Some Traffic Characteristics of Communications Networks with Automatic Alternate Routing

By J. H. WEBER

(Manuscript received August 23, 1961)

As a first step in the investigation of communications networks with automatic alternate routing, a simulator has been prepared using the IBM 7090 high-speed digital computer. The simulator is capable of being applied to a large class of networks, the principal restrictions being that blocked calls are cleared, and no congestion or delay is encountered at the switching points. Although the first version of the simulation program requires that the alternate routing plan be fixed in advance (i.e., before a run), the program design is such that traffic-dependent alternate routing doctrines can easily be provided.

The simulator has so far been used to examine the behavior of small networks of various sizes, configurations, and alternate routing doctrines under normal and abnormal conditions of load. Several criteria are introduced and used to evaluate the relative performance of different networks, leading to conclusions regarding the merits of certain alternate routing procedures and the areas of profitable application of the networks studied. The overload capabilities of these networks are of particular interest and are examined in some detail.

I. INTRODUCTION

The recent rapid expansion of long-distance communications facilities to serve increasing civilian and military demands, along with the evolution of cheaper trunking facilities and more sophisticated switching techniques, continues to bring the problem of network design and engineering to the attention of communications engineers. Although methods have been developed for engineering certain types of networks for the most economical distribution of trunking facilities, several critical problems remain.

One of these is the lack of understanding of the behavior of alternate routing networks under overload conditions, whether the overload is local or system wide. Local disasters, such as storms, earthquakes, etc., have caused severe deterioration of service in certain regions owing to increased loads directed toward the affected area. At other times, such as Christmas Day in the United States, the pattern of traffic shifts radically, again causing serious overloads and long delays in completing calls. Finally, some concern is felt for the behavior of the system under the impact of some widespread disaster, where overloads may appear everywhere simultaneously.

Such considerations lead in turn to two questions. First, how shall networks be designed to be efficient during normal operation and yet not deteriorate catastrophically under overloads, and second, given a network design, can the switching pattern be altered for the duration of an overload to improve performance, and if so, how?

Another problem is our present inability to engineer any but the limited class of alternate routing networks of a "hierarchical" nature which have been widely used in the Bell System and elsewhere.

Since no analytic techniques appeared to be available or soon forthcoming to answer these questions, a simulation study was undertaken in the hope that some insights might be provided into the operation of such networks which would be helpful in their design and in the development of theoretical models to predict their characteristics.

Accordingly, a program was written for the IBM 7090 computer which enables various alternate routing philosophies to be simulated and compared. In line with the general nature of the problem being studied, the program was designed to accept a large variety of networks and be easily expandable to encompass more sophisticated alternate routing procedures as they evolve.

The capabilities and limitations of the simulator are outlined in some detail in the next section, followed by a description of the first experiment using the program. Finally the results are presented and analyzed, and some general characteristics of alternate routing networks of the types studied are set forth.

II. SIMULATOR CHARACTERISTICS

Although many of the problems which arise when alternate routing networks are overloaded are caused by switching delays and shortages, it was decided, as a first step, to consider only the effects of trunking, since the switching problems are unique to particular systems, and would in any event considerably complicate both the program and interpretation of the results. Accordingly, the program was constructed under the following restrictions:

(1) No blocking or delay is introduced by any switching point.

(2) Calls which do not receive service immediately are cleared from the system and do not return. (If setup time is assumed negligible, and there are no delays, the lack of retrials is not likely to materially affect the nature of the results.)

If the network is considered to consist of nodes (corresponding to switching centers) and links connecting them (corresponding to direct trunk groups), then each link may be assigned an originating traffic, a trunk group size and an alternate routing pattern. In addition, calls which overflow the direct route and are to be alternate routed may be assigned a directionality, or originating node, which allows one of two alternate routing configurations to be hunted over, depending upon the direction of the call. Every trunk group is a "two way" group, so no direction need be assigned to calls which are carried on the direct route.

The simulator will accept systems which contain as many as 63 switching points, each of which may be connected to any other switching point by up to 511 trunks. Calls which do not find an available trunk in the direct route may overflow through one of two sets of up to 63 specified routes, depending upon the direction of the call. Each alternate route may contain as many as 7 links, which implies switching through up to 6 intermediate nodes. (A modification of the program allows the alternate routes to be chosen on a "step by step" basis, where the first node in the alternate route chain is specified, and the call proceeds from node to node according to the alternate routing specification at the last node through which the call was switched. "Ring-around-the-Rosy" and "Shuttling" are prevented by keeping records of where the call has already been switched and not allowing it to use the same node twice.) It should be emphasized that the program as described here requires that the entire alternate route be specified at the originating link, and failure to connect on any link of an alternate route allows an entirely new route to be selected.

An over-all maximum size of the system, set by the limitations of the computer memory, is

$$(13 \times \text{Number of links}) + \text{Total number of trunks} + \text{Total number of alternate routes} = 22{,}013.$$

For example, if a system has 40 nodes, (and therefore 780 links), and if there is a total of 3900 trunks in the system, (corresponding to an average of five trunks per link), then 7973 alternate routes, or about 10 per link, can be specified. This is a rather large system, and in fact the simulator allows for experimentation with a wide range of possible trunk, node, and alternate route configurations.

