Advances in Complex Systems, Vol. 12, No. 1 (2009) 45–71

World Scientific
www.worldscientific.com

TOPOLOGICAL ANALYSIS OF SPECIFIC SPATIAL COMPLEX NETWORKS

JUN WANG* and GREGORY PROVAN

Department of Computer Science,
University College Cork, Ireland
**jw8@cs.ucc.ie*

Received 17 December 2007
Revised 20 July 2008

Based on analyses of specific spatial networks, we compare the accuracy of three models in capturing topologies of two types of spatial networks: electronic circuits and brain networks. The models analyzed are an optimization model trading off multiple-objective constraints, an extended preferential attachment model with spatial constraints, and the generalized random graph model. First, we find that the optimization model and the spatial preferential attachment model can generate similar topological structures under appropriate parameters. Second, our experiments surprisingly show that the simple generalized random graph model outperforms the two proposed models. Third, we find that a series of spatial networks under global optimization of wire length, including the electronic circuits, brain networks, neuronal networks and transportation networks, have high s-metric values close to those of the corresponding generalized random graph models. These s-metric observations explain why the generalized random graph model can match the electronic circuits and the brain networks well from a probabilistic viewpoint, and distinguish their structures from self-organized spatial networks, such as the Internet.

Keywords: Complex networks; topology; modeling.

1. Introduction

Research on complex networks has exploded in recent years, to cover the observation, interpretation, and modeling of such networks, with a current emphasis on modeling. The *small-world graph* model (SWG), introduced by Watts and Strogatz [77], and the *preferential attachment* model (PA), proposed by Barabasi and Albert [7], are two of the best-known models of complex networks. Since these two models were introduced, the expectations for understanding of complex networks have soared, and reports and studies have appeared covering a range of domains, such as the WWW, the Internet, citation networks, social networks, metabolic networks, protein interaction networks, genetic networks, brain networks, transportation networks, power grids, electronic circuits, software component networks, food webs, and sexual networks [10, 13, 59].

The PA model and its variants have been widely used to describe various real-world complex networks [10], and preferential attachment is often assumed to be the underlying mechanism at work when power law degree distributions are observed in data sets [7, 3, 79, 38, 8, 2, 49]. However, some researchers have suggested that these models are not fundamental, and/or that they do not have the subtle properties necessary for describing empirical observations in real networks; these researchers proposed optimization as an alternate mechanism which gives rise to power laws in degree distributions [25, 47, 42, 74, 15, 78]. Actually, power law degree distributions can be generated by a variety of mechanisms, and in many cases model parameters can be tuned such that multiple models of widely varying mechanisms fit the motivating real network [10, 47]. This situation leads to a long-standing theoretical controversy are universally applicable underlying mechanisms of complex networks [58].

In order to further clarify the above problem, we compare the modeling fidelity of two models on two different types of spatial networks: electronic circuits and human brain anatomical networks. The models are an optimization model that trades off multiple-objective constraints, and an extended preferential attachment model with spatial constraints. To evaluate and validate the two types of models, we use as the null model the generalized random graph model with the same degree sequence as real circuits [61, 19, 75].

This paper makes three main contributions. First, we find that, under appropriate parameters, an optimization model trading off multiple constraints (wire length and characteristic path length) and a spatial preferential attachment model can generate graphs with similar structure. Our results show that the models can be consistent with each other under certain conditions.

Second, our experiments using machine learning and subgraph frequency analysis show that the generalized random graph model outperforms both the optimization model and the spatial preferential attachment model, and best matches real networks of both the electronic circuits and the brain network.

Third, we find that a series of spatial networks under global optimization of wire length, including the electronic circuits, brain networks, neuronal networks, and transportation networks, have high s-metric values close to those of the corresponding GRG models. These observations explain why the GRG model can match the electronic circuits and the brain network very well from a different angle, and distinguish their structures from self-organized networks, such as the Internet.

We organize the remainder of the article as follows. Section 2 analyzes and summarizes recent work on various preferential attachment models and optimization models of complex networks. Sections 3 and 4 elaborate the proposed models for electronic circuits and the brain anatomical network, and corresponding experimental analysis, respectively. Section 5 summarizes our contributions and discusses proposed future research work.

2. Related Work

About half a century ago, a series of exchanges between Mandelbrot and Simon published in the journal *Information and Control* launched a still-unresolved public

controversy about whether preferential attachment or optimization is the more appropriate explanation for the emergence of power laws [25, 57, 60].

In 1999 the PA model, proposed by Barabasi and Albert, sparked a flurry of research activity on complex networks. In their model, nodes are added sequentially and attach to existing nodes with probability proportional to the existing node's current degree. The PA model and its variants have been used extensively to explain many classes of networks [10, 13, 59]. Spatial restrictions are integrated with the preferential attachment to describe various transportation and technology networks in which spatial constraints are important, such as airport networks, power grids, road networks, electronic circuits, and the Internet [5, 11, 31, 32, 79].

Fabrikant *et al.* proposed a plausible explanation of the power law degree distributions observed in the graphs arising in the Internet topology, based on a toy model of Internet growth in which two objectives are optimized simultaneously: "last mile" connection costs, and transmission delays measured in hops [27]. Their results seem to suggest that power laws tend to arise as a result of complex, multiple-objective optimization. Fabrikant's model can only generate the tree structure, which obviously deviates from real Internet topology. Li *et al.* provided a healthy jumping-off point for a richer, more refined view of various models of complex networks and their relation to real-world systems [47]. They invalidated the PA model by an examination of the Internet's topology at the router level, and clarified that there are many different types of models that yield power law degree distributions, such that graphs generated by these models differ with respect to other topological properties, such as the s-metric [47, 46]. They designed a heuristically optimal toy model which is consistent with real design considerations in a single Internet service provider (ISP) network, and showed that the designed model provides much higher communication performance than networks generated by the PA model. They proved that the structure of the designed toy model is highly unlikely to be present in networks generated by the generalized random graph model with the same degree distribution [47, 46, 19]. Li *et al.* tied the shaping process of the Internet router topology to a mechanism which typically involves optimizing functional objectives of the system as a whole, subject to constraints on their components [47, 46, 12]. Although, Li *et al.* illustrated the technological and economic constraints in the design of a single ISP network, they only qualitatively analyzed the impact of constraints on network growth and did not formulate the corresponding constraints. The toy model fitting real-world networks was created manually, instead of through a formal optimization process; in other words, the major part of Li's paper focuses on invalidating the PA model and does not provide an alternative formulated optimization model which can generate network structure close to the real-world ISP network.

Some other researchers also discussed the optimization mechanism in technology and transportation systems. Gastner and Newman proposed a model that can adjust the network structure by introducing a preference for measuring distance in terms of the geometric length or the number of hops [30]. This work was limited to qualitative analysis and did not verify the model with data from real systems.

Gastner and Newman presented another model for simulating the Boston commuter system, which is very similar to Fabrikant's model [29]. A fundamental difference between Gastner and Newman's model and Fabrikant's model is that geometric positions of nodes are predecided in Gastner and Newman's model, but randomly selected in Fabrikant's model; further, Gastner and Newman focus on the impact of the geographic layout instead of the optimization process itself.

In addition, Mathias *et al.* showed that the small-world topology arises from a tradeoff between maximal connectivity and minimal wiring [53]. Cancho *et al.* used an evolutionary algorithm involving minimization of link density and average distance, and generated different types of networks, including sparse networks with power law degree distributions [36].

Souza *et al.* introduced a family of one-dimensional geometric growth models, constructed iteratively by locally optimizing the tradeoffs between two competing metrics, and showed that this family is equivalent to a family of preferential attachment random graph models with upper cutoffs [25, 9]. This is the first explanation of how preferential attachment can arise from a more basic underlying optimization mechanism of local competition. Their optimization model involves only a one-dimensional sequence of nodes without spatial positions, and the equivalent tempered preferential attachment model does not use the spatial positions of nodes. The authors presented a collection of empirical observations from social, biological, physical, and technological networks, for which such degree distributions give excellent fits. But obviously it is not convincing enough to validate a model by only fitting a degree distribution.

In the electronic engineering domain, a number of efforts have been made to automatically generate synthetic benchmark circuits, since there currently exists a shortage of high-quality public domain benchmark circuits that can be used to test the next generation of computer-aided design (CAD) algorithms [73, 45]. The realism of the generated circuits is assessed by comparing properties of real circuits and generated clones of the real circuit after placement and routing [45, 73]. In contrast, our proposed circuit models (see Sec. 3) focus mainly on topological and organizational principles of circuits instead of concrete physical parameters, and can be used for a wider variety of applications, including benchmark circuits for diagnosis [63].

3. Models of Spatial Complex Networks

Many complex networks occupy some physical space, such that their nodes occupy a precise position in two- or three-dimensional Euclidean space, and their edges are real physical connections [10]. Some important examples are the Internet, electronic circuits, neural and brain networks, electric power grids, transportation networks, ant networks of galleries [10]. It is not surprising that the topology of spatial networks is strongly constrained by their geographical embedding. In this paper, we focus on a particular class of spatial networks in which economical wiring costs

are counterbalanced by a need for global minimization of processing steps across the network. Typical examples are electronic circuits and brain networks. This section describes an optimization model under multiple-objective constraints and an extended preferential attachment model under spatial constraints. We also introduce the generalized random graph model, which is used to validate the proposed models.

