confidence that the network model chosen for characterization is reasonably accurate. Further, and more critically to the national security analyst, it provides a sensitive tool to determine if a network has evolved or devolved. If the distribution of a network property has changed, then the network is also changing. Focusing analytical efforts on the regions of the network in which this change was experienced can provide insight into the social dynamics driving these changes. For example, an inquiry of interest may be whether certain agents in a real world network are disappearing or are growing connections to other agents. Such network changes may be indicative of future planned actions of the agents within the network.

Although the field of SNA is rooted in graph theory utilizing such techniques as graph matching and classification, the ability of statistical tests to scale accurately and compute more quickly make tools in this arena a viable option for future analyst needs. However, to use such tools, distributions must be traceable, or nonparametric testing techniques must be leveraged. The test of hypothesis approaches described in this chapter focused on statistical tests to determine if the degree distribution follows that expected for either a Barabási–Albert network model or an Erdös–Rényi network

model. Although many real data networks exhibit the scale-free property of a Barabási–Albert network (Barabási & Albert, 1999; Zhao et al., 2013), recent findings suggest that real-data networks do not solely possess this property (Small, Judd & Zhang, 2014; Zhang, Small & Judd, 2015). However, with the onset of methods to describe network data as a mixture of network properties, statistical tests can be developed to test both the appropriate mixing proportion of these properties for a network of interest and the overall fit of the mix to the observed data. For instance, the degree distribution may be a mixture of Barabási–Albert network properties (Pareto distribution for the scale-free property) and Erdös–Rényi network properties (Binomial distribution for the random re-wiring property). Further, statistically based methods extend to multilayer models which may be used to describe higher and lower order functioning in a network of interest (Hamill et al., 2008). In short, these methods for testing may provide efficient solutions to assure adequate characterization of a network through a network model even for complex social networks. Ideally, by capturing all the facets of a social network, via a mixture of attributes or functional layers, a better understanding and representation of the network can be created for analysis and monitoring. The same type of approach can be applied to characterizing infrastructure networks for further analysis.

There are many outstanding questions that are of primary interest to those analyzing networks. The detection of fabricated synthetic networks is of particular importance within the operations research community as it identifies the need to uncover the true relations and intent with the network of interest. Knowing when and how a network may grow, change, or otherwise metamorphosize may facilitate planning and response from parties of interest. The statistical tests discussed previously are an efficient means to help answer these questions by identifying when real-world networks are exhibiting characteristics that are, or are not, expected.

Finally, all methods of network analysis help to shape how networks are visualized. Despite the growing use and dependence on data and analysis tools, human analysts require a means to visualize the behaviors of interest. Characterization of real-world network data to a network model is paramount for this visualization, and statistical tests on network properties help maintain that these properties are appropriately represented within the network model. Understanding the social network, how it is structured, and how it may change to shape the world around us are key tasks of SNA. Therefore, network characterization, monitoring, and the translation to an interpretable visualization are tasks of vital importance to the analyst. The tools discussed in this chapter are a means for the analyst to accomplish these tasks for the security of people, places, and resources vital to our survival.

References

Albert, R., Jeong, H., & Barabási, A. (1999). Internet: Diameter of the world-wide web. *Nature*, 401, 130–131.

Barabási, A., & Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439), 509–512.

Barat, A., & Weigt, M. (2000). On the properties of small-world network models. *European Physical Journal B-Condensed Matter and Complex Systems*, 13(3), 547–560.

Blaha, L.M., Arendt, D.L., & Mohd-Zaid, F. (2014). More bang for your research buck: Toward recommender systems for visual analytics. Proceedings of the Fifth Workshop on Beyond Time and Errors: Novel Evaluation Methods for Visualization, BELIV'14, 126–133.

Blasio, B.F, Seierstad, T.G., & Aalen, O.O. (2011). Frailty effects in networks: Comparison and identification of individual heterogeneity versus preferential attachment in evolving networks. *Journal of the Royal Statistics Society: Series C (Applied Statistics)*, 60(2), 239–259.

Borgatti, S.P., Mehra, A., Brass, D.J., & Labianca, G. (2009). Network analysis in the social sciences. *Science*, 323(5916), 892–895.

