Generating Artificial Social Networks

Tobias Johansson

Kristianstad University

Tobias.Johansson@hkr.se

The study of complex social networks is an inherently interdisciplinary research area with applications across many fields, including psychology. Social network models describe, illustrate and explain how people are connected to each other and can, for example, be used to study information spread and interconnectedness of people with different kinds of traits. One approach to social network modelling, originating mainly in the physics literature, is to generate targeted kinds of social networks using models with specialized mechanisms while analyzing and deriving features of the models. Surprisingly though, and despite the popularity of this approach, there is no available functionality for generating a wide variety of social networks from these models. Thus, researchers are left to implement and specify these models themselves, restricting the applicability of these models. In this article, I provide a set of Matlab functions enabling the generation of artificial social networks from 22 different network models, most of them explicitly designed to capture features of social networks. Many of these models originate in the physics literature and may therefore not be familiar to psychological researchers. I also provide an illustration of how these models can be evaluated in terms of a simulated model averaging approach and how they can be applied to psychological research. With the already existing network functionality available in Matlab and other languages, this should provide a useful extension to researchers.

Keywords: Social Network; Model; Psychology; Matlab; Function; Clustering; Community.

DRAFT 2017-01-13

Introduction

The analysis of complex networks, or network science, is an interdisciplinary field emerging mainly from sociology and from graph theory in mathematics (Watts, 2004). Basically, it concerns the understanding of relational phenomena through studying complex networks of interconnected components and its applications include problems in physics (Bianconi & Barabási, 2001), bioinformatics (Barabási & Oltvai, 2004), epidemiology (Meyers, Pourbohloul, Newman, Skowronski, & Brunham, 2005), economics (Mayer, 2009), sociology (McPherson, Smith-Lovin, & Cook, 2001), psychology (Perry-Smith & Shalley, 2003) as well as many other areas (Borgatti, Mehra, Brass, & Labianca, 2009; Newman, 2003). Among the many network models available, the subset of *social network models* focus on how people interconnect in various domains. As such, these models attempt to capture the structure of social networks in different ways. In some contexts such models are highly relevant for psychological research, because the structure of a social network can impose fundamental constraints on how psychological variables impact individuals across the network (Brass, Butterfield, & Skaggs, 1998; Fowler & Christakis, 2008).

One possible dimension along which to classify social network models is "Generic-Specialized". In the one end of this dimension we find generic statistical network models that focus on issues of statistical inference. These models typically do not focus on the mechanisms by which networks evolve, but rather on identifying network probability distributions. For example, one such framework is offered by exponential random graph models (Robins, Pattison, Kalish, & Lusher, 2007). In the other end of the "Generic-Specialized"-dimension we find specialized network models that contain explicit specifications of the mechanisms by which networks evolve. These models are often introduced with a specific purpose in mind, such as suggesting mechanisms by which community structure can develop in a network (Li & Maini, 2005). Models on the Generic end of the dimension are well represented and available in different computational packages and functions (Handcock, Hunter, Butts, Goodreau, & Morris, 2003; Ripley, Snijders, Boda, Vörös, & Preciado, 2015). On the contrary, models on the Specialized end of the dimension are, on the whole, not nearly as available in the form of program functions.

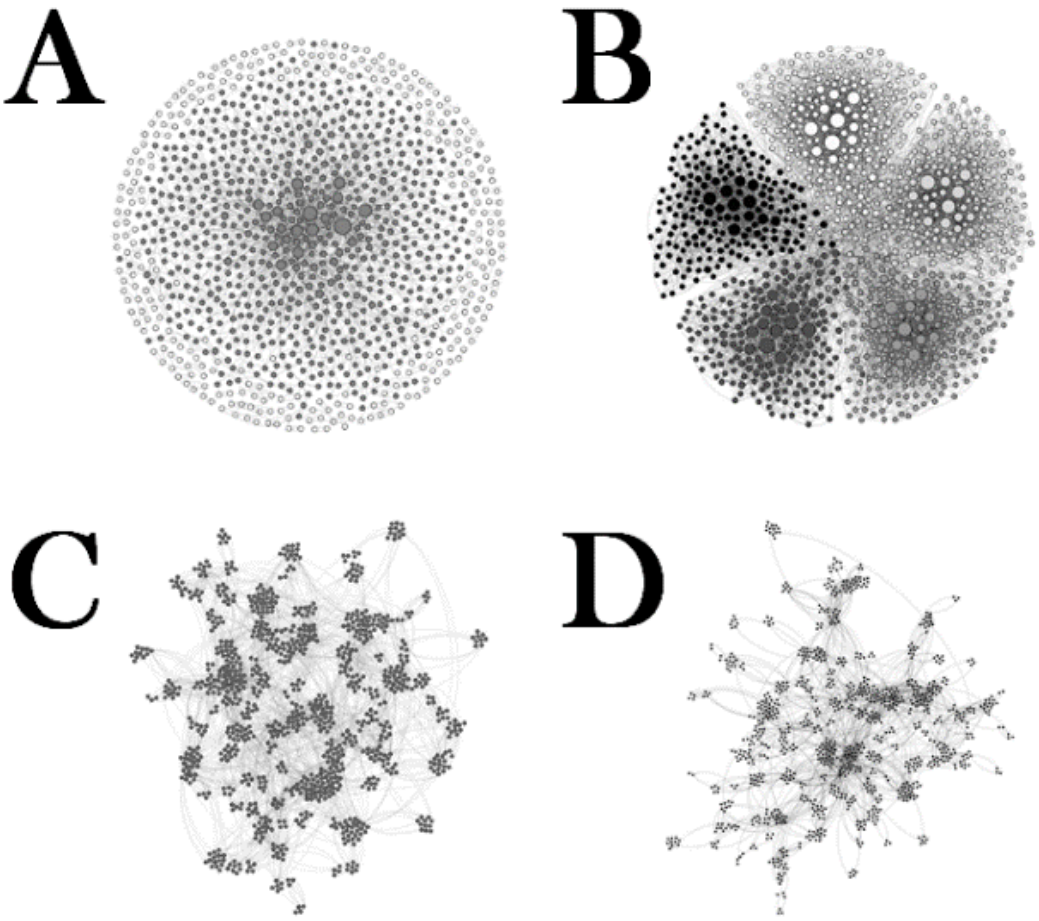

Figure 1. **Artificial Social Networks. A: Singer model** (Singer, Singer, & Herrmann, 2009)**. Grey vertices belong to largest component. B: Li model** (Li & Maini, 2005)**. Greyscale brightness is based on community structure according to the algorithm in** (Blondel, Guillaume, Lambiotte, & Lefebvre, 2008)**. C: Toivonen model** (Toivonen, Onnela, Saramäki, Hyvönen, & Kaski, 2006)**. D: Leskovec model** (Leskovec, Lang, Dasgupta, & Mahoney, 2008)**. A and B visualized using Fruchterman-Reingold algorithm** (Fruchterman & Reingold, 1991) **and C and D using OpenOrd algorithm** (Martin, Brown, Klavans, & Boyack, 2011) **in Gephi 0.9.1** (Bastian, Heymann, & Jacomy, 2009)**. Vertex size is proportional to degree in A and B.**

For example, both the SNAP library (Leskovec & Sosič, 2016) and the *igraph* R package (Csardi & Nepusz, 2006) implement models for generating networks, but very few of these are explicit social network models. On the one hand, this may not be surprising, because specialized models often attract the interest of specialized researchers who write their own functions. On the other hand, this is unfortunate, because it means that some researchers who might benefit from using a variety of these models end up not doing so because of the work required in implementing them. Furthermore, many of the specialized social network models originate in the physics literature. This means that these models may not always be familiar to researchers in other disciplines, such as psychology.

In order to make specialized social network models more readily available, this article provides Matlab (*Matlab*, 2016) functions for generating artificial networks from 22 different network models. This set of models, although not exhaustive, is representative of the kinds of social network models originating from the physics literature. The structure of the rest of this article is as follows. First, I consider some common social network terminology and features. Second, I describe different classes of social network models. Third, I introduce the Matlab functions implementing the network models and illustrate their use for a few of the models. Fourth, by way of examples, I illustrate some cases for which social network modelling may be relevant to psychological research. Fifth and finally, I discuss the relevance of different kinds of modelling approaches, including an illustration of how to compare different models. The overarching aim of this article is to enhance the availability and applicability of specialized social network models.

