Classes of Optimal Network Topologies under Multiple Efficiency and Robustness Constraints

Sanket Patil, Srinath Srinivasa
Open Systems Laboratory
International Institute of Information Technology
Bangalore - 560100, INDIA
{sanket.patil, sri}@iiitb.ac.in

Venkat Venkatasubramanian
Laboratory for Intelligent Process Systems
Department of Chemical Engineering
Purdue University
West Lafayette, IN 47906, USA
venkat@ecn.purdue.edu

***Abstract*—We address the problem of designing optimal network topologies under arbitrary optimality requirements. Using three critical system parameters, efficiency, robustness and cost, we evolve optimal topologies under different environmental conditions. Two prominent classes of topologies emerge as optimal: (1) Star-like (or scale-free) topologies, with high efficiency, high resilience to random failures and low cost, and (2) "Circular Skip Lists" (CSL), with high robustness to random failures as well as targeted attacks, and high efficiency at moderate cost. We analyze CSLs further to observe that they show several structural motifs that are optimal with respect to a variety of metrics.**

***Index Terms*—Network design, optimal networks, optimal structural motifs**

I. Introduction

Designing optimal network topologies is an important problem across various application domains such as: distributed information systems, supply chain networks, content delivery networks and network-centric warfare. The requirements of optimality vary with the purpose for which a network is built. Further, there are conflicting optimality requirements within a network that need to be balanced.

There is a lot of recent interest in developing a philosophical understanding of the optimal network design problem. Motter and Toroczkai [1] identify four classes of network optimization problems in increasing order of complexity: (1) optimizing a property such as diameter, given a static property such as degree sequence (structural optimization) (2) optimizing flow on the edges in a static graph to minimize a global property such as congestion (dynamics optimization on a static graph) (3) finding the optimal structure given a property such as flow dynamics (strucutral optimization for dynamics) and finally, (4) designing an optimal structure in situations where dynamics, such as failures, can affect performance (dynamics-driven network optimization).

Classes of networks have been studied in search of optimal properties especially following the work on complex networks by Strogatz, Barabasi and Albert [2]–[4]. Scale-free networks, with their power law degree distributions are shown to have low diameters [5] and high resilience to random failures. Watts and Strogatz study networks with small-world properties such as low average path length and high clustering, leading to fast propagation of information and high synchronizability [6]. Kleinberg [7] shows that short paths can be found using purely local (node-level) information in a sub-class of small-world networks. Newman studies presence of community structures in disparate complex networks [8].

Cohen et al. show that while scale-free networks are robust in the face of random failures [9], they are easily disrupted by targeted attacks [10]. Valente et al. [11] report that the most robust structure in the face of random failures or targeted attacks is one with at most three distinct node degrees in the network. Dekker and Colbert [12] study targeted attacks in the context of Network Centric Warfare (NCW) [13]. They propose node connectivity (that is the minimum number of nodes whose deletion partitions a network) as the most suitable metric to measure robustness. They study different classes of networks and report that *vertex-transitive* networks are the most robust. In a vertex-transitive network one node cannot be distinguished from another based on its neighbourhood. Thus, the disruption due to an attack does not depend on the target.

Scale-free networks are also unsuitable for large traffic flows since a small number of nodes handling most of the traffic load can lead to congestion. Unless the load handling capacity of a node is directly proportional to its degree, random-regular graphs and Cayley trees are shown to be better suited for designing traffic flow networks [14]. Guimera et al. [15] study the problem of "searchability", i.e. the ability to find short paths based on local information, in the presence of congestion, and report that only two classes of optimal topologies exist: highly polarized (star-like) networks, when the load on the network is low; and *homogeneous isotropic networks* with symmetric node betweenness, as the load increases. Mondragon [16] reports the sharp transition of network structures from star-like to decentalized as the load on the network increases. Donetti et al. [17], [18] study optimal structures for high synchronizability and low first-passage times for random walkers, and propose *entangled networks* with highly homogeneous structural properties.

The topology design problem has received significant interest in the domain of data-centric peer-to-peer networks. Agneeswaran and Janakiram [19] study distributed hash tables (DHT) from the complex systems perspectives of Self-Organized Criticality (SOC) [20] and Highly Optimized Tolerance (HOT) [21]. SOC models robustness in open world systems where events can occur unexpectedly, whereas HOT

978-1-4244-2794-9/09/$25.00 © 2009 IEEE

provides a generative model for highly efficient networks such as power-law networks. They report that a two-layered DHT architecture, with a robust structured overlay on a highly efficient unstructured layer, satisfies both SOC and HOT considerations.

There are studies on the structural properties of DHTs with respect to lookup complexity and resilience to frequent random failures [22]–[24]. Cayley graphs have been proposed as a unified group-theoretic model for design and analysis of optimal DHT topologies [25]. Cayley graphs are characterized by properties such as symmetric degree centrality, hamiltonicity, and robustness. Gummadi et al. [26] report that ring structures are most suitable for DHTs. They also argue that the presence of hamiltonian circuits facilitates design of dynamic DHT routing schemes. Wrapped butterfly networks [27] are characterized by one or more hamiltonian circuits. A hamiltonian decomposition allows load to be equally distributed, also making a network robust. In an earlier work [28], we proposed *circular skip lists* with nearly symmetric degree centrality and multiple Hamiltonian circuits as optimal structures for balancing load in DHTs.

