Comput Math Organ Theory
DOI 10.1007/s10588-016-9232-2

MANUSCRIPT

Generating and analyzing spatial social networks

Meysam Alizadeh[1,2] · Claudio Cioffi-Revilla[1,2] · Andrew Crooks[1,2]

Abstract In this paper, we propose a class of models for generating spatial versions of three classic networks: Erdös-Rényi (ER), Watts-Strogatz (WS), and Barabási-Albert (BA). We assume that nodes have geographical coordinates, are uniformly distributed over an $m \times m$ Cartesian space, and long-distance connections are penalized. Our computational results show higher clustering coefficient, assortativity, and transitivity in all three spatial networks, and imperfect power law degree distribution in the BA network. Furthermore, we analyze a special case with geographically clustered coordinates, resembling real human communities, in which points are clustered over k centers. Comparison between the uniformly and geographically clustered versions of the proposed spatial networks show an increase in values of the clustering coefficient, assortativity, and transitivity, and a lognormal degree distribution for spatially clustered ER, taller degree distribution and higher average path length for spatially clustered WS, and higher clustering coefficient and transitivity for the spatially clustered BA networks.

Keywords Spatial social networks · Network properties · Random network · Small-world network · Scale-free network

✉ Meysam Alizadeh
alizadeh.meysam@gmail.com

1 Computational Social Science Program, Department of Computational and Data Sciences, George Mason University, Fairfax, VA, USA

2 Center for Social Complexity, Krasnow Institute for Advanced Studies, George Mason University, Fairfax, VA, USA

Published online: 03 September 2016

1 Introduction

Generating social networks has attracted significant interest in network science and agent-based modeling, given the considerable impact of network topology and its properties on the behavior and output of models (Alam and Geller 2012; Alizadeh et al. 2014). However, the generation algorithms of social networks have often omitted important empirical features found in real-world social networks. One such regularity often missing from formation processes of social networks is embedded location in physical geographic space (Barthélemy 2011). Indeed, social relations are constrained by and benefit from human ties as well as they are affected by geography (Batty 2012; Onnela et al. 2011). For example, the importance of an individual's location is significant for modeling individuals' behavior and associated opinion dynamics (Soboll et al. 2011), and racial segregation (Edmonds 2006).

Three well-known and fundamental models for generating social network structures are: (1) random networks (Erdös and Rényi 1959, 1960; Gilbert 1959; Rapoport 1957), (2) small-world networks (Watts and Strogatz 1998), and (3) scale-free networks (Barabasi and Albert 1999). Random network construction is only based on stochasticity, whereas a scale-free network is generated through individuals' preferential attachment and generates a network with degree distribution of nodes following a power law distribution. Although many algorithms have been proposed for the generation of social networks (see Newman 2003b for an early review), not all of them are based on empirically valid social dynamics (Borgatti et al. 2013; Cioffi-Revilla 2014, Chap. 4). For examples of empirically based social network construction algorithms, see Cointet and Roth (2007) and Yoshida et al. (2008).

For social networks consisting of human interactions—the focus of this study—empirical findings have repeatedly demonstrated that the probability of multiple modes of interaction (e.g., travel, communication, marriages, commerce, conflict, and many others) is strongly correlated to distance between the individuals or groups. For example, social influence measured by frequency of memorable interactions is inversely correlated with distance, resulting in a hyperbolic function of the form $p(k) = d^{-1}$ (Latane et al. 1995). Analyzing mobile phone data in Belgium, Lambiotte et al. (2008) identified the connection probability of $p(k) = d^{-2}$, replicating inverse-square or so-called social gravity laws first documented a century ago and found in every human and social geography textbook. More recently, Onnela et al. (2011) investigated a large dataset of voice calls and text messages from an anonymous European country and detected that the tie probability decayed with distance as a power law with exponent $\alpha = 1.58$ for voice ties and $\alpha = 1.49$ for text ties.

Although determining the acceptable range of the exponent α demands more empirical studies, the necessity for including spatial factors in social networks is clear. Several algorithms have been proposed for embedding geographic space into network models (see Boccaletti et al. (2006) and Barthélemy (2011) for review). Most of these models have been developed based on real-world spatial networks such as transportation networks, the Internet, and power grid networks, and they are mainly concerned with changes in degree distribution as a result of embedding the spatial dimension.

Our focus here is on human social networks, in which people constitute nodes and their social ties are edges of the network. More specifically, we are interested in examining how the inclusion of geospatial location and proximity in a given generative algorithm for social networks affects the statistical properties and structure of networks in terms of empirical regularities that have been found in real-world human social networks. Our goal is twofold in this study: (1) to emphasize the relevance and importance of the geographical proximity on the formation of social networks; (2) to introduce the concept of spatial social networks to the agent-based modeling community and encourage the use of spatial networks to explore the geography of emergent and collective behaviors. This is particularly important where the behavior of a model depends on geographical properties of agents.

We propose simple methods to construct spatial extensions of the Rapoport–Erdös–Rényi (ER), Watts–Strogatz (WS), and Barabási–Albert (BA) network models and measure and compare their network characteristics. Assuming there is an embedding space representing Cartesian coordinates of the agents, we associate long-range connections with a cost produced by a distance function, such that the probability of connection decays as a power law. In order to better capture the effect of distance selection on network properties, we examine two cases where agents are either uniformly distributed across geographical space, or are spatially clustered.

The rest of the paper is organized as follows. Before going through the proposed spatial network generation algorithms, we review empirical regularities of real-world large social networks and describe the ER, WS, and BA networks in detail in Sect. 2. In Sect. 3, we present our proposed models for generating spatial extensions of ER, WS, and BA models. We present statistical properties of the proposed spatial networks in Sect. 4 and compare them with the non-spatial original ER, WS, and BA networks. Section 5 contains a discussion of main points and conclusions.

2 Overview of social network generation algorithms

We begin by reviewing empirical regularities of large social networks. Next, we present an overview of the three well-known social network generation models and their properties.

2.1 Empirical regularities of large social networks

According to the literature, real-world large social networks exhibit a set of observable empirical regularities. Not all of these regularities are found in all networks and different types of social networks exhibit a subset of these properties which are outlined below. We use network property measures to capture these regularities in Sect. 4 and compare the proposed socio-spatial networks with the classic non-spatial versions.

1. *Low network density* among all possible edges in a network, only a few of them should exist (Wong et al. 2006).

2. *A limit on the nodes' degree* individuals' degree of connectivity should not exceed a reasonable limit (Gilbert 2006; Barthélemy 2003). We capture this characteristic by computing the maximum degree in the network.
3. *Difference in nodes' degree* individuals' degree of connectivity should not be the same throughout the network (Fischer 1982). We compute the standard deviation of the degree to account for this property.
4. *A heavy-tailed distribution of nodes' degree* some individuals in the network should have a large number of immediate neighbors (Fischer 1982; Redner 1998; Albert et al. 1999; Barabasi and Albert 1999; Dorogovtsev and Mendes 2002).
5. *Degree assortativity* individuals with high number of connections tend to be connected with other individuals with many connections (Newman 2002, 2003a). In other words, the assortativity of a network measures the correlation between the nodes' degree. Here we measure this correlation based on the Pearson correlation coefficient of the degrees (Callaway et al. 2001), which varies between -1.0 (disassortative) and $+1.0$ (perfectly assortative).
6. *High clustering* some members of an individual's personal network should know each other (Bruggeman 2008). In other words, clustering measures the conditional probability that friends of an individual are themselves friends.
7. *Presence of communities* there should be some groups in which members are highly connected with their in-group fellows and loosely connected with out-groups (Wong et al. 2006). By tradition, these groups are called *communities* (Newman 2004).
8. *Short path length* individuals in the network should be able to reach each other via a relatively small number of connections (Milgram 1967; Garfield 1979; Chung and Lu 2002). This characteristic reflects the "small-world" effect.

2.2 Erdös–Rényi network

Erdös and Rényi (1959, 1960) produced two random network models, similar to an earlier model independently published by Rapoport (Rapoport 1957) and Cioffi-Revilla (2014, Chap. 4). The first Erdös-Rényi model (ER, for short) produces a graph which is uniformly randomly selected from the set of all possible graphs. The second model connects each node pair with probability p. While working separately, Gilbert (1959) also produced the second model. Here our focus is on the second version. The following pseudo-code generates the second ER network.

```
For each node i
  For each node j = i + 1
    Chance = a random number between 0 and 1
      If p > Chance
        Create an edge between node i and node j
```

It has been shown that the average path length in an ER network is close to ln N, which is small enough to exhibit the short path length property (Watts and Strogatz 1998; Albert and Barabási 2002). This property has important implications in many applications, such as a disease's speed of diffusion (Watts and Strogatz 1998). The average degree for any random network is $p(N - 1)$, which implies that we have to choose a small p to obtain a low-density network (Barthélemy 2011). The main shortcoming of the ER network is that it fails to generate high clustering (Hamill and Gilbert 2009).

2.3 Small-world network

Although the small-world phenomenon (also known as "six degrees of separation") was discovered as early as 1967 by Milgram, it was not until 1998 that Watts and Strogatz proposed a simple but powerful algorithm to generate a social network that exhibits long-range correlations, high clustering, and average short path length. The algorithm starts from a regular lattice and then rewires each pair of nodes with probability p. The following pseudo-code can be used to generate the Watts-Strogatz (1998) small-world network (WS).

```
Create a regular lattice of N nodes each connected to its K nearest neighbors
For each node i
  For each node j = i + 1, …, i + K/2
    If j > N
    j = j - N
    Chance = a random number between 0 and 1
    If p > Chance
        Node m = a uniformly chosen node at random
        Create an edge between node i and node m
        Remove the edge between node i and j
```

The small-world network has a similar topology to that of the ER network. Indeed, WS → ER for $p \to 1$. Although the WS network has a significantly higher clustering coefficient relative to the ER, in general the clustering coefficient and the average path length are highly dependent on the rewiring probability p.