3.1. *Optimization model under multiple design constraints*

This subsection introduces two typical spatial networks, electronic circuits and brain networks, and shows that they possess some similar organizational principles and topological properties, which we model.

Electronic circuits can be viewed as networks in which nodes (or vertices) are electronic components (e.g. logic gates in digital circuits and resistors, capacitors, diodes, and so on in analogic circuits), and edges (or connections) are wires in a broad sense [11]. In circuit design, wire length has been treated as the prime parameter for performance evaluation, since it has a direct impact on several design parameters, such as delay, area occupancy, power dissipation, channel density, and routability [71, 72]. Another driving force underlying circuit design is timing. Achieving timing closure, i.e. the global constraint on timing, has been a major headache for design engineers for over a decade [23]. Many design cost metrics can be treated as technological parameters that can be optimized by trading off delay and wire length [23]. If the signal is transmitted from one component to another, the logic depth is the number of components in the path of signal transmission, and the total delay on a combinational path is proportional to the logic depth of the path [23]. The average logic depth can be approximately simplified as the characteristic path length of the graph. So we can simplify the optimization of circuit design as trading off the total wiring cost and the characteristic path length.

Most structural analyses of brain networks have been carried out on datasets describing the large-scale connection patterns of the cerebral cortex regions [70]. Research showed that the optimal design of the brain anatomical network might have important implications for understanding how functional brain states emerge from their structural underpinnings [35]. In the brain, energy is consumed for establishing fiber tracts between areas, and for propagating action potentials over these fibers. Thus, the total cost of all wires should be kept as low as possible [39]. Although the exact origin of the wiring cost is not completely known, the farther apart two neurons are, the more costly the connection between them is. The wiring cost can therefore be expressed as a function of distance between neurons, and consequently minimized [14]. Minimizing the average number of processing steps (characteristic path length) — i.e. reducing the number of intermediate transmission steps in neural integration pathways — has several functional advantages [39]. Therefore, it is plausible that neural systems are adapted to more than just one design constraint, and that their observed organization is the outcome

of an optimization of multiple parameters, which may be partly opposed to each other. The organization of neural networks is shaped by tradeoffs from multiple constraints, among them the total wiring cost and the characteristic path length [41].

Based on the above analysis, we propose a plausible optimization model with computing constraints. Because each node in an electronic circuit or a brain network has a physical position, we put the electronic circuit's components or the brain's cortex regions on a geometric substrate. We assign each component or cortex region as a vertex, and uniformly put it on a two-dimensional square grid. The wire length between adjacent nodes i and j is defined using the Euclidean or Manhattan distance WL_{ij} on the grid. The wiring cost between nodes i and j can be defined as a function, $f(\mathrm{WL}_{ij})$. The average wiring cost WC is defined as $\mathrm{WC} = \frac{1}{m}\sum_{\mathrm{edge}(i,j)} f(\mathrm{WL}_{ij})$, where m is the number of edges. The average processing steps can be defined as the characteristic path length of the graph: $\mathrm{CL} = \frac{2}{n(n-1)}\sum_{i<j}\mathrm{SL}_{ij}$, where n is the number of nodes, and SL_{ij} is the length of the shortest path between nodes i and j. The objective function F for the optimization is formulated as follows: $F = \lambda\mathrm{CL} + (1-\lambda)\mathrm{WC}$, where the parameter λ is varied depending on the relative importance of the minimization of WC and CL.

According to principles of design and implementation of electronic circuits, wiring cost can be represented by wire length [71, 72], so the objective function F can be simplified as

$$F = \frac{2\lambda}{n(n-1)}\sum_{i<j}\mathrm{SL}_{ij} + \frac{1-\lambda}{m}\sum_{\mathrm{edge}(i,j)}\mathrm{WL}_{ij}.$$

Starting from a connected random network, we use simulated annealing to search for the minimum cost of the objective function [44, 62]. In each annealing rearrangement step, an edge is randomly selected and rewired. In rearrangements, duplicated edges and self-loops are not allowed, to ensure that no node will be disconnected or isolated. The simulated annealing is expensive computationally. It is shown for arbitrary graphs that a degenerate form of the basic annealing algorithm (obtained by letting "temperature" be a suitably chosen constant) produces matchings with nearly maximum cardinality in polynomial average time [67]. So we fixed the temperature of simulated annealing at zero in order to speed up the optimization process, and the algorithm stops when the modifications are not accepted 5000 times in a row. Even for this simplified method, it takes hours to generate a graph corresponding to a network with several hundred nodes.

3.2. *Spatial preferential attachment model*

Cancho *et al.* found *small-world graph* patterns for a small collection of electronic circuits, and observed power law tails with cutoffs in degree distributions [11]. Existing analysis has conjectured that the cutoffs in power law degree distributions

might result from the presence of constraints limiting the number of links when connections are costly [5], and this has been confirmed by further studies on diverse networks such as the Internet, power grids, and transport networks [5, 11, 31, 32, 79, 35]. Recent research on circuit placement showed that the wire length of real circuits exhibits power law distributions [71, 18, 23]. Similarly, new studies on human brain networks have showed that corresponding degree distributions and anatomical distance distributions can be well fitted by an exponentially truncated power law [35].

Based on the evidence of power laws in degree distribution and wire length, and the disadvantage of long-range links, we consider a PA model constrained by the spatial layout, i.e. the *spatial preferential attachment* model (SPA). In the SPA model, the node position is chosen randomly in a two-dimensional square space with uniform density. Connections of a new node i with each existing node j are established with probability $\mathrm{P}(i,j) \propto d_j \mathrm{WL}_{ij}^{-\alpha}$, where WL_{ij} is the spatial (Euclidean or Manhattan) distance between the node positions, d_j is the degree of the node j, and $\alpha \geq 0$ is a tunable parameter used to adjust the spatial constraint and shape the connection probability in the preferential attachment process.

When $\alpha = 0$, the model corresponds to the basic PA model. The complexity will increase with the increase of α, but in general the α fitting real circuits is small, and the graph corresponding to a network with thousands of nodes can be generated in a few minutes.

3.3. *Generalized random graph model*

The Erdos and Renyi random graph model (ER) [26] can be extended in a variety of ways to make random graphs better represent real networks. In particular, one of the simplest properties to include is a prescribed degree sequence. The random graphs with an arbitrary degree distribution are called the generalized random graphs (GRGs). At first a modified *configuration model* is used to generate connected and simple graphs (with no self-loops or multiedges) [61, 59, 10]. But this matching or stub-pair algorithm cannot generate graphs uniformly and introduces a bias which becomes more prominent on networks in which degree sequences are heavy-tailed. Recently a Markov chain Monte Carlo (switching) algorithm has been used to implement the GRG model for network analysis tasks such as motif detection [43, 37, 75]. In addition, Chung *et al.* proposed a stochastic implementation of the generalized random graph model, in which the connection between the nodes i and j is chosen independently with probability p_{ij}, which is proportional to the product of the degree of i and j [19]. This approach is convenient for theoretical analysis, since rigorous proofs for a random graph with extract degree sequences are rather complicated and usually require additional "smoothing" conditions because of the dependency among the edges [19]. It is believed that this stochastic implementation and switching implementations are "basically asymptotically equivalent, subject to bounding error estimates" [1, 47]. In practice the stochastic approach

was found to lead to statistical variance [51], so we select the switching approach in our experiments in this paper [75].

The algorithm of the GRG model is simpler and computationally more efficient than for the other two models. A graph corresponding to a network with thousands of nodes can be generated in a few seconds by the GRG model in an ordinary PC.

4. Experimental Analysis

In this section, we first apply the optimization model, SPA model and GRG model for circuits in the ISCAS-85 benchmark, and then compare global and local properties of these models with real circuits. We also analyze the human brain networks in a similar fashion, and discuss the corresponding experimental results.

4.1. *Electronic circuit networks*

Over the years, there have been many attempts to create and use neutral benchmark circuits for evaluating different tools and algorithms. These benchmark sets are surrogate circuits chosen to represent the kinds of problems a tool will encounter in real use [34, 33]. The widely accepted ISCAS-85 combinatorial benchmark suite has been in use ever since being introduced at the International Symposium of Circuits and Systems in 1985. The ISCAS-85 benchmark circuits are presented in netlists of fundamental logic gates, which provide a standard, nonhierarchical representation specifying both network topology and functionality (in terms of the functionality of primitive gates like NOT, AND, OR) [33].

4.1.1. *Degree distribution*

The most basic topological characterization of a complex network can be obtained in terms of the degree distribution. Figure 1 shows cumulative degree distributions for the full suite of ISCAS-85 benchmark circuits on the log–log scale. We can see that most circuits exhibit long tails with cutoffs in degree distributions.

