Low Mean Internodal Distance Network Topologies and Simulated Annealing

Christopher Rose, *Member, IEEE*

Abstract—Networks have been found through simulated annealing with mean internodal distances lower than any previously reported for a given number of nodes N with a maximum of p outgoing links per node. These networks form the closest known approximations to Moore networks. However, the improvements in mean internodal distance obtained are relatively small (a few percent) and this improvement decreases rapidly with increasing p. Simulated annealing also seems to promote uniformity (over individual nodes) in mean distance to the rest of the network. This side-effect of the annealing process might be useful in regularizing other network attributes.

The annealing process also revealed that there exist neighborhoods of low mean internodal networks. This structure of the network space might be used in designing improved search algorithms for low mean internodal distance networks.

The relation between average nodal switching complexity and mean internodal distance was briefly explored through simulated annealing. Initial sharp decreases in mean internodal distance could be had through moderate increases in switching complexity. However, further decreases required much larger increases in average complexity.

I. INTRODUCTION

THE previous paper [1] has shown that semi-randomly constructed networks have surprisingly small mean distances between nodes. This property is useful in that smaller mean internodal distance implies smaller aggregate link capacity. Thus, in principal, less capacity need be allocated for networks with small mean internodal distances. This intriguing property of semi-random networks, however, suggests more subtle issues. Given a set of nodes with p or fewer outgoing links per node, what is the minimum mean internodal distance achievable. Or equivalently, how closely may the Moore bound [4], [1] be approached. Likewise, what is a good set of rules for constructing such a network?

One approach to this problem would be an exhaustive search of all possible networks with given N and L. For any reasonably sized network, however, the large number of possible networks which must be evaluated and the computational complexity of providing such an evaluation renders this method futile. Thus, a simulated annealing approach was taken whereby candidate networks were modified incrementally. At

each step, the modification was either adopted or not adopted with some probability depending upon whether the change lowered or increased the mean internodal distance. Of course, such a technique cannot be guaranteed to converge to an absolute minimum. Nonetheless the net effect of this process was to produce networks with lower mean internodal distances than the initial candidate.

Using simulated annealing, networks have been found with mean internodal distances lower than any previously reported for a given number of nodes N with a maximum of p outgoing links per node ($N \le 160, p \le 5$). These networks form the closest known approximations to Moore networks. However, the improvements in mean internodal distance obtained are relatively small (a few percent) and this improvement decreases rapidly with increasing p. This phenomenon was also observed in [6] and is therefore of primarily theoretical interest.

A possibly practical by-product of the simulated annealing process is its promotion of uniformity (over individual nodes) in mean distance to the rest of the network. This side-effect of the annealing process is useful in that it promotes "access fairness" over the network nodes even when the properties of the initial network are grossly unfair. Furthermore, this result suggests that an annealing approach might be useful when a given network attribute must be made uniform over the network nodes.

The results also suggested other theoretically interesting properties which were briefly investigated.

1) The space of all possible networks seems to have some structure with respect to mean internodal distance. Specifically, the annealing process revealed that networks with low mean internodal distance lie in local minima "basins". This result is rooted in the observation that the annealing process did not greatly alter either the mean internodal distance or the network structure. Conversely, networks with larger mean internodal distances were modified dramatically. These structural results might be used to comb the search space more efficiently.

2) There appears to be an inverse relationship between aggregate switching complexity of the network (as defined by the average number of switching crosspoints per node) and the mean internodal distance. Mean internodal distance decreases with increasing switching complexity. Starting from a uniform network with identical numbers of incoming and outgoing links, slight increases in complexity yield large decreases in mean internodal distance. However, further decreases are increasingly expensive in terms of complexity.
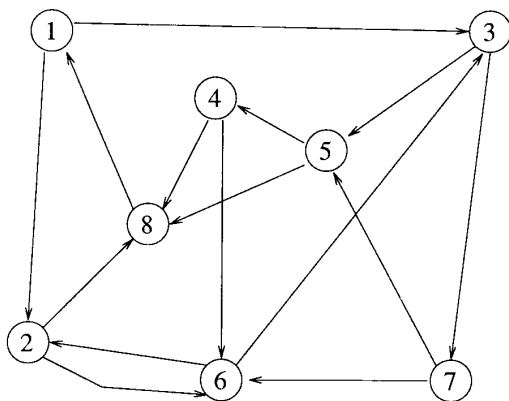
Fig. 1. Model network with 8 nodes and 15 directed links.

II. BACKGROUND

A. Network Representation

There are a number of ways to represent the connectivity of a network [1], [4]. The one which proves most useful here is the *link table*.

Consider a network of N nodes and L links. Any such network may be represented by numbering the nodes from 1 to N and then providing a table listing the connections, each line of the table corresponding to a node. So, for example, if node 1 has directed links to nodes 10 and 99, then the line corresponding to node 1 will have the entry 10,99. The total number of lines in the link table will be N and the total number of table entries will be L. The number of links emanating from a node (*out-degree*) is given by the number of link entries on the corresponding line of the table. The number of links impinging upon a node (*in-degree*) is given by the number of times the corresponding number appears in the table. This notation is compact and is amenable to computer manipulations.

