Fu G, Wilkinson S, Dawson R.

A Spatial Network Model for Infrastructure System Development.

Computer-Aided Civil and Infrastructure Engineering 2016

DOI: http://dx.doi.org/10.1111/mice.12204

Copyright:

©2016 The Authors Computer-Aided Civil and Infrastructure Engineering published by Wiley Periodicals, Inc. on behalf of Editor

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

DOI link to article:

http://dx.doi.org/10.1111/mice.12204

Date deposited:

01/04/2016

This work is licensed under a Creative Commons Attribution 4.0 International License
A Spatial Network Model for Civil Infrastructure System Development

Gaihua Fu*, Sean Wilkinson & Richard J. Dawson
School of Civil Engineering and Geosciences, Newcastle University, Newcastle upon Tyne, UK

Abstract: Infrastructure networks play an important role in improving economic prosperity, enabling movement of resources, and protecting communities from hazards. As these networks serve population, they evolve in response to social, economic, environmental, and technological changes. Consideration of these interactions has thus far been limited by use of simplified data sets and idealized network structures, and is unable to explain the complexity and suboptimal structures displayed by real infrastructure networks. This article presents a new computational model that simulates the growth and evolution of infrastructure systems. Empirical evidence obtained from analysis of nontrivial real-world data sets is used to identify the mechanisms that guide and govern system-scale evolution of infrastructure networks. The model investigates the interplay of three key drivers, namely network demand, network efficiency, and network cost in shaping infrastructure network architectures. The validity of the model is verified by comparing key topological and spatial properties of simulated networks with real-world networks from six infrastructure sectors. The model is used to develop and explore different scenarios of infrastructure network futures, and their resilience is shown to change as a result of different infrastructure management policies. The model can therefore be used to identify system-wide infrastructure engineering strategies to reduce network costs, increase network efficiency, and improve the resilience of infrastructure networks to disruptive events.

1 INTRODUCTION

Infrastructure networks enable the flow of goods and services to support a modern economy and improve quality of life. As the global population becomes increasingly urbanized (UN, 2014), these networks become increasingly important at ensuring the delivery of crucial services such as lighting, water, and mobility that support human settlements (Castillo et al., 2014; Dobbs et al., 2013). The downside to this is that our reliance on these systems is now so great that their disruption can lead to disproportionate consequences for the communities that rely on them (Bocchini et al., 2014; Cavalieri et al., 2014; Cavallaro et al., 2014; Duenas-Osorio and Rojo, 2011; Futurechi and Miller-Hooks, 2014; Wilkinson et al., 2013; Wilkinson et al., 2012a). Although it is well recognized that resilient networks are crucial to communities, the features of a resilient network are only partially understood (Argyroudis et al., 2015; Buldyrev et al., 2010; Esposito et al., 2015; Franchin and Cavalieri, 2015; Fu et al., 2014a; O’Rourke et al., 2005).

As infrastructure networks serve populations, they must evolve in response to evolving populations, new developments and demographics shifts. This evolution is believed to have been governed by a few fundamental rules, that optimize their ability to provide services to the populations that rely on them (Bettencourt et al., 2007), resulting in emergent behavior where infrastructure networks evolve to possess particular architectures, and therefore particular vulnerabilities (Carvalho et al., 2009; Crucitti et al., 2004). To assess vulnerabilities in future networks, it is necessary to understand the key drivers behind this network evolution and to develop models that can simulate the growth of infrastructure networks.

Among techniques developed in the literature to better understand systems behavior, complex network theory has made significant contribution to quantitatively represent and model civil infrastructure systems.
The majority of this research focuses on investigating topological aspect of infrastructure systems (Albert et al., 2004; Cohen et al., 2001; Crucitti et al., 2004; Crucitti et al., 2006; Dunn et al., 2013; Fu et al., 2014a; Rosas-Casals et al., 2007; Sridhar and Sheth, 2008). However, like many other network systems, physical space plays a significant role on the characterization of these networks (Barthelemy, 2011; Strano et al., 2012). For example, distance affects the time and cost it takes for goods to be transported in a road network (Ford et al., 2015), and the spatial distribution of demand in an electricity system can influence where a particular cable or substation needs to be constructed to serve population. Ignoring the spatial aspect of networks, and how they might evolve, is therefore to miss some important features of these networks.

Some models have been developed with spatial constraints in mind, and many of them formulate network development as an optimization problem (Barthelemy and Flammini, 2006; Gastner and Newman, 2006a, b; Xie et al., 2007). That is, the formation of an infrastructure network is considered to be satisfying some optimal policies, for example, minimum network construction cost function as studied in Barthelemy and Flammini (2008), minimum network construction and travelling cost function as investigated in Gastner and Newman (2006a). However, the deterministic nature of these models leads to idealized network structures that are optimized against a snapshot of the wider drivers of infrastructure evolution and therefore fail to explain the variations and suboptimal layouts displayed by real infrastructure networks. An alternative branch of research is based on probability theory (Boas et al., 2009; Guimera and Amaral, 2004; Lennartsson et al., 2012; Wilkinson et al., 2012b; Yook et al., 2002). A model of this type drives network growth by observing probability distributions of some key network features, and it enables considerable complexity and variations of infrastructure networks to be represented. Many models proposed in the literature work well for networks where physical layout of network nodes are predetermined (Guimera and Amaral, 2004), or they target at synthetic networks or specific class of infrastructure networks (Barthelemy and Flammini, 2008; Boas et al., 2009; Wilkinson et al., 2012b). Empirical analysis of basic mechanisms that govern infrastructure growth and development is still missing. There is also a lack of quantitative evaluation of proposed models against real-world data.

To address this, this article presents a computational model with the aim to: (1) reproduce networks that exhibit the spatial and topological structure and properties observed in a wide range of real infrastructure networks and (2) simulate the evolution of existing networks under scenarios representing different future network management strategies. The model draws upon large amounts of empirical evidence we have obtained from the analysis of nontrivial real-world data sets. By accounting for geographical constraints in network development, the model explores the interplay of three key drivers, namely network demand, network efficiency, and network cost in governing network development. The validity of the model is verified by comparing simulated networks with real-world networks. Our experimental results demonstrate that the model can reproduce networks that share topological and spatial properties with their real-world counterparts. We further demonstrate how the model can be employed to generate a range of future network simulations that reflect the influence of different drivers of network development.

This article reports on substantial methodological development and analysis of large engineering systems made following initial proof of concept demonstration in Fu et al. (2014b). First, empirical data analysis includes a number of data sets across multiple infrastructure domains. Second, although we present our preliminary conceptualization of the model in Fu et al. (2014b), here the model has been defined and described mathematically, and three regulating parameters have been introduced into the model so that to generate networks of different structure and properties. Third, a rigorous analysis has been performed and a system boundary has been identified to show the types of networks that can be generated with the model. Finally, in this article, we have validated the model by comparing simulated networks against real infrastructure networks using matrix from both classical graph theory and spatial network theory.