At first, we select a typical circuit, C432, in the ISCAS-85 benchmark suite as an example, and implement the above three models with a number of nodes and edges identical to that of C432. In the following section, we will discuss the global properties of the implemented models and compare them with the circuit C432. Since all ISCAS-85 benchmark circuits have similar degree distributions, the analytical results for C432 hold for other ISCAS-85 circuits.

Figure 2 shows the cumulative degree distributions of the SPA model for C432 at various values of α. By continually increasing α of the SPA model, the modularity of the generated graph will keep increasing, and fewer long-range links will appear. Finally, the degree distribution will degrade from the power law distribution to the exponential or normal distribution.

The optimization process is looking for a solution that minimizes the above objective function at an appropriate value of λ. Figure 3 shows the cumulative

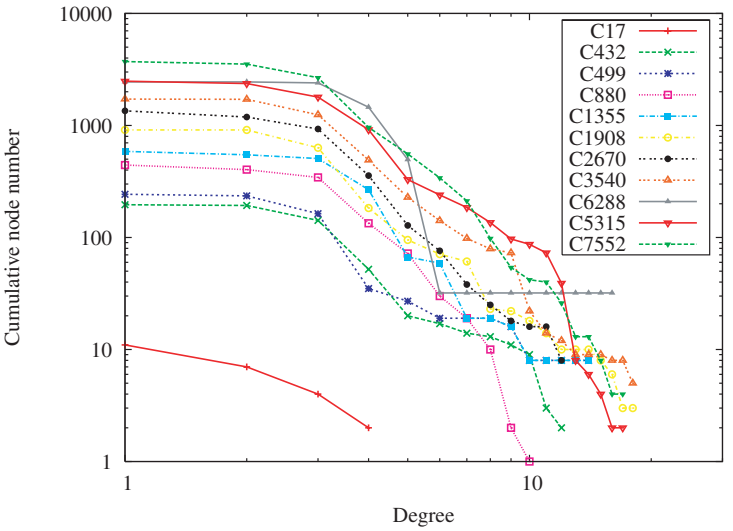

Fig. 1. Cumulative degree distributions of ISCAS-85 benchmark circuits on the log–log scale.

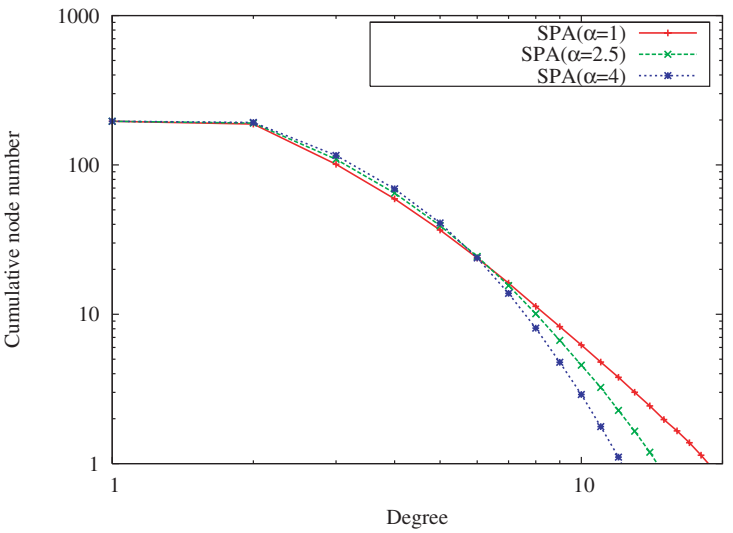

Fig. 2. Cumulative degree distributions of the SPA model having the same number of nodes and edges of C432 at various values of α. All three curves are averaged over 1000 graphs, respectively.

degree distributions of the optimization model for C432 at various values of λ, demonstrating that the degree distribution of the optimization model with $\lambda = 0.2$ exhibits a power law tail with cutoff. From Fig. 4 we can further find that the corresponding wire length distribution approximately displays a power law distribution as well. This result is consistent with placement experiments on real

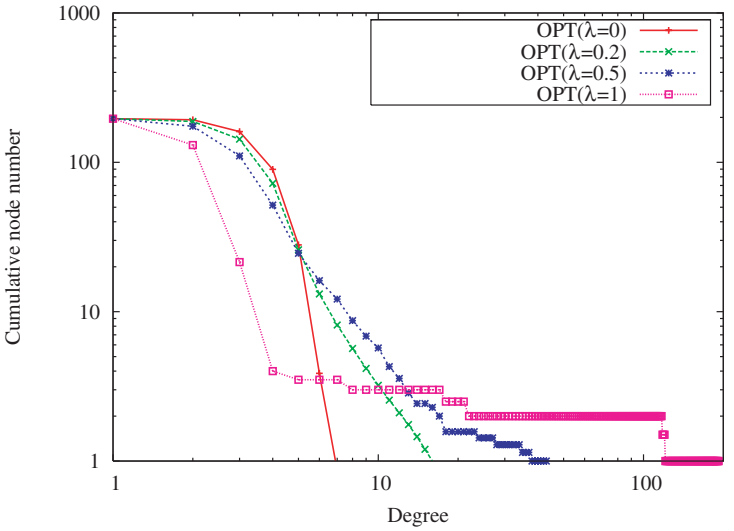

Fig. 3. Cumulative degree distributions of the optimization model for C432 at various values of λ. All five curves are averaged over 1000 graphs.

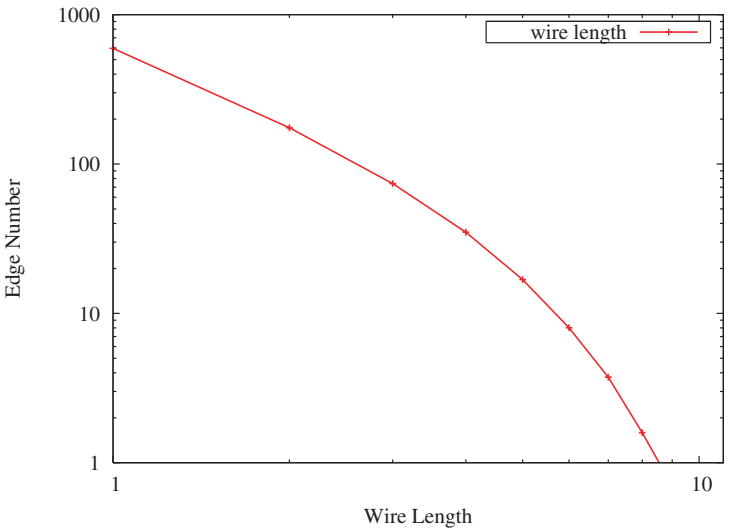

Fig. 4. Wire length distribution of the optimization model having the same number of nodes and edges as C432 with $\lambda = 0.2$.

circuits [71, 18, 23]. Obviously, according to its definition, the SPA model can also generate the wire length distribution with a power law tail.

We can tune the α parameter in the SPA model, or the λ parameter in the optimization model, to fit real circuits. Figure 5 shows that both models can fit the degree distribution of C432 with appropriate parameters. Since most ISCAS-85

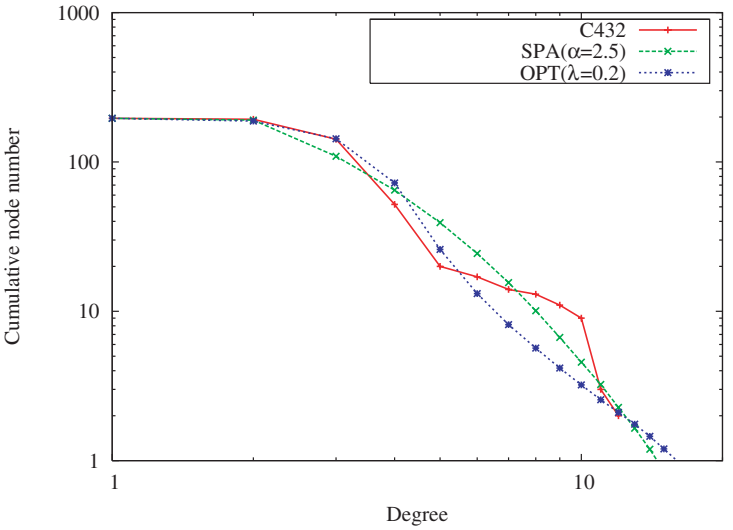

Fig. 5. Cumulative degree distribution of the C432 circuit, and cumulative degree distributions of graphs generated by the SPA model and the optimization model (averaged over 1000 graphs, respectively).

circuits display degree distributions following a power law with cutoff, we can also tune parameters and match other ISCAS-85 circuits using the SPA model and the optimization model in terms of degree distributions.

4.1.2. *s-metric and characteristic path length*

In addition to degree distributions, we compare some other global properties of real circuits and graphs generated by the SPA model and the optimization model. Table 1 shows that the characteristic path lengths and s-metric values of these models are very close to those of the real circuits. The GRG model maintains the degree sequence of real circuits [75].

Shortest paths play an important role in transport and communication within a network. A measure of the typical separation in the network is given by the characteristic path length, defined as the mean shortest length over all pairs of nodes [10]. The characteristic path length is also an important factor in the optimization model proposed in Subsec. 3.1.