B. Mean Internodal Distance

The formalisms of this topic are covered in detail in [1]. A short example is provided here to illustrate the calculation of mean internodal distance.

Consider the network depicted in Fig. 1. Its mean internodal distance $\bar{h}$ may be calculated as follows. Starting at node 1,[1] nodes 2 and 3 may be reached in one hop. Nodes 5, 6, 7 and 8 may in turn be reached from nodes 2 and 3 in one hop. Thus, 5, 6, 7, and 8 are reachable in two hops from node 1. Finally, nodes 1, 2, 3, 4, 5, 6, 7, and 8 are reachable from nodes 5, 6, 7, and 8. However, only node 4 has not been previously visited. Therefore only node 4 is called reachable in three hops.

To calculate the mean internodal distance first define h_{ik} as the number of nodes reachable in k hops starting at node i. Then define d_i as the number of hops necessary to reach the node(s) farthest from node i. Then, in the preceding example

[1]Node 1 itself is considered reachable in 0 hops.

we have

$$h_{10} = 1$$
$$h_{11} = 2$$
$$h_{12} = 4$$
$$h_{13} = 1$$
$$d_1 = 3. \tag{1}$$

The mean distance from node 1 to the rest of the network is then

$$\bar{h}_1 = \frac{h_{11} + 2h_{12} + 3h_{13}}{\sum\limits_{k=1}^{d_1} h_{1k}} = \frac{13}{8} = 1.625. \tag{2}$$

C. Why Simulated Annealing?

Initially, a direct approach to the problem of finding networks with the lowest mean internodal distances might be contemplated. Consider a set of N nodes and L links with out-degree fixed at $p = L/N$. The total number of networks T which may be constructed is

$$T = \binom{N}{p}^N. \tag{3}$$

Notice that even for small N and p, T is very large.[2] Thus, for any reasonable number of nodes, an exhaustive search for the best network is futile given modern computing methods.

III. METHODS: SIMULATED ANNEALING

Simulated annealing is a process whereby candidate solutions to a problem are repeatedly evaluated according to some objective function and incrementally changed to achieve better solutions [2]. The nature of each individual change is probabilistic in that there is some probability it worsens the solution. However, the probability is skewed so that changes which serve to minimize the objective function occur with higher frequency. In addition, an "annealing schedule" is followed whereby the probability of allowing a change which worsens the solution is gradually reduced to zero.

Thus, simulated annealing is similar to gradient descent algorithms wherein each change to the candidate solution is made such that the objective function is always minimized. The fundamental difference is that simulated annealing assumes no initial knowledge of the gradient. In addition, it allows escape from local minima by allowing changes which increase the objective function. These two properties form the allure of simulated annealing techniques.

The topography of the objective function is the primary determinant of whether simulated annealing will be effective. For example, simulated annealing is no better than random search on a surface which is completely featureless except for the minima such as shown in Fig. 2(a). In addition, surfaces for which the local minima have deep basins of attraction and the global minima shallow ones [Fig. 2(b)] will also confound simulated annealing. The surface topography best suited to

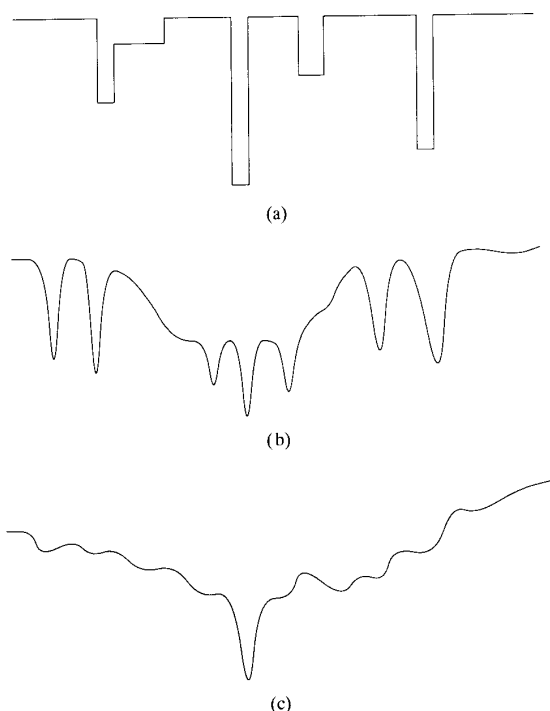[2]For $N = 8$ and $p = 2$, $T \approx 4 \times 10^{11}$.

Fig. 2. Examples of objective function variation for simulated annealing; (a) A "golf course" surface with no features except for objective function minima, (b) objective function with deep basins of attraction about suboptimal minima, (c) Objective function with shallow basins about suboptimal minima and a deep basin about the optimum.

simulated annealing is one in which the basins of attraction grow deeper with more globally optimal solutions as depicted in Fig. 2(c).