The article is structured as follows. Section 2 presents our empirical analysis of real-world networks from six infrastructure sectors. Section 3 describes a spatial network model that simulates the growth and evolution of infrastructure systems. Section 4 analyzes the properties of the simulated networks. Section 5 validates the model and demonstrates how the model can be employed to evolve an existing infrastructure network into different futures. Section 6 concludes the article and suggests some potential future extension of this work.

2 EMPIRICAL ANALYSIS OF REAL INFRASTRUCTURE NETWORKS

We have studied a number of real infrastructure networks to ascertain what characteristics they share and what may be the underlying drivers that have resulted in these attributed emerging. From the complex
An infrastructure network model

systems perspective, a network is a set of nodes, which are connected by edges/links. The degree of a node is the number of connections it has to other nodes, and the degree distribution is the probability distribution of these degrees over the whole network. The reader is referred to Table A1 of Appendix A for the list of terms and notations used in this paper.

We first investigated how infrastructure facilities (nodes) are distributed in space and found that the distribution of network nodes is driven by the demand for the service that an infrastructure system delivers. This correlation between demand (proxied as population density here) and network nodes is summarized and plotted in Figure 1, which shows network node density as a function of population density for a range of infrastructure types.

Several different forms of correlation between infrastructure nodes and demand can be observed. A strong correlation exists between population density and electricity substation density (Figure 1a), that is, a densely populated area has more substations than less populated areas. The same happens to road networks. The near straight lines in the log-log plot indicate that network node density does not correlate with the population density in a linear form, but rather follows a power law relationship. Because demand/population usually has a nonuniform or heterogeneous distribution in space, the resulting distribution of infrastructure nodes is spatially clustered as shown in Figure 2. The correlation becomes much weaker between telecommunications masts and demand (Figure 1b). This weak correlation results in a rather dispersed infrastructure distribution, as shown in Figure 3. This dispersed distribution was also shown in water distribution networks and a weak negative correlation between network nodes and demand was identified for this network (i.e., hot spots of water network facilities, e.g., reservoirs, are not located in the most populated areas).

Having studied infrastructure node distribution, we then examined how these nodes are connected to each other. We found in all of the networks studied, there was a clear bias toward connecting a network node to ones that are in its geographical proximity. Here, we normalize the physical length of a network link/edge to show the percentage of the geographical space it spans, that is,

\[ D = \frac{d}{d_{\text{max}}} \]  

(1)

where \( d \) is the geographical length of a network edge, and \( d_{\text{max}} \) is maximum possible length that an edge can span in a space \( S \). As shown in Figure 4, the range of edges of the railway network of England and Wales only spans 0.01%–4% of network space. This is not surpris-

Fig. 1. Infrastructure facility density versus population density on a log-log plot (a) strong correlation observed on electricity and road data (b) weak correlation observed on telecommunication and water distribution data. Here, density is calculated for per census ward of England and Wales.

Population data are sourced from the 2011 Census of England and Wales (ONS, 2012), and infrastructure data are sourced from the data sets provided by Ordnance Survey (OS, 2015).
a comparable sized network with shorter links. We assume that the cost of a network is primarily related to the physical length of its network links, because an edge length determines both the cost of building a network link, as well as the cost of moving the network service from one location to another. Many other factors can influence the cost of a network link. Although some are specific to a particular type of network, for example, airport landing tax for air traffic network, the cost incurred from geographical distance is applicable to all network types, providing a universal metric for investigating and assessing network costs.

More specifically, we calculate the cost, \( C \), of a network in terms of the length of all its edges. To enable comparison of networks operating over different spatial scales, we normalize \( C \) against the cost of its Minimum Spanning Tree (MST; Nesetril et al., 2001). The MST is the subgraph of a network that links all the nodes by the shortest physical distance. Thus, our cost is defined relative to the minimum length of edges required to satisfy demand. In all cases, \( C \geq 1 \) and a network with a larger \( C \) is more costly than a network with a smaller \( C \).

\[
C = \frac{\sum_{i<j} d_{ij}}{\sum_{i<j} d^{\text{MST}}_{ij}} \tag{2}
\]

We further observed that networks dominated with short-range links have a relatively uniform node degree...
distribution, for example, each node is connected to a small but similar number of other nodes. On the other hand, a network with heterogeneous edge length usually exhibits heterogeneous node degree distributions, that is, there are a large number of nodes with a small degree and a small number of nodes with a high degree (Figure 6).

A high-degree node serves as a hub (e.g., a hub airport in an air traffic network) in network traffic and their existence can greatly reduce the number of steps for a network service to travel from one network node to another, and therefore improves network efficiency. Following the definition in Latora and Marchiori (2001), network efficiency, $E$, is defined and measured as the inverse of the shortest topological path length (Watts and Strogatz, 1998), $p_{ij}$, between all possible node pairs $i$ and $j$ for a network with $N$ number of nodes.

$$E = \frac{1}{N \times (N - 1)} \sum_{i \neq j}^{\text{all pairs}} \frac{1}{p_{ij}}$$

where a topological path is defined as an alternating sequence of nodes and edges (a sequence of adjacent nodes) that begins with $i$ and ends with $j$ but with no node visited more than once. The length of the path, $p_{ij}$, is defined as the number of edges in the sequence. Equation (3) can be reformulated as

$$E = \frac{2}{N \times (N - 1)} \sum_{i < j} \frac{1}{p_{ij}}$$

for a network where the weight of edges is symmetric, that is, $d_{ij} = d_{ji}$.

This efficiency measure assesses how well the nodes communicate over the network (i.e., how well a network mobilizes the flow of service, such as power, goods, or passengers, it delivers). A network with a larger $E$ is more efficient at communicating between network nodes than a network with a small $E$. The existence of a high-degree node in a network reduces $p_{ij}$, and hence increases $E$. In Figure 7, we demonstrate how $E$ increases with $K_{\text{max}}$, where $K_{\text{max}}$ is the degree of the most connected node in a network.

It is important to note that network efficiency can be defined by using other measures, for example, the geographical path length. The topological path length is employed in this article for the reasons: (1) it has been traditionally used to measure network efficiency (Latora et al., 2001); (2) for networks which are designed to perform efficiently, such as air traffic network and Internet, their efficiency measure is better captured by topological path length (For example, most air passengers prefer to travel with a direct flight and avoid ones with many intermediate stops, due to the increased travelling time and inconvenience caused by the latter) (Grosche, 2009); (3) using the topological path length in network efficiency calculation enables us to keep a balance of both topological and geographical influence in network design/evolution (as geographical length has been used to calculate network cost).