2.4 Scale-free network

Barabasi and Albert (1999) showed that networks such as World Wide Web, academic citation patterns, actor collaboration networks, and power grids are scale-free (SF), because they have degree distributions where the probability that a node has links with k other nodes decays as a power law probability density function, $P(k) \sim k^{-\alpha}$, where α is a constant and $k = 1, 2, \ldots, N$. Such a power law is also a hybrid function, consisting of a continuous (α) and a discrete (k) dimension, so analyzing its associated fields requires special operators such as nabladot and related methods (Cioffi-Revilla 2015). Barabasi and Albert (1999) hypothesized that *growth* and *preferential attachment* are two mechanisms that explain the scale invariance behavior ("scaling") of real networks. Their proposed generative algorithm combined the two mechanisms by adding a new node at each time step and assigning each existing node a probability (of connecting to the new node being added) that is proportional to a node's current degree. This is described by the following pseudo-code:

```
Create a fully connected initial network of m0 nodes
For each node i = m0+1,…,N
   Current degree = 0
   While Current degree < m
       Node j = a uniformly randomly selected node from all existing
                  nodes except the node i and nodes adjacent to node i
       p = degree of node j / number of all edges in the network
       chance = a random number between 0 and 1
       if p > chance
              Create an edge between node i and node j
```

where m_0 is the initial size of the network, m is the number of nodes to which any new added node will be connected, and N is the final network size. In this SF network, the degree distribution is a power law with exponent $\alpha = 3$, average path length grows as $l \sim \log(N)$ (Newman 2003b), assortativity decreases as $\log^2 N/N$ (Newman 2002), and clustering decreases as $1/N$ (Klemm and Eguiluz 2002).

3 Generating spatial social networks

We consider an $m \times m$ Cartesian space and randomly generate N points drawn from a uniform distribution to represent the spatial coordinates of agents. Since the total density of points in space might have an effect on results, we define geographic density (GD) as a modeling parameter and will perform sensitivity analysis on it.

3.1 Generating a spatial ER network

Hamill and Gilbert (2009) proposed a generative process for network construction based on social circles in which nodes connect when their circles reach a given radius. Building on this idea, Holzhauer et al. (2013) used geo-referenced data to produce a network with properties of the SW model. Here, we randomly assign each agent to one point on the $m \times m$ Cartesian space. Then, similar to the ER network-formation process, each distinct pair of agents connects with probability given by the following power law or hyperbolic function:

$$p(d) = Cd^{-\alpha}, \tag{1}$$

where C is a normalizing coefficient, d denotes the spatial distance between nodes, and α is a distance-decay exponent. In Eq. (1), constant C normalizes the distribution, such that:

$$\int_{d_{min}}^{d_{max}} p(d)\mathrm{d}d = 1, \tag{2}$$

which yields the following solution:

$$1 = \int_{d_{min}}^{d_{max}} Cd^{-\alpha}\mathrm{d}d = C\int_{d_{min}}^{d_{max}} d^{-\alpha}\mathrm{d}d = \frac{C}{1-\alpha}\left[d^{-\alpha+1}\right]_{d_{min}}^{d_{max}}, \tag{3}$$

We can see that Eq. (3) makes sense only if $\alpha > 1$. As a result, Eq. (3) yields:

$$C_1 = \frac{1-\alpha}{d_{max}^{1-\alpha} - d_{min}^{1-\alpha}}, \tag{4}$$

The following pseudo-code can be used to produce the proposed spatial Erdös–Rényi (SER) network:

```
For each node i:
  For each node j = i+1 … N:
     p = Cd^-α
     Chance = a random number between 0 and 1
     If p > Chance:
        Create an edge between i and j
```

3.2 Generating spatial WS networks

Given that each agent is randomly assigned to a point on an $m \times m$ Cartesian space, similar to Watts and Strogatz (1998), we begin with a regular lattice where each agent is connected to its k nearest neighbors. Here, to incorporate geographical proximity into the model, two scenarios are plausible. First, following Kleinberg (2000), rewiring occurs with probability p and the selection of a new node is conditioned on the spatial distance between nodes:

$$p(d) = C_1 d_{ij}^{-\alpha}, \tag{5}$$

where d_{ij} is the Euclidean distance between agent i and agent j in Cartesian space, α is a fixed clustering exponent, and C_1 is a normalization factor. Constant C_1 in Eq. (5) is given by the normalization requirement (similar to Eq. 2). The advantage of this approach is that we can compare the resulting network with that of WS model since they share the rewiring probability p. We denote this version of the model as SWS-1. However, our proposed approach is to condition the rewiring step on inter-node distance, replacing probability p in Kleinberg's (2000) approach with Eq. (4), and then selecting the new node uniformly at random:

$$p(d) = C_2 d_{ij}^{\alpha}, \tag{6}$$

Constant C_2 in Eq. (6) can be obtained by an approach similar to Eq. (2). The following pseudo-code describes this second version of the spatial Watts–Strogatz (SWS-2) model:

```
Create a regular lattice of N nodes each connected to its k nearest neighbors
For each node i:
  For each node j = i + 1, …, i + k/2:
    If j > N:
      j = j - N
      d_ij = Euclidean distance between node i and node j
      p = Cd_ij^α
      Chance = a random number between 0 and 1
      If Chance < 1-p:
           Node m = a uniformly randomly selected node from all nodes except
                    the node i and nodes connected to node i
           Create an edge between node i and node m
           Remove the edge between node i and j
```

3.3 Generating a spatial Barabási–Albert network

In a BA network, the probability with which a new vertex connects to the existing vertices is not uniform and there is a propensity to connect to well-connected vertices. Here, to incorporate the geographical proximity of agents into the algorithm, we assume that there is a higher probability that a new added node will be linked to a vertex that (1) already has a large number of connections, and (2) located geographically closer. This means that nearby hubs are more likely to be chosen by new nodes than are distant hubs. After randomly assigning all agents to a point in the $m \times m$ Cartesian space, starting with a small number (m_0) of vertices, at every time step we add a new vertex with m ($\leq m_0$) edges that link the new vertex to m different vertices already present in the system. To incorporate preferential attachment and spatial proximity, we assume that the probability $P(k_j, d_{ij})$ that a new vertex i will be connected to vertex j depends on the connectivity k_j and the geographical distance d_{ij} of that vertex:

$$p(k_j, d_{ij}) \propto k_j f(d_{ij})$$

where $f(d_{ij})$ denotes a function of distance between node i and node j. While several functions have been proposed for the $f(d_{ij})$ (e.g., Yook et al. 2002; Xulvi-Brunet and Sokolov 2002; Manna and Sen 2002; Barthélemy 2003), in this study, we use a power law, based on Onnela et al.'s (2011) finding. Accordingly, we define the connection probability for new nodes as:

$$p(k_j, d_{ij}) = Ck_j d_{ij}^{-\alpha}, \tag{7}$$

where C is the normalization constant determined in way to fulfill the requirement of:

$$\iint Ck_j d_{ij}^{-\alpha} = 1, \tag{8}$$

The following pseudo-code elaborates our proposed approach to construct spatial Barabási-Albert (SBA) networks:

```
Create a fully connected initial network
For each node i = m0+1,…,N:
  Current degree = 0
  While Current degree < m:
     Node j = a uniformly randomly selected node from all existing
              nodes except the node i and nodes adjacent to node i
     p = C × (degree of node j) × (distance between node i and j)^-alpha
     chance = a random number between 0 and 1
     if p > chance:
          Create an edge between node i and node j
```

4 Results

In this section we present our computational results of the properties of the proposed socio-spatial network models and compare them with those of ER, SW, and BA networks when it is appropriate. More specifically, we measure clustering coefficient, network density, average path length, assortativity, transitivity, and min, max, mean, and standard deviation degree. While several algorithms have been proposed for some of these measures, for simplicity of reproducing the results, we use the algorithms that have been implemented in the Python's *Networkx* library. All results are averaged over 25 independent model runs with different random seeds. It should be noted that in all analysis in this study we only consider undirected networks. Moreover, since the focus of this study is on the human social network, we limit our analysis to Onnela et al's (2011) findings that $\alpha \approx 1.5$ and only consider three values of $\alpha = 1.2$, 1.5, and 2 for our sensitivity analysis purposes.

4.1 Properties of the SER network

In this subsection we compute and report the characteristics of our proposed SER social network in terms of empirical regularity measures. ER and SER networks have different parameters, so pairwise comparisons are not possible. However, computing and comparing their network properties can provide valuable insights about their behavior under considered parameter settings. Here we report network properties for three different values of connection probability p in the ER model and three different exponents α for the SER model, as shown in Table 1. We consider 1000 nodes in this section, however, our further analysis (not reported) showed that increasing the number of nodes to 5000 or 10,000 does not change the findings. The main results is that the SER model shows significantly larger clustering coefficient in its typical parameter range compared to the typical parameter range in the ER model. Moreover, the assortativity is significantly greater in the SER model. In fact, for $\alpha = 1.5$, assortativity equals to 0.22, consistent with empirical measurements for various networks within the range [0.2, 0.36] (Newman 2003a). In addition, degree in the SER model shows smaller values of maximum and standard deviation. Note that, for $\alpha \geq 2$, the SER model often becomes disconnected and, therefore, the average path length cannot be computed.

Table 1 Network properties of the ER and SER models

	ER			SER		
	$p = 0.01$	$p = 0.02$	$p = 0.05$	$\alpha = 1.2$	$\alpha = 1.5$	$\alpha = 2$
Mean degree	9.96	19.96	50.08	10.19	6.04	3.56
Clustering coefficient	0.01	0.02	0.05	0.09	0.2	0.41
Density	0.01	0.02	0.05	0.01	0.006	0.003
Average path	3.26	2.68	2.027	2.17	2.88	N.A.
Assortativity	−0.004	−0.003	−0.001	0.12	0.22	0.46
Transitivity	0.01	0.02	0.05	0.09	0.2	0.47
Min degree	2	7.16	28.68	1.28	0	0
Max degree	21.72	34.52	73.84	22.16	15.44	11.04
Std. degree	3.14	4.39	6.86	3.4	2.53	1.92

Node degree k is the number of links connected to a node in a given network, which is a fundamental network property correlated with or driving other network patterns and types (Cioffi-Revilla 2014, Chap. 4). A network's degree distribution shows the behavior of relative frequency of degree values, a distribution that is often denoted by p_k, Pr (k), or $p(k)$ and defined as a fraction of nodes that have k number of links in the network. This function can be interpreted as the probability that a uniformly randomly selected node in the network has degree k.