The s-metric is a summary statistic of node interconnectivity, and is linearly related to the assortativity coefficient: assortative (disassortative) networks are those where nodes with similar (dissimilar) degrees tend to be tightly interconnected [47]. The s-metric was successfully applied to validate or invalidate various models for the Internet topology at the router level [47]. The s-metric of the graph g is defined as $s(g) = \sum_{\text{edge}(i,j)} d_i d_j$, where (i, j) is the edges in the graph, and d_i and d_j are the degrees of the nodes i and j, respectively.

Table 1. Global properties of the real circuits, the GRG model, the SPA model, and the optimization model (OPT). All values of the three models are averaged over 1000 graphs, respectively.

Model	Characteristic path length	s-metric
C17	2.38	84
SPA ($\alpha = 3$)	2.44 ± 0.03	85.4 ± 4.88
OPT ($\lambda = 0.3$)	2.37 ± 0.04	86.8 ± 6.1
GRG	2.52 ± 0.13	85.86 ± 2.69
C432	4.53	6986
SPA ($\alpha = 2.5$)	4.52 ± 0.13	7349.05 ± 523.84
OPT ($\lambda = 0.2$)	4.5 ± 0.07	7097.8 ± 263.47
GRG	4.33 ± 0.05	6875.99 ± 143.46
C499	4.65	9848
SPA ($\alpha = 2$)	4.56 ± 0.1	9162.6 ± 654.98
OPT ($\lambda = 0.22$)	4.59 ± 0.07	9025.9 ± 371.33
GRG	4.4 ± 0.06	10491.57 ± 306.78

Because the implementation of the selected optimization model is very time-consuming, we do not present statistics on all ISCAS-85 circuits. The ISCAS-85 circuits are created by standard cell libraries and follow similar design principles, so it is reasonable to generalize the current experimental results to the other ISCAS-85 circuits.

In the above experiments, the three models display similar global properties. Recently, new systematic measures of a complex network's local structure were introduced and successfully applied to evaluate and validate models of protein–protein interaction networks [54, 65, 64]. In contrast to focusing on motifs, and the corresponding significance profile (SP) [56, 55], Przulj *et al.* argued that it is as important to understand the organization of infrequently observed subgraphs as the frequently observed ones, and new bottom-up approaches based on subgraph frequency statistics have been successfully applied to analyze complex networks [54, 65, 64].

4.1.3. *Machine learning approach*

Middendorf *et al.* exploited discriminative classification techniques recently developed in machine learning to classify a given real protein interaction network (as one of many proposed network models) by enumerating local substructures [54]. They presented a predictive approach using labeled graphs of known growth models as training data for a discriminative classifier. This classifier is a generalized decision tree called an alternating decision tree (ADT) [28], which uses the Adaboost algorithm [54]. Presented with a new graph of interest, it can reliably and robustly predict the growth mechanism that gave rise to that graph [54].

We use the same classifier to evaluate the models for real circuits. The classifier quantifies the topology of a network by conducting an exhaustive subgraph census up to a given subgraph size, and tries to identify network mechanisms by using the

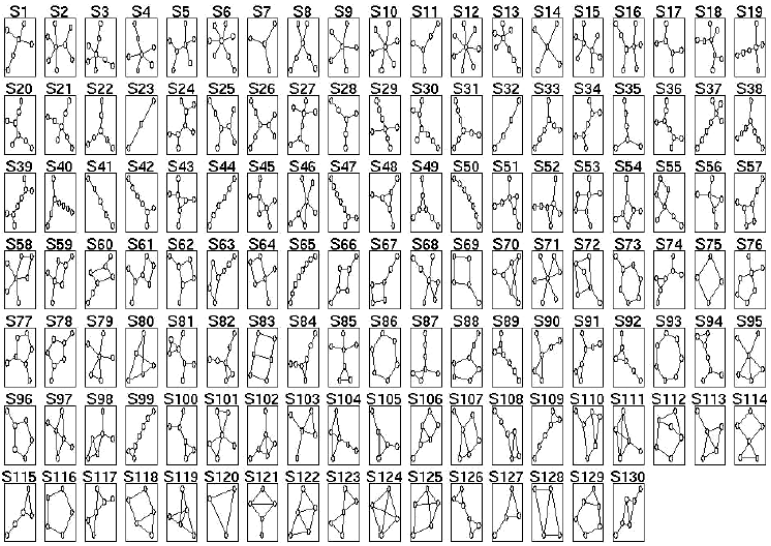

Fig. 6. Full list of 130 7-edge subgraphs.

raw subgraph counts. Two different ways are used to count subgraphs in order to show robustness of the experiments. We first count all subgraphs containing up to 7 edges (130 nonisomorphic subgraphs), and then we count all subgraphs that can be constructed by a walk of length 8 (148 nonisomorphic subgraphs). All 130 7-edge subgraphs are shown in Fig. 6 [54]. We generated 3000 graphs — 1000 graphs for each of the three models we analyzed as experimental data. As mentioned before, every graph is generated with the same number of nodes and of edges as in the real circuit. The counts of subgraphs are input features for the classifier. We used a 5-fold cross-validation boosting process to evaluate the models.

Tables 2, 3, and 4 give the 7-edge subgraph prediction scores of several ISCAS-85 circuits for each of the three models, averaged over folds. Since the prediction based on 8-walk subgraphs shows the same results, the corresponding experimental data are omitted.

Table 2. Prediction scores for C17 (7-edge subgraphs).

Model	GRG	OPT ($\lambda = 0.3$)	SPA ($\alpha = 3$)
Fold = 0	39.86	−39.86	−39.86
Fold = 1	27.73	−27.74	−26.63
Fold = 2	27.97	−25.03	−25.52
Fold = 3	31.38	−28.73	−29.94
Fold = 4	28.13	−25.8	−27.03
Average	31.01	−29.43	−29.8
STDEV	5.17	6.01	5.86

Table 3. Prediction scores for C432 (7-edge subgraphs).

Model	GRG	OPT ($\lambda = 0.2$)	SPA ($\alpha = 2.5$)
Fold = 0	36.98	−39.72	−38.51
Fold = 1	21.52	−28.36	−29.31
Fold = 2	30.88	−31.08	−37.8
Fold = 3	40.19	−40.23	−40.16
Fold = 4	39.79	−39.85	−39.67
Average	33.87	−35.85	−37.16
STDEV	7.84	5.68	4.45

Table 4. Prediction scores for C499 (7-edge subgraphs).

Model	GRG	OPT ($\lambda = 0.22$)	SPA ($\alpha = 2$)
Fold = 0	24.8	−24.77	−24.8
Fold = 1	25.16	−25.27	−25.19
Fold = 2	28.08	−28.08	−28.07
Fold = 3	24.41	−24.6	−24.08
Fold = 4	20.6	−25.67	−18.69
Average	24.61	−25.8	−24.17
STDEV	2.67	1.41	3.42

A given network's subgraph counts determine paths in the ADT dictated by inequalities specified by the decision nodes. For each class, the ADT classifier outputs a real-valued prediction score, which is the sum of all weights over all paths of the decision tree. The class with the highest score wins. The prediction score for a specific class is related to the probability for the tested network to be in this class [28, 54]. The GRG model is the only model having a positive prediction score in every case. Also, the comparatively small standard deviations over different folds indicate robustness of the classification against data subsampling, and make sure that the GRG model is clearly separated from the other two models by the machine learning approach. The scores of the optimization and SPA models overlap, hence the classier cannot clearly differentiate them.

4.1.4. *Subgraph frequency distribution analysis*

Based on the above dataset of subgraphs, we further studied the local structure of the electronic circuits in terms of the subgraph frequency distribution. The subgraph frequency distribution analysis provides another effective bottom-up method for comparing the similarity of network models [65, 64].

We compared the correlation coefficients of the subgraph frequency distribution for different models and real circuits. Table 5 lists the correlation coefficients of subgraph frequency distributions, which are counted in two different ways. The subgraph frequency distributions of the GRG model are closer to real circuits than

Table 5. Correlation coefficients of 7-edge subgraph frequency distributions between real circuits and different models (averaged over 1000 graphs, respectively).

Circuit	GRG	OPT ($\lambda = 0.2$)	SPA ($\alpha = 2.5$)
C17	0.966	0.957	0.947
C432	0.995	0.857	0.935
C499	0.961	0.700	0.878

those of other models, and the results are consistent with experimental results from the machine learning studies. Since the experiments based on eight-walk subgraphs show the same results, the corresponding data are omitted due to limited space.

4.1.5. *Analysis of experimental results on electronic circuits*

In the above experiments, both the machine learning and the subgraph frequency distribution analysis show that the GRG model outperforms the optimization model and the SPA model. This is not really surprising. Since there are more constraints in the GRG model, the input to the GRG model is the degree sequence of the original real circuit, whereas the other models only get the number of nodes and the number of edges as the input. Our discovery is not the only case in which the GRG model outperforms other models in matching real-world networks. In a recent study on modeling protein–protein interaction networks, Przulj *et al.* [66] showed that the stochastic implementation of the GRG model by Chung *et al.* [19] outperforms other models.