Unfortunately, the general topography of the objective function space is usually unknown. This is especially true in the case of mean internodal distance for which a small change in network topology such as moving one link could possibly render a network unconnected thereby producing an infinite change in the objective function.

Fortunately, for larger networks, the probability of producing a disconnected network by the moving of one link is small. In addition, good networks do tend to be "near" other good networks as measured by number of links changed. Empircally, simulated annealing seems to work reasonably well to find low mean internodal distance networks. The algorithm outlined below was used in this study.

IV. Annealing Algorithm: Link Redirection:

1) Choose a node (a line in the link table) at random.

2) Choose a link emanating from this node (an entry in the link list) at random.

3) Choose a target node at random and tentatively modify the network so that the chosen link impinges on the new target node.

4) Evaluate $\bar{h}$ for this tentative modification.

5) Conditionals:

- If the network is disconnected, restore the unmodified network and go to 1.
- If $\bar{h}$ is increased, adopt the change with probability inversely proportional to the size of the increase times some factor, $\kappa > 0$ which decreases toward 0 with increasing number of trials (corresponding to some "annealing schedule"), then go to 1.
- If $\bar{h}$ is decreased, adopt the change, then go to 1.

In all cases the following annealing schedule was used to reduce κ to near zero. This *ad hoc* schedule was chosen after some experimentation.

Annealing Schedule: Ad hoc:

1) $\kappa = .01/50{,}000$ trials.
2) $\kappa = .001/20{,}000$ trials.
3) $\kappa = .0001/15{,}000$ trials.
4) $\kappa = .00001/10{,}000$ trials.
5) $\kappa = .000001/5{,}000$ trials.
6) STOP.

IV. Results

A. Minimum $\bar{h}$ Networks

A reasonable heuristic for the choice of starting network is to use one with known good mean internodal distance properties. One network with mean internodal distances approaching the Moore limit [4] is the perfect shuffle [5] as shown in Fig. 3. For $N = 160$ and $p = 2^3$ a shufflenet [6] achieves an $\bar{h}$ of 6.03125 and a similar semi-random network [1] achieves an average $\bar{h}_{ave}$ of 6.4. The perfect shuffle, however, has an $\bar{h}$ of 5.76; substantially lower than either the semirandom network or the shufflenet. Thus, the perfect shuffle provides a good heuristic starting point for the annealing process.

The histogram of the mean distances from each node is provided in Fig. 4(a)–(c) for the Shufflenet, semirandom network and for the perfect shuffle. Since all nodes have identical characteristics in shufflenet, the associated histogram is an impulse at $\bar{h} = 6.03$. The histogram of the semi-random network is more diffuse and has a mean width of $\sigma_{\bar{h}} = 0.194$ about $\bar{h} = 6.26$. The histogram of the perfect shuffle is broader still with a mean width of $\sigma_{\bar{h}} = 0.262$. This relatively large variation across nodes is characteristic of the unannealed perfect shuffle and to a lesser extent, the unannealed semirandom network. Thus, for the perfect shuffle and semi-random topologies, access to the network varies from node to node.[4] For Shufflenet, access from each node is identical.

Five annealing trials were run starting from the perfect shuffle network with $N = 160$ $p = 2$. Another five control trials were run starting from semirandom networks. Using the semirandom network, the lowest mean internodal distance $\bar{h}_{min}$ achieved was 5.74. Using the perfect shuffle network, an $\bar{h}_{min} = 5.61$ was achieved and represents the lowest known mean internodal distance for a network with $N = 160$ and $p = 2$ and a reduction in $\bar{h}$ of 2.6%. This result seems

---

[3] A single annealing run for the $N = 160$ $p = 2$ network required many hours on a CRAY X-MP computer system. The computation time varied empirically as N^2. Thus, this study is limited to networks of size $N \leq 160$.

[4] This variation is most pronounced for $p = 2$ and decreases quickly with increasing p.
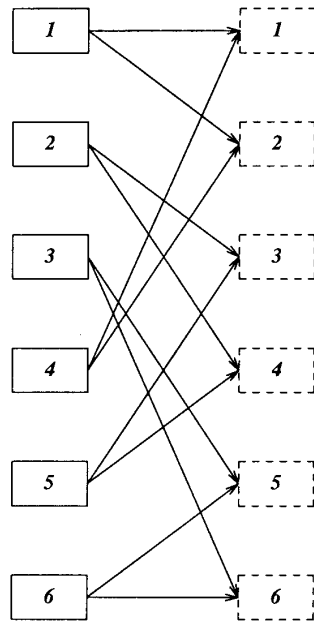
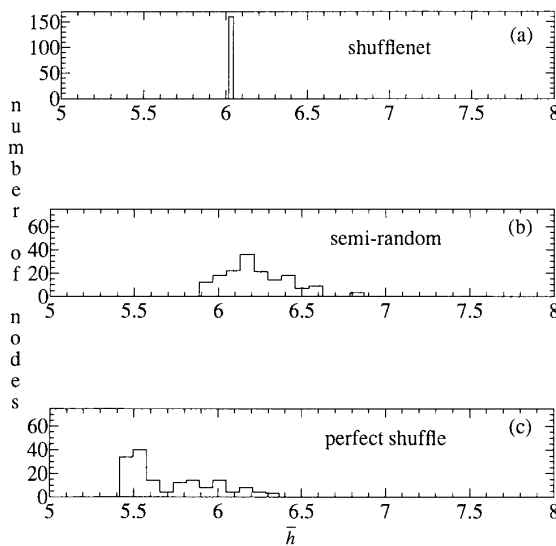Fig. 3.   The perfect shuffle network with $N = 6$ nodes and $L = Np = 12$ links.