Social Network Features

A network or graph $G(V,E)$ consists of n vertices $v \in V$ and m edges $e \in E$ connecting some of the vertices. In a social network, the vertices typically represent people and the edges represent relations or interconnectedness in some sense. For example, an edge connecting two vertices could represent an existing friendship relation between two people. The networks considered here are all undirected, meaning that the edges represent symmetric relations. Most of the networks considered here are also unweighted, meaning that all existing edges have the same weight. Vertices in a network form components, so that all vertices within a component are reachable via paths through other vertices. Figure 1A shows an example of an artificial social network. The grey vertices represent the largest component and vertex size is proportional to its degree k: how many other vertices it is directly connected to.

There are a number of characteristics often exhibited by social networks. For example, the *degree distribution* $P(k)$ of social networks is often highly skewed, so that there is a small number of vertices with large k and a large number of vertices with small k (as in Figure 1) and can often be approximated by power-law or exponential distributions (or a combination thereof (Clauset, Shalizi, & Newman, 2009)): $P(k) \sim k^{\gamma}$ or $P(k) \sim e^{\lambda k}$ respectively.

Another feature typical of social networks is *short average shortest distance* $\bar{D}$ (Travers & Milgram, 1969; Watts & Strogatz, 1998). The shortest summed distance $D_{i,j}$ between vertices v_i and v_j is the smallest number of edges in G through which v_j can be reached from v_i, so that each traversed edge equals a distance of 1. The average shortest distance $\bar{D}_{i,j}$ is then simply the average of all $\binom{n}{2}$ shortest distances in G (assuming G has only one component).

Social networks also show a relatively *high average clustering*, which can be measured through the average clustering coefficient $\bar{C}$ (Watts & Strogatz, 1998). The clustering coefficient C_i for each vertex v_i ranges from 0 to 1 and is a measure of the extent to which vertices connected to v_i ("neighbors" to v_i) are connected to each other. Clustering in the entire network can then be measured by the average of these, so that $\bar{C} = \frac{\sum_{i=1}^{n} C_i}{n}$.

Many social networks also reveal distinct *community structure*. There are various definitions of what community structure means and many different associated algorithms (Blondel et al., 2008; Newman, 2011, 2016; Palla, Derényi, Farkas, & Vicsek, 2005), but a common denominator is that community structures in some sense form relatively dense subnetworks within a larger network. For example, people in a workplace have different assignments and this puts natural constraints on which and how many people they interact with. People within a community are closer to each other than they are to people in other communities on average. Figures 1B, 1C and 1D show artificial networks with clear community structure. The number of communities will depend on one´s definition, but the network can clearly be partitioned into a number of different groups with much fewer connections between than within them.

A final typical feature of social networks is *assortativity* or homophily. This property designates similarity between vertices and neighbors. In the case of social networks, this means that people tend to connect with people similar to themselves. For example, social networks often exhibit assortativity with respect to degree k, so that increases in k is associated with increases in average k for neighboring vertices (Newman, 2002).

Classification of Network Models

Network models can be classified in several ways. Here, I focus on the mechanisms underlying the models and use the classification given in (Toivonen et al., 2009), which includes three categories of models: *Growing*, *Dynamical*, and *Spatial*. For completeness, I add a fourth category: *Static*, even though these models are not explicitly social network models. Table 1 summarizes these four types of models.

Static models typically involve some probabilistic procedure performed on each element of the network. For example, the classic Erdős–Rényi model (Erdős & Rényi, 1959; Gilbert, 1959) independently forms edges between vertices in a network with probability p. There is no growth mechanism and no interaction between the elements of the network incorporated into the network formation procedure. Thus, the model is static.

Growing models start with a number of $n0$ vertices and then implements a growth procedure whereby the network accumulates both vertices and edges until the network has n vertices. The classic Barabási-Albert model (Barabási & Albert, 1999) is an example of a growing network model. Here, the network starts with a number of vertices (m+1 in the current implementation). On each time step a new vertex and m edges are added to the network. The m edges connect the new vertex to m existing vertices selected using preferential attachment, so that the probability of being selected is proportional to degree.

In dynamical models no vertices are added to the network. Instead, the network starts out with the final number of vertices while adding, removing, and/or modifying edges iteratively. For example, in the Davidsen model (Davidsen, Ebel, & Bornholdt, 2002) we start with n vertices. On each iteration an edge is formed between two neighbors of a random vertex and, with probability p, a random vertex and its edges are removed and replaced with a new vertex with one random edge.

Table 1. *Classification of Network Models.*

Type	*Description*
Static	Independent formation of edges via pre-defined probability.
Growing	Network grows from $n0$ to n vertices, acquiring edges along the way.
Dynamical	Network contains n vertices, acquiring edges iteratively.
Spatial	Vertices are distributed in space, updating their position or acquiring edges as a function of distance and attributes associated with vertices.

Finally, in spatial models vertices are distributed in space. These models may also involve attributes other than spatial position associated with each vertex. Edge formation or updating of vertex position is then typically a function of distance and vertex attributes. For example, in the Boguna model (Boguñá, Pastor-Satorras, Díaz-Guilera, & Arenas, 2004) vertices are distributed uniformly in a social space. Edges are then formed as a function of distance in social space.

In a few cases a model may belong to several categories. This is especially so for the Qiao model (Qiao, Huang, Li, & Fan, 2014) which is explicitly designed to serve as a bridge between different types of social network models. Depending on parameter values, the model can behave as both a growing and a dynamical network model.

It is important to recall that different models proposed in the literature have been proposed with different purposes and scope in mind. As a result, some of them capture features of social networks more adequately than others. For example, the Barabási-Albert model (Barabási & Albert, 1999) produces highly skewed (power-law) degree distributions in line with real social networks, but the clustering, assortativity and community structure associated with the model is far from that of real social networks. This is not surprising, because the model was not designed to capture those features. Other models have extended the mechanisms in the Barabási-Albert model to capture specific features of social networks (Catanzaro, Caldarelli, & Pietronero, 2004; Li & Maini, 2005).

A comparison of models explicitly designed to capture features of social networks was reported in (Toivonen et al., 2009). This comparison involved seven specialized network models (all of which are included among the functions in this article) and also an exponential random graph model. In this comparison, the spatial models produced adequate community structure and assortativity, but unrealistic clustering and degree distributions. The latter two features were better captured by growing and dynamical network models. The exponential random graph model produced instable results as well as weak community structure. In the next section I describe the use of the current Matlab functions implementing different types of social network generation.

Matlab Functions for Social Network Generation

Associated with this article is a set of 22 Matlab functions for generating artificial networks. Table 2 describes each of these models and related parameters. For the most part, the functions utilize the core functionality of Matlab. Some functions utilize the **randsample** function though, which is included in the Statistics and Machine Learning Toolbox available from Mathworks. For most of the included network models, the output returned when calling a function is the n^*n adjacency symmetric zero-diagonal matrix **M** describing the generated network of n vertices. If there is an edge between vertex i and j then $M_{ij} = M_{ji} = 1$, otherwise $M_{ij} = M_{ji} = 0$.

Running these functions in Matlab is very simple. Download the function files and put them in the same folder. Set this folder as the current folder in Matlab or add the folder to the search path (type **help cd** or **help addpath** in the Command Window, or set current folder directly by point and click when opening Matlab). To view the help text associated with each function, simple write **help <function name>**. For example, to view the help text associated with the Barabási-Albert model (Barabási & Albert, 1999) write **help barabasi** and the following output appears:

```
barabasi  Generate Barabasi-Albert network

  M = barabasi(N,M) computes adjacency matrix M according to the
  Barabasi-Albert model. N is the number of vertices and M is the
  number of edges added each time-step as the network grows. The
  starting network consists of M+1 vertices and 0 edges.
  See DOI: 10.1126/science.286.5439.509

  Example
     n = 1e3;
     m = 4;
     M = barabasi(n,m);
     Computes adjacency matrix M for the Barabasi-Albert model with
     1000 vertices, adding 4 edges each time-step starting from an
     unconnected network of 5 vertices.
```

In order to run this function we need to specify its two parameters, n and m. In order to generate a network with 200 vertices and $m = 3$, we write **M = barabasi(2e2,3)** in the command window. Matlab has some graph and network algorithms in its core functionality (see **help graph**) that can be useful. One can also visualize

networks easily in Matlab, although for larger networks there are better programs available. For a quick and raw visualization where vertex color is matched to degree use the code in Figure 2A.

As noted, for most of the functions the returned output is the symmetric unweighted adjacency matrix **M**. There are three exceptions to this, namely for the Singer, Axelrod, and Parravano model functions. The Singer function returns **M** and, if requested, a combined friendship and interest function (see **help singer**); the Axelrod function returns an edge-weighted adjacency matrix **M** (see **help axelrod**); finally, the Parravano function returns a list of [*x*, *y*]-coordinates for the vertices (see **help parravano**). The Axelrod and Parravano functions also visualize the network formation process iteratively during simulation.