Figure 1 (left) shows the degree distribution of ER and SER networks for various values of p and α. In the ER model, links between nodes are randomly distributed, generating a normal degree distribution (Ducruet and Beauguitte 2013). The degree distribution of the SER model for the range of exponent values $1 < \alpha \leq 2$ also resembles a normal distribution, but with significantly different moments. Also it appears that as the exponent α increases, the normal curves for SER model grow taller. As discussed in Sect. 2, an important property of real-world social networks is having heavy-tailed degree distributions. We know that an ER network does not exhibit heavy-tailed degree distribution and, therefore, the SER model should show similar behavior. However, the question of interest here is which one has the heavier tail. To determine this, we compare the upper tail degree distribution of the SER network with that of the non-spatial ER. The distribution of a random variable X is said to have a heavy tail if:

$$Pr[X > x] \sim x^{-\alpha} \; as \; x \to \infty, \quad \alpha > 0$$

To compare degree distributions in two networks, we compute the function $Pr[\mathrm{X} > x]$, called the complementary cumulative distribution function (CCDF; Cioffi-Revilla 2014: chap. 6). If the resulting plot of one distribution falls above the other, we may conclude that the upper one has a heavier tail than the lower. Figure 1 (right) shows parametric $Pr[X > x]$ functions corresponding to three different values of connection probability p for ER and SER networks with different values of α. Interestingly, the SER network produces heavier degree distributions with decreasing values of α, meaning that nodes with high degree are more likely to

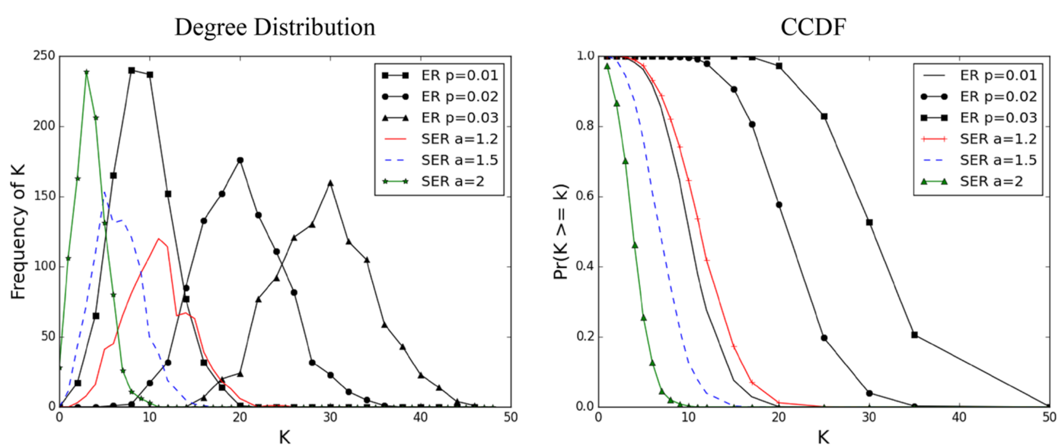

Fig. 1 Comparing degree distribution (*left*) and complementary cumulative density function or CCDF (*right*) of ER and SER networks (GD = 0.8)

exist in such spatial networks. We should have expected this because larger values of α reduces the long-range connections between nodes, which in turn decreases the degree of the hubs.

Next, we explore the effect of the exponent α and geographical density (GD) on the properties of the SER network. Figure 2 plots nine network properties of the SER network for different values of α and geographical density values. Results show that increasing the exponent α increases the average path length, but decreases mean, standard deviation, minimum, and maximum degree values in the SER network. Interestingly, the value of $\alpha = 1.5$ is a tipping point for the clustering coefficient, so the coefficient decreases until $\alpha = 1.5$, increasing afterward. On the other hand, increasing GD increases clustering, network density, mean degree, standard deviation of degree, transitivity, and assortativity, but decreases the average path length in the SER network. The underlying explanation for these results is that increasing the GD leads agents to locate closer together, which, in turn, increases the likelihood that pairs of agents will connect.

So far we have only looked at averages from our simulation runs and have reasoned based on visual inspection of plots. We now take a more detailed look at the variances associated with different settings and examine at how much variance can be attributed to variations in different parameters, based on analysis of variance (ANOVA). We complement Fig. 2 averages of our statistics with Table 2, which shows which proportion of overall variation can be attributed to different parameters and parameter interactions, along with the p value associated with each factor. From the table, we can conclude that most of the variations in the statistics of the SER network can be explained by the exponent α. The only cases where GD has greater

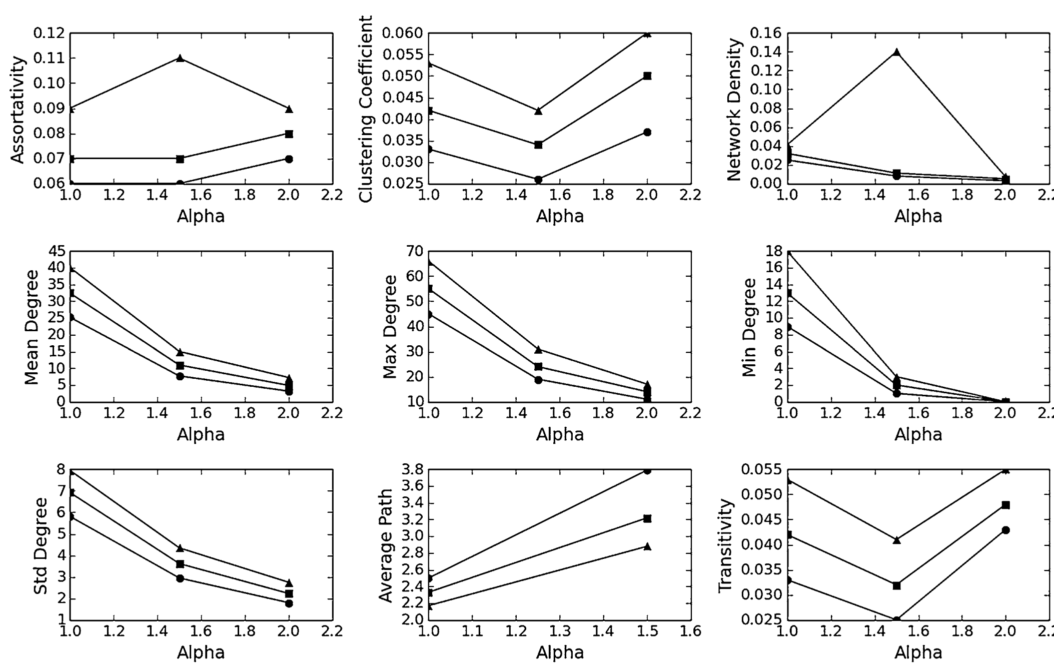

Fig. 2 Effect of exponent α and geographic density (GD) on properties of the SER network (Circle: GD = 0.3; Square: GD = 0.5; Triangle: GD = 0.8)

Table 2 Analysis of variance for SER network properties

	Sources of variation in proportions of sum of squares (p value)		
	GD	α	$GD \times \alpha$
Clustering coefficient	0.56 (0.00)	0.36 (0.00)	0.00 (0.00)
Mean degree	0.08 (0.00)	0.89 (0.00)	0.02 (0.00)
Std. degree	0.09 (0.00)	0.9 (0.00)	0.01 (0.00)
Min degree	0.06 (0.00)	0.85 (0.00)	0.06 (0.00)
Max degree	0.08 (0.00)	0.88 (0.00)	0.02 (0.00)
Assortativity	0.23 (0.00)	0.04 (0.002)	0.02 (0.17)
Transitivity	0.41 (0.00)	0.42 (0.00)	0.00 (0.00)
Density	0.08 (0.00)	0.88 (0.00)	0.03 (0.00)
Average path	0.33 (0.00)	0.59 (0.00)	0.03 (0.00)

effect are the clustering coefficient and assortativity. The results also illustrate a significant difference between the various levels of exponent α and GD on average values of all network properties.

4.2 Properties of the SWS network

In this section, we first consider Kleinberg's (2000) version of the SWS model, where the rewiring probability p is the same as in the original WS model, however, the probability of a new connection between two nodes is a function of their geographical distance d, as in Eq. (1). We consider 1000 nodes for both networks, each node connecting to its 20 nearest neighbors ($k = 20$), and geographical density is set to $GD = 0.8$. With respect to node degree, both WS and SWS-1 models produce the same degree distribution, since the main parameter affecting the shape of the distribution is the rewiring probability p, which in this case is the same for both models. Other network properties of the SWS-1 are reported and compared with the WS model in Table 3. Results show that the SWS-1 network has higher clustering coefficient, assortativity, transitivity, and minimum degree compared to the WS network, especially for low values of p. On the other hand, the maximum and standard deviation of degree tend to be lower in SWS-1 model.