Fig. 4.   Histograms of mean internodal distance for each network node. (a) Shufflenet, (b) connected semirandom network, (c) perfect shuffle network.

to corroborate the assumption that networks with low $\bar{h}$ lie in some sense "near" one another in the space of possible networks.[5]

A more important feature is illustrated in Fig. 5(a) and (b) wherein histograms of the mean distances from each node for the annealed semirandom $\bar{h} = 5.74$ network and the $\bar{h} = 5.61$ annealed perfect shuffle network are shown. Notice that the variation between nodes is much reduced from that seen in

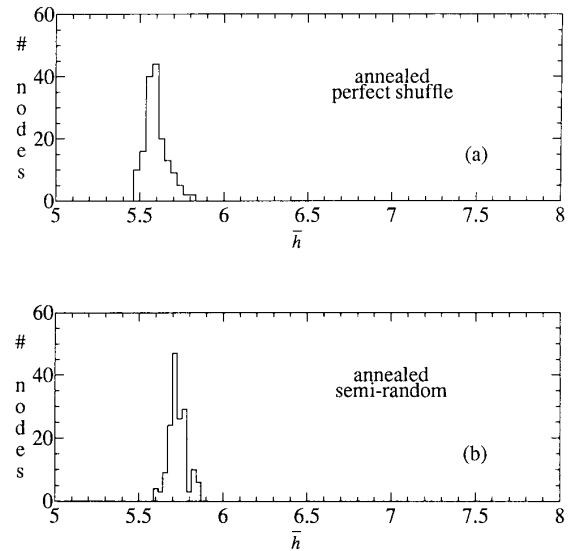[5]For comparison, the Moore bound for such a network is $\bar{h} = 5.456$.

Fig. 5.   Histograms of mean internodal distance for each network node after annealing; (a) Annealed perfect shuffle, (b) annealed connected semirandom network.

Fig. 4(b) and (c). It seems that the annealing process serves to equalize the initially uneven distribution of mean distances.

In Fig. 6 the minimum $\bar{h}$ found through the annealing process is plotted versus network size N for various p. These results are compared to both the comparable perfect shuffle network and the Moore bound.[6] It may be seen that some reduction in mean internodal distance is achieved by the annealing process. For $p = 2$ and $N \leq 160$, an average of 4.6% reduction was noted as compared to the perfect shuffle starting network. The magnitude of these reductions, however, decreases markedly with increasing p. For $p = 3, 4, 5$, the average reduction was only 1.9%, 1.5%, 1.6%, respectively. Thus, for $p > 2$, simulated annealing is unable to achieve much lower $\bar{h}$ than the perfect shuffle network.

However, such a result is seen to be inevitable when $\bar{h}$ is examined as a function of p for both the Moore bound and shufflenet for fixed N (Fig. 7). The $\bar{h}$ of the perfect shuffle and the Moore bound converge with increasing p. Thus, a decrease in the relative efficacy of simulated annealing would be expected when p is large and the starting network is the perfect shuffle. In short, when p is large, the perfect shuffle is near-optimal in mean internodal distance so that further optimization has little effect.

However, the annealing process also promotes uniformity of network properties as may be seen in Fig. 8 where $\sigma_{\bar{h}}$ is plotted as a function of network size N for various p. Once again, the most pronounced reductions in $\sigma_{\bar{h}}$ are seen for $p = 2$ with substantially lower reductions with increasing p.

B. Mean Internodal Distance and Switching Complexity

In the previous section, only networks with constant out-degree (number of links emanating from a node) were studied.