Table 1 lists the cost $C$ and efficiency $E$ for a number of infrastructure networks that we studied. Networks with a large $E$ (such as an air traffic network) usually have a large $C$. These networks have evolved (either by design or through natural evolution) in this way as reachability or accessibility between network nodes is the most important driver. This leads to many direct connections between network nodes or connection to an already highly connected network node. This increases $E$, but at the same time may introduce long-range connections and hence result in a large $C$.

On the other hand, a network with small $E$ (such as road and railway networks) has a cost index very close to that of MST. There are a few reasons for this. First, the consideration of construction cost naturally leads to short distance connections. Second, these networks are
Table 1

<table>
<thead>
<tr>
<th>Network</th>
<th>N</th>
<th>$\bar{k}$</th>
<th>E</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electricity transmission network (England &amp; Wales)</td>
<td>269</td>
<td>3.01</td>
<td>0.146</td>
<td>2.44</td>
</tr>
<tr>
<td>Railway network (England &amp; Wales)</td>
<td>3,547</td>
<td>2.92</td>
<td>0.038</td>
<td>2.86</td>
</tr>
<tr>
<td>Electricity distribution network (Denwick, England)</td>
<td>2,199</td>
<td>2.02</td>
<td>0.025</td>
<td>1.18</td>
</tr>
<tr>
<td>Gas network (England &amp; Wales)</td>
<td>506</td>
<td>2.15</td>
<td>0.059</td>
<td>1.87</td>
</tr>
<tr>
<td>Road network (West Midland England)</td>
<td>2,239</td>
<td>2.97</td>
<td>0.051</td>
<td>2.43</td>
</tr>
<tr>
<td>Air traffic (European)</td>
<td>1,369</td>
<td>4.66</td>
<td>0.385</td>
<td>22.67</td>
</tr>
</tbody>
</table>

fundamentally two-dimensional and the planarity characteristics of these networks largely prohibit cross-edge connection and this reduces the number of direct links between distance points (Gastner and Newman, 2006c). Third, crossing an intermediate node during travelling is much simpler than in an air traffic network, for example, and therefore a route with many intermediate connections is more acceptable in these networks. This leads to a network with a small $C$ and a small $E$.

These observations support the intuitive conclusions that cost and efficiency are two variables in the formation of networks. Depending on the purpose of networks, an infrastructure network may either be designed, or naturally evolve, to be dominated by one variable or may find a balance between two variables.

### 3 A SPATIAL NETWORK MODEL

Section 2 shows that network demand, network efficiency, and network cost are three major factors that characterize network architectures. The demand for a network service influences how network nodes are distributed in space. The linkages between network nodes are formed by balancing network cost and efficiency. We investigate the interplay of these three drivers of network evolution and their influence on network structure and properties.

We assume that a network is embedded in a predefined two-dimensional space $S$ of linear size 1. We use $\psi(a)$ to designate the demand density for a subarea $a$ in $S$ (see Figure 8a, for an example). The probability, $P(i \in a)$, for a subarea $a$ of $S$ to have a network node $i$ is defined as

$$P(i \in a) \sim \omega \cdot \psi(a)^\gamma \quad (4)$$

where $\omega$ is a normalization factor to ensure the sum of probability is 1, and $\gamma$ is a regulating parameter that determines how clustered or dispersed the nodes are or whether they need to be located close to or far away from centers of demand. Because network demand is usually heterogeneously distributed, a uniform or dispersed nodal distribution is generated when $\gamma$ is close or equal to 0 (Figure 8b), whereas a clustered distribution is generated when $\gamma \neq 0$. The greater the value of $|\gamma|$, the more clustered the nodes in the network. When $\gamma > 0$, the model produces node distributions that positively correlate to demand distribution (Figure 8c), with the greater the value of $\gamma$, the more likely that a node is allocated to an area of high demand. The smaller the
value of $\gamma$, the more the chance that a node is placed in an area of low demand. When $\gamma < 0$, the model produces nodal distributions that are negatively correlated with the distribution of demand (Figure 8d). The model does not fix or predetermine the maximum number of nodes in each square in Figure 8a, but distributes nodes to be proportional to the probability density expressed by $\psi(a)\gamma$ in Equation (4).

Once a node $i$ is introduced into $S$, it connects to existing nodes. We have two variables that regulate network linkages. One aims to reduce the cost of the connection, and one aims to improve the efficiency of the connection. An existing node $j$ acquires the connection from the new node $i$ if it balances the cost and efficiency requirements of the connection. We use the physical distance between $i$ and $j$, $d_{ij}$, to proxy the cost of the network connection of $i$, and use the degree of $j$, $k_j$, to proxy the efficiency of the connection. The probability, $\Pi((i, j))$, for building an edge $(i, j)$ between a newly introduced node $i$ and an existing node $j$ is

$$\Pi((i, j)) \sim \frac{f(k_j)}{g(d_{ij})}$$

where $f(k_j)$ and $g(d_{ij})$ are monotonic increasing functions of $k_j$ and $d_{ij}$, respectively. That is, the probability of building an edge $(i, j)$ is jointly determined by $k_j$ and $d_{ij}$, and it is proportional to $f(k_j)$ and is inversely proportional to $g(d_{ij})$. The larger $k_j$, or the smaller $d_{ij}$, the more likely that $i$ is connected to $j$.

In this research, we study networks when $f(k_j)$ and $g(d_{ij})$ take the forms of

$$f(k_j) = k_j^\beta$$

$$g(d_{ij}) = e^{d_{ij}^\alpha}$$

Scaling parameters $\alpha \in (0, 1]$ and $\beta \in [0, 1]$ have been introduced to shape the connection probability. $\alpha$ governs the preference of connecting to a proximal node (the smaller is $\alpha$, the more likely a short link will be established). $\beta$ governs the preference of connecting to high-degree nodes (the larger is $\beta$, the more likely a node with high degree will be connected). The ranges of $\alpha$ and $\beta$ allow the model to adequately generate the classes of networks that share properties of real infrastructure networks. The physical implication and effects of these ranges are further investigated and discussed in Section 4.