The more interesting scenario, however, is when we define the rewiring probability as a function of distance between nodes (i.e., let $p(d) = C_2 d_{ij}^{\alpha}$). The intuition here is that if nodes are already close in the geographical space, we want them to be rewired with less probability than those that are geographically distant. Figure 3 (left) illustrates the degree distribution of the WS and SWS-2 networks for various values of rewiring probability p and distance exponent α. We know that the degree distribution of the WS network does not match most real-world networks very well (Newman 2003b). It seems that incorporating the spatial proximity of individuals into the construction of the small-world network does not fundamentally change its degree distribution; it only makes it taller. Results show that for all three values of α the corresponding degree distribution is similar to that of WS for

Table 3 Comparing network properties in WS and SWS-1 networks

	$p = 0.001$		$p = 0.01$		$p = 0.1$	
	WS	SWS-1	WS	SWS-1	WS	SWS-1
Mean degree	20	20	20	20	20	20
Clustering coefficient	0.69	0.71	0.61	0.69	0.52	0.52
Density	0.02	0.02	0.02	0.02	0.02	0.02
Average path	4.96	11.16	3.58	4.93	3.22	3.22
Assortativity	−0.0002	0.0014	−0.004	−0.0014	−0.006	−0.01
Transitivity	0.69	0.71	0.6	0.69	0.52	0.52
Min degree	17.8	19	16.6	17.96	15.52	15.24
Max degree	22.12	21.04	23.92	22.08	25	24.92
Std. degree	0.44	0.14	0.98	0.44	1.37	1.38

Note $\alpha = 1.5$, $GD = 0.8$, $C_1 = 0.02$

rewiring probability $p = 0.01$. Moreover, it appears that increasing the value of exponent α slightly makes the distribution thinner.

In order to compare the tails of the degree distributions on WS and SWS-2 networks and examine the effect of rewiring probability and distance exponent α on tails, we plot the CCDF of both networks for various values of their parameters in Fig. 3 (right), as before. We can see that the WS for $p = 0.01$ and $p = 0.1$ provide upper and lower bounds of the CCDF respectively. In other words, the WS for $p = 0.01$ has the heaviest tail, while SWS-2 with $\alpha = 2$ has a CCDF very close to it. Another interesting observation for the SWS-2 model is that as the distance exponent α grows, the corresponding degree distributions have heavier tails, meaning that larger hubs are more likely to exist for greater values of α. One explanation could be that SWS model produces less extreme-size hubs because nearby hubs with lower degree might be more attractive than far hubs with higher degree.

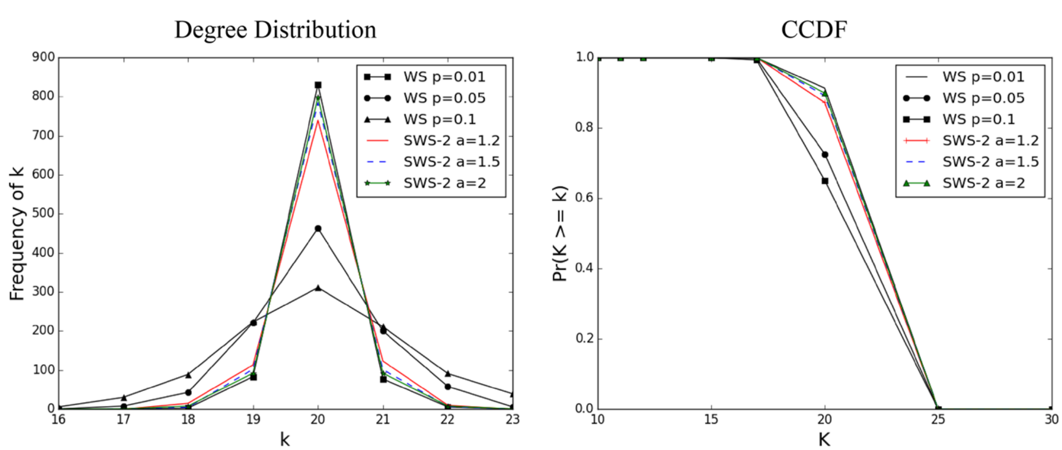

Fig. 3 Comparing degree frequency distribution (*left*) and CCDF (*right*) in WS and SWS-2 networks (GD = 0.8)

To examine the effect of the distance exponent α and GD on SWS-2 network properties, we compute the statistics for various values of α and GD, as shown in Fig. 4. Results show that clustering, average path length, minimum degree, and transitivity increase with α, but the maximum and standard deviation of degree decreases as α increases. In the case of assortativity, results show a tipping value for α. We also observe that increasing GD decreases clustering, average path length, transitivity, and assortativity. However, the standard deviation of degree decreases with GD. The ANOVA test shows that most of the variation in network properties can be attributed to the decay exponent α (Table 4). For example, in the case of clustering, average path length, and transitivity, ANOVA results show that α causes 90, 80, and 90 percent of the variations, respectively. Moreover, Table 4 shows a significant difference between various levels of exponent α and GD on average values of all network properties except for assortativity.

4.3 Properties of the SBA network

Our proposed SBA network and the original BA model have the parameter *m* in common (i.e., the number of links to build from a new node to prior nodes). As a result, we can compare network statistics across the two models and directly examine the effect of adding geography to the model. We compute the network properties for both models and report them in Table 5. Results show that the SBA model always produces networks with equal or higher clustering, assortativity, and transitivity. On the other hand, the SBA network has a lower maximum degree,

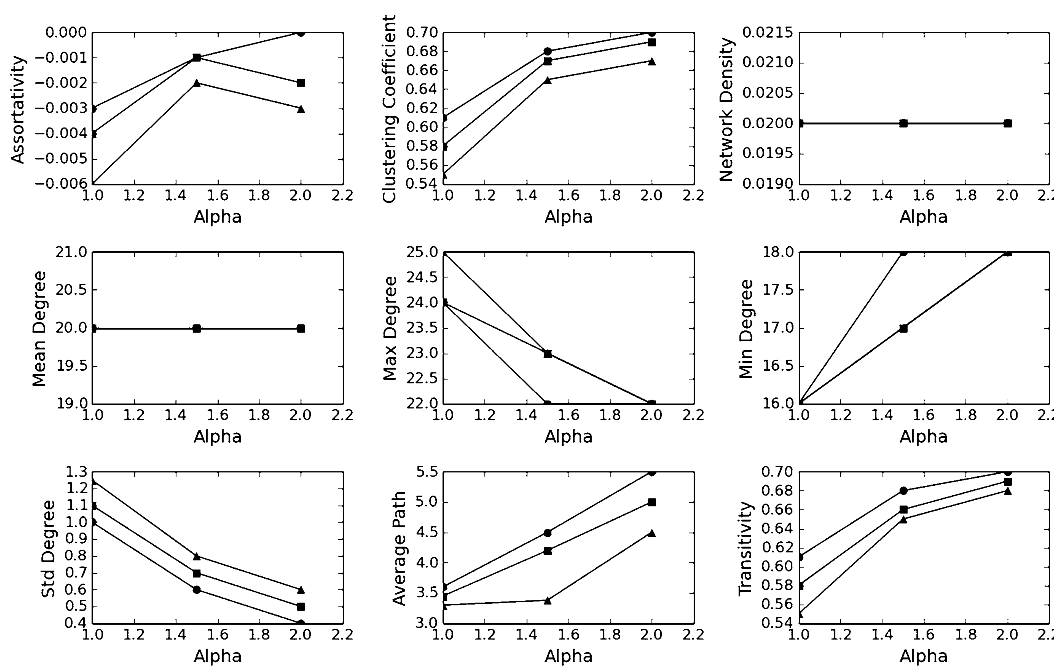

Fig. 4 Effect of exponent α and geographic density GD on properties of the SWS-2 network (Circle: GD = 0.3; Square: GD = 0.5; Triangle: GD = 0.8)

Table 4 Analysis of variance for SWS-2 network properties

	Sources of variation in proportions of sum of squares (p value)		
	GD	α	*GD* $\times$ α
Clustering coefficient	0.08 (0.00)	0.9 (0.00)	0.02 (0.00)
Std. degree	0.1 (0.00)	0.9 (0.00)	0.00 (0.00)
Min degree	0.08 (0.00)	0.6 (0.00)	0.009 (0.00)
Max degree	0.08 (0.00)	0.70 (0.00)	0.00 (0.00)
Assortativity	0.00 (0.40)	0.00 (0.12)	0.00 (0.97)
Transitivity	0.09 (0.00)	0.89 (0.00)	0.01 (0.00)
Density	0.00 (0.00)	0.00 (0.00)	0.00 (0.35)
Average path	0.14 (0.00)	0.8 (0.00)	0.04 (0.00)

minimum degree, and standard deviation of degree. It also has a longer average path length at all corresponding m values compared to the original BA network.

The most important feature of the BA model is that it produces a network with power law degree distribution. Accordingly, an interesting question to investigate is how this degree distribution changes when long-range connections are penalized by $d^{-\alpha}$. To illustrate this, we plot the degree distribution of the SBA model for various values of m and distance exponent α on double logarithmic scales in Fig. 5 (left) and compare it with the degree distribution in the original BA model. The straight line pattern of data resembles that of a power law distribution. However, we still need to test alternative distributions to find the one that best fits the data, because the tails of the data in Fig. 5 (left) are noisy. Another way to identify a power law is to plot the complementary cumulative distribution function on double logarithmic scales to see if a straight line is observed. If a set of values obeys a power law with exponent α, its CCDF also follows a power law, but with exponent α-1 in a log–log plot (see, e.g., Newman 2005; Cioffi-Revilla 2014, Chap. 6).