Generally, the approach to addressing the network optimization problem in literature has been either to analyze classes of networks (such as Cayley graphs or scale-free networks) for specific optimal properties (such as low diameter or symmetric betweenness) or to design networks for domain specific optimality requirements (as in DHTs or NCW). Our approach is to address the problem of designing network topologies under a variety of optimality requirements and trade-offs. In terms of the classification of network optimization problems [1] mentioned above, we address the class 4 problem.

Using genetic algorithm optimization, we design both undirected and directed networks using multiple optimality metrics. We use diameter, average path length and closeness centrality as design metrics for efficiency of communication. For robustness, we use degree centrality, node betweenness and edge connectivity as design metrics. Cost is measured in terms of the number of edges. Different combinations of these metrics are useful in different application domains. We also use two sliders to model application dependent trade-offs between efficiency, robustness and cost. Different types of structures emerge as optimal under different optimality requirements. We observe two prominent optimal topology classes: (1) star-like or scale-free networks, with low diameters at low costs and low resilience to targeted attacks and (2) circular skip lists (CSL), with low diameters at moderate costs and high resilience to both random failures and targeted attacks.

Circular skip lists occur universally except under very low emphasis on robustness and severe restrictions on cost. We observe that the trade-off between efficiency and robustness is pronounced only under severe restrictions on cost. In all other cases, CSLs are optimal topological structures in terms of balancing efficiency, robustness and cost. They are characterized by several motifs that are optimal under a variety of requirements, such as low diameter and low average path lengths; resilience to random failures as well as targeted attacks due to the presence of multiple independent paths; homogeneity with near zero entropy of degree distribution, near zero skew in node and edge betweenness. Circular skip lists also contain multiple Hamiltonian circuits.

Our findings not only corroborate several recent results, but also generalize them. Mondragon [16] reports that the transition from star-like networks to symmetric networks is sharp as the load in a network increases. We observe that this holds under other metrics of robustness as well. Graphs with hamiltonian decompositions are shown to be optimal in DHTs [26], [27]. Valente et al. [11] report networks with a nearly symmetric degree centrality as most robust against failures as well as attacks. Guimera [15] proposes networks with symmetric node betweenness as most optimal to handle congestion in traffic flow networks. Dekker [12] proposes regular graphs with *optimal connectivity* to be ideal under targeted attacks. We show that circular skip lists are characterized by all these properties which are optimal under different circumstances.

II. The Network Design Formalism

Venkatasubramanian et al. [29] study how the structure of a class of complex networks relate to their performance. They propose that a complex network optimizes its structure to maximize its survival fitness (or performance). Performance of a network in turn depends on three critical parameters efficiency, robustness and cost; and an environmental selection pressure variable that decides the trade-off between efficiency and robustness. They define efficiency in terms of the average path length (APL); robustness in terms of the number and size of components that result in a network upon a single node deletion; and cost in terms of the number of edges. Using these parameters in a fitness function, networks of n nodes with n undirected edges are allowed to evolve under varying emphasis on efficiency and robustness. The star topology emerges as most efficient and least robust, whereas the circle topology emerges as least efficient and most robust.

We extend the above formalism to accommodate multiple constraints and optimality objectives to study different classes of networks.

- Diameter, APL and closeness centrality are used as design metrics for efficiency.
- Degree centrality, node betweenness centrality and edge connectivity are used as design metrics for robustness.
- We design topologies with arbitrary number of edges. Further, we consider both undirected and directed graphs.
- An environmental variable α decides the trade-off between robustness and efficiency. We use another variable β as a cost control parameter.

Different combinations of the above design metrics are applicable in different scenarios. In DHTs, minimizing lookup complexity (diameter) while maintaining small and symmetric finger tables across machines (degree centrality) is a design objective. To handle traffic flow, designing networks with low APL while balancing load on the nodes (node betweenness) to avoid congestion is important. In case of NCW and supplychains, having alternate paths when a communication link fails (edge connectivity) is a design requirement.

Thus, we derive different objective functions for different combinations of the above metrics, and design topologies accordingly using a genetic algorithm optimization process. A detailed report can be found in [30]. In this paper, we present only a representative subset of our results using a small number of metrics due to space constraints.

A. *Efficiency*

In this paper, we use efficiency based on diameter in the design process. The worst diameter for a connected graph of n nodes is $n-1$, which is the diameter of a straight line graph in case of undirected graphs, and a circle in case of directed graphs. The best diameter is 1, which is the diameter of a clique (complete graph). In other words, a topology is most efficient if the diameter is 1, and least efficient if it is $n-1$. We map a diameter d that falls in the interval $[1, n-1]$, to a value of efficiency (η_d) in the interval $[0, 1]$, as:

$$\eta_d = 1 - \frac{d-1}{n-2}$$

B. *Robustness*

We use two definitions of robustness.: ρ_p, based on the skew in degree centrality, to cover the symmetric load perspective (as in DHTs and traffic flow networks); and ρ_λ, based on edge connectivity, to cover the targeted attack perspective (as in NCW).