The proposed network model is a growth model. A network is built by starting with a few seed nodes (or an existing network), and new nodes are incrementally introduced into the network. We use $N$ as the total number of nodes that a network will have at the end of network generation, $n$ ($n \ll N$) as the number of seed nodes for network initialization, and $k$ ($k < n$) as the average node degree of the resultant network. The procedure for generating a spatial network is:

1. In space $S$, either:
   a. Initiate the model with an existing network of nodes and links; or
   b. Distribute $n$ seed nodes in $S$ with probability defined by Equation (4) and link these nodes with an MST;
2. Add a new node $i$ in $S$ with probability defined by Equation (4);
3. Connect $i$ to existing $k/2$ existing nodes, with probability determined by Equation (5). Node $i$ is discarded if it does not manage to establish desired connections;
4. Repeat (2) and (3), until the network has $N$ number of nodes.

By varying regulating parameters, our model is capable of generating a wide range of networks. Figures 9 and 10 show two example networks generated with the model. We use the demand distribution presented in Figure 8a for generating these networks, where triangle-shaped nodes represent seed nodes (which are of small number), round nodes represent newly added nodes, and the size of a node is plotted proportionally to its degree. A node is plotted in the same color of its demand area as shown in Figure 8a (online color version only). When $\beta$ is fixed, a large $\alpha$ removes the influence of distance and results in a network with links of mixed length or range. When we decrease
Fig. 10. An example network where $\alpha = 0.2$, $\beta = 1.0$, and $\gamma = 0.7$.

$\alpha$, newly created nodes are connected to nearby nodes and clusters start to emerge (Figure 9). This type of network has a small $C$ and a small $E$, exhibiting properties consistent with those of electricity or road networks.

When $\alpha$ is fixed, decreasing $\beta$ reduces the influence of node degree on network connection. Increasing $\beta$ increases the probability that a node will link to high-degree nodes, and hub nodes start to emerge (Figure 10). This type of network has a large $E$ and a large $C$, consistent with the properties of the Internet or air transport networks. We consider and evaluate this further in Section 5.

In Table A2, we summarize the approaches adopted by 17 spatial network models reported in the literature, as well as the infrastructure networks that were used to study. Key findings from this literature review are summarized below. Much research on infrastructure network generation takes a static approach and assumes the knowledge of the network size and physical layout of network nodes, and global optimization methods have been applied to derive networks that perform best in term of an objective function. These works provide some guideline in network design, but they fail to capture the dynamic growth of infrastructure networks. As a solution, some local optimization methods have been developed (Barthelemy and Flammini, 2006, 2008; Wilkinson et al., 2012b; Xie et al., 2007), or side-stepping spatial issues by using predefined nodal locations (Gastner and Newman, 2006b, c; Guimera and Amaral, 2004). Our model supports the distribution of network nodes that correlate to demand with different scales and forms, and both clustered and dispersed nodal distributions can be generated. Finally, previous research takes a much simplified form of $f(k_j)$, for example, $f(k_j) = 1$ as studied in Kaiser and Hilgetag (2004) and Waxman (1988), and $f(k_j) = k_j$ as investigated in Barthelemy (2003). By defining $f(k_j) = k_j^\beta$ and allowing $\beta \in [0, 1]$, our model considers both linear and sublinear preferential attachment, as well as their interplay with spatial constraints and demand distribution, which has not been sufficiently explored before.

4 NETWORK PROPERTY ANALYSIS

In this section, we analyze key topological features of networks simulated, so that a system boundary can be identified to show the classes of networks that can
be generated with the model. The results presented here are for networks with 1,000 nodes and average degree of 4, and are averaged over 100 realizations for each setting. We have applied our model to systems of different sizes and node degrees, and results exhibit similar trends and patterns to those reported in this section. All simulations described in this article were performed by implementing the model according to object-oriented paradigm, which has long been employed in prototyping computational models and numerical analysis of large engineering systems (Adeli and Yu, 1995; Jiang and Adeli, 2004; Kao and Adeli, 2002; Karim and Adeli, 1999a, b).

4.1 Node degree distribution

$\alpha$ and $\beta$ crucially affect the node degree distribution of the network. By varying $\alpha$ and $\beta$, networks of different connectivity are generated. As seen in Figure 11, a critical threshold of approximately $\alpha = 0.2$ is observed. When $\alpha > 0.2$, distance does not take effect and we obtain networks with pure preference attachment. These networks display a power law degree distribution for linear preferential attachment (i.e., when $\beta \to 1$), agreeing with results reported in Barabasi and Albert (1999). A stretched exponential degree distribution is generated for sublinear preferential attachment (i.e., when $\beta$ takes a value between 0 and 1). An exponential degree distribution is obtained when $\beta \to 0$. This is shown in Figure 12 and agrees with results reported in Krapivsky et al. (2000).

When $\alpha < 0.2$, space becomes a dominant driver and only nodes in close spatial proximity are likely to become connected. This constraint limits the choice of available connections and therefore reduces the chance of forming high-degree nodes. This is demonstrated in Figure 13, where networks with $\alpha = 0.02$ have been generated, and the result is compared against that of networks when $\alpha$ takes a larger value (0.4 in this case). As observed, instead of 57 (when $\alpha = 0.4$), the most connected node has degree 19 (i.e., $K_{\text{max}} = 19$) when $\alpha = 0.02$. 

Fig. 11. The types of networks generated with the model for different settings of $\alpha$ and $\beta$.

Fig. 12. When $\alpha$ is larger than approximately 0.2 (by way of example $\alpha = 0.04$ here), network development is dominated by preferential attachment to high-degree nodes. Generated networks have a power law distribution when $\beta \to 1$; the networks have an exponential distribution when $\beta \to 0$; and have stretched exponential when $\beta \in (0, 1)$.

Fig. 13. When $\alpha$ is small, that is, distance has influence and the degree of the most connected node is limited, $K_{\text{max}}$ is constrained relative to larger values (compared here with $\alpha = 0.4$).
4.2 Edge length distribution

Our analysis shows that $\alpha$ is the only model parameter that influences the edge length distribution. The model generates networks whose edge length follows an exponential distribution, which is a special case of Gamma distribution. Figure 14 shows how $\alpha$ serves as a scale parameter. When $\alpha$ is small, the model generates networks with edge length distribution skewed toward shorter lengths. When $\alpha$ is large, the model generates networks whose edge lengths have a broader distribution.

4.3 Clustering coefficient

The clustering coefficient $CC$ measures the degree of topological clustering of a typical node’s neighborhood. We follow the definition by Watts-Strogatz (Watts and Strogatz, 1998) and use the term here to refer to the average clustering coefficient for all nodes of a network. Our analysis shows that $CC$ is greatly influenced by $\alpha$, which is shown in Figure 15.