Table 5 Comparing the network properties in the BA and SBA models

	$m = 1$		$m = 3$		$m = 5$	
	BA	SBA	BA	SBA	BA	SBA
Mean degree	1.99	1.99	5.98	5.98	9.95	9.95
Clustering coefficient	0.0	0.0	0.03	0.05	0.04	0.06
Density	0.002	0.002	0.006	0.006	0.01	0.01
Average path	6.94	7.43	3.48	3.6	2.97	3.17
Assortativity	−0.12	−0.11	−0.08	−0.05	−0.06	−0.02
Transitivity	0.0	0.0	0.017	0.03	0.03	0.04
Min degree	1	1.1	2.9	3	4.71	5
Max degree	57	30	98	46	124	58
Std. degree	3.18	2.28	7.04	4.96	10.28	7.13

Note $\alpha = 1.5$, $GD = 0.8$, $N = 1000$

The log–log plot of the CCDFs in Fig. 5 (right) show excellent fits to a power law fit in the original BA network. However, in case of SBA networks with different values of decay exponent α, the corresponding CCDF plots behave like a power law mostly for small values of k, dropping more steeply for larger k, which can suggest a power law with exponential cut-off as the best fit. The fact that the BA lies above all other CCDF curves implies that it has heavier tail. The reason can be found in Table 5, where results show that the BA model has a significantly higher maximum degree than the SBA model for all parameter combinations. This means that BA model produces one or more extreme-size hub(s) that are not produced by a SBA network. On the other hand, the SBA model has significantly less standard deviation

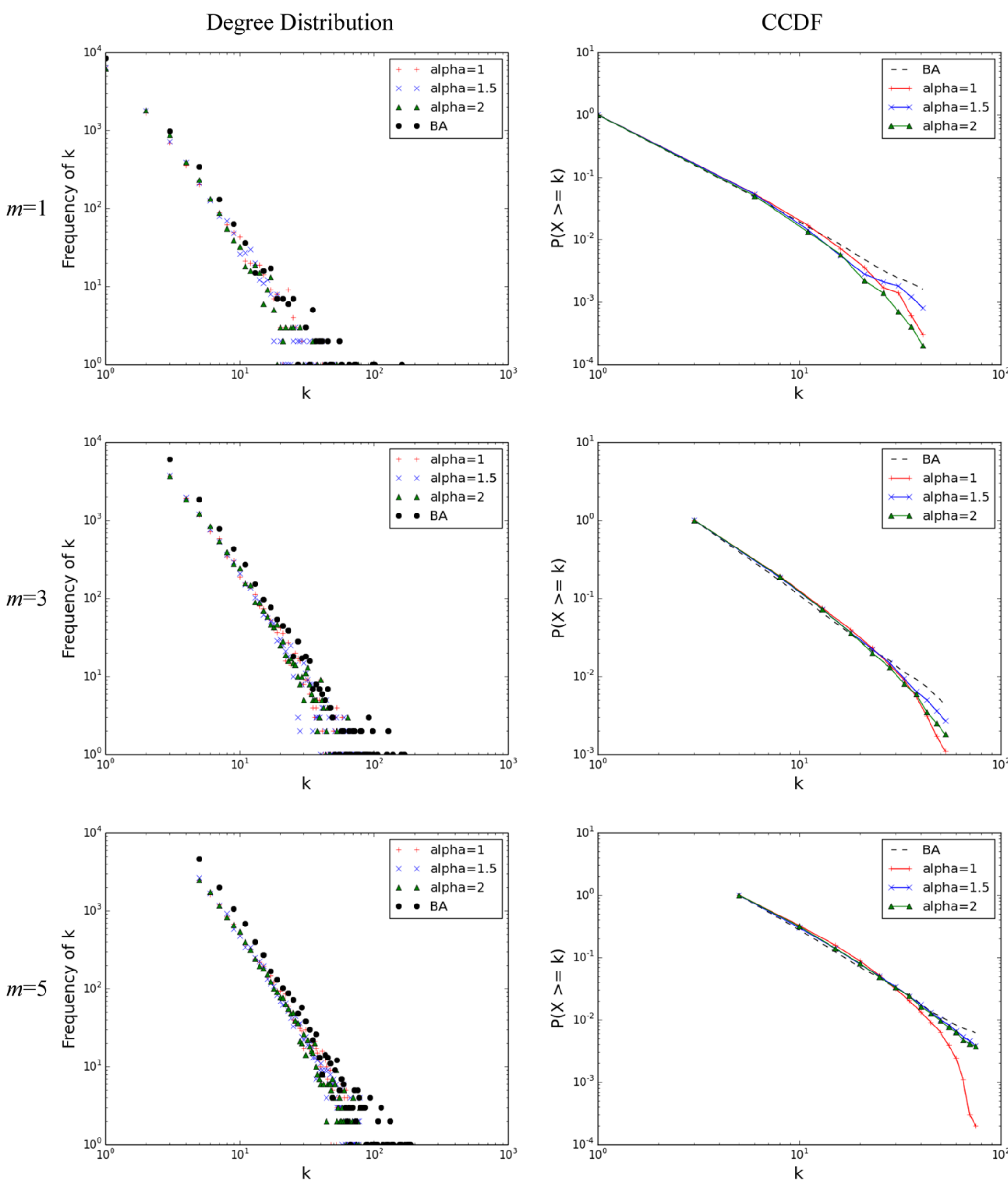

Fig. 5 Comparing degree frequency distribution and associated CCDF in BA and SBA networks for various values of α and m (N = 10,000 and GD = 0.8)

of degree compared to the BA model, meaning that differences between hubs and regular nodes are (on average) less pronounced in the SBA model.

To statistically test the goodness of fit, we compare the fit of the power law with other competing distributions. In doing so, we follow Clauset et al. (2009) in using the likelihood ratio test proposed by Vuong (1989). The test computes the logarithm of the ratio of the likelihoods (LR) or normalized log-likelihood ratio (NLR) of the data under two alternative distributions. Results for various combinations of m and α are reported in Table 6. We consider lognormal, exponential, and power law with exponential cut-off as competing distributions. Positive (negative) values of the NLR or LR indicate that a power law is a better (worse) fit compared to alternative distributions. Therefore, if the associated p value is less than a predefined threshold (e.g., 0.1), we reject the model with the worse fit. However, if the p value is large, we are not able to choose among competing distributions.

The more detailed properties of the SBA model and the effect of the exponent α and m are demonstrated in Fig. 6. These results show that clustering, transitivity, standard deviation of degree, and maximum degree all increase with distance exponent α. Network assortativity is the only feature that decreases with increasing α. Average path is nonlinear and concave in α. Network density, minimum degree, and mean degree are constant for various values of α.

The effect of GD is also shown in Fig. 6. Results show that clustering, standard deviation of degree, transitivity, and maximum degree decrease as GD grows. However, average path length and assortativity increase with GD. ANOVA test results indicate that most of the variation in network properties in the SBA network can be attributed to m (Table 7). For example, 98 percent and 67 percent of variations in average path length and clustering coefficient are due to the role of m. By contrast, GD has the least effect on variations. The results also show that differences between values of m are significant at the 0.01 level for all network

Table 6 Testing a power law against other models for degree distributions in SBA networks

	Lognormal		Exponential		Power law with exponential cut-off		Best fit
	NLR	*p*	*NLR*	*p*	*LR*	*p*	
m = 1							
$\alpha = 1$	−0.21	0.65	2.39	0.38	−0.49	0.32	Power law
$\alpha = 1.5$	−0.36	0.22	20.51	0.003	4.73	0.99	Power law
$\alpha = 2$	−0.90	0.03	35.83	2.17	0.014	0.86	Lognormal
m = 3							
$\alpha = 1$	−1.68	0.22	84.73	6.72	−6.15	0.0004	Power law with exp. cut-off
$\alpha = 1.5$	−0.21	0.68	73.10	1.02	−1.03	0.15	Power law
$\alpha = 2$	−0.01	0.91	47.31	0.000	−0.38	0.37	Power law
m = 5							
$\alpha = 1$	−2.86	0.11	493.3	2.51	−2.64	0.02	Power law with exp. cut-off
$\alpha = 1.5$	−0.65	0.45	18.75	0.01	−3.78	0.006	Power law with exp. cut-off
$\alpha = 2$	−0.51	0.51	80.8	1.04	39	0.37	Power law

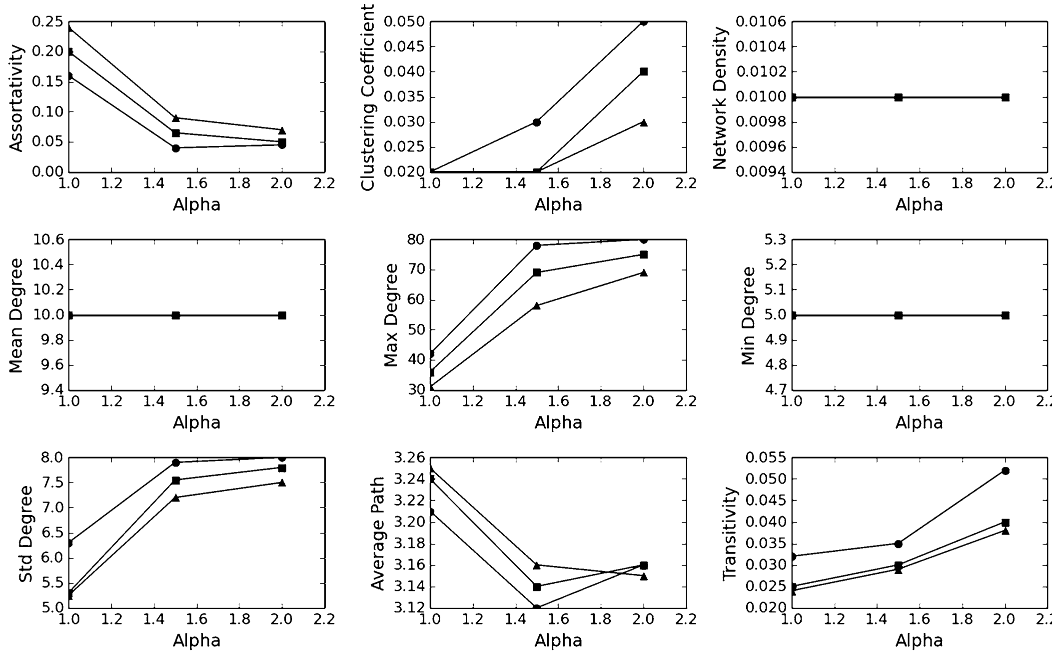

Fig. 6 Effect of exponent α and geographic density GD on properties of the SBA network ($m = 5$) (Circle: GD = 0.3; Square: GD = 0.5; Triangle: GD = 0.8)

properties (p value $= 0.00$), except for the minimum degree, where p value $= 0.52$. Finally, differences between values of α are significant at the 0.01 level (p value <0.01) for all network statistics except the mean, minimum degree, and density, where the p value >0.01. Although most interactions between factors are also significant at the 1 percent significance level, since the interaction effects are too low, it is safe to say that there no/little interdependence between the factors.