[6]Shufflenet was included only for $p = 2$ in this comparison owing to its limited range of network sizes for $N \leq 160$ with $p > 2$.
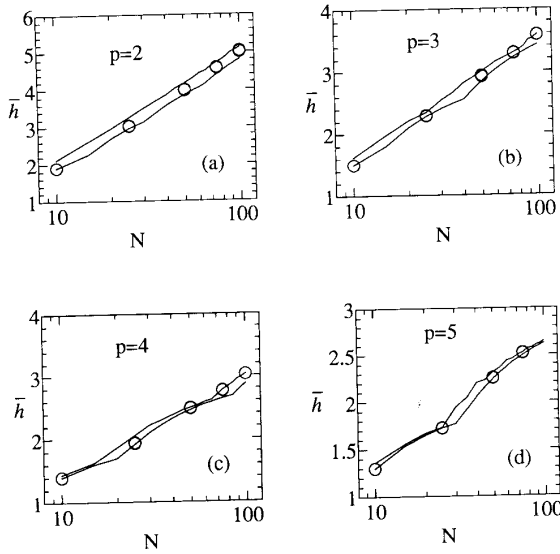
Fig. 6. Superimposed plots of $\bar{h}$ versus N for the perfect shuffle network (solid line, upper bound), annealed perfect shuffle networks (circles) and the Moore bound (solid line, lower bound) for, (a) $p = 2$, (b) $p = 3$, (c) $p = 4$, (d) $p = 5$. Notice the tight distribution of annealed network data (circles).
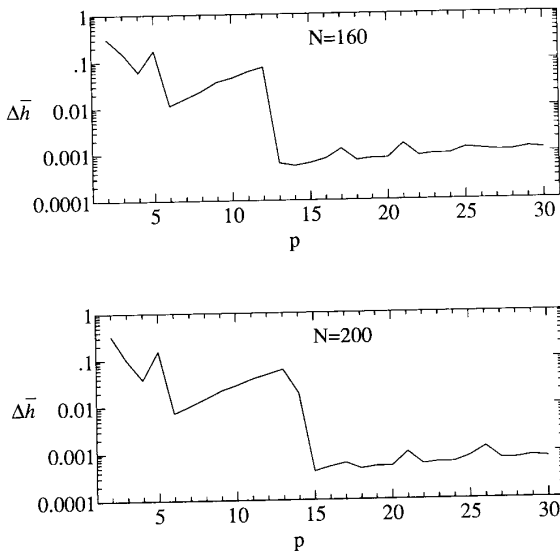

Fig. 8. Superimposed plots of $\sigma_{\bar{h}}$ versus N for the perfect shuffle network (solid lines) and annealed perfect shuffle networks (circles) for, (a) $p = 2$, (b) $p = 3$, (c) $p = 4$, (d) $p = 5$. Notice the tight distribution of annealed network data (circles).

Fig. 7. A plot of the difference in $\bar{h}$ ($\Delta\bar{h}$) between the perfect shuffle network and the Moore bound versus p.

Fig. 9. An N-node star network.

The choice was made for two reasons.

1) The Moore bound is only valid for networks with p or fewer outgoing links per node.

2) Low nodal in/out degree seems to require only a low level of "switching complexity."

Point 1 is structural and obvious. Point 2 deserves further attention.

Consider the so-called "star" network depicted in Fig. 9. Most of the nodes have exactly one incoming and outgoing link. The central node has $N - 1$ incoming and outgoing links. While such a network obviously violates the conditions for
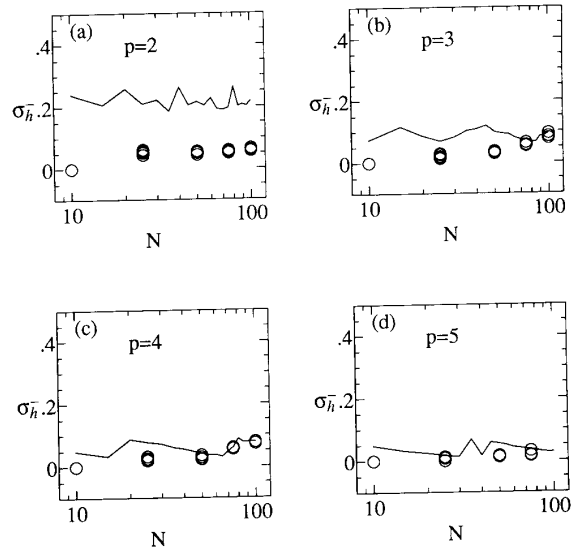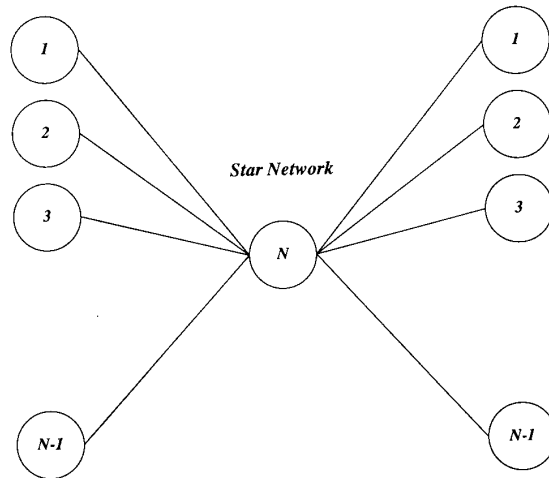
which the Moore bound is valid (unless $p \geq N - 1$), with $L = 2(N - 1)$ links this network topology is able to achieve a mean internodal distance of less than two.[7] In limited trials of an annealing algorithm which allowed the out-degree of the nodes to vary, a star-like pattern was invariably produced with the concomitant impressive decrease in mean internodal distance. However, the central star node must be sufficiently complex to switch the large number of incoming lines to outgoing lines. By comparison, each node in the perfect shuffle network need only switch a few incoming to outgoing lines.