We define skew in degree centrality as the difference in the maximum degree in the graph ($\hat{p}$) and the mean degree of the nodes ($\bar{p}$). (In case of directed graphs, we consider both in and outdegrees.) For a connected graph of n nodes, the worst skew occurs for the star topology. The central node has a degree of $n-1$ and all the nodes surrounding it have a degree of 1. Therefore, the worst skew is $\frac{(n-1)(n-2)}{n}$. The best skew is 0, when all the nodes have the same degree. This occurs when the topologies are *regular* graph topologies as in a circular topology or a clique. This holds for both directed and undirected graphs. Thus,

$$\rho_p = 1 - \frac{n(\hat{p}-\bar{p})}{(n-1)(n-2)}$$

Edge connectivity (λ) is the minimum number of edges whose removal renders a network disconnected. In case of an undirected graph, the tree topologies have the worst connectivity of 1, and the circle has the worst connectivity of 1 in directed graphs. For both cases, the clique has the best connectivity, $n-1$. Thus, robustness, when defined in terms of connectivity is:

$$\rho_\lambda = \frac{\lambda - 1}{n-2}$$

C. *Cost*

We divide cost into two components: (1) infrastructure cost as a function of the number of edges, e, in the network and (2) node level maintenance/"bookkeeping" cost, as a function of the node's degree, p (in case of directed graphs, in and outdegrees, p^{in} and p^{out}). We place upper bounds on both these in our topology design.

Infrastructure Cost: The minimum number of edges (e_{min}) required to have a connected undirected graph is $n-1$ and it is n in case of a directed graph. We associate a cost, $k = 0$, to a minimally connected graph. Any "extra" edge has an associated cost. All extra edges cost the same. An undirected clique has the highest cost, with $\hat{e} = \frac{n(n-1)}{2}$ (and $\hat{e} = n(n-1)$, for a directed clique) number of edges. Thus, the cost of a topology is defined as the ratio of the number of extra edges in a topology to the number of extra edges in the clique with the same number of nodes.

$$k = \frac{e - e_{min}}{\hat{e} - e_{min}}$$

Maximum Permissible Degree: The Maximum Permissible Degree, p, is an upper limit on the number of edges that can be incident on a node. In case of a directed graph, p^{in} and p^{out}, are the upper limits on the number of incoming and outgoing edges. Degree is a measure of the local "book-keeping cost". It can also be thought of as the amount of load a node is handling through edges incident on it.

D. *Fitness*

Fitness of a graph, $\phi(G(V,E))$, is defined in terms of the optimization dimensions: efficiency (η), robustness (ρ) and infrastructure cost (k). The maximum permissible degree is also a design constraint. The general fitness function is as follows.

$$\phi = \alpha\rho + (1-\alpha)\eta - \beta k$$

Here, $0 \le \alpha \le 1$, is an application dependent parameter that acts as a slider between efficiency and robustness. A high value of α indicates that a high emphasis should be placed on the robustness of topologies during the topology breeding process. The parameter β, $0 \le \beta \le 1$ is used for additional cost control (in addition to the upper bound k). When set a high value, β helps the evolutionary process to "squeeze out" the most cost-effective topology that achieves a certain efficiency and robustness (controlled by α) by removing as many superfluous edges as possible.

Thus, the global optimization objective is to find the set of edges to construct the fittest graph:

$$\arg\max_E \phi(G(V,E))$$

We use the following two combinations of efficiency, robustness and cost measures in our "topology breeding" experiments using a standard genetic algorithm procedure: (1) efficiency in terms of diameter (η_d), robustness in terms of degree centrality (ρ_p), cost, k and (2) efficiency in terms of diameter (η_d), robustness in terms of edge connectivity (ρ_λ), cost, k. Thus, we have the following fitness functions:

$$\phi_p = \alpha\rho_p + (1-\alpha)\eta_d - \beta k$$

$$\phi_\lambda = \alpha\rho_\lambda + (1-\alpha)\eta_d - \beta k$$

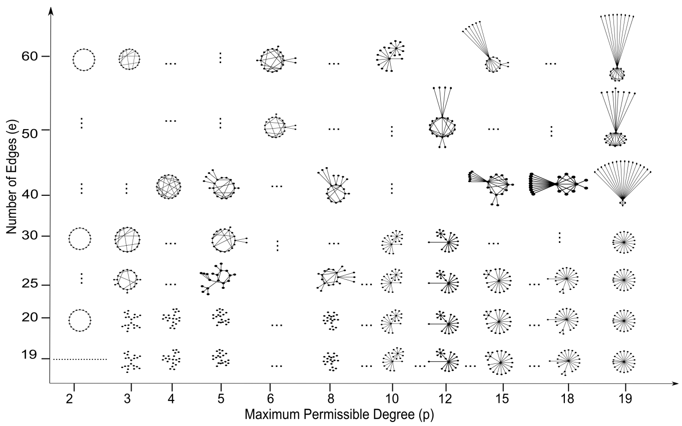

Fig. 1. Undirected Optimal Topology Space with emphasis on robustness, $\alpha = 0$ and cost control parameter, $\beta = 1$, for different number of edges, e and maximum permissible degree, p. Here, the number of nodes, $n = 20$.