4.4 Analyzing the effect of spatial clustering

So far we have assumed that the points representing spatial coordinates of nodes are uniformly distributed across space. However, in the real-world, it is rare to find a region where the population is uniformly distributed. Rather, population is dense in large cities and sparse in rural areas (Bretagnolle and Pumain 2010). In other words, population is spatially clustered in any given country, with cities being the cluster centers. In this sense, it is rewarding to analyze the extent to which having geographically clustered points affects the behavior and properties of our proposed spatial social networks. In doing so, we construct a set of n Gaussian-distributed clusters with random variances. Moreover, we consider different levels of GD, which controls vacant spaces between points and clusters. Figure 7 illustrates how spatially clustered data points differ from the case of uniformly distributed and the effect of GD.

Figure 8 compares the degree distribution of the SER with those of an SER with clustered embedded coordinates (SER-C). Results show that having spatially clustered points shifts the peak of the distribution to the left, tending toward a

Table 7 ANOVA test of network properties in SBA networks with respect to d, m, and α

	Sources of variation in proportions of sum of squares (p value)						
	GD	m	α	$GD \times m$	$GD \times \alpha$	$m \times \alpha$	$GD \times m \times \alpha$
Clustering coefficient	0.01 (0.00)	0.67 (0.00)	0.19 (0.00)	0.01 (0.00)	0.01 (0.00)	0.09 (0.00)	0.01 (0.00)
Mean degree	0.00 (0.00)	1.00 (0.00)	0.00 (0.49)	0.00 (0.00)	0.00 (0.33)	0.00 (0.48)	0.00 (0.95)
Std. Degree	0.01 (0.00)	0.91 (0.00)	0.05 (0.00)	0.00 (0.00)	0.00 (0.00)	0.03 (0.00)	0.00 (0.00)
Min degree	0.00 (0.52)	0.98 (0.00)	0.00 (0.16)	0.00 (0.92)	0.00 (0.58)	0.00 (0.08)	0.02 (0.7)
Max degree	0.04 (0.00)	0.47 (0.00)	0.28 (0.00)	0.01 (0.00)	0.00 (0.4)	0.09 (0.00)	0.12 (0.09)
Assortativity	0.01 (0.00)	0.74 (0.00)	0.09 (0.00)	0.00 (0.00)	0.00 (0.00)	0.11 (0.00)	0.00 (0.00)
Transitivity	0.01 (0.00)	0.87 (0.00)	0.07 (0.00)	0.00 (0.00)	0.00 (0.00)	0.04 (0.00)	0.00 (0.00)
Density	0.00 (0.00)	1.00 (0.00)	0.00 (0.55)	0.00 (0.00)	0.00 (0.16)	0.00 (0.56)	0.00 (0.89)
Average path	0.00 (0.006)	0.98 (0.00)	0.00 (0.00)	0.00 (0.04)	0.00 (0.35)	0.00 (0.00)	0.01 (0.68)

lognormal distribution. This is especially the case for $n = 3$ and $n = 5$. In addition, results show that increasing the value of GD produces heavier tailed distributions, but does not change the overall distribution type. Figure 8 also shows that SER-C results generate a large number of nodes with zero degree, especially for low values of GD, consistent with expectations. With respect to other network properties, results in Table 8 show that incorporating clustered coordinates into the SER model improves most properties. More specifically, spatial clustering increases network clustering, mean degree, density, assortativity, and transitivity. However, it significantly enhances the standard deviation of degree. Increasing the number of clusters n does not significantly induce variation in network properties.

We compute and plot the degree distribution of the SWS-2 and SWS with clustered embedded coordinates (SWS-C) in Fig. 9, which shows the degree distribution of the SWS-C network for three different values of GD and number of clusters n. In all three plots, the SWS-C network shows a taller distribution than in the SWS-2 network, and the number of clusters n and GD show no significant effect on the shape of the distributions. In terms of other network properties, the SWS-C network has a higher average path length compared to the SWS-2 network, and for

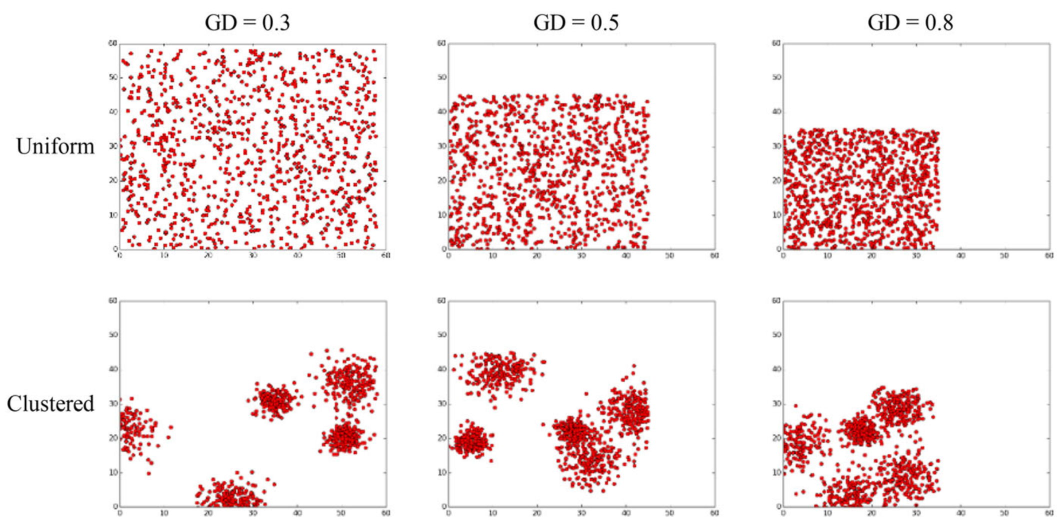

Fig. 7 Uniform versus clustered Points for different geographical densities

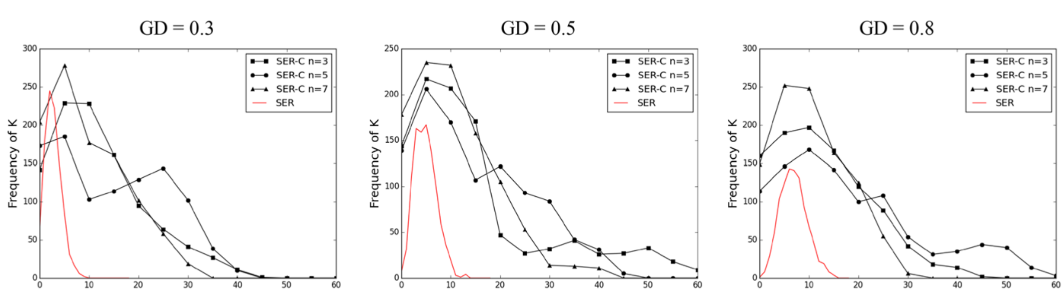

Fig. 8 Degree distributions of the SER-C and SER models for different values of n and GD ($N = 1000$, $\alpha = 1.5$)

$n = 5$ and $n = 7$, it has higher assortativity. Similar to the SER-C model, increasing the number of clusters in geographical space does not affect network properties (Table 9).

Finally, we plot the degree distribution and CCDF of the Spatial Barabási-Albert (SBA) network and SBA with clustered embedded coordinates (SBA-C) in Fig. 10. As in previous cases, we consider three values of GD and the number of clusters n. The log–log plot of the degree distributions resembles the power law distribution. However, due to noisy tails in all three cases, it is best to examine the CCDF plots. Results show that, for higher values of degree, the upper tail of the CCDF plots slightly diverges from a straight line, suggesting that alternative distributions should be considered. To statistically test for the best fit, we compute the NLR, LR, and corresponding p values, as reported in Table 10. Results show that, as the number of clusters increases, the distribution diverges from a perfect power law; at approximately $n = 7$ an exponential cut-off becomes a better fit. With respect to other network properties of interest, results show that the SBA-C network has significantly higher clustering coefficients and maximum degree compared to the

Table 8 Comparing network properties in SER and SER-C networks

	SER	SER-C		
		n = 3	n = 5	n = 7
Mean degree	6.04	9.59	9.53	9.54
Clustering coefficient	0.2	0.26	0.27	0.27
Density	0.006	0.009	0.009	0.009
Average path	N.A.	N.A.	N.A.	N.A.
Assortativity	0.22	0.58	0.59	0.58
Transitivity	0.2	0.32	0.32	0.32
Min degree	0	0	0	0
Max degree	15.44	28.56	28.08	28.8
Std. degree	2.53	5.85	5.84	5.84

Note $\alpha = 1.5$, $GD = 0.8$

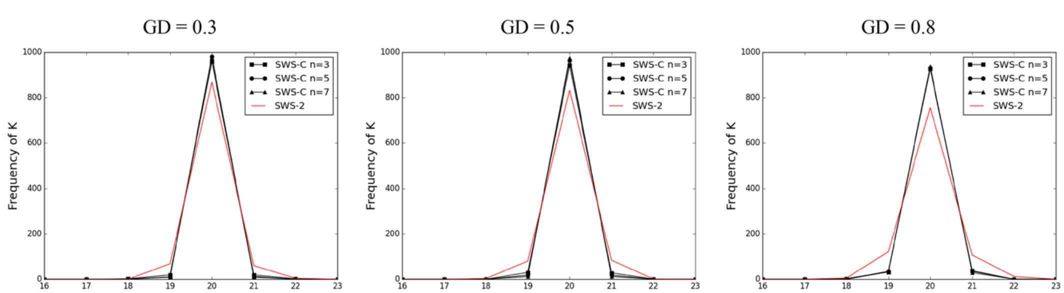

Fig. 9 Degree distributions of the SWS-C and SWS-2 models for different values of n and GD ($N = 1000$, $\alpha = 1.5$)

SBA network (Table 11). The number of clusters is not a source of variation in SBA-C's network properties, as reported earlier.