A null model is used to evaluate and validate the proposed models for real-world networks. Some research on complex networks only compares studied models with the simplest (and quite inaccurate) ER model [69, 77]. If these models are rechecked by the GRG model, many of them will show outcomes similar to those presented here. Both our results and other research [20, 6] show that one needs to be very careful and compare measured properties to the appropriate null model. Although the GRG model is very successful in our experiments, it still fails to produce motifs in real circuits [56, 76]. This situation presents a big challenge to generating synthetic but realistic networks, and we still have a lot of work to do on this topic.

The GRG model reproduces the degree distribution of real networks directly, and is different from models such as the optimization model and the SPA model, which try to model and explain the general underlying mechanism of networks. Although the GRG model can better capture topological structure, due to the number of constraints imposed, we cannot use it to discover laws governing the topology growth process of a particular system: it lacks the predictive and rescaling power necessary for benchmark generation.

We have shown that, under appropriate parameters, the optimization mechanism, which trades off wire length and characteristic path length, can generate network structure similar to that of networks generated using the extended PA model with spatial constraints, although the two models represent completely different strategies for generating the topological structure. The results provide us with a deeper understanding of the underlying mechanisms in complex networks.

The computational complexity of the SPA model is much lower than that of the optimization model; as shown in Table 5, the SPA model matches the circuits C432 and C499 better than the optimization model in terms of the subgraph frequency distribution, so we can use the SPA model as an efficient alternative to the expensive optimization model for synthesizing larger benchmark circuits. For instance, the optimization model takes many days to generate 1000 graph instances matching a larger circuit (C3540) for the machine learning experiments, but we can generate corresponding graph instances by the SPA model in a few hours, and quickly find the appropriate parameter, $\alpha = 5$, for fitting the real circuit. The machine learning results also show that the GRG model fits C3540 better than does the SPA model.

4.2. *Brain networks*

The human brain is a large complex network with nontrivial topological properties. Characterizing the underlying architecture of such a network is an important issue in neuroscience: it can reveal general principles of structural and functional organization in the human brain, and increase our understanding of how the human brain is capable of generating and integrating information from multiple sources in real time [70]. Recently researchers investigated a large-scale anatomical network of the human cerebral cortex using cortical thickness measurements from magnetic resonance images [35]. As shown in Sec. 3, brain networks and electronic circuits share similar organizational principles, so we applied the same approaches used in electronic circuit analysis on the giant component of the brain anatomical network discovered in Ref. 35.

4.2.1. *Degree distribution*

The degree distribution of the brain anatomical network follows an exponentially truncated power law [35]. Although we can regulate the degree distributions of the optimization model and the SPA model in the previous section by adjusting corresponding parameters, the exponent of the power law in the degree distributions is limited in a narrow range near 3. Figure 7 shows that the optimization model and the SPA model can capture the general tendency, but they do not match the degree distribution of the human brain anatomical network with high fidelity.

4.2.2. *s-metric and characteristic path length*

We also investigate the s-metric and characteristic path length of the optimization model and the SPA model, and compare them with those of the brain anatomical

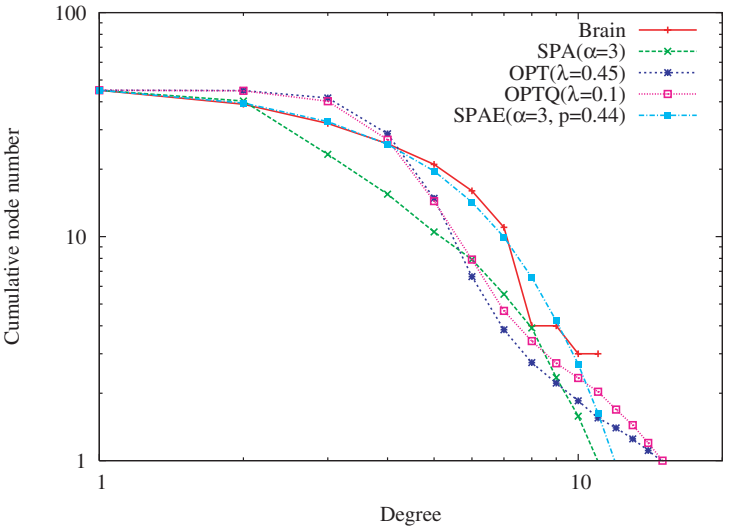

Fig. 7. Cumulative degree distribution of the brain anatomical network, and cumulative degree distributions of graphs generated by the SPA model, the optimization model (OPT), the updated optimization model with the quadratic wiring cost metric (OPTQ), and the extended SPA model (SPAE). All models are averaged over 100 graphs, respectively.

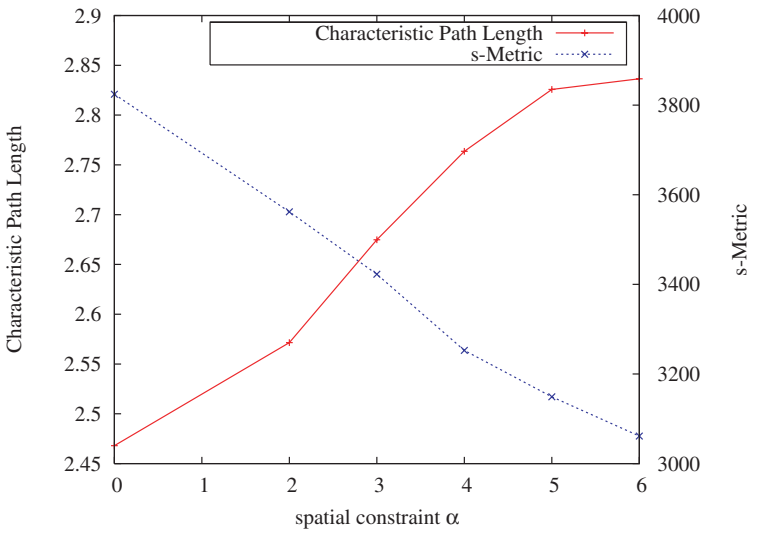

Fig. 8. The statistics of the *s*-metric and characteristic path length in the SPA model. All values are averaged over 100 graphs.

network. As shown in Figs. 8 and 9, both the *s*-metric and the characteristic path length are monotonic functions of the parameters of the optimization model and the SPA model. The data in Table 6, Figs. 8 and 9 show that both the optimization model and the SPA model cannot find appropriate parameters to satisfy values of

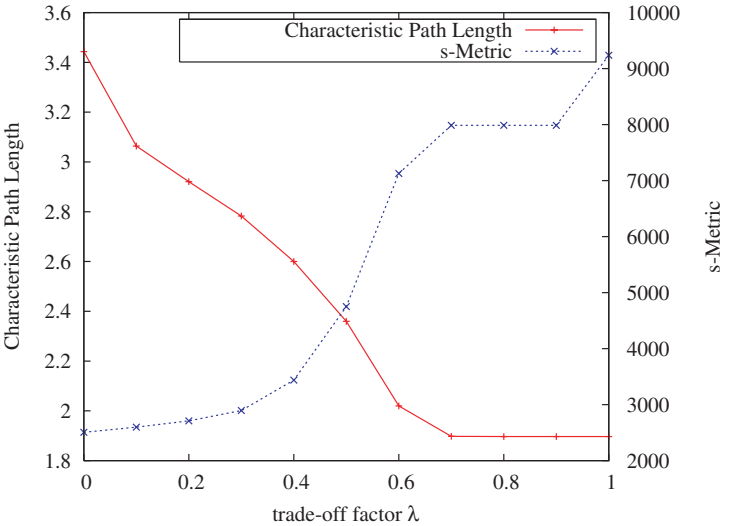

Fig. 9. The statistics of the s-metric and characteristic path length in the optimization model. All values are averaged over 100 graphs.

Table 6. Global properties of the brain anatomical network, the GRG model, the SPA model, the optimization model (OPT), the updated optimization model with the quadratic wiring cost metric (OPTQ), and the extended SPA model (SPAE). All values are averaged over 100 graphs, respectively.

Model	Characteristic path length	s-metric
Brain	3.05	3957
GRG	2.65 ± 0.05	3819.35 ± 65.61
SPA ($\alpha = 0$)	2.47 ± 0.01	3920.84 ± 368.14
SPA ($\alpha = 3$)	2.68 ± 0.08	3388.17 ± 241.74
SPA ($\alpha = 5$)	2.83 ± 0.08	3136.1 ± 196.23
OPT ($\lambda = 0.1$)	3.06 ± 0.03	2598.72 ± 77.96
OPT ($\lambda = 0.4$)	2.6 ± 0.05	3438.16 ± 314.55
OPT ($\lambda = 0.45$)	2.50 ± 0.07	3854.29 ± 533.48
OPTQ ($\lambda = 0.01$)	3.07 ± 0.03	2605.45 ± 83.08
OPTQ ($\lambda = 0.05$)	2.84 ± 0.04	2892.84 ± 161.67
OPTQ ($\lambda = 0.1$)	2.53 ± 0.05	3851.8 ± 348.78
SPAE ($\alpha = 3; p = 0.44$)	2.82 ± 0.14	4031.9 ± 681.12

the s-metric and characteristic path length of the real brain network simultaneously, and the GRG model apparently fits the brain network better than the optimization model and the SPA model. In Table 6, each model lists three different values of the parameter which corresponds to a synthetic network which best fits the real brain network in terms of the characteristic path length and s-metric, by trading off the characteristic path length and s-metric, respectively.