When $\alpha$ is large, $CC$ is small. This is because a large $\alpha$ encourages random linkages between nodes in the network, and as such the resultant networks show the $CC$ of a random network, which is in the range of $O(\frac{1}{N})$ (Watts and Strogatz, 1998). The smaller we make $\alpha$, the larger is the $CC$. This is because a small $\alpha$ encourages neighborhood linkages and therefore increases the likelihood that any two nodes with a common neighbor are themselves connected. This increases $CC$. When $\alpha$ is sufficiently small, the resultant networks exhibit the $CC$ of a Small World network (Watts and Strogatz, 1998), which is significantly higher than a random network (Erdős and Rényi, 1959). The $\beta$ is another parameter that influences $CC$. For a given $\alpha$, increasing $\beta$ obtains a network with a slightly higher $CC$. This agrees with the results obtained in Bollobás (2003) and Klemm and Eguiluz (2002), that is, degree-based preferential models generate networks with a higher $CC$ than that of a random network.

5 MODEL VALIDATION AND NETWORK GROWTH SIMULATION

In Section 5.1, we describe how we have validated our model by regenerating a number of networks and comparing their properties with the real-world counterparts. In Section 5.2, we apply the model to simulate the growth and future evolution of existing networks under scenarios of alternative infrastructure policies.

5.1 Model validation

The network model has been employed to simulate a number of real infrastructure networks that are listed in Table 1. The parameter values we calibrated and used for deriving each network are given in Table 2, and these were obtained by fitting the observed frequency and distribution of empirical data to Equations (4) and (5) using a Maximum likelihood method (Aldrich, 1997). Data in Table 2 show that most of these networks have their node distribution strongly correlated with demand (i.e., a large $\gamma$); exception is the gas network whose nodes have weak negative correlation with demand (i.e., $\gamma$ has a negative value close to 0). The air traffic network is more strongly driven by a need for efficiency, that is, $\beta$ is close to 1. Railway, power, and road networks are more driven by the need to reduce costs (i.e., $\alpha$ takes a value close to 0); these networks have a rather uniform node degree distribution (i.e., efficiency is not a major driver and $\beta = 0$).

Two sets of experiments have been performed. In the first set of experiments, the coordinates of network nodes were predefined at the location of existing railway
Table 2
Parameter values taken for generating networks

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.034 0.0 0.579</td>
<td>0.009 0.0 0.807</td>
<td>0.007 0.0 0.728</td>
<td>0.042 0.0 −0.103</td>
<td>0.008 0.0 0.716</td>
<td>0.287 0.9 0.991</td>
</tr>
</tbody>
</table>

stations. The model described in Section 3 was used only to generate the network links. In Figure 16, we present a simulated railway network, alongside its real-world counterpart. It shows that the model algorithm subsequently generates network structures that are similar to the real network. Furthermore, networks produced in this way are shown to have a good distribution fit, including spatial distribution of nodes and edge length distribution, as reported in Fu et al. (2014b).

In the second set of experiments, we generated networks without the knowledge of network nodes, that is, networks were built from scratch. Our experimental results show that this simulation was not able to generate visually similar networks. This is expected as the model does not capture geographical constraints such as mountains and floodplains which might impose limits on where to/not to allocate infrastructure facilities. However, the model can reproduce networks that share nontrivial properties with their real-world counterparts. This validation is useful in that as most networks of a same infrastructure sector possess and display particular properties (Barthelemy, 2011; Boas et al., 2009; Guimera and Amaral, 2004; Rui et al., 2013; Yook et al., 2002), the fact that the network model is able to capture and reproduce measurable characteristics and properties of real networks, presents evidence that we can obtain some unique insights about the fundamental mechanisms and underlying principles that drive that evolution and behavior.

We analyzed topological characteristics of simulated networks, including cluster coefficient \( \text{CC} \), average topological path length \( APL_T \), and network diameter \( D_T \), as defined and described in Boccaletti et al. (2006) and Watts and Strogatz (1998). We also evaluated the spatial properties of simulated networks. The following measures are used. The first is \( APL_G \), which is the average geographical path length and is calculated as

\[
APL_G = \frac{1}{N(N-1)} \sum_{i \neq j} P_{ij}^G
\]
where $p^G_{ij}$ is the shortest physical distance that connects node $i$ to node $j$ along a topological path. $APL_G$ is similar to $APL_T$ but different from it in that $APL_T$ measures the topological average path length of a network by considering the minimum steps along a path, and $APL_G$ measures the physical average path length of a network by looking at the shortest physical distance along a path.

The second measure is spatial diameter of a network, $D_G$, which is the largest in the set of all shortest geographical paths between any two network nodes, that is,

$$D_G = \text{Max}_{i \neq j} \left\{ p^G_{ij} \right\} \quad (9)$$

Table 3 presents the results of this analysis. The results are averaged over 100 realizations and compared against real networks, and standard deviation (SD) and coefficient of variation (CV) are calculated to indicate the dispersion of simulation results. Indicators, $C$ and $E$, which drive network generation, achieved good accuracy, for example, 95% of simulated results for $C$ having a value within 5% of the actual measure, and the same accuracy has also been achieved for $E$ in the air traffic network simulations where $E$ has a significant influence. Though other measures have greater dispersion, many fall within one SD from the simulated mean, and all fall within two SDs from the simulated mean.

For electrical transmission and air traffic networks, the difference between real and simulated networks is less than 15%, and the difference for rail and road networks is less than 20%. This indicates that using just demand, cost, and efficiency as drivers, the model is capable of reproducing networks that share spatial and topological characteristics with many example real-world counterparts.

However, as there are many other factors that can influence the structure and properties of an infrastructure, Table 3 highlights a couple of instances where particular network characteristics have a difference of over 30%. This can be for a number of reasons, for example, planning policy may limit infrastructure development in greenfield sites or floodplains, or promote infrastructure in a socially deprived area, and the crossing of infrastructure links is prohibited or uncommon for many systems. Also, a major node such as a shopping mall (and its associated roadway infrastructure) might develop in one location but draw demand from a much wider geographical area. Although cost, efficiency, and demand are always important, physical constraints and such policies inevitably influence detailed network design.

Though these considerations are important, they are included in a second-order process in our model, in that the drivers of population, efficiency, and cost also influence policy and therefore $\alpha$, $\beta$, and $\gamma$. In addition, the biggest differences between our simulated networks and the real networks are the $APL_T$ and $D_T$ of the rail network (network No. 2) and gas network (network No. 4). This is mainly due to the linear growth of such a network in its spatial domain, and networks possessing this property usually have larger $APL_T$ and $D_T$ when compared to other networks of same size and degree. As the network model does not capture linear spatial growth of networks, the resultant networks have a smaller $APL_T$ and $D_T$ when compared to that of their real counterparts.