Brass, D.J. (1995). A social network perspective on human resources management. *Research in Personnel and Human Resources Management*, 13, 39–79.

Casella, G., & Berger, R.L. (2002). *Statistical Inference*. Duxbury, Thompson Learning.

Conte, D., Foggia, P., Sansone, C., & Vento, M. (2004). Thirty years of graph matching in pattern recognition. *International Journal of Pattern Recognition and Artificial Intelligence*, 18, 265–298.

Daudin, J.J., Picard, F., & Robin, S. (2008). A mixture model for random graphs. *Statistics and Computing*, 18(2), 173–183.

Durante, D., Dunson, D.B., & Vogelstein, J.T. (2017). Nonparametric bayes modeling of populations of networks. *Journal of the American Statistical Association*, 112, 1516–1530.

Erdös, P., & A. Rényi. (1959). On Random Graphs I. *Publicationes Mathematicae Debrecen*, 6, 290–297.

Freeman, L.C. (1978). Centrality in social networks conceptual clarification. *Social Networks*, 1(3), 215–239.

Gilbert, J., Valveny, E., & Bunke, H. (2012). Graph embedding in vector spaces by node attribute statistics. *Pattern Recognition*, 45(9), 3072–3083.

Gonzalez, J.A., Holder, L.B., & Cook, D.J. (2002). Graph-based relational concept learning. Proceedings of the Nineteenth International Conference on Machine Learning (ICML), 219–226. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

Guzman, J.D., Deckro, R.F., Robbins, M.J., Morris, J.F., & Ballester, N.A. (2014). An analytical comparison of social network measures. *IEEE Transactions on Computational Social Systems*, 1(1), 35–45.

Hamill, J.T., Deckro, R.F., Chrissis, J.W., & Mills, R. (2008). Layered Social Networks. *IO Sphere*, pp. 27–33.

Havig, P.R., McIntire, J.P., Geiselman, E., & Mohd-Zaid, F. (2012). Why social network analysis is important to Air Force applications. Proceedings of SPIE, volume 8389, 83891E–83891E-9.

Inokuchi, A., Washio, T., & Motoda, H. (2000). An apriori-based algorithm for mining frequent substructures from graph data. Proceedings of the 4th European Conference on Principles of Data Mining and Knowledge Discovery (PKDD), 13–23. Springer-Verlag, London, UK.

Jin N., Young, C., & Wang, W. (2009). Graph classification based on pattern co-occurrence. Proceedings of the 18th ACM Conference on Information and Knowledge Management (CIKM), 573–582. New York, NY, USA.

Ketkar, N.S., Holder, L.B., & Cook, D.J. (2009). Empirical comparison of graph classification algorithms. Proceedings of the 2009 IEEE Symposium on computational Intelligence and Data Mining CIDM, 259–266.

Lagraa, S., Seba, H., Khennoufa, R., MBaya, A., & Kheddouci, H. (2014). A distance measure for large graphs based on prime graphs. *Pattern Recognition*, 47(9), 2993–3005.

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

Li, G., Semerci, M., Yener, B., & Zaki, M.J. (2012). Effective graph classification based on topological and label attributes. *Statistical Analysis and Data Mining*, 5(4), 265–283.

Li, P., Zhang, J., & Small, M. (2011). Emergence of scaling and assortative mixing through altruism. *Physica A: Statistical Mechanics and its Applications*, 390, 2192–2197.

Livi L., & Rizzi, A. (2013). The graph matching problem. *Pattern Analysis and Applications*, 16(3), 253–283.

Macindoe, O., & Richards, W. (2010). Graph comparison using fine structure analysis. SOCIALCOM'10 *Proceedings of the 2010 IEEE Second International Conference on Social Computing*, Minneapolis, MN, 193–200.

Mohd-Zaid, M.F. (2016). A statistical approach to characterize and detect degradation within the Barabási-Albert network, PhD diss., Air Force Institute of Technology (AFIT-ENC-DS-16-S-003).

Mohd-Zaid, F., Schubert Kabban, C.M., & Deckro, R.F. (2017). A test on the L-moments of the degree distribution of a Barabási -Albert network for detecting nodal and edge degradation. *Journal of Complex Networks*, 6(1), 24–53.