4.2.3. *Analysis of discrepancies*

We can account for the deviation between the proposed models and the actual brain network structure as follows. Firstly, the wiring diagram in the current data may be still somewhat incomplete. Secondly, the wiring cost function does not fully represent costs associated with cortical placement, or constraints other than connectivity need to be taken into consideration [14]. Although the first reason remains a possibility and can be addressed by future reconstructions, its exploration lies beyond the framework of the optimization approach. Here, we will focus on the merits of the second reason.

In the above optimization model, wiring cost is simplified as total wire length. This wiring cost metric fits electronic circuits well, since it is consistent with design and implementation principles of electronic circuits. But for brain networks or neuronal networks, the exact origin of the wiring cost is not completely known, and one can only guess at a functional relationship of wiring cost based on wire length between cortical regions or neurons. Recently, Chklovskii *et al.* [17, 14] argued that the wiring cost may scale as the wire length squared, reducing the optimal layout problem to a constrained minimization of a quadratic form. For biologically plausible constraints, this problem has exact analytical solutions, which give reasonable approximations to actual layouts in the brain. Using this proposal we update the objective function in Subsec. 3.1 as

$$F = \frac{2\lambda}{n(n-1)} \sum_{i<j} \mathrm{SL}_{ij} + \frac{1-\lambda}{m} \sqrt{\sum_{\mathrm{edge}(i,j)} \mathrm{WL}_{ij}^2}.$$

The machine learning approach described in Subsec. 4.1.3 can also be used to select the optimal values of the parameters of specific models. The results of the machine learning show that the optimization model (when $\lambda = 0.45$) and the updated optimization model (when $\lambda = 0.1$) match the brain anatomical network best, respectively. As shown in Table 6, the values of the s-metric in both cases are very close to those of the actual brain network. This suggests that the s-metric seems an important factor for fidelity of the optimization models.

We further compare the various optimization models, the SPA model, and the GRG model, in Table 7: the results show that the updated optimization model's topology better models the real brain network's topology, but still cannot match the s-metric and characteristic path length of the real brain network simultaneously; also, it underperforms the GRG model. As shown in Fig. 7, the degree distribution of the updated optimization model does not improve much. Obviously, in addition to wiring cost, there are some other important constraints on the structure and layout of the brain anatomical network that must be incorporated.

In the optimization model proposed in Subsec. 3.1, the nodes are treated as small points and uniformly placed in a square grid. Since the ISCAS-85 circuits are constituted by basic logic gates (AND, NAND, OR, NOT, BUFF, etc.), it is reasonable to simplify gates as small squares with the same size, and put them

Table 7. Prediction scores of the GRG model, the different optimization models (OPT and OPTQ), and the SPA model fitting the brain anatomical network (7-edge subgraphs).

Model	GRG	OPT ($\lambda = 0.45$)	OPTQ ($\lambda = 0.1$)	SPA ($\alpha = 3$)
Fold = 0	8.14	−16.96	−8.40	−8.00
Fold = 1	13.13	−17.56	−11.68	−16.98
Fold = 2	15.14	−17.62	−14.45	−13.95
Fold = 3	18.21	−19.54	−21.81	−19.43
Fold = 4	14.33	−19.07	−15.40	−11.17
Average	13.79	−18.15	−14.35	−13.91
STDEV	3.67	1.10	4.98	4.54

on a grid evenly. But brain networks are much more complicated, because cortical regions have significantly different sizes and irregular shapes [35, 41]. Recent studies have shown that cortical region sizes have significant influence on the structure and placement of brain networks, and that the size constraint substantially restricts the number of permissible rearrangements [41, 22]. In addition, recent studies on the cerebral cortex layout have discovered common layout principles shared by electronic circuits and brain networks [14, 16, 39, 35]. Actually, current VLSI design processes always reuse existing, well-established modules with various sizes instead of building systems from scratch, so the corresponding optimization algorithms need to consider the size constraint of different components or modules in circuit layout and placement [21, 14]. In future work, we hope to improve the current optimization model using the powerful placement and layout algorithms borrowed from VLSI placement optimization.

The SPA model can also be updated by incorporating specific mechanisms for flexibly adjusting the power law exponent of the degree distribution. Dorogovtsev *et al.* [24] extended the PA mechanism by adding links between existing nodes, with probability proportional to the product of their degrees. This model can generate a power law degree distribution with a wide range of exponents, and actually some researchers have incorporated the spatial constraints mentioned in Subsec. 3.2 and applied this approach to the analysis of the airport network's structure [31, 32]. We introduce a new parameter, p, to extend the current SPA model: the proportions of edges created by the PA and connecting existing nodes are p and $1-p$, respectively. As can be seen in Table 6, the experiments show that the extended SPA model can match the brain network well when $\alpha = 3$ and $p = 0.44$. Figure 7 also shows that the extended SPA model can match the degree distribution of the brain network almost perfectly. The results of machine learning in Table 8 further demonstrate that the extended SPA model outperforms all other models, including the GRG model.

As shown in the examples of matching the brain network, although we can always extend topology models and achieve higher fidelity by introducing the richer sets of domain-specific parameters, a simple descriptive model like the GRG is still

Table 8. Prediction scores of the GRG model, the updated optimization model (OPTQ), the SPA model, and the extended SPA model (SPAE) fitting the brain anatomical network (7-edge subgraphs).

Model	GRG	SPA ($\alpha = 3$)	SPAE ($\alpha = 3; p = 0.44$)	OPTQ ($\lambda = 0.1$)
Fold = 0	−24.74	−24.91	24.41	−24.18
Fold = 1	−25.04	−25.74	24.48	−25.28
Fold = 2	−24.99	−23.27	24.91	−25.30
Fold = 3	−32.56	−32.48	32.44	−32.42
Fold = 4	−16.87	−32.18	17.31	−25.29
Average	−24.84	−27.71	24.71	−26.49
STDEV	5.55	4.31	5.36	3.35

very useful. Discovering new mechanisms and developing new models with higher fidelity are not easy, and adding parameters and fitting corresponding values generally lead to dramatically increased computational complexity. Our experiments on electronic circuits and brain networks show that the GRG model can often generate acceptable results.

4.3. *Analysis based on the s-metric*

Tables 9 and 10 show an interesting phenomenon: the s-metric values of the ISCAS-85 benchmark circuits and the human brain anatomical network [35] are very close to those of the corresponding GRG models. The s-metric potentially unifies many aspects of complex networks, because it is closely related to betweenness, degree correlation, and graph assortativity. In addition, it has a direct interpretation as the relative log-likelihood of a graph synthesized by the GRG model, which can only produce graphs with high s-metric values [47]. The above observations of the s-metric explain why the GRG model can match the electronic circuits and the brain network well from a probabilistic viewpoint.

Table 9. s-metric statistics of ISCAS-85 circuits and the corresponding GRG models. All values of the GRG models are averaged over 1000 graphs, respectively.

Circuit	s-metric (real)	s-metric (GRG)
C17	84	85.86 ± 2.69
C432	6986	6875.99 ± 143.46
C499	9848	10491.57 ± 306.78
C880	11241	11274.8 ± 75.63
C1355	21216	21626.84 ± 218.73
C1908	36326	29133.64 ± 442.54
C2670	31708	30968.24 ± 182.42
C3540	53768	58300.16 ± 493.41
C5315	96277	97379.29 ± 600.77
C6288	97702	100072.6 ± 565.04
C7552	111876	108620.1 ± 484.71

Table 10. s-metric statistics of the human brain anatomical network, the Macaque cortical connectivity network, the *C. elegans* local network, the *C. elegans* global network, the German highway network, the Chinese airport network, and the corresponding GRG models. All values of the GRG models are averaged over 100 graphs, respectively.

Network	s-metric (real)	s-metric (GRG)
Human	3957	3819.35 ± 65.61
Macaque	2368861	2375055.27 ± 4136.29
C. elegans (local)	127622	126103.72 ± 591.41
C. elegans (global)	916807	911946.68 ± 9739.35
German highway	8025	7904.1 ± 61.45
Chinese airport	1728592	1716900.08 ± 3647.17

We further analyzed the Macaque cortical connectivity network [22], the *C. elegans* neuronal networks [22], the German highway network (autobahn) [40], and the Chinese airport network [48], and found the same observations of the s-metric as listed in Table 10. All these man-made and biological networks share a common planning principle: wire length optimization over the entire network [22, 40]. Our findings indicate that the global and optimal planning of these networks might have important implications for understanding how structural and functional organization emerges from underlying driving forces.