In order to obtain some quantitative feel for the relationship between switching complexity and mean internodal distance a limited set of experiments were performed. First, the com-
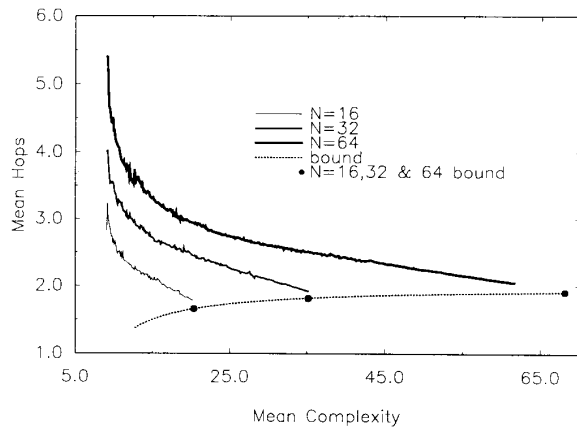
[7] $\bar{h} = 2(N - 1)^2/N^2$.

Fig. 10. Minimum mean internodal distance $\bar{h}$ versus average nodal complexity, $\bar{c}$ for networks with $N = 16, 32$ and $L = 2N$ links. See text for further description.

plexity of a switching node was defined to be the product of its *total* in-degree and its out-degree.[8] Second, several different networks types with $2N$ links were subjected to the modified annealing process wherein the outdegree of each node was unconstrained. At each step of this annealing process the average nodal complexity was recorded. The annealing parameter κ was varied manually so as to obtain a large number of networks covering a wide range of $\bar{h}$ and $\bar{c}$. Between 10 000 and 25 000 trials were performed on each candidate network.

A distillation of the results are shown in Fig. 10 where *minimum* $\bar{h}$ is plotted against the average network complexity, $\bar{c}$. These minimum values are defined as the smallest values of $\bar{h}$ obtained by the annealing process at a given level of complexity. The asymptotic minimum $\bar{h}$ versus complexity (parameterized by N) of star networks is also shown.

As expected, $\bar{h}$ decreases with increasing switch complexity, $\bar{c}$. In each case, after an initial rapid decline in $\bar{h}$, the rate at which $\bar{h}$ declines approaches linear. These results suggest that initially the $\bar{h}$ of a low $\bar{c}$ network can be efficiently decreased by increasing $\bar{c}$. However, after this initial sharp decline in $\bar{h}$, further decreases require larger increases in complexity.

C. Contiguity of Good Solutions

The use of simulated annealing assumes some gross form of continuity in the search space. Specifically, it is assumed that small changes to the network produce small changes in the objective function by which the network is evaluated. Although this is not strictly true, it is sufficiently true that the simulated annealing process is effective in reducing the $\bar{h}$ of different candidate networks.

An implication of the small-modification/small-change assumption is the "nearness" of good solutions to other good solutions. Some evidence in support of this contention is provided visually in Fig. 11 where the graphs of several

[8] Each node is assumed to have a single source and sink attached. Thus, the nodes of a unidirectional ring network have effective in-degrees/out-degrees of 2 rather than 1 (and therefore a complexity of 4).
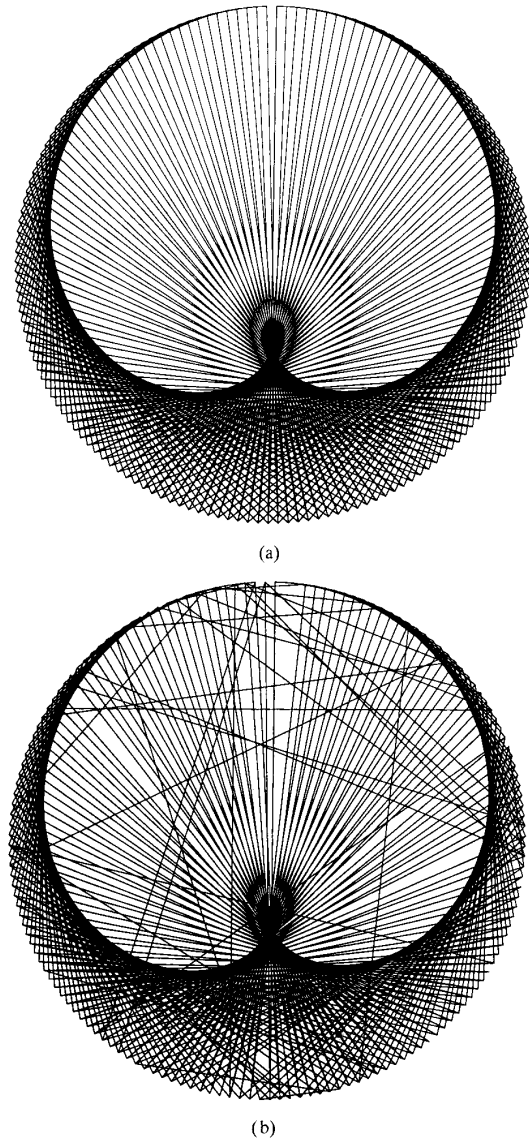