Moonesinghe, H.D.K., Valizadegan, H., Fodeh, S., & Tan, P-N. (2007). A probabilistic substructure-based approach for graph classification. 19th IEEE International Conference on Tools with Artificial Intelligence (ICTAI), 1, 346–349.

Moreno, J.L. (1934). *Who Shall Survive? A new Approach to the Problem of Human Interrelations*. Beacon House. ISBN 978-9992695722.

Morris, J.F., & Deckro, R.F. (2013). SNA data difficulties with dark networks. *Behavioral Sciences of Terrorism and Political Aggression*, 5(2), 70–93.

Morris, J.F., ONeal, J.W., & Deckro, R.F. (2014). A random graph generation algorithm for the analysis of social networks. *The Journal of Defense Modeling and Simulation: Applications, Methodology, Technology*, 11(3), 265–276.

Newman, M.E.J. (2010). *Networks: An Introduction*. Oxford University Press, Cambridge.

Nguyen, P.C., Ohara, K., Mogi, A., Motoda, H., & Washio, T. (2006). Constructing decision trees for graph-structured data by chunkingless graph-based induction. *Proceedings of the 10th Pacific-Asia Conference on Advances in Knowledge Discovery and Data Mining* (PAKDD), Springer-Verlag, Berlin, Heidelberg, 390–399.

Office of the Joint Chiefs of Staff. (2016). *Joint Concept for Human Aspects of Military Operations*. Washington, DC: Department of Defense.

Richards, W., & Wormald, N. (2009). Representing small group evolution. *Proceedings of the 2009 International Conference on Computational Science and Engineering, IEEE Computer Society*, Washington, DC, USA, 4, 159–165.

Seshadhri, C., Kolda, T., & Pinar, A. (2012). Community structure and scale-free collections of Erdos-Renyi graphs. *Physical Review E*, 056–109.

Small, M., Judd, K., & Zhang, L. (2014). How is that complex network complex? *IEEE International Symposium on Circuits and Systems (ISCAS)*, 1263–1266.

Smith, R. (2007). *The Utility of Force: The Art of War in the Modern World*. Alfred A. Knopf, New York.

Thomas, C., Stoica, A., & Beuscart, J-S. (2010). Online social network popularity evolution: An additive mixture model. *Proceedings of the 2010 International Conference on Advances in Social Network Analysis and Mining* (ASONAM), 346–350.

Ugander, J., Backstrom, L., & Kleinberg, J. (2013). Subgraph frequencies: Mapping the empirical and extremal geography of large graph collections. *Proceedings of the 22nd International Conference on World Wide Web*, 1307–1318.

Wasserman, S., & Faust, K. (2009). *Simple Distributions. Social Network Analysis*. New York: Cambridge University Press.

Watts, D.J., & Strogatz, S.H. (1998). Collective dynamics of 'small-world' networks. *Nature*, 393(6684), 440–442.

West, D.B. (2001). *Introduction to Graph Theory*. Prentice Hall.

Yan, X., Cheng, H., Han, J., & Yu, P.S. (2008). Mining significant graph patterns by leap search. *Proceedings of the 2008 ACM SIGMOD International Conference on Management of Data* (SIGMOD), 433–444.

Zanghi, H., Ambroise, C., & Miele, V. (2008). Fast online graph clustering via Erdös–Rényi mixture. *Pattern Recognition*, 41(12), 3592–3599.

Zeng, Z., Tung, A.K.H., Wang, J., Feng, J., & Zhou, L. (2009). Comparing Stars: On Approximating Graph Edit Distance. *Proceedings of the VLDB Endowment*, 2(1), 25–36.

Zhang, L., Small, M., & Judd, K. (2015). Exactly scale-free scale-free networks. *Physica A: Statistical Mechanics and its Applications*, 433, 182–197.

Zhao, Z.D., Yang, Z., Zhang, Z., Zhou, T., Huang, Z.G., & Lai, Y.C. (2013). Emergence of scaling in human-interest dynamics. *Scientific Reports*, 3, 3472.