Some researchers proposed the HOT (highly/heuristic optimized/organized tolerance/tradeoffs) conceptual framework to model the mechanisms driving design and growth of technological and biological networks, which involves optimizing functional objectives of networks subject to constraints on their components [47, 27]. Li *et al.* applied the concept of the HOT framework to the Internet, manually creating a deterministic model of the physical topology of the router-level Internet of a single ISP, which follows the economic and technological constraints in the design of the real network [47, 46]. As shown in Table 11, the Internet (HOT) topology [47] has an s-metric value much lower than that of the corresponding GRG model. The organizing principles of the Internet are completely different from the above electronic circuits, neuronal networks, and transportation networks, although the design or evolutionary principles of these networks are also consistent with the HOT conceptual framework. Hence, in the context of highly engineered systems and

Table 11. s-metric statistics of the router-level Internet, the gene transcriptional regulatory networks of *E. coli*, and the corresponding GRG models. All values of the GRG models are averaged over 100 graphs, respectively.

Network	s-metric (real)	s-metric (GRG)
Internet (HOT)	28442	52232.9 ± 3387.75
E. coli TRN (Shen-Orr)	26621	42402.61 ± 1782.63
E. coli TRN (Ma)	1301244	2375893.92 ± 43876.86

complex biological systems, it is largely impossible to understand any nontrivial network properties while ignoring domain-specific details. For instance, the Internet is designed in a decentralized and self-organized manner, and despite embedded spatial constraints, the deployment focuses on optimizing local connection costs at the edge of the network, known as the "last mile" [4, 46, 27, 25]. High-degree nodes can exist, but are found only within local networks on the far periphery of the network, and would not appear anywhere close to the backbone [4, 46]. This pattern can result in high performance (traffic flow) and robustness to failures [47, 46]. In contrast, in the networks shown in Tables 9 and 10, the high-degree nodes are likely to connect to each other and appear in the cores of the networks [4, 46], and so these networks have high s-metric values close to those of the corresponding GRG models. These domain-specific features can be important factors which distinguish the Internet from the networks in Tables 9 and 10. Furthermore, we analyzed the giant components of two *E. coli* transcriptional regulatory networks (TRNs), collected by Shen-Orr *et al.* [68] and Ma *et al.* [50], respectively. As shown in Table 11, their structures have patterns similar to the Internet: links between high-degree nodes are systematically suppressed, whereas those between high-degree nodes and low-degree nodes are favored, so they naturally have much lower s-metric values than those of the GRG model. Some researchers think that this effect decreases the likelihood of crosstalk between different functional modules of the cell, and increases the overall robustness of a network by localizing effects of deleterious perturbations [52].

5. Summary and Future Work

In this article, we have proposed and analyzed the optimization model and the SPA model for generating topologies of electronic circuits and brain networks, based on the observed properties and organizational principles of actual networks. The simulation results showed that these two models can generate similar structures under appropriate parameters. Our experiments also showed that even a simple optimization model can be very computationally expensive, so the SPA model can be used as an efficient alternative in practical applications. Furthermore, we compared topological properties of the proposed models with those of the real networks and the GRG model. The results showed that the GRG model outperforms both the optimization model and the SPA model, and closely matches the benchmark circuits and the brain network. In addition to electronic circuits and brain networks, we found that many other spatial networks under global optimization of wire length, including neuronal networks and transportation networks, also have high s-metric values close to those of the corresponding GRG models. From a probabilistic viewpoint, the GRG model is a simple and acceptable approach to modeling topologies of this class of spatial networks. In the future, it will be very interesting to verify the model on more real-world spatial networks.

Based on domain analysis, we can extend the existing models and achieve higher fidelity by introducing richer sets of domain-specific parameters. For

example, although the GRG model produces higher-fidelity results than the optimization model and the SPA model for the brain network, the extended SPA model with an additional parameter can outperform the GRG model and match the brain network even better. Similarly, if we want to model electronic circuits more precisely, it is necessary to take advantage of domain-specific knowledge of circuits, such as the well-known Rent's rule [73, 18], which displays the relationship between the number of external signal connections to a logic block (i.e. the number of "pins") and the number of logic gates in the logic block. The exponent of Rent's rule can provide a useful clue to improving future circuit topology models.

References

[1] Aiello, W., Chung, F. and Lu, L., A random graph model for massive graphs, in *STOC '00: Proc. Thirty-Second Annual ACM Symposium on Theory of Computing* (ACM, New York, 2000), pp. 171–180.
[2] Albert, R., Scale-free networks in cell biology, *J. Cell Sci.* **118** (2005) 4947.
[3] Albert, R., Jeong, H. and Barabasi, A.-L., Error and attack tolerance of complex networks, *Nature* **406** (2000) 378–382.
[4] Alderson, D. L., Technological and economic drivers and constraints in the internet last mile, Tech report, CaltechCDSTR **004** (2004), http://caltechcdstr.library.caltech.edu/S3
[5] Amaral, L. A., Scala, A., Barthelemy, M. and Stanley, H. E., Classes of small-world networks, *Proc. Natl. Acad. Sci. USA* **97** (2000) 11149–11152.
[6] Artzy-Randrup, Y., Fleishman, S. J., Ben-Tal, N. and Stone, L., Comment on "network motifs: simple building blocks of complex networks" and "superfamilies of evolved and designed networks," *Science* **305** (2004) 1107.
[7] Barabasi, A. L. and Albert, R., Emergence of scaling in random networks, *Science* **286** (1999) 509–512.
[8] Barabasi, A. L. and Oltvai, Z. N., Network biology: understanding the cell's functional organization., *Nat. Rev. Genet.* **5** (2004) 101–113.
[9] Berger, N., Borgs, C., Chayes, J. T., D'Souza, R. M. and Kleinberg, R. D., Degree distribution of competition-induced preferential attachment graphs, *Comb. Probab. Comput.* **14** (2005) 697–721.
[10] Boccaletti, S., Latora, V., Moreno, Y., Chavez, M. and Hwang, D.-U., Complex networks: structure and dynamics, *Phys. Rep.* **424** (2006) 175–308.
[11] Cancho, R. F. I., Janssen, C. and Solé, R. V., Topology of technology graphs: small world patterns in electronic circuits, *Phys. Rev. E* **64** (2001) 046119.
[12] Carlson, J. M. and Doyle, J., Highly optimized tolerance: a mechanism for power laws in designed systems, *Phys. Rev. E* **60** (1999) 1412.
[13] Chakrabarti, D. and Faloutsos, C., Graph mining: laws, generators, and algorithms, *ACM Comput. Surv.* **38** (2006).
[14] Chen, B. L., Hall, D. H. and Chklovskii, D. B., Wiring optimization can relate neuronal structure and function, *Proc. Natl. Acad. Sci. USA* **103** (2006) 4723–4728.
[15] Chen, Q., Chang, H., Govindan, R., Jamin, S., Shenker, S. and Willinger, W., The origin of power-laws in Internet topologies revisited, in *INFOCOM* (2002).
[16] Cherniak, C., Mokhtarzada, Z., Rodriguez-Esteban, R. and Changizi, K., Global optimization of cerebral cortex layout, *Proc. Natl. Acad. Sci. USA* **101** (2004) 1081–1086.

[17] Chklovskii, D. B., Exact solution for the optimal neuronal layout problem, *Neural Comput.* **16** (2004) 2067–2078.
[18] Christie, D. and Stroobandt, P., The interpretation and application of Rent's rule, *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.* **8** (2000) 639–648.
[19] Chung, F. and Lu, L., The average distances in a random graph with given expected degrees, *Internet Math.* **1** (2003) 91C113.
[20] Colizza, V., Flammini, A., Serrano, M. A. and Vespignani, A., Detecting rich-club ordering in complex networks, *Nature Phys.* **2** (2006) 110.
[21] Cong, J., Shinnerl, J. R., Xie, M., Kong, T. and Yuan, X., Large-scale circuit placement, *ACM Trans. Des. Autom. Electron. Syst.* **10** (2005) 389–430.
[22] Costa, L., Kaiser, M. and Hilgetag, C., Predicting the connectivity of primate cortical networks from topological and spatial node properties, *BMC Syst. Biol.* **1** (2007) 16.
[23] Dambre, J., Prediction of interconnect properties for digital circuit design and technology exploration. Ph.D. dissertation, Ghent University (2003).
[24] Dorogovtsev, S. N. and Mendes, J. F. F., Scaling behaviour of developing and decaying networks, *Europhys. Lett.* **52** (2000) 33–39.
[25] D'Souza, R. M., Borgs, C., Chayes, J. T., Berger, N. and Kleinberg, R. D., Emergence of tempered preferential attachment from optimization, *Proc. Natl. Acad. Sci. USA* **104** (2007) 6112–6117.
[26] Erdors, P. and Renyi, A., On random graphs I, *Publicationes Mathematicae* **6** (1959) 290–297.
[27] Fabrikant, A., Koutsoupias, E. and Papadimitriou, C. H., Heuristically optimized trade-offs: a new paradigm for power laws in the Internet, in *ICALP '02: Proc. 29th Int. Colloquium on Automata, Languages and Programming* (Springer-Verlag, London, 2002), pp. 110–122.
[28] Freund, Y. and Mason, L., The alternating decision tree learning algorithm, in *ICML* (1999), pp. 124–133.
[29] Gastner, M. T. and Newman, M. E. J., Shape and efficiency in spatial distribution networks, *J. Stat. Mech.* (2006) P01015.
[30] Gastner, M. T. and Newman, M. E. J., The spatial structure of networks, *The European Physical Journal B — Condensed Matter and Complex Systems* **V49** (2006) 247–252.
[31] Guimera, R. and Amaral, L., Modeling the world-wide airport network, *Eur. Phys. J. B: Condensed Matter* **38** (2004) 381–385.
[32] Guimera, R., Mossa, S., Turtschi, A. and Amaral, L. A. N., The worldwide air transportation network: anomalous centrality, community structure, and cities' global roles, *Proc. Natl. Acad. Sci. USA* **102** (2005) 7794.
[33] Hansen, M. C., Yalcin, H. and Hayes, J. P., Unveiling the ISCAS-85 benchmarks: a case study in reverse engineering, *IEEE Des. Test* **16** (1999) 72–80.
[34] Harlow, J. E., Overview of popular benchmark sets, *IEEE Design Test Comput.* **17** (2000) 15–17.
[35] He, Y., Chen, Z. J. J. and Evans, A. C. C., Small-world anatomical networks in the human brain revealed by cortical thickness from MRI, *Cereb. Cortex* **17** (2007) 2407–2419.
[36] i Cancho, R. F. and Sole, R. V., *Optimization in Complex Networks*, Lecture Notes in Physics, Vol. 625 (Springer, Berlin/Heidelberg, 2003), pp. 114–126.
[37] Itzkovitz, S. R., Milo, N. K., Newman, M. E. J. and Alon, U., Reply to "comment on 'subgraphs in random networks'," *Phys. Rev. E* **70** (2004) 058102.
[38] Jeong, H., Tombor, B., Albert, R., Oltvai, Z. N. and Barabsi, A. L., The large-scale organization of metabolic networks, *Nature* **407** (2000) 651–654.