III. Optimal Topology Families

We conducted topology breeding experiments using different design metrics for networks with up to 200 nodes. Different types of structures emerge at different points in an "optimal topology space" defined by the optimization parameters. Below we present our main findings.

A. Undirected Topologies

Figure 1 shows a sample breed of optimal undirected topologies when the emphasis on robustness, $\alpha = 0$, and the cost control parameter, $\beta = 1$, thus effecting the most efficient (i.e. diameter optimal) as well as the most cost-effective topologies for the given upper bounds on degree, p, and cost, k. We show topologies for a small number of nodes, $n = 20$, so that the structural features can be clearly seen. We make the following observations which hold for any n.

When the number of edges, $e = n - 1$ (i.e. cost, $k = 0$), the only possible connected networks are trees. When $e = n - 1$ and the maximum permissible degree, $p = 2$, the only possible network is a straight line, with diameter $n - 1$, which corresponds to an efficiency, $\eta_d = 0.0$. The degree centrality of a straight line is nearly symmetric, thus it has $\rho_p \approx 1$, for large n. However, a straight line has $\rho_\lambda = 0$, since there is exactly one path between any pair of nodes in the graph.

With $e = n - 1$, if we increase p, without regard to robustness (i.e. $\alpha = 0.0$), diameter decreases, and networks with multiple hubs emerge as optimal. This process converges at the star topology. A star is the most efficient topology when $(n-1) \leq e < \frac{n(n-1)}{2}$, with a diameter 2. However, it is the least robust topology in terms of degree centrality. All trees are equivalent in terms of robustness when we consider edge connectivity, as all trees have $\lambda = 1$. Failure of any edge in a star has the same effect, that of partitioning a single node from the rest of the network. On the other hand, in case of trees with multiple hubs, the edges in the middle act as "bridges", and their failure causes a greater disruption than a failure of an edge in the periphery. Failure of a node causes maximum disruption in a star: in the worst case, when the central node fails, the network is partitioned into $n-1$ single node components. Thus, star and star-like networks are highly susceptible to targeted attacks.

When $e = n$ and $p = 2$, the circle emerges as the optimal topology, in terms of both efficiency and robustness, regardless of the value of α or β. An undirected circle has a diameter of $\lceil \frac{n}{2} \rceil$. It is 2-regular. Further, being a hamiltonian circuit, is has an edge connectivity of 2. A circle also has a node connectivity of 2. When $e = n$ and $p > 2$, either trees or circle emerge as optimal structures, depending on the value of α.

When $e > n$, we observe topologies that we call *circular skip lists* (CSL). A CSL is a topology in which each node has edges to one or more other nodes at random distances ("skips") on a logical circle. A circle is the minimal CSL whereas a complete graph is the maximal CSL. When the maximum permissible degree, p is low, CSLs emerge as optimal. As p increases, CSLs start giving way to topologies that are a "hybrid" betweeen CSLs and trees, with a large central loop and small peripheral hubs, as can be seen in figure 1. For higher p, the central loops start growing smaller and hubs larger; the topologies start "shedding" redundant edges (owing to a high value of β); eventually scale-free or star-like networks emerge.

With increasing e, we observe that the onset of hybrid tolpogies starts for increasingly higher p. Circular skip lists prevail until the following relation holds, $p <= \frac{2e}{n}$. In fact, the best CSLs occur at the boundary, $p = \frac{2e}{n}$, where topologies utilize nearly all the edges to achieve low diameters. They are also robust with symmetric degree centrality and high edge connectivity, despite not being optimized for robustness. For p beyond this, CSLs start unfolding and shedding edges to form hybrid structures.

With increasing e, we see scale-free networks towards the far p end in figure 1, with a small number of hubs forming a highly clustered core. Under the conditions ($\alpha = 0$ and $\beta = 0$), ideally the optimization should converge to the star. However, we see clusters formed as the evolutionary process is unable to throw away all redundant edges, possibly due to local minima.

Topologies shown in figure 1 occur when $\alpha = 0$. As we increase α, (while keeping $\beta = 1$), we see hubs expanding into smaller hubs leading to a straight line, when $e = n - 1$. For higher e, CSLs prevail regardless of p. And scale-free networks start transforming into circles trying to achieve a symmetric degree centrality (or a higher edge connectivity). As the cost control parameter β is relaxed, CSLs emerge everywhere except at low values of e, regardless of α. We also see the clusters becoming stronger when p and e increase, eventually forming a clique (complete graph), which is the most efficient as well as the most robust topology, when cost is not a constraint.