(a)

(b)

Fig. 11. Graphical comparison of networks before and after annealing for $N = 160$ and $p = 2$. The nodes are equally spaced around the circumference of a circle and the links are as drawn (without regard to direction); (a) perfect shuffle network $(\bar{h} = 5.76)$, (b) annealed perfect shuffle network $(\bar{h} = 5.61)$, (c) Shufflenet $(\bar{h} = 6.03)$, (d) annealed Shufflenet $(\bar{h} = 5.75)$, (e) bidirectional ring network $(\bar{h} = 40)$, (f) annealed bidirectional ring network $(\bar{h} = 5.77)$.

networks are shown both before and after annealing. However, since such visual depictions are highly subjective, a table showing the number of links modified by the annealing process for each network is given in addition (Table I).

Although each network underwent changes which noticeably decreased $\bar{h}$, the end result is strikingly similar to the initial network in the cases of Shufflenet and the perfect shuffle. In contrast, consider the bidirectional ring network with large $\bar{h}$. The annealing process grossly changed the appearance of the connection graph. Thus, simulated annealing, as applied
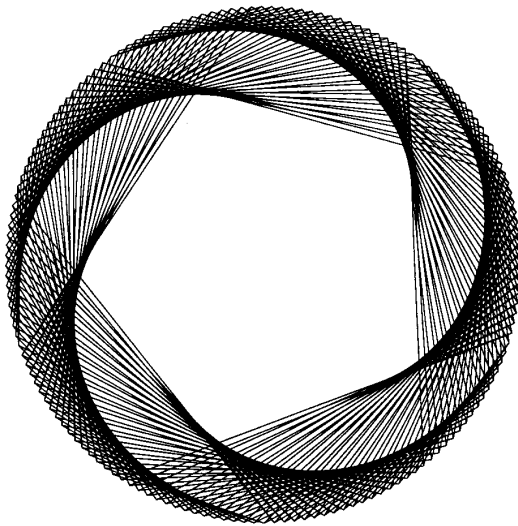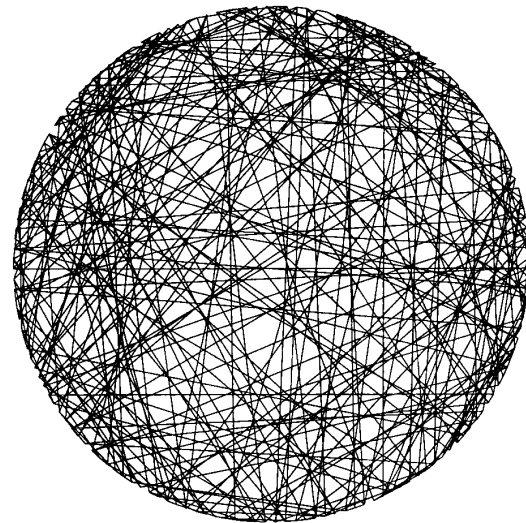
(c)
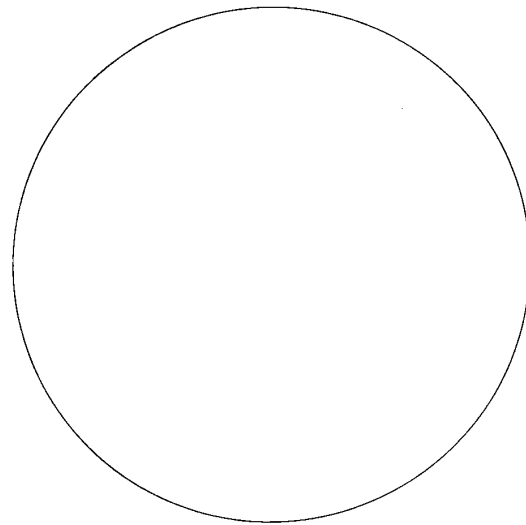
(e)
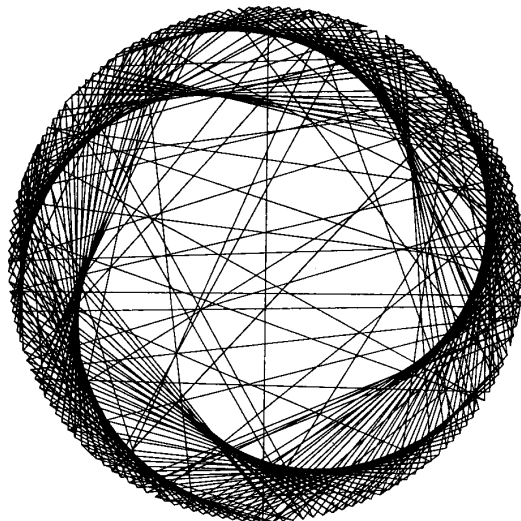
(d)

Fig. 11. (cont.)