5 Discussion and conclusion

Location is a key feature of agents that has been empirically proven to have significant impact on the probability of connections. Availability of large data sets of human interactions, such as mobile calls, text messages, and social media data, among many others, has enabled researchers to discover regularities in human social networks, including the strong correlation of individuals' geography and their friendship and interaction patterns. These findings, along with the need for spatial

Table 9 Comparing network properties in SWS-2 and SWS-C networks

	SWS-2	SWS-C		
		n = 3	n = 5	n = 7
Mean degree	20	20	20	20
Clustering coefficient	0.68	0.7	0.7	0.7
Density	0.02	0.02	0.02	0.02
Average path	4.65	6.96	7.55	7.6
Assortativity	0.001	0.001	0.0035	0.004
Transitivity	0.68	0.7	0.7	0.7
Min degree	17.72	18.28	18.48	18.68
Max degree	22.32	21.56	21.24	21.28
Std. degree	0.51	0.27	0.24	0.23

Note $\alpha = 1.5$, $GD = 0.8$

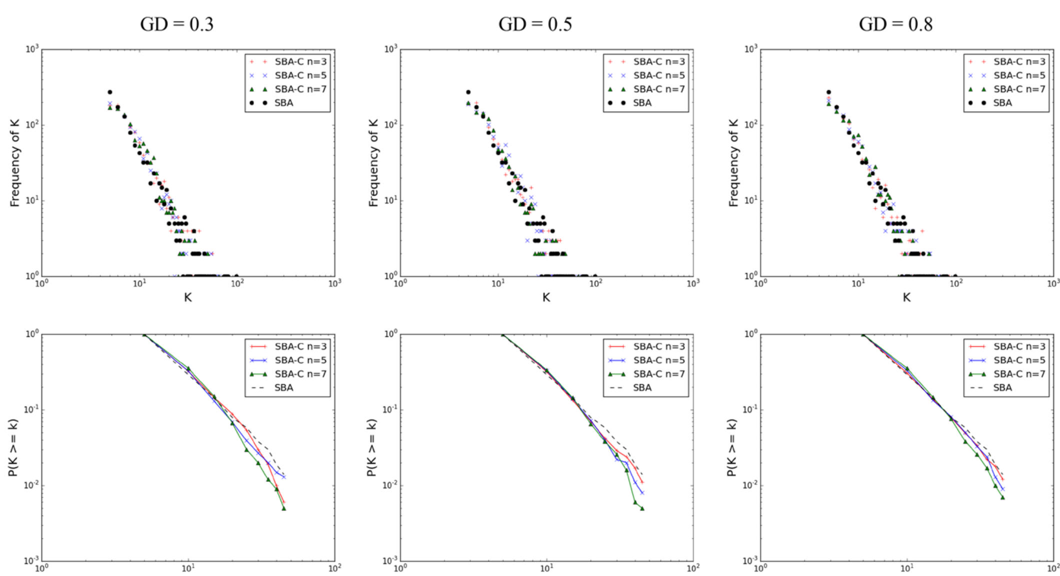

Fig. 10 Degree distributions in SBA-C and SBA models for different values of n and GD ($N = 1000$, $\alpha = 1.5$, $m = 5$)

Table 10 Testing a Power Law against other Distributions for the SBA-C Networks

	Lognormal		Exponential		Power Law with exponential cut-off		Best Fit
	NLR	*p*	*NLR*	*p*	*LR*	*p*	
SBA	−0.84	0.21	59.62	1.03	−0.16	0.56	Power law
SBA-C							
$n = 3$	0.01	0.37	71.00	5.63	−0.20	0.51	Power law
$n = 5$	−0.45	0.55	8.25	0.12	−0.95	0.16	Power law
$n = 7$	0.003	0.92	119.22	0.001	−1.62	0.07	Power law with exp. cut-off

Note $\alpha = 1.5$, $GD = 0.8$, $m = 5$

Table 11 Comparing Network Properties in SBA and SBA-C Networks

	SBA	SBA-C		
		n = 3	n = 5	n = 7
Mean degree	9.74	9.95	9.95	9.95
Clustering coefficient	0.02	0.06	0.06	0.06
Density	0.01	0.01	0.01	0.01
Average path	3.17	3.26	3.26	3.27
Assortativity	0.05	0.007	0.005	0.008
Transitivity	0.02	0.05	0.05	0.055
Min degree	5	4.28	3.96	4.04
Max degree	58	71.44	71.52	71.2
Std. degree	7.13	7.14	7.17	7.08

* $\alpha = 1.5$, $GD = 0.8$, $m = 5$

analysis of computational results, have motivated the development of spatial social network generation algorithms.

Based on classical principles and theory of human and social geography (e.g., inverse-square interaction laws), combined with network science models, we proposed a class of spatial social network generation algorithms. Our proposed network construction methods build on and extend the three well-known ER, WS, and BA graphs. Assuming that the probability of connection between two agents decays with their distance as a power law (i.e., as $p(d) \sim d^{-\alpha}$ with $1 < \alpha \leq 2$), we penalized long-range spatial interactions, consistent with empirically validated social theory. That is, we allow a pair of agents to be connected with probability proportional to the inverse of the distance between them on an $m \times m$ Cartesian space. To explore the network characteristics in each model, we used measures that account for empirical regularities found in real-word social networks. We also compared our proposed models with their associated originals to assess the effect of geographic space on each models' structural features.

As a general finding, our results demonstrated that including a geographic distance selection mechanism in the formation of networks enhances clustering coefficient, assortativity, and transitivity in all three classic networks (Table 12) which are empirically grounded desired characteristics. One common way to compute the clustering coefficient is to measure the fraction of triads with a complete triangle (Newman 2003b; Cioffi-Revilla 2014). In fact, incorporating the spatial proximity in the network formation model increase the probability of finding triangles in networks. This happens because if node A is connected to node B and node B is connected to node C, this implies that the probability that nodes A and C are geographically close to node B is high, which in turn implies that the probability that nodes A and C are spatially close is high. As a result, and due to the importance of distance in creating links, the probability that A and C are connected is increased in spatial networks. Assortativity is higher in spatial social networks, indicating that the probability that hubs are connected is increased when distance is taken into account in the formation of social networks. This has immediate implication for networks security properties, such as vulnerability, reliability, and resilience, as these and others have been shown to be sensitive to the network properties (e.g., Modarres et al. 2010; Myers 2010; Newman 2003a).

Numerical analysis of the spatial Erdös-Rényi (SER) network reveals that, similar to the original ER model, it generates a normal degree distribution, but with much lower peaks (i.e., platykurtic). Moreover, increasing the distance-decay exponent α while holding geographic density constant causes the degree distribution to become taller and thinner (leptokurtic). In other words, for lower values of α the SER model has heavier tails. Pairwise comparison of the ER and SER models is not possible because they have different parameters, but, for considered values of p and distance exponent α, the ER with $p = 0.05$ has degree distribution with the heaviest tail. Graphical sensitivity analysis and analysis of variance showed that both geographic density and the distance-decay exponent α have statistically significant effects on average values of all measured network properties, with most of the variations attributed to the distance exponent α.

With respect to the spatial Watts-Strogatz (SWS) model, we analyzed two different cases: (1) the rewiring probability p is fixed and the connection probability is a function of distance between nodes (SWS-1); and (2) the rewiring probability is a function of distance between nodes, decaying as a power law, and the probability of connection between nodes is uniformly distributed (SWS-2). Results showed that the SWS-1 network has the same degree distribution to that of the WS network, but with greater clustering, assortativity, and average path length for small values of p. For the SWS-2 network we found that as the exponent α grows, the degree distribution becomes taller (leptokurtic), meaning a decrease in the standard deviation of the degree distribution. Moreover, the degree distribution of the SWS network has heavier tails for higher values of α, meaning that larger size hubs are more likely to exist in spatial networks. The ANOVA test revealed significant differences between values of geographical density and exponent α on all measured network characteristics at the 0.05 significance level, except for the assortativity, with most of the variations explained by α.

The proposed SBA model and the original BA model have in common the parameter m, so we could compare their properties. Numerical analysis showed that the spatial model always produces networks with lower clustering, maximum degree, standard deviation of degree, and minimum degree. However, the SBA has

Table 12 Summary of the Comparison between Original and Spatial Networks' Properties

	Degree distribution	Mean degree	Clustering coefficient	Density	Average path length	Assortativity	Transitivity	Min degree	Max degree	Std. degree
SER vs. ER	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
SWS-1 vs. WS		Same	$\geq$	Same	$\geq$	$\geq$	$\geq$	$\geq$	$<$	$\leq$
SBA vs. BA	Mostly power law	Same	$\geq$	Same	$>$	$>$	$>$	$>$	$<$	$<$

N.A. Not applicable

longer average path length and higher assortativity compared to BA model at corresponding m values. For the values of α that we analyzed (i.e. $1 < \alpha \leq 2$, grounded in recent empirical findings about real-world social networks), the degree distribution of the SBA model tends to deviate from a perfect power law. That is, in some cases, especially for lower values of exponent α, we observe a crossover from power law behavior to exponential behavior. The tail analysis of degree distributions showed that for small values of k, the SBA model has heavier tail than the original BA model. However, for medium and large value of k, the original BA model has heavier tails. This is because the SBA model has significantly less maximum degree compared with the BA model. Moreover, the ANOVA test revealed that m and geographic density have the most and least effect on network characteristics, respectively. For future use of the proposed network models, Table 13 should be consulted regarding the effect of distance exponent α and geographical density on network properties.

We analyzed an empirically relevant case where agent coordinates are clustered in space. This is an important extension of our analysis because in most of real-world cases the spatial distribution of the individuals is clustered at multiple regional scales such as counties and countries. Our results demonstrated that, in the presence of spatially clustered coordinates, the SER network shows greater clustering, assortativity, and transitivity. Moreover, its degree distribution tilts toward the lognormal distribution as the number of zero degree agent-nodes (social isolates) significantly increases. In the SWS network, results showed taller (leptokurtic) degree distributions, higher assortativity, and average path length. Finally, having clustered coordinates increases the clustering coefficient and transitivity in the SBA network. It also imposes some deviations from the perfect power law for large values of degrees, making the power law with exponential cut-off the better fit for degree distribution.