The list of models in Table 2 encompasses models of many different kinds and the reader is referred to original sources in order to gain a fuller understanding of the types of networks generated by each model. I offer one example here using the Toivonen model (Toivonen et al., 2006), or T model for short. The T model is a growing network model able to generate networks with clear community structure, relatively high clustering, relatively high assortativity and highly skewed degree distributions. Thus, it captures many of the features of social networks (Johansson, 2016; Toivonen et al., 2009). The model has four parameter, n, n_0, p and *lim*. The network grows from a chain of n_0 vertices to a total of n vertices, adding one vertex each time step. On each time step, the new vertex is attached to one random vertex with probability p and to two vertices with probability 1-p. These random vertices are called primary contacts. The neighbors of these primary contacts are called secondary contacts and the new vertex is attached to m secondary contacts of each primary contact, where m is a uniformly distributed integer with lower and upper limits given by *lim*. Figure 2B shows a realization of this network revealing clear community structure, with parameter values n = 200, n_0 = 10, p = .95 and *lim* = [0, 3]. Given that we can generate artificial social networks one might naturally ask how social networks may be relevant to psychological research. In the next section I consider two examples in detail to illustrate this relevance.

Social Networks in Psychological Research

In this section I consider two fictitious examples of how social network models may be applied to psychological research. Of course, there are many examples using real data illustrating this basic point in many different ways (Brass et al., 1998; Feiler & Kleinbaum, 2015; Fowler & Christakis, 2008; Krause, James, & Croft, 2010; Mollgaard & Mathiesen, 2016). Nevertheless, by using fictitious data we can tailor the data to highlight central aspects of the examples more clearly. One example illustrates how personality traits can be incorporated into specialized network models and the other example illustrates how the structure of a network could affect opinion formation.

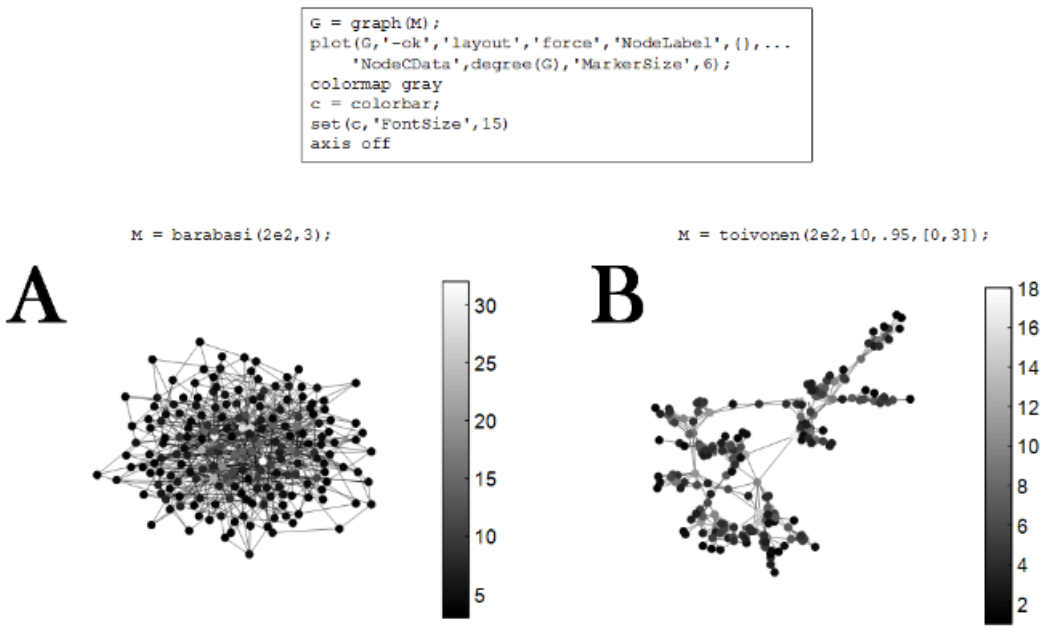

Figure 2. Visualization of Barabasi (A) and Toivonen (B) Network in Matlab. Greyscale bar indicate degree *k*.

Table 2. *Description of Network Models.*

Name	*Type*	*Parameters*	*Description*
Axelrod (Axelrod, 1997)	Spatial	n_c, n_f, n_t, T, D	n_c^2 vertices ("cultural sites") are connected in a square lattice. Each vertex is associated with n_f cultural features, each taking a uniformly random value from 1 to n_t. On each of T iterations a vertex v_i and one of its neighbors v_j are randomly selected. If they differ on any of their n_f features then, with a probability matching the percentage of shared features between them, one of the differing features is selected from v_i and substituted for the corresponding value of v_j. The returned edge-weighted adjacency matrix M indicates the percentage of shared features plus a small number (floating point relative accuracy, for drawing edges). D is a drawing parameter. If D is 1 then the network is visualized over time (wider edges means more shared features) and if D is 0 the network is not visualized and computed much faster.
Barabasi (Barabási & Albert, 1999)	Growing	n, m	n is the number of vertices and m is the number of edges added each time-step as the network grows. The starting network consists of m+1 vertices and 0 edges. New edges are formed with preferential attachment.
Boguna (Boguñá et al., 2004)	Spatial	n, a, b	n is the number of vertices distributed uniformly in a one-dimensional space from 0 to 1. Edges are formed with probability $p = 1/(1+(d/b)^a)$, where d is the distance in space.
Caldarelli (Caldarelli, Capocci, De Los Rios, & Muñoz, 2002)	Static/spatial	n, k	Uses a power-law fitness distribution $p(x) \sim x^{-k}$ where x is uniformly distributed in the interval [0,1]. n is the number of vertices and k is the power-law exponent. The edge probability function is $f(x_i,x_j) = (x_i * x_j)/\max(x)^2$.
Callaway (Callaway, Hopcroft, Kleinberg, Newman, & Strogatz, 2001)	Growing	n	n is the number of vertices. One edge is added each time-step with random attachment.
Catanzaro (Catanzaro et al., 2004)	Growing	n, p, f	n is the number of vertices, p is the probability of adding one new vertex on the current iteration and connecting it to an existing vertex selected with a probability proportional to degree (same as Barabasi model). 1-p is the probability of connecting two unconnected existing edges, v_1 and v_2, on the current iteration, where v_1 is selected with a probability proportional to its degree and v_2 is selected with a probability proportional to the relationship between its own degree and the degree of v_1. f specifices the functional form of this relationship, where f = 1 specifies an inverse functional form and f = 2 specifies an exponential functional form.
David (Davidsen et al., 2002)	Dynamical	n, p, it	n is the total number of vertices and it is the number of iterations. On each iteration 1) an edge is formed between two neighbors of a random vertex and 2) with probability p, a random vertex along with its edges is removed and replaced with a new vertex with one random edge.
Erdos (Erdős & Rényi, 1959)	Static	n, p	n is the number of vertices and p is the probability of an edge between two random vertices.
Grabowski (Grabowski & Kosiński, 2006)	Spatial	l, σ, Q, t, k_{max}	The network consists of n vertices arranged in a square lattice with sides of length l, so that $n = l^2$. Each vertex is associated with Q normally distributed features with mean 0 and standard deviation σ, with the features rounded to integers. Over the course of t iterations the features of a vertex change as a function of its current features and those of its neighbors. Feature similarity between vertices is then used to define social distance, which in turn is used to update edges. The edges in the original lattice always remain, but others may form and disappear. k_{max} sets the maximum degree of a vertex. If k_{max} is reached the edge to the most socially distant vertex is deleted. If k_{max} is not reached an edge is formed to the socially closest vertex within the neighbors of neighbor. If there are several then the spatially closest is formed. If there are still several, then a random one of these is formed.
Grindrod (Grindrod, 2002)	Static/Spatial	n, a, b	n is the number of vertices. Edges are formed with probability $p = a^* b^{-k}$, where k is $\lvert i-j \rvert$ for the natural ordering of the vertices from i = 1,2,...n, and a and b are constrained to the range [0,1].
Kumar (Kumar, Novak, & Tomkins, 2006)	Growing	t, m, g, $p1$, $p2$, vi	On each of t iterations one new vertex is first added. This vertex is a "passive" with probability $p1$, a "linker" with probability $p2$, or an "inviter" with probability 1-$p1$-$p2$.