Figure 2 shows the sharp transition from scale-free networks to CSLs as an effect of α on the breeding process. Here, $n = 20$, $p = n - 1$ and $\beta = 1.0$. The robustness metric used is that of degree centrality, ρ_p. When $\alpha = 0$, we observe highly efficient scale-free networks. As α increases, scale-free networks give way to lesser fit hybrid networks. The point

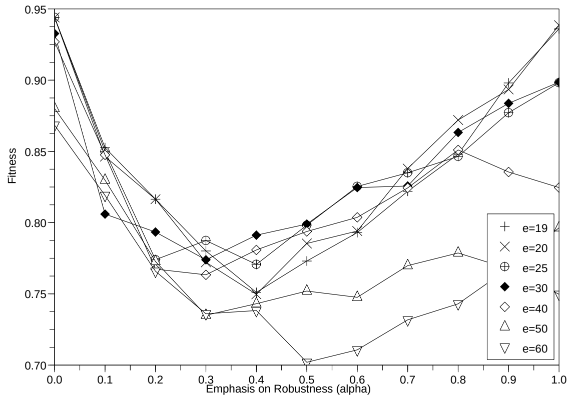

Fig. 2. Except for very small number of of edges (e), scale-free networks start "folding" into CSLs even at a low emphasis on robustness (α), despite the maximum permissible degree, $p = n - 1$. Here, $n = 20$ and cost control parameter, $\beta = 1.0$.

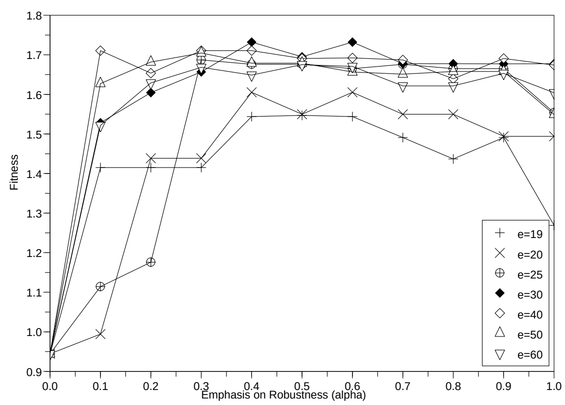

Fig. 3. The *value* of a network, $v_p = \eta + \rho_p - \beta k$ rises sharply. Despite the value of p being $n - 1$ and $\beta = 1$, we observe that circular skip lists set in even for small values of α. Hence, the "flattening" of the value curves.

where the fitness curves start rising is the value of α at which circular skip lists start emerging. As soon as $e > n$, we observe the emergence of CSLs even for a small emphasis on robustness, despite p being $n - 1$. A similar trend occurs when edge connectivity is used as the robustness metric.

We define the *value* of an optimal network as, $v_p = \eta_d + \rho_p - \beta k$. (When edge connectivity is used as the robustness measure in the breeding process, the corresponding definition of value is, $v_\lambda = \eta_d + \rho_\lambda - \beta k$.) Figure 3 plots v_p of optimal networks that emerge at different α. Again, as e increases, We can see a sharp rise in v_p due to the emergence of CSLs. Further, the "flattening" of the v_p curves indicate that the trade-off between efficiency and robustness is pronounced only under severe restriction on cost. That is where scale-free networks with high efficiency and low robustness occur.

Figure 4 shows undirected optimal topologies when the

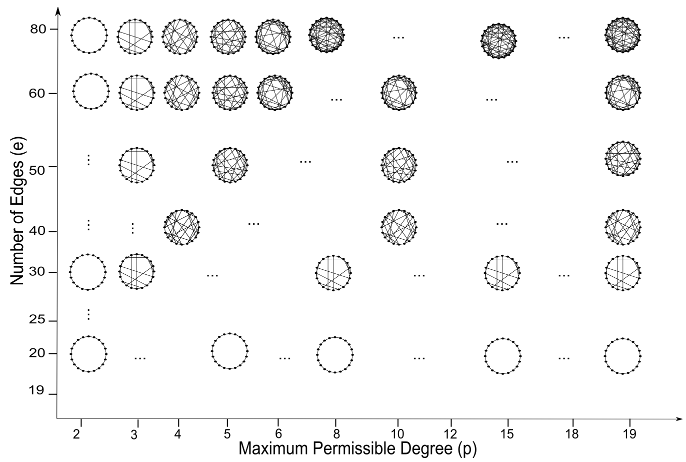

Fig. 4. Undirected Optimal Topology Space with emphasis on robustness, $\alpha = 1$ and cost control parameter, $\beta = 0$, for different number of edges, e and maximum permissible degree, p. Here, the number of nodes, $n = 20$. Robustness metric used is edge connectivity, ρ_λ.

emphasis on robustness is maximum ($\alpha = 1$) and the cost control parameter is relaxed ($\beta = 1$). Edge connectivity is used as the metric of robustness. Thus, the breeding process evolves topologies with as high an edge connectivity, λ, as possible, constrained only by the maximum permissible degree, p and the number of edges, e. When $e >= \frac{np}{2}$, we observe that CSLs emerge with λ of $p - 1$. Further, only a few node pairs have $p - 1$ edge independent paths between them, with most node pairs having p edge independent paths, which is the theoretical upper bound. (The maximal possible connectivity of a graph is equal to the minimal node degree in the graph.) Again, trees can be observed only when α is reduced drastically and β increased.