The results and validation presented here use data sets sourced mainly (for reasons of practicality and accessibility) from the United Kingdom. However, the model is still valid for many other infrastructure networks because: (1) the validation is made on nontrivial real-world data sets and (2) the validation was made on multiple infrastructure sectors including road, electricity distribution, electricity transmission, rail, gas, and airline. Existing literature shows that networks of a same infrastructure sector share similar network structure and properties (Barthelemy, 2011; Boas et al., 2009; Guimera and Amaral, 2004; Yook et al., 2002). This provides confidence that the model is valid for application to infrastructure networks in other regions and countries. However, we do expect that second-order processes, such as planning constraints for floodplain or greenfield sites, will lead to different types of local variation.

5.2 Network growth simulation

Infrastructure development policy and regulation change over time, as a result of economic, social, environmental, and technological drivers. For example, an economic downturn could limit the amount of funds that a government would invest on infrastructure systems; hence reducing cost will be a priority during network development. On the other hand, in periods of rapid economic growth, the emphasis might be on network efficiency and less on cost. Similarly, locations of demand change through time. Altering the model parameters enables a range of different network evolution scenarios to be explored that reflect alternative policies. We use Great Britain’s railway network (Figure 16a) to demonstrate this.

In the first case, we assume a Business as Usual (BAU) scenario, that is, no significant changes in technology, economics, or policies, and network demand increases at a steady rate. We configure the model to take current Great Britain’s railway network (Figure 16a) as the initial network, and then evolve this with the same parameter set, that is, $\alpha = 0.009$, $\beta = 0.0$, and $\gamma = 0.807$ (as listed in Table 2). A simulated future network under this scenario is illustrated in Figure 17.
Table 3
Comparison of topological and spatial properties between simulated and real networks (R indicates a real network, S indicates a simulated network, and numbers in brackets are SD and CV for each calculated measure)

<table>
<thead>
<tr>
<th></th>
<th>CC</th>
<th>$D_T$</th>
<th>$APL_T$</th>
<th>$D_G$</th>
<th>$APL_G$</th>
<th>C</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>R</td>
<td>0.19</td>
<td>23</td>
<td>9.13</td>
<td>7.24</td>
<td>2.56</td>
<td>2.44</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>0.21</td>
<td>21.14</td>
<td>8.34</td>
<td>7.69</td>
<td>2.87</td>
<td>2.47</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.015, 0.071)</td>
<td>(0.67, 0.080)</td>
<td>(0.55, 0.071)</td>
<td>(0.22, 0.076)</td>
<td>(0.070, 0.028)</td>
<td>(0.0167, 0.098)</td>
</tr>
<tr>
<td>2</td>
<td>R</td>
<td>0.20</td>
<td>174</td>
<td>45.59</td>
<td>11.98</td>
<td>2.97</td>
<td>2.86</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>0.18</td>
<td>147.37</td>
<td>40.36</td>
<td>11.09</td>
<td>2.46</td>
<td>2.93</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.017, 0.094)</td>
<td>(17.67,0.12)</td>
<td>(4.36, 0.11)</td>
<td>(0.76, 0.068)</td>
<td>(0.31, 0.12)</td>
<td>(0.11, 0.037)</td>
</tr>
<tr>
<td>3</td>
<td>R</td>
<td>0.62</td>
<td>150</td>
<td>59.23</td>
<td>0.615</td>
<td>0.23</td>
<td>1.18</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>0.67</td>
<td>187.27</td>
<td>72.97</td>
<td>0.932</td>
<td>0.37</td>
<td>1.20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.041, 0.061)</td>
<td>(23.25,0.12)</td>
<td>(8.3, 0.11)</td>
<td>(0.18, 0.19)</td>
<td>(0.070, 0.19)</td>
<td>(0.040, 0.033)</td>
</tr>
<tr>
<td>4</td>
<td>R</td>
<td>0.03</td>
<td>76</td>
<td>27.45</td>
<td>10.31</td>
<td>3.56</td>
<td>1.87</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>0.04</td>
<td>53.39</td>
<td>18.91</td>
<td>9.27</td>
<td>3.27</td>
<td>1.90</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.006, 0.15)</td>
<td>(13.1,0.24)</td>
<td>(4.9, 0.25)</td>
<td>(0.76, 0.008)</td>
<td>(0.22, 0.067)</td>
<td>(0.062, 0.032)</td>
</tr>
<tr>
<td>5</td>
<td>R</td>
<td>0.14</td>
<td>67</td>
<td>26.81</td>
<td>1.74</td>
<td>0.57</td>
<td>2.43</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>0.13</td>
<td>73.24</td>
<td>29.53</td>
<td>2.62</td>
<td>0.69</td>
<td>2.49</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.009, 0.115)</td>
<td>(5.1, 0.069)</td>
<td>(2.19, 0.074)</td>
<td>(0.46, 0.17)</td>
<td>(0.075, 0.11)</td>
<td>(0.091, 0.036)</td>
</tr>
<tr>
<td>6</td>
<td>R</td>
<td>0.14</td>
<td>7</td>
<td>2.87</td>
<td>51.84</td>
<td>19.71</td>
<td>22.67</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>0.16</td>
<td>7.46</td>
<td>3.11</td>
<td>56.77</td>
<td>22.83</td>
<td>23.56</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.014, 0.087)</td>
<td>(0.51,0.068)</td>
<td>(0.22, 0.071)</td>
<td>(4.87, 0.085)</td>
<td>(2.13, 0.093)</td>
<td>(0.92, 0.038)</td>
</tr>
</tbody>
</table>

As we can see, the newly added network nodes have shorter linkages and therefore maintain the low cost.

In the second case, we assume an accelerated growth scenario. The investments in infrastructure are boosted due to rapid economic development that grows 50% more demand than the BAU scenario. There is a high expectation on network efficiency and restriction on network cost is relaxed. To generate a future network under this scenario, we increase $\alpha$ and $\beta$ such that $\alpha = 0.1$ (having a value greater than that of the current network so that to relax cost restriction), and $\beta = 1.0$ (having a value greater than that of the current network so that to improve network efficiency). We keep $\gamma = 0.807$ as we assume that future network maintains its correlation with demand distribution at its current form. Comparison of the original network (Figure 16a) and the new accelerated growth network (Figure 18) shows that many newly added network nodes possess long-range links and high-degree nodes start to emerge in the network. This results in a more efficient network but with the price of higher construction cost, this network evolution may be comparable to construction of high-speed railways.

The evolution scenarios can alter the network characteristics. Under the BAU scenario the topological characteristics of the network, such as degree distribution, are similar to the current values. Under the accelerated growth scenario, a higher network efficiency leads to more high-degree nodes, substantially

Fig. 17. A future network under a BAU scenario. Round nodes are original network nodes and triangle nodes are newly added nodes.
Fig. 18. A future network under an accelerated growth scenario. Round nodes are original network nodes and triangle nodes are newly added nodes, and the size of a node is proportional to its degree.