Name	*Type*	*Parameters*	*Description*
			Then, m edges are added. One of the vertices (v_s) in each edge pair (v_s, v_t) is sampled based on degree from existing linkers and inviters. If v_s is a linker, then v_t is sampled based on degree from linkers and inviters, but favoring linkers by a factor of g. If v_s is an inviter, then v_t is a new vertex added to the network and designated a "passive". $v1$ is the number of unconnected initial vertices.
Kumpula (Kumpula, Onnela, Saramäki, Kaski, & Kertész, 2007)	Dynamical	n, pd, pe, pr, it	n is the number of vertices. For each of it iterations, pd is the probability of connecting a random vertex v_r with one other vertex along a local weighted updated search path, pr is the probability of connecting v_r with another random vertex, and pe is the probability of deleting a vertex and its edges and replacing it with a new unconnected one. The model involves additional fixed parameters: w_0 (initial vertex strength = 1), Δ (weight increase = .5).
Leskovec (Leskovec, Kleinberg, & Faloutsos, 2007; Leskovec et al., 2008)	Growing	n, m	The model uses a burning algorithm and generates directed networks in its original formulation. Here, I have adapted it to generate undirected networks. The original model contains 3 parameters, while this only contains 2. The network grows from a single vertex to n vertices. On each time step a new vertex v_i is added and connected to a random existing vertex v_j. Then, x neighbors to v_j are randomly selected, where x is a random number from a geometric distribution with mean m. If $x = 0$ or if there are no neighbors then proceed to the next time step. Otherwise, if v_j has fewer than x neighbors then all neighbors are selected, Then repeat the sampling procedure from the neighbors of the x neighbors, and so on, conditioned on not revisiting the same vertex on the current time step. Then connect v_i to all visited vertices.
Li (Li & Maini, 2005)	Growing	n, Mc, $m0$, me, a	The model starts with Mc internally fully connected communities of $m0$ vertices. Edges are then formed between communities from random vertices within communities so that all communities are connected by an edge. The network then grows to a total of n vertices by the following procedure: On each time step, a new vertex v_i is added and assigned to a community randomly. Vertex v_i is then connected to $1 <= me <= m0$ vertices within its community using preferential attachment based on within-community degree. Then, with probability a, v_i is connected to me vertices in other communities using preferential attachment based on between-community degree.
Marsili (Marsili, Vega-Redondo, & Slanina, 2004)	Dynamical	n, a, b, c, it	n is the total number of vertices, a is the probability of connecting two random vertices, b is the probability of connecting to secondary contact, c is the probability of deleting a random edge and it is the number of iterations.
Parravano (Parravano & Reyes, 2008)	Spatial	n, ρ, γ, R, T, D	This is a gas like social network model, consisting of n particles. At first, the particles are distributed uniformly random in a square with side l = sqrt(n/ ρ), where ρ is the density. Each particle has an unchanging internal state x, whose distribution is specified through parameter D. D = 1: uniformly random from 0 to 1; D = 2: uniformly random from -1 to 1; D = 3: all 1; D = 4: half -1, half 1. Using periodic boundaries the particles then move as they interact with particles in their neighborhood. The neighborhood is defined as a circle with radius l/R with the particle in the center, where $R \geq 2$. The movement takes place over T iterations using Eq. 2 in (Parravano & Reyes, 2008), depending on the parameter γ.
Qiao (Qiao et al., 2014)	Growing/ Dynamical	m_0, N, E, p_s, p_a, p_t, p_n, p_p, p_r	m_0 is the number of initial unconnected vertices, N is the final number of vertices and E is the final number of edges. The model can be either growing or dynamic depending on configuration and includes four types of mechanisms for connecting vertices: random attachment, transitive attachment (connecting via neighbors) within social groups, preferential attachment and non-preferential attachment. The model algorithm is relatively involved. See (Qiao et al., 2014) for further explication of the algorithm and probability parameters p_s, p_a, p_t, p_n, p_p, and p_r.
Singer (Singer et al., 2009)	Spatial	n, σ, it	The model starts with n unconnected vertices and forms connections over it iterations. On each iteration a randomly selected vertex seeks an encounter, either with

Name	*Type*	*Parameters*	*Description*
			a connected vertex, or with an unconnected vertex, with probabilities as a function of both degree and a parameter governing the tendency of each vertex to seek new acquaintances. The σ parameter specifies the standard deviation of the normally distributed acquaintance parameter with a mean of 1 and truncated at 0. Encounters with connected vertices are governed by a combined friendship and interest function, while encounters with unconnected vertices are governed by degree and a threshold related to σ. Connections can be replaced if suggested connections reveal gain according to the interest function. The combined friendship and interest function, scaled between 0 and 1, constitutes a combination of how many times two vertices have encountered each other and how well they match each other. The interest function itself is uniformly distributed between 0 and 1.
Toivonen (Toivonen et al., 2006)	Growing	*n, n₀, p, lim*	*n* is the total number of vertices, n_0 is the number of initial vertices, *p* is the probability of a new node attaching to one random vertex (1-*p* is the probability of attaching to two random vertices), and *lim* is a vector with two elements indicating the lower and upper boundaries of a uniform distribution used to select the number of secondary contacts.
Vazques (Vázquez, 2003)	Growing	*n, u, m₀, m*	*n* is the total number of vertices, 1-*u* is the probability of connecting a new vertex to m_0 random vertices, and *u* is the probability of converting *m* potential edges to actual edges.
Wong (Wong, Pattison, & Robins, 2006)	Spatial	*n, p, pb, h*	*n* is the number of vertices distributed uniformly in a two-dimensional unit space, *p* and *pb* are probabilities and *h* is a threshold.
Watts (Watts & Strogatz, 1998)	Static	*n, k, β*	*n* is the number of vertices, 2^*k is the mean degree, and β is the probability of rewiring the edges of each node starting from a regular ring lattice.

Personality Traits and Networks

Personality traits, or "relatively enduring styles of thinking, feeling, and acting." [50, p. 509], have a long-standing place in the psychological literature (Allport & Allport, 1921). Traditionally, personality traits are conceptualized either as theoretical terms in a nomological network (Cronbach & Meehl, 1955)or as latent variables causing observed responses (Borsboom, Mellenbergh, & van Heerden, 2004). A relatively new approach links personality traits directly to network theory by viewing indicators of personality traits as directly and indirectly linked interacting components in a network (Cramer et al., 2012). Although this latter approach is promising in reducing the abstract components of personality theory to concrete causal relations, we shall for the sake of simplicity stick to a simple conception of personality when relating personality traits to *social* networks.

One commonly investigated trait dimension is *extraversion/introversion* (Eysenck & Eysenck, 1963; McCrae & Costa, 1997), where extraverted individuals are more sociable and outgoing and introverted individuals are more solitary. Thus, on average extraverts have more friends than introverts. Furthermore, if two individuals have similar levels of extraversion they are more likely to become friends (Feiler & Kleinbaum, 2015). Suppose our only objective is to set up a social network model that captures these two features: 1) degree related to extraversion and 2) assortativity with respect to extraversion. To this end, we can use a spatial model and associate each vertex with a realization of extraversion (*E*). For example, in order to incorporate assortativity with respect to extraversion we can use the Boguna model (Boguñá et al., 2004), where the probability of an edge is a function of distance in 1D social space. If we replace this social space with *E* then the probability of an edge will reflect distance in *E*. We can also add an additional influence reflecting the mean *E* with respect to two vertices and specify the weight of each influence. Thus, in the original Boguna model, edges between vertices *i* and *j* are formed with probability

$$p_{i,j} = \frac{1}{1+\left(\frac{dS_{i,j}}{b}\right)^{a}}, \quad (1)$$

where *dS* is distance in social space in the interval [0, 1] and *a* and *b* are parameters. In our new model, edges are formed with probability

$$p_{i,j} = w\left[\frac{1}{1+\left(\frac{dE_{i,j}}{b}\right)^{a}}\right] + (1-w)\left[\frac{1}{1+\left(\frac{1-mE_{i,j}}{b}\right)^{a}}\right], \quad (2)$$