B. Directed Topologies

We conducted similar experiments for directed graphs. Figure 5 shows the optimal directed graph topologies for $n = 20$, $\alpha = 0.0$ and $\beta = 1.0$.

Circular skip lists occur more naturally in directed graphs. For small number of edges, CSLs help achieve strong connectivity. The first directed hubs occur around $e = 2(n-1)$, when p is very high. Similar to the case of undirected graphs, large hubs with highly clustered cores can be seen for high values of e and p. However, CSLs are prevalent everywhere else. This effect is even more complete as α increases.

Figure 6 shows directed optimal topologies when edge connectivity, λ, is used as the robustness metric. When the emphasis on robustness is maximum ($\alpha = 1$) and the cost control parameter is relaxed ($\beta = 1$), we see topologies very similar to the corresponding case of undirected graphs (figure 4). CSLs with high edge connectivity prevail. Similarly, when $e > np$, CSLs with λ of $p - 1$ emerge.

C. Two Classes of Optimal Topologies

Thus, in general, two prominent classes of topologies emerge as optimal under different optimality requirements: (1) scale-free or star-like networks and (2) circular skip lists.

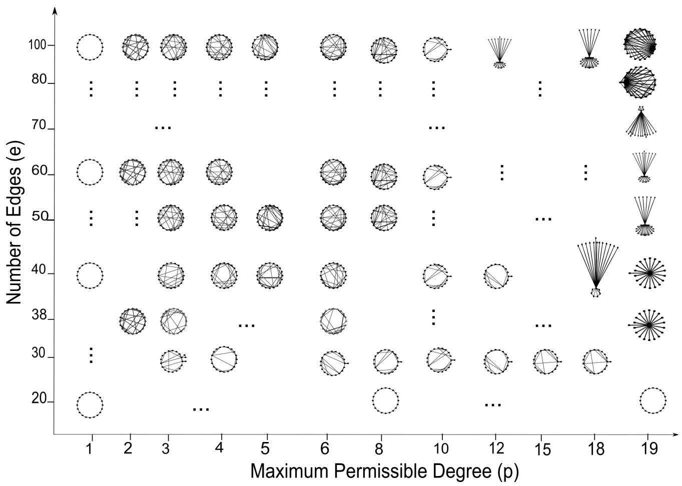

Fig. 5. Directed Optimal Topology Space with emphasis on robustness, $\alpha = 0$ and cost control parameter, $\beta = 1$, for different number of edges, e and maximum permissible degree, p. Here, the number of nodes, $n = 20$.

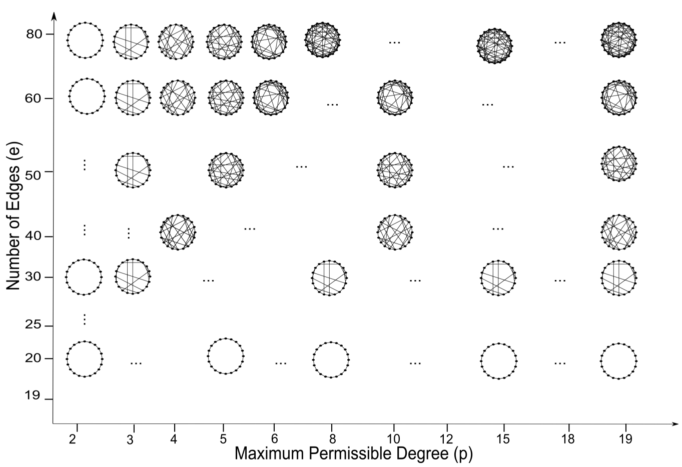

Fig. 6. Directed Optimal Topology Space with emphasis on robustness, $\alpha = 1$ and cost control parameter, $\beta = 0$, for different number of edges, e and maximum permissible degree, p. Here, the number of nodes, $n = 20$. Robustness metric used is edge connectivity, ρ_λ.

Further, we can observe that scale-free networks are optimal only when both the following design requirements are satisfied: (1) very low emphasis on robustness (or very high emphasis on efficiency) and (2) severe restrictions on cost. In all other cases CSL are optimal topologies in terms of balancing efficiency, robustness and cost.

IV. Optimal Structural Motifs of CSLs

Both directed and undirected circular skip lists show several structural features that are potentially optimal under varied applications. Below we discuss some of them using an example directed CSL of 32 nodes, shown in figure 7.

Hamiltonian Circuits: Circular skip lists contain multiple hamiltonian circuits. Presence of hamiltonian circuits is an optimal feature because, a graph containing a hamiltonian circuit is at least biconnected (both in terms of nodes and edges). Presence of hamiltonian circuits is also considered important in applications such as DHTs [26], [27] as they facilitate design of dynamic DHT routing schemes; also, a hamiltonian circuit helps in load balancing.