[39] Kaiser, M., Brain architecture: a design for natural computation, *Philos. Trans. R. Soc. A* **365** (2007) 3033–3045.
[40] Kaiser, M. and Hilgetag, C. C., Spatial growth of real-world networks, *Phys. Rev. E* **69** (2004) 036103.
[41] Kaiser, M. and Hilgetag, C. C., Nonoptimal component placement, but short processing paths, due to long-distance projections in neural systems, *PLoS Comput. Biol.* **2** (2006) e95.
[42] Keller, E. F., Revisiting "scale-free" networks, *Bioessays* **27** (2005) 1060–1068.
[43] King, O. D., Comment on "subgraphs in random networks," *Phys. Rev. E* **70** (2004) 058101.
[44] Kirkpatrick, S., Gelatt, C. D. and Vecchi, M. P., Optimization by simulated annealing, *Science* **220, 4598** (1983) 671–680.
[45] Kundarewich, P. D. and Rose, J., Synthetic circuit generation using clustering and iteration, in *FPGA '03: Proc. 2003 ACM/SIGDA Eleventh Int. Symp. Field Programmable Gate Arrays* (ACM, New York, 2003), pp. 245–245.
[46] Li, L., Alderson, D., Willinger, W. and Doyle, J., A first-principles approach to understanding the Internet's router-level topology, in *SIGCOMM '04: Proc. 2004 Conf. Applications, Technologies, Architectures, and Protocols for Computer Communications* (ACM, New York, 2004), pp. 3–14.
[47] Li, L., Doyle, J. C. and Willinger, W., Towards a theory of scale-free graphs: definition, properties, and implications, *Internet Math.* **2**(4) (2006) 431–523.
[48] Li, W. and Cai, X., Statistical analysis of airport network of China, *Phys. Rev. E* **69** (2004) 046106.
[49] Light, S., Kraulis, P. and Elofsson, A., Preferential attachment in the evolution of metabolic networks, *BMC Genom.* **6** (2005) 159.
[50] Ma, H.-W., Buer, J. and Zeng, A.-P., Hierarchical structure and modules in the *Scherichia coli* transcriptional regulatory network revealed by a new top-down approach, *BMC Bioinform.* **5** (2004) 199.
[51] Mahadevan, P., Krioukov, D. V., Fall, K. R. and Vahdat, A., Systematic topology analysis and generation using degree correlations, in *SIGCOMM* (2006), pp. 135–146.
[52] Maslov, S. and Sneppen, K., Specificity and stability in topology of protein networks, *Science* **296** (2002) 910–913.
[53] Mathias, N. and Gopal, V., Small worlds: how and why, *Phys. Rev. E* **63** (2001) 021117.
[54] Middendorf, M., Ziv, E. and Wiggins, C. H., Inferring network mechanisms: the *Drosophila melanogaster* protein interaction network, *Proc. Natl. Acad. Sci. USA* **102** (2005) 3192–3197.
[55] Milo, R., Itzkovitz, S., Kashtan, N., Levitt, R., Shen-Orr, S., Ayzenshtat, I., Sheffer, M. and Alon, U., Superfamilies of evolved and designed networks, *Science* **303** (2004) 1538–1542.
[56] Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D. and Alon, U., Network motifs: simple building blocks of complex networks, *Science* **298** (2002) 824–827.
[57] Mitzenmacher, M., A brief history of generative models for power law and lognormal distributions, *Internet Math.* **1** (2003).
[58] Mitzenmacher, M., Editorial: The future of power law research, *Internet Math.* **2** (2005) 525–534.
[59] Newman, M. E. J., The structure and function of complex networks, *SIAM Rev.* **45** (2003) 167.
[60] Newman, M. E. J., Power laws, Pareto distributions and Zipf's law, *Contemp. Phys.* **46** (2005) 323.

[61] Newman, M. E. J., Strogatz, S. H. and Watts, D. J., Random graphs with arbitrary degree distribution and their applications. Working Papers 00-07-042, Santa Fe Institute (2000).
[62] Press, W. H., Teukolsky, S. A., Vettering, W. T. and Flannery, B. P., *Numerical Recipes in C++: The Art of Scientific Computing*, 2nd ed. (Cambridge University Press, 2002).
[63] Provan, G. M. and Wang, J., Automated benchmark model generators for model-based diagnostic inference, in *IJCAI* (2007), pp. 513–518.
[64] Przulj, N., Biological network comparison using graphlet degree distribution, *Bioinformatics* **23** (2007) e177–e183.
[65] Przulj, N., Corneil, D. G. and Jurisica, I., Modeling interactome: scale-free or geometric? *Bioinformatics* **20** (2004) 3508–3515.
[66] Przulj, N. and Higham, D. J., Modelling protein–protein interaction networks via a stickiness index, *J. R. Soc. Interface* **3** (2006) 711–716.
[67] Sasaki, G. H. and Hajek, B., The time complexity of maximum matching by simulated annealing, *J. ACM* **35** (1988) 387–403.
[68] Shen-Orr, S. S., Milo, R., Mangan, S. and Alon, U., Network motifs in the transcriptional regulation network of *Escherichia coli*, *Nat. Genet.* **31** (2002) 64–68.
[69] Sole, R. V., Pastor-Satorras, R., Smith, E. and Kepler, T. B., A model of large-scale proteome evolution, *Adv. Complex Syst.* **5** (2002) 43.
[70] Sporns, O., Chialvo, D. R., Kaiser, M. and Hilgetag, C. C., Organization, development and function of complex brain networks, *Trends Cogn. Sci.* **8** (2004) 418–425.
[71] Stroobandt, D., Analytical methods for *a priori* wire length estimates in computer systems. Ph.D. dissertation, Ghent University (2001).
[72] Stroobandt, D., *A priori* wire length distribution models with multiterminal nets, *IEEE Trans. Very Large Scale Integr. Syst.* **11** (2003) 35–43.
[73] Stroobandt, D., Verplaetse, P. and van Campenhout, J., Towards synthetic benchmark circuits for evaluating timing-driven CAD tools, in *ISPD '99: Proc. 1999 Int. Symp. Phys. Design* (ACM, New York, 1999), pp. 60–66.
[74] Tanaka, R., Csete, M. and Doyle, J., Highly optimised global organisation of metabolic networks, *Syst. Biol. (Stevenage)* **152** (2005) 179–184.
[75] Viger, F. and Latapy, M., Efficient and simple generation of random simple connected graphs with prescribed degree sequence, in *COCOON* (2005), pp. 440–449.
[76] Wang, J. and Provan, G., On motifs and functional blocks in technological networks, in *ECCS 2007: Proc. 2007 European Conference on Complex Systems* (2007).
[77] Watts, D. J. and Strogatz, S. H., Collective dynamics of "small-world" networks, *Nature* **393** (1998) 440–442.
[78] Willinger, W., Alderson, D., Doyle, J. C. and Li, L., More "normal" than normal: scaling distributions and complex systems, in *WSC '04: Proc. 36th Conf. Winter Simulation* (Winter Simulation Conference, 2004), pp. 130–141.
[79] Yook, S.-H., Jeong, H. and Barabasi, A.-L., Modeling the Internet's large-scale topology, *Proc. Natl. Acad. Sci. USA* **99** (2002) 13382.