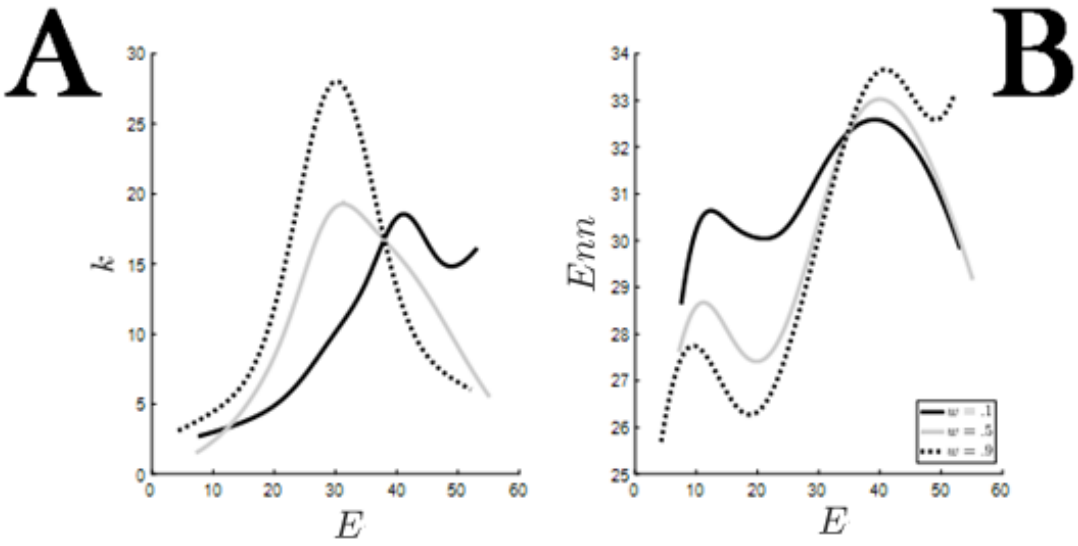

Figure 3. **Fitted three-term Gaussian models for *k* as a function of *E* (panel A) and *Enn* as a function of *E* (panel B). Each curve is fitted to 10⁴ realizations of the network model.** (Curves based on binned expected values and on smoothing splines appeared very similar to the Gaussian curves, although of course more jagged.)

where *dE* is the distance in *E*, *mE* is the mean of *E* for a pair of vertices, and *w* is a weight parameter in the interval [0, 1]. Essentially, what we have done is simply to make the model reflect two influences on edge formation: 1) pair-wise similarity in *E* and 2) pair-wise mean value of *E*, along with a parameter governing the weight of each influence. Changing the boguna Matlab function to reflect these two influences is straightforward. The original and new function code is

Original Code

```
function M = boguna(n,a,b)
v = repmat(rand(1,n),n,1);
M = triu(1./(1+(abs(v-v')./b).^a),1)>rand(n);
M = sparse(M+M');
end
```

New Code

```
function M = boguna(E,a,b,w)
n = length(E);
E = (E-min(E))./(max(E)-min(E));
v = repmat(E,n,1);
M = triu(w.*(1./(1+(abs(v-v')./b).^a))...
    +(1-w).*(1./((1+(1-((v+v')./2))./b).^a)),1)>rand(n);
M = sparse(M+M');
end
```

In the New Code *E* is a row vector with extraversion values. These are then rescaled by the function to the interval [0, 1]. Figure 3A plots curves fitted to degree *k* and *E* while Figure 3B plots curves fitted to *E* of the nearest neighbors (*Enn*) and *E* for $n = 50$, $a = 2$, $b = .2$ and $w = \{.1, .5, .9\}$ when $E \sim N(30, 5$ for this adapted Boguna model. The fitted curves are based on Gaussian models with three terms (see *Gaussian Models* in the internal Matlab documentation). Figure 3A reveals that the more impact of the first term in Equation 2 (higher *w*) the more symmetric becomes the relationship between degree *k* and *E*. This is expected because when distance in *E* governs edge formation the boundedness and normality of *E* favors edge formation in the middle of *E*. When *w* is lower (.1) the second term of Equation 2 has more impact and edge formation is governed to a larger extent by the pair-wise mean values in *E*. That is, higher values of *E* is associated with greater probability of edge formation. The first term in Equation 2 governs assortativity with respect to *E* (similar individuals connect), while the second term in Equation 2 governs degree *k* with respect to *E*.

One interesting feature revealed by Figure 3A is that as *w* becomes smaller *k* grows with *E* except for very high levels of *E*. This happens because as *w* decreases, the pair-wise mean values of *E* govern edge formation increasingly. As a result, vertices with higher values of *E* increasingly connect with other vertices. However, as long as $w > 0$ the distance in *E* will have some impact as well and this is manifested in decrease in *k* for very high values of *E*. Another related interesting feature revealed by Figure 3B is that as *w* becomes larger *Enn* grows with *E* except for very high levels of *E* (and a region of low values as well). This happens because as *w* increases, the distance in *E* governs edge formation increasingly. As a result, vertices close to each other increasingly connect with each other, so that high *E* vertices increasingly connect with other high *E* vertices. However, as long as $w < 1$ the pair-wise mean values of *E* will have some impact as well, which is manifested by the decrase in *Enn* for very high values of *E*. Put briefly, the two processes governing edge formation in this model, corresponding to the two terms of Equation 2, compete with each other. One prediction emerging from this model is that if a network exhibits both 1) higher degree *k* with increasing *E* on average and 2) higher *Enn* with

increasing E on average, we should always observe the dips for very high levels of E as shown in Figure 3. One of the key points of modeling is to generate testable predictions and this example illustrates how we can do so with specialized social network models.

Opinion Formation and Networks

Opinion formation and persuasion research concerns information transmission between individuals and how the resulting dynamics affect the formation of opinions. As such, this type of research fits naturally within a network framework (Watts & Dodds, 2007; Weenig & Midden, 1991). In the following example we consider the effect of network structure on opinion change. In this context, we will make a non-trivial observation, namely that an aggressive opinion change approach can have different effects on opinion change depending on the structure of network communities.

The model we will use for opinion change is as follows. Assume we have a network $G(V,E)$. Every vertex v has the same fixed persuasion probability p and an initial opinion o which can take one of two values. We then 1) select a random vertex v and its neighbors vn with opposing views. Then, 2) for each of the neighbors vn change the opinion to that of v with probability p. We call this changed set of neighbors vnc. Next, 3) for each member of vnc, let each of its neighbors vnn independently change the opinion back to the original with a probability equal to proportion of neighbors of the specific vnn that have changed its opinion. If the set vnc is not empty return to step 1 and replace v with vnc.

In more concrete terms, the algorithm proceeds by changing the opinion of one individual, say John. John then persuades some of his friends, say Sue, Bill and Janet. The friends of Sue, Bill and Janet react to this by trying to change their opinions back again. Suppose James is a friend of both Bill and Janet. Then James will try harder changing their opinions than if James knew only Bill. Likewise, if James has few friends he will try harder persuading Bill and Janet than if James has many friends. We could say that the more threatened James own view is by the opinion change the more effort he will spend restoring the original opinion. If Sue, Bill and Janet are all persuaded to go back to their original opinions then we stop. Otherwise, Sue, Bill and/or Janet will try to persuade their friends. Some of these may be persuaded, and then their friends will react, and so on.

Figure 4 provides an illustration of the opinion change algorithm. Here, at time step 1, we have a small network with opinions distributed in white and black. At time step 2, the enlarged vertex attempts to change the opinions of its neighbors with opposing views, successfully changing the opinion of one of them to white. At time step 3, the opposing black neighbor of the newly changed vertex tries to change it back to black, but fails. At time step 4, the newly changed vertex tries to convince its opposing neighbor and succeeds in doing so, upon which the algorithm stops.

In order to simulate this type of opinion change in a social network with clear community structure we use the *Li* model (Li & Maini, 2005). This model is similar to the Barabási-Albert model (Barabási & Albert, 1999) but is able to generate distinct communities. The number of communities is set through the Mc parameter and the distinctiveness of these is set through the a probability parameter. When $a = 0$ there are no edges between communities (complete community structure) and when $a = 1$ there are as many edges between as within communities (no community structure). For the simulations we will compare $a = .1$ with $a = .9$ for $n = 100$ vertices. The number of edges added each time-step was set to $m = 3$ in the former case and $m = 2$ in the latter case, in order to keep mean degree k similar across the two values of a. The number of communities was $Mc = 2$ and one opinion was associated with vertices in one community and the other opinion with the other community.

The p parameter can be interpreted as an aggressiveness parameter, because it designates the probability of changing the opinion of a neighbor. Aggressiveness, however, comes with a potential cost, because the more neighbors are changed the more the others will potentially strike back. This pattern can be seen for both clear community structure ($a = .1$) and for no community structure ($a = .9$) in Figure 5, because there is an optimal p for opinion change in the middle of the p distribution. Being more aggressive than this optimal level results in less net opinion change. Furthermore, when there is clear community structure associated with the original opinions the least amount of opinion change is achieved when aggressiveness is at its highest ($p = 1$), which is not the case when there is no community structure. In the latter case, the least amount of opinion change occurs when there is very little aggressiveness ($p = .1$). This example illustrates the basic point that depending on the network structure, aggressiveness in persuading others pays off differently.