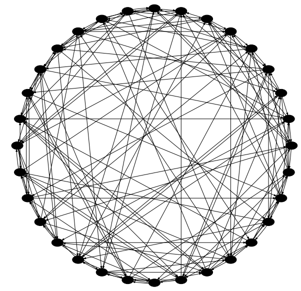

Fig. 7. A directed circular skip list with $n = 32$, $p = 5$ and $e = 160$. CSLs are highly symmetric and possess several optimal properties.

TABLE I
Structural Properties of an example directed CSL of $n = 32$, $p = 5$, $e = 160$

Property	Value
Diameter	3
APL	2.53
Eccentricity (min, max)	(3, 3)
Radius	3
No. of centers	32
Skew in Closeness	0.0016
Entropy of Indegree	0.0
Entropy of Outdegree	0.0
Skew in Node Betweenness	0.005
Skew in Edge Betweenness	0.0025
Edge Connectivity (min, max)	(4, 5)

Symmetric Centrality Measures: Centrality measures are used to determine the importance of nodes and edges in a graph. If a small number of nodes/edges have a significantly higher centrality value than the rest, the network is susceptible to congestion. Also, since such nodes/edges are presumably more important than the rest of the network, they are susceptible to targeted attacks.

We measure the skew in node (or edge) betweenness as the difference between the maximum node (or edge) betweenness in the network and the average betweenness. A (normalized) value of 1.0 for the skew indicate that there is only one node (or edge) in the network through which all the traffic flows; whereas a value of 0.0 indicate a uniform load distribution. Table I shows typical values of betweenness skews for CSLs, which are near 0. Networks with such homogeneity are not easily congested under heavy traffic. Also, such networks are resilient to targeted attacks.

Entropy of Degree Distribution: Entropy of degree distribution is indicative of the heterogeneity of a network. It is measured as, $H = -\sum_p P(p) log P(p)$, where $P(p)$ is the fraction of nodes with a degree p. Typically, CSLs are regular or nearly regular graphs. Thus, have an entropy of nearly 0.

Symmetric Distances: CSLs have low diamter and APL in the graph. The distances between node pairs are also homogeneous, thus leading to a very low skew in closeness centrality (which measures the per node average path length). Node eccentricities (which measures the greatest separation a node suffers from another node in the network) are also symmetric. Radius (which is the smallest eccentricity) is typically very close to the diameter, thus there are very few "peripheral" nodes which incur more communication cost than the rest of the nodes.

Multiple Independent Paths: Circular skip lists are nearly *optimally connected* [12]. Given a CSL with minimum degree, p_{min}, its edge connectivity, λ is at least, $p_{min} - 1$, when $k >= \frac{np}{2}$ (and $k >= np$, for directed graphs).

Recent studies have reported optimal networks with features such as: homogeneous degree [11], [14], betweeness [15], [16] and optimal connectivity [12]. We observe that CSLs are characterized by all these optimal features.

V. Future Directions

We study the design of optimal network topologies under different optimality requirements. Two prominent classes of topologies emerge as optimal: (1) scale-free networks and (2) circular skip lists. We analyze CSLs to show that they have structural properties that are optimal under a variety of requirements. Thus, CSLs potentially form the underpinnings of optimal network design.

This work is part of a larger vision, which is to develop a deeper theoretical understanding of network design in terms of general design principles. In that light, this work has lead to at least three future directions: (1) identifying structural signatures from the optimal networks that emerge, which can be used by autonomous agents to *snap* or quickly (re)construct networks in the face of perturbations (2) developing efficient determinisic techniques to construct topologies that are equivalent to the optimal circular skip lists (3) developing theoretical results that can provide guarantees about circular skip lists [31] in the phase of perturbations.

References

[1] A. E. Motter, Z. Toroczkai. *Introduction: Optimization in networks*. Chaos, Volume 17, Issue 2, pp. 026101-026101-3, 2007.

[2] A.-L. Barabsi. *Scale-Free Networks*. Scientific American, 288, pp. 60-69, May 2003.

[3] Steven H. Strogatz. *Exploring complex networks*. Nature 410, pp. 268276. doi:10.1038/35065725, 2001.

[4] R. Albert and A.-L. Barabsi. *Statistical mechanics of complex networks*. Review of Modern Physics, 74, pp. 4797, 2002.

[5] R. Albert, H. Jeong, and A.-L. Barabasi. *Diameter of the world-wide web*. Nature 401, pp. 130-131, 1999.

[6] D. J. Watts and S. H. Strogatz.*Collective dynamics of 'small-world' networks.*. Nature 393 (6684): 40910, doi:10.1038/30918, 1998.

[7] J. M. Kleinberg. *Navigation in a small world*. Nature 406: 845. doi:10.1038/35022643, 2000.

[8] M. E. J Newman. *Modularity and community structure in networks*. Proceedings of the National Academy of Sciences, 2006.