Fig. 19. Degree distribution of the railway network at present, under the BAU scenario, and under the accelerated growth scenario.

altering the degree distribution (Figure 19). It is well known that networks with these characteristics are more resilient to random attack, but more vulnerable to intentional attacks (Boccaletti et al., 2006; Cohen et al., 2000, 2001). Thus, alternative infrastructure network policies have altered the resilience of the networks.

6 CONCLUSIONS AND FUTURE RESEARCH

Urbanization and other socioeconomic drivers are increasing society’s reliance on infrastructure networks. Although this enables growth and provides important services, communities are more susceptible to disruptions of these systems, further aggravated by drivers such as intensifying global environmental change. Understanding how these networks evolve and what drives this evolution is an important first step to developing future scenarios of infrastructure provision and subsequently testing the implications of different scenarios for prosperity, resilience to hazards, or other factors.

In this article, we have quantitatively studied the properties of real-world infrastructure networks for six infrastructure sectors. Our results revealed a varied influence of network demand on the distribution of infrastructure facilities. We further demonstrated that cost and efficiency are two crucial variables that characterize the structure of these systems, and depending on the purpose of networks, an infrastructure network may either be designed or naturally evolve to bias toward one variable or may consider to strike a balance between these variables.

Based on this empirical analysis, we have developed a computational network model that can reproduce networks with the structures and properties of real infrastructure networks. In contrast to models that draw upon either small or simplified data sets, or are focusing on generating idealized network structures, our computational model enables considerable complexity and variations of infrastructure networks to be generated. We have demonstrated, through synthetic and real examples, how the model can be employed to grow/evolve an existing infrastructure network into different futures, through prioritizing and configuring the drivers of network development. The model is parameterized using large amounts of empirical evidence obtained from analysis of nontrivial real-world data sets. The validity of the model is verified by comparing key topological and spatial properties of simulated networks with real-world networks, which, we consider, is important but has rarely been undertaken in the literature.

Through empirical identification and validation, this article provides an approach to identifying, and formalizing in a model, key mechanisms that guide and govern infrastructure network development. The model can be used to explore cost-effective policies and network design choices that will enable possible future spatial development patterns that meet changes in populations, service provision while ensuring continued or even enhanced network resilience. Whereas the model considers demand, cost, and efficiency in infrastructure network development, decisions on infrastructure network development, decisions on infrastructure
provision can often be influenced by other issues such as planning constraints. These are more difficult drivers to include within a general model, as many of these policies are sector, or regionally, specific.

Further validation against historical time series might reveal drivers that relate to changes in infrastructure policy and regulation. However, attributing and isolating changes in infrastructure network form to a particular policy or regulatory change is challenging because the lag between policy and regulatory change often exceeds many decades.

Further development is exploring some of these issues considered above including, for example, research to extend the model to simulate evolution of electricity networks with respect to carbon reduction policies. By simulating a number of sequential time steps, the model can be used to explore the process, and necessary changes in policy, to achieve desired outcomes, such as carbon reduction, from infrastructure network transitions.

Finally, we recognize that infrastructure networks do not evolve independently and further research is to study the coevolution of interdependent networks, for example, the addition of a rail link might require electricity supply, or obviate the need for a road link.

**ACKNOWLEDGMENTS**

This research is supported by the UK EPSRC grants EP/I035781/1 and EP/K012398/1. Richard Dawson is funded by an EPSRC fellowship (EP/H003630/1). We are grateful to the journal editor and anonymous reviewers for their very useful and constructive comments that have greatly contributed to improving the quality of the article.

**REFERENCES**


An infrastructure network model

from M w 7.6 Padang earthquake of 30 September 2009, Natural Hazards, 63(2), 521–47.


APPENDIX A: NOTATIONS AND PREVIOUS NETWORK MODELS

Table A1

<table>
<thead>
<tr>
<th>Notations</th>
<th>Abbreviations</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>MST</td>
<td>–</td>
<td>Minimum spanning tree for a given set of network nodes</td>
</tr>
<tr>
<td>C</td>
<td>–</td>
<td>The cost of a network as defined in Equation (2)</td>
</tr>
<tr>
<td>E</td>
<td>–</td>
<td>The efficiency of a network as defined in Equation (3)</td>
</tr>
<tr>
<td>ki</td>
<td>–</td>
<td>The degree of a network node i</td>
</tr>
<tr>
<td>dij</td>
<td>–</td>
<td>The physical distance between nodes i and j</td>
</tr>
<tr>
<td>K</td>
<td>–</td>
<td>The average node degree of a network</td>
</tr>
<tr>
<td>kmax</td>
<td>–</td>
<td>The degree of the most connected node in a network</td>
</tr>
<tr>
<td>α</td>
<td>–</td>
<td>The regulating parameter that governs the preference of a low-cost connection</td>
</tr>
<tr>
<td>β</td>
<td>–</td>
<td>The regulating parameter that governs the preference of a high-efficiency connection</td>
</tr>
<tr>
<td>γ</td>
<td>–</td>
<td>The regulating parameter that determines the spatial distribution and layout of network nodes</td>
</tr>
<tr>
<td>D</td>
<td>–</td>
<td>The degree of a network node</td>
</tr>
<tr>
<td>CC</td>
<td>–</td>
<td>The normalized physical length of a network edge/link, as defined in Equation (1)</td>
</tr>
<tr>
<td>APLT</td>
<td>–</td>
<td>The global cluster coefficient of a network</td>
</tr>
<tr>
<td>Dk</td>
<td>–</td>
<td>The average topological path length of a network</td>
</tr>
<tr>
<td>Dk</td>
<td>–</td>
<td>The topological diameter of a network</td>
</tr>
<tr>
<td>Dk</td>
<td>–</td>
<td>The average geographical path length of a network</td>
</tr>
<tr>
<td>DG</td>
<td>–</td>
<td>The geographical diameter of a network</td>
</tr>
<tr>
<td>plij</td>
<td>–</td>
<td>The length of the shortest topological path between network nodes i and j</td>
</tr>
<tr>
<td>plij</td>
<td>–</td>
<td>The shortest geographical distance that connects node i to node j along a topological path</td>
</tr>
</tbody>
</table>
Table A2
Spatial network models studied in the literature (\(k_i\) is the degree of a network node \(i\), and \(d_{ij}\) is the physical distance between nodes \(i\) and \(j\))