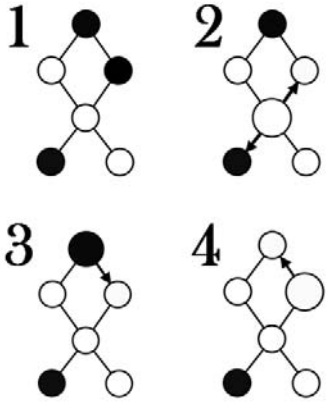

Figure 4. Illustration of opinion change algorithm. See text for explanation.

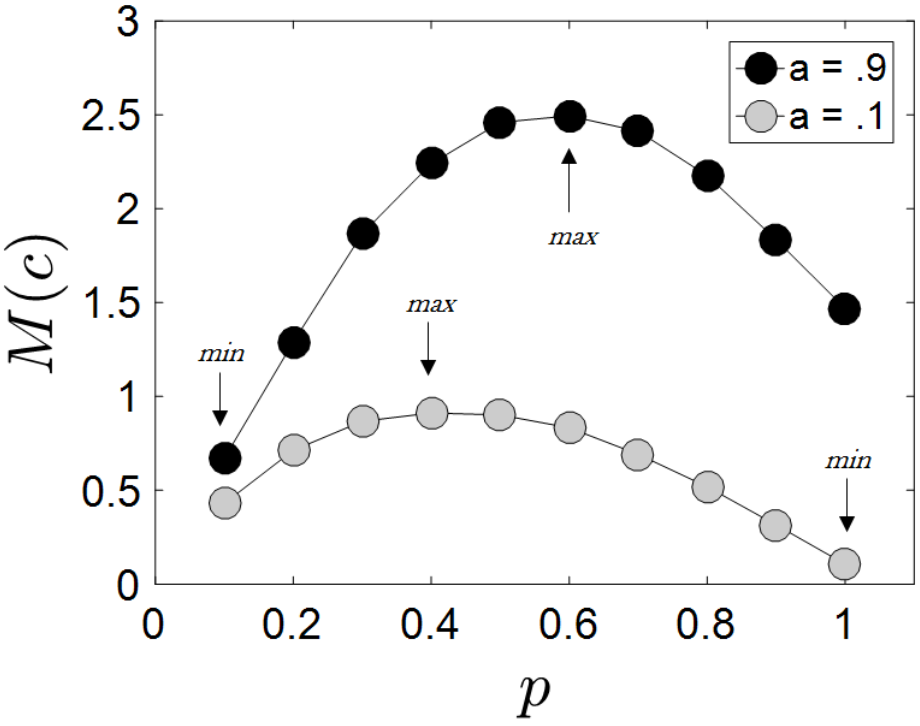

Figure 5. Mean number of changed opinions $M(c)$ as a function of p in the Li model for $a = .1$ and $a = .9$. The means are averaged over 10^3 realizations for $n = 100$ vertices.

The complexity emerging from interacting network components makes it difficult to generate predictions purely based on informal reasoning. Here, we have considered the effect of extraversion on network structure and the effect of aggressiveness and network structure on opinion change. These effects were explored by embedding the questions in specialized network models along with theoretical assumptions. Psychological variables, such as extraversion and aggressiveness, occur within individuals. However, they mostly achieve their effects by interacting with other individuals who are situated in social networks. Therefore, psychological research could quite plausibly benefit from considering explicit specialized social network models when investigating questions related to network structure. The next final section considers this issue more broadly.

Different Modeling Approaches

The "Generic-Specialized" dimension of network models considered in this article is not unique to the analysis of complex networks. In fact, the same dimension can be applied to virtually any substantive research area where modeling has been applied. For example, in psychology there is an abundance of mechanistic computational memory (Hintzman, 1984; Hopfield, 1982; Ratcliff, 1978; Shiffrin & Steyvers, 1997; Stewart, Brown, & Chater, 2005), learning (Elman, 1990; McClelland, McNaughton, & O'Reilly, 1995; Servan-Schreiber & Anderson, 1990) and categorization (Kruschke, 1992; Nosofsky, 1986; Vanpaemel & Storms, 2008) models. These models incorporate theoretical assumptions about the mechanisms underlying psychological processes and fall in the specialized end of the "Generic-Specialized" dimension. For example, a memory model could involve assumptions about how information is stored, how it is updated, how it is forgotten, and how it is used. In the generic end of the "Generic-Specialized" dimension we find statistical models with a more general purpose, that are much broader in scope and that have a more explicit statistical foundation, such as latent variable models (Loehlin, 1987), time series analysis (Jebb, Tay, Wang, & Huang, 2015) and generalized linear models (Hedeker, 2005). These models account for relationships, covariance and dependencies in the data in different ways, but typically do not involve specification of underlying processes at the detailed level of specialized models.

At the risk of belaboring the obvious I assert that both Generic and Specialized models have their use in psychology, and this includes network models. It is obvious that both types of models have been applied successfully on the whole in the past in psychology and in network science. Furthermore, sometimes a model or class of models moves along the dimension, often from the Specialized end in the Generic direction. For example, a model may be proposed in a relatively

detailed, yet heuristic and algorithmic fashion initially, in order to model some specific phenomenon. Then, further on, the model may be placed on a more solid statistical or mathematical foundation (Bollobás & Riordan, 2004; Goldenberg, 2009).

Even though Specialized models are often not expressed in a form directly amenable to analytic statistical inference, most of the time numerical or simulation strategies are applicable in order to evaluate the models. For example, in order to engage in model comparison we could fit the models with respect to a combined error function containing as many quantities as the number of free parameters, and then assess the resulting models with its optimal parameters with respect to some quantity of interest (Johansson, 2016; Toivonen et al., 2009). Alternatively, in line with a Bayesian approach (Wasserman, 2000) we could compare the marginal likelihoods of the models through simulation. For example, suppose we have a model with two parameters, we could then simulate the model over an informative grid of parameter value combinations. For each combination of parameter values, we compute the probability of observing the empirical data in question (the likelihood) with respect to some quantity of interest (Rohrmeier & Cross, 2014). The average of these likelihoods is then the marginal likelihood of the model with respect to a uniform prior on the model parameters. The ratio of two marginal model likelihoods is the Bayes Factor and denotes the factor by which the data supports one model over the other (Dienes, 2011). We will illustrate this approach by an example along with the Matlab code required to implement it.

Consider the Boguna model, previously mentioned in the Section 5.1. This model has two parameters a and b. Suppose we want to compute the marginal likelihood for this model with respect to mean degree k of an observed network with $n = 50$ vertices. First, we define a grid of parameter values of interest. Note that the resulting likelihood will depend on this choice. Suppose we select a from 1 to 3 in steps of 0.1 and b from 0.05 to 1 in steps of 0.05. We then define a distance D from the observed k and let D define the boundaries of what we count as the observed data. Suppose that $k = 10.5$, and we set $D = 0.5$. Then, degrees k in the interval $I = [10, 11]$ will count as reproducing the observed data. We then simulate the network over s iterations and compute the proportion of simulated mean degrees falling in the interval I for each of the parameter value combinations. The mean of all these is our marginal likelihood. Simulating this in Matlab is straightforward, although of course potentially time consuming depending on the model, parameter values and s, because of the brute force nature of this way of approximating a likelihood through Monte Carlo simulation. However, the virtue of this method is its applied simplicity. The Matlab code for simulating this using a parallel loop and returning the marginal likelihood for $s = 10^3$ is as follows.

Marginal Likelihood Boguna

```
a = 1:.1:3;
b = .05:.05:1;
np = length(a)*length(b);
p1 = reshape(repmat(a,length(b),1),1,np);
p2 = repmat(b,1,length(a));
n = 50;
s = 1e3;
kobs = 10;
D = 1;
I1 = kobs-D;
I2 = kobs+D;
L = zeros(1,np);
c = np;
parfor i = 1:np
   k = zeros(1,s);
   for m = 1:s
      k(s) = mean(sum(boguna(n,p1(i),p2(i))));
   end
   L(i) = mean(k>I1 & k<I2);
end
ML = mean(L);
disp(ML)
```

This gives a marginal likelihood $L_B = 1.43*10^{-5}$. Suppose we wish to compare the Boguna model to the Grindrod model (Grindrod, 2002), which has two parameters constrained to the interval [0, 1]. This model has a marginal likelihood $L_C = 2.25*10^{-5}$. The ratio L_B/L_C gives the Bayes Factor BF quantifying the factor by which the observed data supports one model over the other. In this case, $BF = 0.64$ which is close to 1, indicating that none of the two models is distinctly supported over the other ((Dienes, 2011).