[9] R. Cohen, K. Erez, D. ben-Avraham and S. Havlin. *Resilience of the internet to random breakdowns*. Physical Review Letters, Vol. 85, No. 21, pp. 4626-4628, November 2000.

[10] R. Cohen, K. Erez, D. ben-Avraham and S. Havlin. *Breakdown of the internet under intentional attack*. Physical Review Letters, Vol. 86, Issue 16, pp. 3682-3685, April 2001.

[11] A. X. Valente, A. Sarkar and H. A. Stone. *Two-Peak and Three-Peak Optimal Complex Networks*. Physical Review Letters, Vol. 92, Issue 11, pp. 118702.1-118702.4, 2004.

[12] A. H. Dekker and B. D. Colbert, *Network Robustness and Graph Topology*. In proceedings of the 27th Australasian Computer Science Conference, 2004.

[13] D.S. Alberts, J. J. Garstka and F. P. Stein. *Network Centric Warfare: Developing and Leveraging Information Superiority*. CCRP Publ., 2nd Edition (Revised). Aug 1999, Second Print Feb 2000.

[14] L. Zhao, Y-C Lai, K. Park and N. Ye. *Onset of traffic congestion in complex networks*. Physical Review E, Vol. 71, Issue 2, 026125, 2005.

[15] R. Guimera, A. Arenas, A. Diaz-Guilera, F. Vega-Redondo and A. Cabrales. *Optimal network topologies for local search with congestion*. Physical Review Letters, Vol. 89, Issue 24, 2002.

[16] R. J. Mondragon C. *Optimal Networks, Congestion and Braess Paradox*. In proceedings of Inter-Perf, 2006.

[17] L Donetti, P.I Hurtado and M. A. Muoz. *Entangled Networks, Synchronization, and Optimal Network Topology*. Physical Review Letters, Vol. 95, Issue 18, id. 188701, 2005.

[18] L. Donetti, F. Neri and M. A. Muoz. *Optimal network topologies: Expanders, Cages, Ramanujan graphs, Entangled networks and all that*. Journal of Statistical Mechanics: Theory and Experiment, 2006.

[19] V. S. Agneeswaran and D. Janakiram. A Complex System Inspired Theoretical Formalism for Data Management in Peer-to-Peer Grids. In proceedings of IEEE International Symposium on Parallel and Distributed Systems, 2008.

[20] P. Bak. *How Nature Works: The Science of Self-Organized Criticality*. Springer Verlag, 1996.

[21] J. M. Carlson and J. Doyle. *Highly Optimized Tolerance: A Mechanism for Power Laws in Designed Systems*. Physical Review E, Vol. 60, Issue 2, pp. 14121427, 1999.

[22] D. Loguinov, J. Casas and X. Wang. *Graph-Theoretic Analysis of Structured Peer-to-Peer Systems: Routing Distances and Fault Resilience*. IEEE/ACM Transactions on Networking, Vol. 13, No. 5, pp. 1107-1120, October 2005.

[23] J. Xu, A. Kumar and X. Yu. *On the fundamental tradeoffs between routing table size and network diameter in peer-to-peer networks*. IEEE Journal of Selected Areas in Communication, Vol. 22, No. 1, pp. 151-163, January 2004.

[24] S. Ratnasamy, S. Shenker and I. Stoica. *Routing algorithms for DHTs: Some open questions*. In proceedings of IPTPS, 2002.

[25] C. Qu, W. Nejdl and M. Kriesell. *Cayley DHTs-A Group-Theoretic Framework for Analyzing DHTs Based on Cayley Graphs*. International Symposium on Parallel and Distributed Computing, 2004.

[26] K. Gummadi, R. Gummadi, S. Gribble, S. Ratnasamy, S. Shenker and I. Stoica. *The impact of dht routing geometry on resilience and proximity*. In proceedings of ACM SIGCOMM, Karlsruhe, Germany, August 2003.

[27] J. C. Bermond, E. Darrot, O. Delmas and S. Perennes. *Hamilton circuits in the directed wrapped Butterfly network*. Discrete Applied Mathematics, 1998.

[28] S. Patil, S. Srinivasa, S. Mukherjee, A. R. Rachakonda and V. Venkatasubramanian. *Breeding Diameter-Optimal Topologies for Distributed Indexes*. Complex Systems, Vol. 18, Issue 2(c), pp 175-194, 2009.

[29] V. Venkatasubramanian, S. R. Katare, P. R. Patkar and F. Mu. *Spontaneous emergence of complex optimal networks through evolutionary computation*. Computers and Chemical Engineering, Vol. 28, pp.1789-1798, 2004.

[30] S. Patil and S. Srinivasa. *Design and Analysis of Optimal Network Topologies*. Technical Report, Open Systems Laboratory, International Institute of Information Technology-Bangalore, March 2009.

[31] S. Patil and S. Srinivasa. *Theoretical Notes on Regular Graphs as Applies to Optimal Network Design.* Technical Report, Open Systems Laboratory, International Institute of Information Technology-Bangalore, February 2009.