The exponential truncation we found may have critical effects on the dynamics of the system, especially for domains in which nodes with large degrees have important implications such as opinion dynamics and radicalization (Alizadeh et al. 2015). Our finding of mechanisms leading to an exponential truncation indicates that the number of major hubs will be smaller than predicted for scale-free networks where degree distribution follows a perfect power law. This finding has immediate implications for understanding and managing onset and spread of epidemics and social contagion.

Finally, the proposed spatial network models offer a simple way to the agent-based modelling community to model the geography of individuals wherever it is desired and relevant. Our class of network generation algorithms enables the modelers to effectively generate a network that simultaneously captures the spatial proximity of individuals and the desired topology. The advantage of our models is that the geography and topology are linked together and not separate. This would make the researchers able to study the bipartite networks of physical space and any arbitrary conceptualizations of social networks (e.g. friends, liking, conversations, etc.) as well.

Table 13 Summary of the Effect of Distance Exponent α and Geographical Density on Spatial Networks Properties

		Degree distribution	Mean degree	Clustering coefficient	Density	Average path length	Assortativity	Transitivity	Min degree	Max degree	Std. degree
α↑	SER	Taller	↓	NP	NP	↑	NP	NP	↓	↓	↓
	SWS-2	Thinner	Same	↑	Same	↑	NP	↑	↑	↓	↓
	SBA	Mostly power law	Same	↑	Same	NP	↓	↑	Same	↑	↑
GD↑	SER		↑	↑	↑	↓	↑	↑	↑	↑	↑
	SWS-2		Same	↓	Same	↓	↓	↓	↓	↑	↑
	SBA		Same	NP	Same	NP	NP	NP	Same	NP	NP

NP No pattern

Acknowledgments This study was supported in part by the Center for Social Complexity and the Computational Social Science program within the Department of Computational and Data Sciences at George Mason University. M. Alizadeh is funded by a GMU Presidential Fellowship and, together with C. Cioffi-Revilla, by ONR-Minerva Grant No. N00014130054.

References

Alam SJ, Geller A (2012) Networks in Agent-based Social Simulation. In: Heppenstall A, Crooks AT, See LM, Batty M (eds) Agent-based models of geographical systems. Springer, New York, pp 199–218

Albert R, Barabási A-L (2002) Statistical mechanics of complex networks. Rev Mod Phys 74(1):47

Albert R, Jeong H, Barabási AL (1999) Internet: Diameter of the world-wide web. Nat 401(6749):130–131

Alizadeh M, Coman A, Lewis M, Cioffi-Revilla C (2014) Intergroup conflict escalation leads to more extremism. J Artif Soc Soc Simul 17(4):4

Alizadeh M, Cioffi-Revilla C, Crooks A (2015) The effect of in-group favoritism on the collective behavior of individuals'opinions. Adv Complex Syst 18(01n02):1550002

Barabasi A-L, Albert R (1999) Emergence of Scaling in Random Networks. Science New Ser 286(5439):509–512

Barthélemy M (2003) Crossover from scale-free to spatial networks. Europhys Lett 63(6):915

Barthélemy M (2011) Spatial networks. Phys Rep 499(1):1–101

Batty M (2012) The new science of cities. MIT Press, Cambridge

Boccaletti S, Latora V, Moreno Y, Chavez M, Hwang DU (2006) Complex networks: structure and dynamics. Phys Rep 424(4–5):175–308

Borgatti S, Everett MG, Johnson JC (2013) Analyzing social networks. Sage, Los Angeles

Bretagnolle A, Pumain D (2010) Simulating urban networks through multiscalar space-time dynamics (Europe and United States, 17th–20th centuries). Urban Studies 47(13):2819–2839

Bruggeman J (2008) Social networks. Routledge, London

Callaway DS, Hopcroft JE, Kleinberg JM, Newman MEJ, Strogatz SH (2001) Are randomly grown graphs really random? Phys Rev E 64:041902

Chung F, Lu L (2002) The average distances in random graphs with given expected degrees. Proc Natl Acad Sci USA 99:15879–15882

Cioffi-Revilla C (2014) Introduction to computational social science: principles and applications. Springer, London

Cioffi-Revilla C (2015) A unified framework for convergence of social, engineering, and natural sciences. In: Bainbridge WS, Roco MC (eds) Handbook of science and technology convergence. Springer, London (in preparation)

Clauset A, Shalizi CR, Newman ME (2009) Power-law distributions in empirical data. SIAM review 51(4):661–703

Cointet JP, Roth C (2007) How realistic should knowledge diffusion models be?. J Artif Soc Soc Simul 10(3):5

Dorogovtsev SN, Mendes JFF (2002) Evolution of networks. Adv Phys 51:1079–1187

Ducruet C, Beauguitte L (2013) Spatial science and network science: review and outcomes of a complex relationship. Netw Spat Econ 14(3–4):1–20

Edmonds B (2006) How are physical and social spaces related? In: Billari FC, Fent T, Prskawetz A, Scheffran J (eds) Agent-based computational modelling. Springer, New York. (Downloaded on 10 March 08 from http://cfpm.org/cpmrep127.html)

Erdös P, Rényi A (1959) On random graphs. Publ Math 6:290–297

Erdös P, Rényi A (1960) On the evolution of random graphs. Publ Math Inst Hung Acad Sci 5:17–61

Fischer CS (1982) To dwell among friends. University of Chicago Press, Chicago

Garfield E (1979) It's a small world after all. Curr Contents 43:5–10

Gilbert EN (1959) Random graphs. Ann Math Stat 30:1141–1144

Gilbert N (2006) Putting the social into social simulation. Keynote address to the first world social simulation conference, Kyoto

Hamill L, Gilbert N (2009) Social circles: a simple structure for agent-based social network models. J Artif Soc Soc Simul 12(2):3

Holzhauer S, Krebs F, Ernst A (2013) Considering baseline homophily when generating spatial social networks for agent-based modelling. Comput Math Org Theory 19(2):128–150

Kleinberg JM (2000) Navigation in a small world. Nature 406:845

Klemm K, Eguiluz VM (2002) Growing scale-free networks with small-world behavior. Phys Rev E 65(5):057102

Lambiotte R, Blondel VD, de Kerchove C, Huens E, Prieur C, Smoreda Z, Van Dooren P (2008) Geographical dispersal of mobile communication networks. Physica A 387(21):5317–5325

Latane B, Liu JH, Nowak A, Bonevento M, Zheng L (1995) Distance matters: physical space and social impact. Pers Soc Psychol Bull 21(8):795–805

Manna SS, Sen P (2002) Modulated scale-free network in Euclidean space. Phys Rev E 66(6):066114

Milgram S (1967) The small world problem. Psychol Today 2:60–67

Modarres M, Kaminskiy M, Krivtsov V (2010) Reliability engineering and risk analysis: a proctical guide, 2nd edn. CRC Press, Boca Raton

Myers A (2010) Complex system reliability: multichannel systems with imperfect fault coverage, 2nd edn. Springer, London

Newman MEJ (2002) Assortative mixing in networks. Phys Rev Lett 89(20):208701

Newman MEJ (2003a) Mixing patterns in networks. Phys Rev E 67:026126

Newman MEJ (2003b) The structure and function of complex networks. SIAM Rev 45(2):167–256

Newman MEJ (2004) Fast algorithm for detecting community structure in networks. Phys Rev E 69:066133

Newman MEJ (2005) Power laws, Pareto distributions and Zipf's law. Contemp phys 46(5):323–351

Onnela JP, Arbesman S, Gonzalez MC, Barabási AL, Christakis NA (2011) Geographic constraints on social network groups. PLoS One 6(4):e16939

Rapoport A (1957) Contribution to the theory of random and biased nets. Bull Math Biophys 19:257–277

Redner S (1998) How popular is your paper? An empirical study of the citation distribution. Eur Phys J B 4:131–134

Soboll A, Elbers M, Barthel R, Schmude J, Ernst A, Ziller R (2011) Integrated regional modelling and scenario development to evaluate future water demand under global change conditions. Mitig Adapt Strateg Glob Change 16(4):477–498

Vuong QH (1989) Likelihood ratio tests for model selection and non-nested hypotheses. Econometrica 57(2):307–333

Watts DJ, Strogatz SH (1998) Collective dynamics of 'small-world' networks. Nature 393:4

Wong LH, Pattison P, Robins G (2006) A spatial model for social networks. Physica A 360:99–120

Xulvi-Brunet RI, Sokolov M (2002) Evolving networks with disadvantaged long-range connections. Phys Rev E 66(2):026118

Yook S-H, Jeong H, Barabási A-L (2002) Modeling the Internet's large-scale topology. Proc Natl Acad Sci 99(21):13382–13386

Yoshida T, Tomizawa N, Gotoh T, Iguchi H, Sugioka K, Ikeda, KI (2008) Consumer phase shift simulation based on social psychology and complex networks. In: 2008 IEEE Congress on Services-Part I, pp 289–296

Meysam Alizadeh is a post-doctoral researcher at the Center for Complex Networks and Systems Research at Indiana University. He received a PhD in Computational Social Science from George Mason University. Currently, his research interests are agent-based modeling, social media data analysis, natural language processing, intergroup relations, machine learning, and opinion dynamics.

Claudio Cioffi-Revilla is the Professor of Computational Social Science, founding and former Chair of the Department of Computational Social Science, and founding and current Director of the Mason Center for Social Complexity at George Mason University. He holds two doctoral degrees in Political Science and International Relations and his areas of special interest include quantitative, mathematical, and simulation models applied to complex human and social systems. Dr. Cioffi's research on conflict, international relations, and social complexity has been funded by DARPA, NSF, ONR, NATO, and European research agencies.

Andrew Crooks is an assistant professor in the Department of Computational Social Science at George Mason University. He received his PhD from the Centre for Advanced Spatial Analysis (CASA) at University College London (UCL) where he focused on combining ABM and GIS for the studying residential dynamics. His research interests include ABM, GIS and visualisation focused upon city systems.