The aim of this article is not to go into details with respect to model comparison, but rather to enhance the applicability of specialized social network models in psychology. This requires getting to know the models under consideration. In many cases, it is likely that implementing psychological variables into these models requires careful thinking and adjustment of the model functions. In fact, the careful thinking required is precisely one of the main advantages of modeling. A model parameter implemented in a particular way corresponds to some particular way of conceptualizing a process. When a parameter is distributed a particular way it corresponds to a distributional assumption, and so on. Without making these choices, the model produces no output. As such, modeling naturally offers the potential of insight and theory development, along with theory testing. Hopefully, the included social network Matlab functions will assist researchers in achieving these goals with some degree of success.

References

Allport, F. H., & Allport, G. W. (1921). Personality Traits: Their Classification and Measurement. *The Journal of Abnormal Psychology and Social Psychology*, *16*(1), 6–40. https://doi.org/10.1037/h0069790

Axelrod, R. (1997). The Dissemination of Culture: A Model with Local Convergence and Global Polarization. *Journal of Conflict Resolution*, *41*(2), 203–226. https://doi.org/10.1177/0022002797041002001

Barabási, A.-L., & Albert, R. (1999). Emergence of Scaling in Random Networks. *Science*, *286*(5439), 509–512. https://doi.org/10.1126/science.286.5439.509

Barabási, A.-L., & Oltvai, Z. N. (2004). Network biology: understanding the cell's functional organization. *Nature Reviews Genetics*, *5*(2), 101–113. https://doi.org/10.1038/nrg1272

Bastian, M., Heymann, S., & Jacomy, M. (2009). Gephi: An Open Source Software for Exploring and Manipulating Networks. https://doi.org/10.13140/2.1.1341.1520

Bianconi, G., & Barabási, A.-L. (2001). Bose-Einstein Condensation in Complex Networks. *Physical Review Letters*, *86*(24), 5632–5635. https://doi.org/10.1103/PhysRevLett.86.5632

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, *2008*(10), P10008. https://doi.org/10.1088/1742-5468/2008/10/P10008

Boguñá, M., Pastor-Satorras, R., Díaz-Guilera, A., & Arenas, A. (2004). Models of social networks based on social distance attachment. *Physical Review E*, *70*(5). https://doi.org/10.1103/PhysRevE.70.056122

Bollobás, B., & Riordan, O. M. (2004). Mathematical results on scale-free random graphs. In S. Bornholdt & H. G. Schuster (Eds.), *Handbook of Graphs and Networks* (pp. 1–34). Weinheim, FRG: Wiley-VCH Verlag GmbH & Co. KGaA. Retrieved from http://doi.wiley.com/10.1002/3527602755.ch1

Borgatti, S. P., Mehra, A., Brass, D. J., & Labianca, G. (2009). Network Analysis in the Social Sciences. *Science*, *323*(5916), 892–895. https://doi.org/10.1126/science.1165821

Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2004). The Concept of Validity. *Psychological Review*, *111*(4), 1061–1071. https://doi.org/10.1037/0033-295X.111.4.1061

Brass, D. J., Butterfield, K. D., & Skaggs, B. C. (1998). Relationships and Unethical Behavior: A Social Network Perspective. *Academy of Management Review*, *23*(1), 14–31. https://doi.org/10.5465/AMR.1998.192955

Caldarelli, G., Capocci, A., De Los Rios, P., & Muñoz, M. A. (2002). Scale-Free Networks from Varying Vertex Intrinsic Fitness. *Physical Review Letters*, *89*(25). https://doi.org/10.1103/PhysRevLett.89.258702

Callaway, D. S., Hopcroft, J. E., Kleinberg, J. M., Newman, M. E. J., & Strogatz, S. H. (2001). Are randomly grown graphs really random? *Physical Review E*, *64*(4). https://doi.org/10.1103/PhysRevE.64.041902

Catanzaro, M., Caldarelli, G., & Pietronero, L. (2004). Assortative model for social networks. *Physical Review E*, *70*(3). https://doi.org/10.1103/PhysRevE.70.037101

Clauset, A., Shalizi, C. R., & Newman, M. E. J. (2009). Power-Law Distributions in Empirical Data. *SIAM Review*, *51*(4), 661–703. https://doi.org/10.1137/070710111

Cramer, A. O. J., van der Sluis, S., Noordhof, A., Wichers, M., Geschwind, N., Aggen, S. H., … Borsboom, D. (2012). Dimensions of Normal Personality as Networks in Search of Equilibrium: You Can't Like Parties if You Don't Like People: Dimensions of normal personality as networks. *European Journal of Personality*, *26*(4), 414–431. https://doi.org/10.1002/per.1866

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281–302. https://doi.org/10.1037/h0040957

Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal, Complex Systems*, (1695). Retrieved from http://igraph.org

Davidsen, J., Ebel, H., & Bornholdt, S. (2002). Emergence of a Small World from Local Interactions: Modeling Acquaintance Networks. *Physical Review Letters*, *88*(12). https://doi.org/10.1103/PhysRevLett.88.128701

Dienes, Z. (2011). Bayesian Versus Orthodox Statistics: Which Side Are You On? *Perspectives on Psychological Science*, *6*(3), 274–290. https://doi.org/10.1177/1745691611406920

Elman, J. L. (1990). Finding Structure in Time. *Cognitive Science*, *14*(2), 179–211. https://doi.org/10.1207/s15516709cog1402_1

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

Eysenck, S. B. G., & Eysenck, H. J. (1963). On the Dual Nature of Extraversion. *British Journal of Social and Clinical Psychology*, *2*(1), 46–55. https://doi.org/10.1111/j.2044-8260.1963.tb00375.x

Feiler, D. C., & Kleinbaum, A. M. (2015). Popularity, Similarity, and the Network Extraversion Bias. *Psychological Science*, *26*(5), 593–603. https://doi.org/10.1177/0956797615569580

Fowler, J. H., & Christakis, N. A. (2008). Dynamic spread of happiness in a large social network: longitudinal analysis over 20 years in the Framingham Heart Study. *BMJ*, *337*(dec04 2), a2338–a2338. https://doi.org/10.1136/bmj.a2338

Fruchterman, T. M. J., & Reingold, E. M. (1991). Graph drawing by force-directed placement. *Software: Practice and Experience*, *21*(11), 1129–1164. https://doi.org/10.1002/spe.4380211102

Gilbert, E. N. (1959). Random Graphs. *The Annals of Mathematical Statistics*, *30*(4), 1141–1144. https://doi.org/10.1214/aoms/1177706098

Goldenberg, A. (2009). A Survey of Statistical Network Models. *Foundations and Trends® in Machine Learning*, *2*(2), 129–233. https://doi.org/10.1561/2200000005

Grabowski, A., & Kosiński, R. A. (2006). Evolution of a social network: The role of cultural diversity. *Physical Review E*, *73*(1). https://doi.org/10.1103/PhysRevE.73.016135

Grindrod, P. (2002). Range-dependent random graphs and their application to modeling large small-world Proteome datasets. *Physical Review E*, *66*(6). https://doi.org/10.1103/PhysRevE.66.066702

Handcock, M. S., Hunter, D. R., Butts, C. T., Goodreau, S. M., & Morris, M. (2003). *statnet: Software tools for the Statistical Modeling of Network Data.* Retrieved from http://statnetproject.org

Hedeker, D. (2005). Generalized Linear Mixed Models. In B. S. Everitt & D. C. Howell (Eds.), *Encyclopedia of Statistics in Behavioral Science*. Chichester, UK: John Wiley & Sons, Ltd. Retrieved from http://doi.wiley.com/10.1002/0470013192.bsa251

Hintzman, D. L. (1984). MINERVA 2: A simulation model of human memory. *Behavior Research Methods, Instruments, &amp Computers*, *16*(2), 96–101. https://doi.org/10.3758/BF03202365

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences of the United States of America*, *79*(8), 2554–2558.