<table>
<thead>
<tr>
<th>Model classification</th>
<th>Model description</th>
<th>Networks studied</th>
<th>Simulation and validation</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Static model with global optimization</td>
<td>Node distribution correlates to population Network linkages generated by minimizing the sum of network construction and travelling cost</td>
<td>Road networks/ Internet/airline networks</td>
<td>Evaluation on interstate transportation networks of U.S.</td>
<td>(Gastner and Newman, 2006a)</td>
</tr>
<tr>
<td>2 Static model with global optimization</td>
<td>Node positions predefined Network linkages generated by minimizing the sum of network path length</td>
<td>Road networks/ Internet/airline networks</td>
<td>Synthetic networks and analysis on edge length distribution</td>
<td>(Gastner and Newman, 2006c)</td>
</tr>
<tr>
<td>3 Static model with global optimization</td>
<td>Nodes randomly distributed Network linkages generated by minimizing the average travel cost from one node to another</td>
<td>Transportation networks, with a focus on air traffic networks</td>
<td>Synthetic networks under different settings</td>
<td>(Barthelemy and Flammini, 2006)</td>
</tr>
<tr>
<td>4 Growing model with local optimization</td>
<td>Random node distribution A node connects to an existing node which minimizes the construction cost and travel efficiency of new nodes</td>
<td>Infrastructure networks</td>
<td>Synthetic networks and analysis on average edge length and route factor</td>
<td>(Gastner and Newman, 2006b)</td>
</tr>
<tr>
<td>5 Growing model with local optimization</td>
<td>A new node adds to the networks if it maximizes/minimizes sum degree of its neighborhood A node connects to an existing node that has a minimum value of (d_{ij})</td>
<td>Urban road networks</td>
<td>Synthetic networks and analysis on topological path length, geographical path length, topological efficiency, geographical efficiency, betweenness centrality, and other urban street-related properties</td>
<td>(Rui et al., 2013)</td>
</tr>
</tbody>
</table>

Continued
<table>
<thead>
<tr>
<th>Model classification</th>
<th>Model description</th>
<th>Networks studied</th>
<th>Simulation and validation</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>6 Growing model with local optimization</td>
<td>Nodes distributed randomly A node connects to an existing node that has a minimum value of $d_{ij}$</td>
<td>Transportation networks</td>
<td>Synthetic networks and analysis on degree distribution, topology distance distribution, edge length distribution, and geographical distance distribution</td>
<td>(Xie et al., 2007)</td>
</tr>
<tr>
<td>7 Growing model with local optimization</td>
<td>Nodes randomly and uniformly distributed A node connects to an existing node that has minimum value of $d_{ij}$</td>
<td>Urban road networks</td>
<td>Evaluation against empirical data in terms of total edge length, distribution of area, and perimeter of cells</td>
<td>(Barthelemy and Flammini, 2008)</td>
</tr>
<tr>
<td>8 Growing model with probability method</td>
<td>Node distribution is linearly correlated with population density Linkages established with a probability $p(i \to j) \propto \frac{k_i}{d_{ij}}^{\alpha}$</td>
<td>With Internet in mind but can apply to other infrastructure networks</td>
<td>Synthetic networks and analysis on degree and edge length distribution Evaluation on Internet networks</td>
<td>(Yook et al., 2002)</td>
</tr>
<tr>
<td>9 Static model with probability method</td>
<td>Nodes randomly distributed with uniform density Links formed by considering physical distance and node degree</td>
<td>General spatial networks</td>
<td>Synthetic networks and analysis on degree distribution, cluster coefficient, assortativity, and network construction cost</td>
<td>(Barthelemy, 2003)</td>
</tr>
<tr>
<td>10 Growing model with probability method</td>
<td>Assume the availability of nodes and their locations</td>
<td>Air traffic networks</td>
<td>Synthetic networks and analysis on degree and betweenness distribution Evaluation on world airports and North America airports</td>
<td>(Guimera and Amaral, 2004)</td>
</tr>
<tr>
<td>Model classification</td>
<td>Model description</td>
<td>Networks studied</td>
<td>Simulation and validation</td>
<td>Reference</td>
</tr>
<tr>
<td>----------------------</td>
<td>-------------------</td>
<td>------------------</td>
<td>---------------------------</td>
<td>-----------</td>
</tr>
<tr>
<td>11 Growing model with probability method</td>
<td>Nodes distributed in space by following a predefined form. A new node connects to an existing node with a probability that is proportional to the sum degree of this node’s neighboring nodes.</td>
<td>Air traffic networks</td>
<td>Evaluation on European air traffic network on degree and spatial degree distribution</td>
<td>(Wilkinson et al., 2012b)</td>
</tr>
<tr>
<td>12 Growing model with probability method</td>
<td>Random node distribution. Linkages established with a probability that is exponentially inverse to physical distance.</td>
<td>General spatial networks</td>
<td>Synthetic networks and analysis on average path length and cluster coefficient. Evaluation on a German highway network, and a few biological networks.</td>
<td>(Kaiser and Hilgetag, 2004)</td>
</tr>
<tr>
<td>13 Static model with probability method</td>
<td>Nodes distributed randomly. Linkage probability of two nodes is aggregation of social distances of multiple dimensions/aspects.</td>
<td>Social networks</td>
<td>Evaluation on Web of Trust</td>
<td>(Boguna et al., 2004)</td>
</tr>
<tr>
<td>14 Static model with probability method</td>
<td>Nodes predefined either randomly or in clusters. Linkage probability of two nodes is $P(i \rightarrow j) \propto e^{-\alpha d_{ij}}$.</td>
<td>General spatial networks</td>
<td>Synthetic networks and analysis on assortativity, clustering, and fragmentation. Evaluation on two Swedish swine transport networks.</td>
<td>(Lennartsson et al., 2012)</td>
</tr>
<tr>
<td>15 Static model with probability method</td>
<td>Assume the availability of nodes and their locations. Linkage probability of two nodes is $P_{ij} = e^{-\alpha d_{ij}}$.</td>
<td>Highway networks</td>
<td>Evaluation on four highway networks.</td>
<td>(Boas et al., 2009)</td>
</tr>
<tr>
<td>16 Static model with probability method</td>
<td>Nodes uniformly distributed. A node connects to its nearest neighbor within a radius $r$ which has a power law distribution.</td>
<td>Epidemics networks/disease transmission networks</td>
<td>Synthetic networks and analysis on percolation thresholds under different settings.</td>
<td>(Warren et al., 2002)</td>
</tr>
<tr>
<td>17 Static model with probability method</td>
<td>Assume the availability of nodes and their linkages in form of a 2-D lattice. A network edge is obtained with the probability equal to a predefined edge density.</td>
<td>Infrastructure networks</td>
<td>Evaluation on real infrastructure networks.</td>
<td>(Dueñas-Osorio, 2005)</td>
</tr>
</tbody>
</table>