(f)

Fig. 11. (cont.)

here, incrementally changes initially good solutions (low $\bar{h}$ networks), whereas initially poor solutions (high $\bar{h}$ networks) are grossly rearranged.

V. DISCUSSION AND CONCLUSIONS

A. Approaching the Moore Bound

For networks whose nodes have a maximum of p outgoing links extremely low mean distance between nodes were found using simulated annealing. The annealing process is most effective in reducing $\bar{h}$ for $p = 2$ and its efficacy decreases monotonically with increasing p since the perfect shuffle starting network becomes near-optimal with increasing p. Regardless, the networks found represent the lowest $\bar{h}$ networks currently known. Equivalently, these networks also

represent the closest known approximations to the Moore network [4], [1]. Since low $\bar{h}$ minimizes both the delay and aggregate link capacity in multihop store and forward networks [1] these networks are almost optimal in their utilization of aggregate capacity.

B. Properties of Simulated Annealing

Since there are a very large number of possible networks given N nodes and L links, an exhaustive search for the networks with lowest mean internodal distance is futile using currently available computers. Simulated annealing provided a systematic method for the identification of low mean internodal distance networks, and indirectly gave some information about the structure of the solution space. As such, an initial exploration using simulated annealing could in principle lead to

TABLE I
A TABULATION OF THE NUMBER OF LINKS MODIFIED BY THE
ANNEALING PROCESS FOR THE PERFECT SHUFFLE, SHUFFLENET, AND
THE BIDIRECTIONAL RING NETWORKS WITH $N = 160$ AND $p = 2$

Network Type	$\Delta \bar{h}$	% links changed
Perfect Shuffle	2.6%	9.7%
Shufflenet	4.5%	22.5%
Bidirectional Ring	85.5%	80.6%

better search algorithms. Furthermore, since all that is required is the evaluation of some objective function, the simulated annealing process could find utility in the optimization of other network attributes such as reliability or throughput/delay performance. As an added benefit, the annealing process makes network properties more uniform over the nodes.

However, there are limitations to the annealing process. Primarily, a globally optimal solution cannot be guaranteed owing to the stochastic nature of the process. In addition, the types of solution obtained are heavily dependent on the modification rule used. Another drawback of simulated annealing is the increasing complexity of evaluating the objective function $\bar{h}$ with increasing number of nodes and links. For this reason, only networks with $N \le 160$ and smaller $p = (N/L) \le 4$ were considered here.

C. Complexity versus $\bar{h}$

This abbreviated study showed that the mean internodal distance of a regular network with uniform in/out-degree can be initially lowered by increasing the average switching complexity slightly. Switching complexity is defined as the average number of switching crosspoints per node (assuming crossbar type switches). Further decreases in $\bar{h}$ require larger increase in complexity where these increases in complexity are caused by enlarging the in-degrees and out-degrees of some nodes. Thus, the highest complexity/lowest $\bar{h}$ networks are star-like in structure.

Of course, these tentative observations and conclusions are highly dependent upon the measure of "switch complexity" used. However, the basic idea of using spatially diffuse structures so that high concentrations of switch complexity may be avoided at single nodes is one of the basic tenets of distributed switching. Thus, even if for some measure of complexity the *aggregate switching complexity* of a centralized network (such as the star) is *smaller* than that of a more distributed topology (such as the perfect shuffle), the issue may be one of realizability. If it is much simpler to build many smaller switching units than it is to build one large unit, then a diffuse architecture where complexity is distributed over all the network nodes is preferable. Maintaining the in/out-degree of each node at some manageable level is then certainly in order. Regardless, the relationship of $\bar{h}$ to maximum nodal switch complexity is a subject worthy of further investigation.

REFERENCES

[1] C. Rose, "Mean internodal distance in regular and random multihop networks," *IEEE Trans. Commun.*, see this issue, pp. 1310–1318.
[2] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," *Sci.*, vol. 220, pp. 671–680, 1983.
[3] J. L. Gross and T. W. Tucker, *Topological Graph Theory.* New York: Wiley-Interscience, 1987.
[4] W. G. Bridges and S. J. Toueg, "On the impossibility of directed Moore graphs," *Combinatorial Theory (B)*, vol. 29, pp. 339–341, 1980.
[5] H. S. Stone, "Parallel processing with the perfect shuffle," *IEEE Trans. Comput.*, vol. C-20, pp. 153–161, 1971.
[6] A. S. Acampora, M. J. Karol, and M. G. Hluchyj, "Terabit lightwave networks: The multihop approach," *AT&T Tech. J.*, vol. 66, no. 6, pp. 21–34, 1987.

Christopher Rose (M'92) for a photograph and biography, see this issue, p. 1318.