Jebb, A. T., Tay, L., Wang, W., & Huang, Q. (2015). Time series analysis for psychological research: examining and forecasting change. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00727

Johansson, T. (2016). Gossip spread in social network Models. *Physica A: Statistical Mechanics and Its Applications*. https://doi.org/10.1016/j.physa.2016.11.132

Krause, J., James, R., & Croft, D. P. (2010). Personality in the context of social networks. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1560), 4099–4106. https://doi.org/10.1098/rstb.2010.0216

Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review*, *99*(1), 22–44. https://doi.org/10.1037/0033-295X.99.1.22

Kumar, R., Novak, J., & Tomkins, A. (2006). Structure and evolution of online social networks (p. 611). ACM Press. https://doi.org/10.1145/1150402.1150476

Kumpula, J. M., Onnela, J.-P., Saramäki, J., Kaski, K., & Kertész, J. (2007). Emergence of Communities in Weighted Networks. *Physical Review Letters*, *99*(22). https://doi.org/10.1103/PhysRevLett.99.228701

Leskovec, J., Kleinberg, J., & Faloutsos, C. (2007). Graph evolution: Densification and shrinking diameters. *ACM Transactions on Knowledge Discovery from Data*, *1*(1), 2–es. https://doi.org/10.1145/1217299.1217301

Leskovec, J., Lang, K. J., Dasgupta, A., & Mahoney, M. W. (2008). Statistical properties of community structure in large social and information networks (p. 695). ACM Press. https://doi.org/10.1145/1367497.1367591

Leskovec, J., & Sosič, R. (2016). SNAP: A General-Purpose Network Analysis and Graph-Mining Library. *ACM Transactions on Intelligent Systems and Technology*, *8*(1), 1–20. https://doi.org/10.1145/2898361

Li, C., & Maini, P. K. (2005). An evolving network model with community structure. *Journal of Physics A: Mathematical and General*, *38*(45), 9741–9749. https://doi.org/10.1088/0305-4470/38/45/002

Loehlin, J. C. (1987). *Latent variable models: an introduction to factor, path, and structural analysis*. Hillsdale, N.J: L. Erlbaum Associates.

Marsili, M., Vega-Redondo, F., & Slanina, F. (2004). The rise and fall of a networked society: A formal model. *Proceedings of the National Academy of Sciences*, *101*(6), 1439–1442. https://doi.org/10.1073/pnas.0305684101

Martin, S., Brown, W. M., Klavans, R., & Boyack, K. W. (2011). OpenOrd: an open-source toolbox for large graph layout (p. 786806). https://doi.org/10.1117/12.871402

Matlab. (2016). (Version Release 2016a). Natick, Massachusetts, United States: The Mathworks, Inc.

Mayer, A. (2009). Online social networks in economics. *Decision Support Systems*, *47*(3), 169–184. https://doi.org/10.1016/j.dss.2009.02.009

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, *102*(3), 419–457.

McCrae, R. R., & Costa, P. T., Jr. (1997). Personality trait structure as a human universal. *American Psychologist*, *52*(5), 509–516. https://doi.org/10.1037//0003-066X.52.5.509

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology*, *27*(1), 415–444. https://doi.org/10.1146/annurev.soc.27.1.415

Meyers, L. A., Pourbohloul, B., Newman, M. E. J., Skowronski, D. M., & Brunham, R. C. (2005). Network theory and SARS: predicting outbreak diversity. *Journal of Theoretical Biology*, *232*(1), 71–81. https://doi.org/10.1016/j.jtbi.2004.07.026

Mollgaard, A., & Mathiesen, J. (2016). The Dynamics of Initiative in Communication Networks. *PLOS ONE*, *11*(4), e0154442. https://doi.org/10.1371/journal.pone.0154442

Newman, M. E. J. (2002). Assortative Mixing in Networks. *Physical Review Letters*, *89*(20). https://doi.org/10.1103/PhysRevLett.89.208701

Newman, M. E. J. (2003). The Structure and Function of Complex Networks. *SIAM Review*, *45*(2), 167–256. https://doi.org/10.1137/S003614450342480

Newman, M. E. J. (2011). Communities, modules and large-scale structure in networks. *Nature Physics*, *8*(1), 25–31. https://doi.org/10.1038/nphys2162

Newman, M. E. J. (2016). Equivalence between modularity optimization and maximum likelihood methods for community detection. *Physical Review E*, *94*(5). https://doi.org/10.1103/PhysRevE.94.052315

Nosofsky, R. M. (1986). Attention, similarity, and the identification–categorization relationship. *Journal of Experimental Psychology: General*, *115*(1), 39–57. https://doi.org/10.1037/0096-3445.115.1.39

Palla, G., Derényi, I., Farkas, I., & Vicsek, T. (2005). Uncovering the overlapping community structure of complex networks in nature and society.

Nature, *435*(7043), 814–818. https://doi.org/10.1038/nature03607

Parravano, A., & Reyes, L. M. (2008). Gaslike model of social motility. *Physical Review E*, *78*(2). https://doi.org/10.1103/PhysRevE.78.026120

Perry-Smith, J. E., & Shalley, C. E. (2003). The social side of creativity: a static and dynamic social network perspective. *Academy of Management Review*, *28*(1), 89–106. https://doi.org/10.5465/AMR.2003.8925236

Qiao, J., Huang, H.-Q., Li, G.-Y., & Fan, Y. (2014). Bridging the gap between different social networks. *Physica A: Statistical Mechanics and Its Applications*, *410*, 535–549. https://doi.org/10.1016/j.physa.2014.05.067

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*(2), 59–108. https://doi.org/10.1037/0033-295X.85.2.59

Ripley, R., Snijders, T. A. B., Boda, Z., Vörös, A., & Preciado, P. (2015). *Manual for SIENA version 4.0*. Oxford: University of Oxford, Department of Statistics; Nuffield College.

Robins, G., Pattison, P., Kalish, Y., & Lusher, D. (2007). An introduction to exponential random graph (p*) models for social networks. *Social Networks*, *29*(2), 173–191. https://doi.org/10.1016/j.socnet.2006.08.002

Rohrmeier, M. A., & Cross, I. (2014). Modelling unsupervised online-learning of artificial grammars: Linking implicit and statistical learning. *Consciousness and Cognition*, *27*, 155–167. https://doi.org/10.1016/j.concog.2014.03.011

Servan-Schreiber, E., & Anderson, J. R. (1990). Learning artificial grammars with competitive chunking. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(4), 592–608. https://doi.org/10.1037/0278-7393.16.4.592

Shiffrin, R. M., & Steyvers, M. (1997). A model for recognition memory: REM—retrieving effectively from memory. *Psychonomic Bulletin & Review*, *4*(2), 145–166. https://doi.org/10.3758/BF03209391

Singer, H. M., Singer, I., & Herrmann, H. J. (2009). Agent-based model for friendship in social networks. *Physical Review E*, *80*(2). https://doi.org/10.1103/PhysRevE.80.026113

Stewart, N., Brown, G. D. A., & Chater, N. (2005). Absolute Identification by Relative Judgment. *Psychological Review*, *112*(4), 881–911. https://doi.org/10.1037/0033-295X.112.4.881

Toivonen, R., Kovanen, L., Kivelä, M., Onnela, J.-P., Saramäki, J., & Kaski, K. (2009). A comparative study of social network models: Network evolution models and nodal attribute models. *Social Networks*, *31*(4), 240–254. https://doi.org/10.1016/j.socnet.2009.06.004

Toivonen, R., Onnela, J.-P., Saramäki, J., Hyvönen, J., & Kaski, K. (2006). A model for social networks. *Physica A: Statistical Mechanics and Its Applications*, *371*(2), 851–860. https://doi.org/10.1016/j.physa.2006.03.050

Travers, J., & Milgram, S. (1969). An Experimental Study of the Small World Problem. *Sociometry*, *32*(4), 425. https://doi.org/10.2307/2786545

Vanpaemel, W., & Storms, G. (2008). In search of abstraction: the varying abstraction model of categorization. *Psychonomic Bulletin & Review*, *15*(4), 732–749.

Vázquez, A. (2003). Growing network with local rules: Preferential attachment, clustering hierarchy, and degree correlations. *Physical Review E*, *67*(5). https://doi.org/10.1103/PhysRevE.67.056104

Wasserman, L. (2000). Bayesian Model Selection and Model Averaging. *Journal of Mathematical Psychology*, *44*(1), 92–107. https://doi.org/10.1006/jmps.1999.1278

Watts, D. J. (2004). The "New" Science of Networks. *Annual Review of Sociology*, *30*(1), 243–270. https://doi.org/10.1146/annurev.soc.30.020404.104342

Watts, D. J., & Dodds, P. S. (2007). Influentials, Networks, and Public Opinion Formation. *Journal of Consumer Research*, *34*(4), 441–458. https://doi.org/10.1086/518527

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of "small-world" networks. *Nature*, *393*(6684), 440–442. https://doi.org/10.1038/30918

Weenig, M. W., & Midden, C. J. (1991). Communication network influences on information diffusion and persuasion. *Journal of Personality and Social Psychology*, *61*(5), 734–742. https://doi.org/10.1037/0022-3514.61.5.734

Wong, L. H., Pattison, P., & Robins, G. (2006). A spatial model for social networks. *Physica A: Statistical Mechanics and Its Applications*, *360*(1), 99–120. https://doi.org/10.1016/j.physa.2005.04.029