In order to estimate the reliability of the simulation results, outputs may be printed out at subintervals of the run. Furthermore, since the system starts empty, equilibrium may be attained before records are kept by running the program for a number of subintervals and discarding their records. The number of subintervals to be printed out, as well as the number to be discarded, can be specified in the input, along with the average holding time per call, the total time the simulation should be run in holding times, and indications as to what sort of alternate routing scheme is to be used. The holding times of all calls are assumed to be exponentially distributed with identical means, and traffic levels are varied by altering the average time between calls offered to each link. Pseudo-random numbers to specify the input are generated by a multiplicative congruential technique which gives a cycle of 2^{33} numbers before a repeat. The random number generator is not cleared after every simulation, so that if several experiments are made successively, they will not utilize the same sequence of random numbers. Thus runs can be repeated identically if desired, or, alternatively, a different set of random numbers can be used for the same system configuration by the simple expedient of reordering the experiments.

The information which is printed out, in addition to that derived directly from the inputs, (number of nodes, number of trunks and loads per link, alternate routing patterns, etc.) is as follows:

(1) An estimate of the probability of loss (blocking) from each link; i.e., the proportion of calls directly offered to a specific link which are unable to be served at all.

(2) An estimate of the probability of direct overflow; i.e., the proportion of calls which overflow the direct route, although they may be served on an alternate route.

(3) Number of calls offered to each link, both directly and as an alternate route.

(4) Load in erlangs carried by each link, both from direct and alternate routed traffic.

(5) Calls carried by each link, both from direct and alternate routed traffic.

(6) Over-all average blocking; i.e., $\sum_{i=1}^{n} a_i p_i / \sum_{i=1}^{n} a_i$ where a_i is the load offered to link i, and p_i is the blocking experienced by a_i.

(7) Total carried load, which is really total calls multiplied by holding times. That is, a call is assumed to provide the number of erlangs its holding time would represent, regardless of how many links are used. This quantity is derived indirectly, by multiplying the total offered load by one minus the overall average blocking.

For moderate sized systems this program will process calls at the rate of about 1,200,000 calls per hour. A sample output is shown in Fig. 1.

NETWØRK SIMULATIØN

RESULTS FØR FIRST 5 5 THS

NUMBER ØF NØDES	AVERAGE HØLDING TIME	ALTERNATE RØUTING PLAN	NUMBER ØF HØLD TIMES	INITIALIZING SUBGRØUPS
5	1000.00	1	200.	1

LINK NUM	ØFFERD LØAD	NUM TKS	LØSS PRØB	PR DIR ØVRFLØ	C ØFF CALLS	A ØFF CALLS	T ØFF CALLS	DIRECT CAR LD	ALTERN CAR LD	TØTAL CAR LD	D CAR CALLS	A CAR CALLS	T CAR CALLS
1 2	8.750	3	0.2336	0.7045	1841	0	1841	2.66	0.	2.66	544	0	544
1 3	61.250	57	0.2889	0.2889	12648	5851	18459	43.59	11.17	54.76	8994	2326	11320
1 4	26.250	15	0.3330	0.5645	5442	2665	8107	11.57	2.79	14.36	2370	616	2986
1 5	17.500	9	0.2777	0.5836	3475	985	4460	7.26	1.10	8.36	1447	248	1695
2 3	35.000	19	0.2837	0.5342	7273	1030	8303	16.54	1.62	18.16	3388	297	3685
2 4	43.750	25	0.3478	0.5479	8829	4348	13177	19.67	4.62	24.29	3992	965	4957
2 5	70.000	64	C.3395	0.3395	14454	9256	23710	46.58	15.55	62.13	9547	3193	12740
3 4	78.750	75	0.3714	0.3714	16409	14051	30460	50.97	22.42	73.39	1C315	4548	14863
3 5	52.500	31	0.3514	0.5482	1C598	5725	16323	23.61	6.58	30.20	4788	1389	6177
4 5	87.500	81	0.3583	0.3583	18C43	15297	33340	56.17	23.23	79.41	11578	4705	16283

ØVERALL AVERAGE BLØCKING = 0.335161

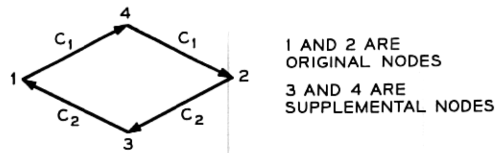
TØTAL CARRIED LØAD = 319.95

ALTERNATE RØUTE PATTERN

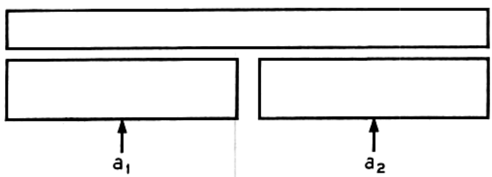
```
                         LINK NUMBER  1  2, FIRST DIRECTIØN TRAFFIC  50 PER CENT
5 0 0 0 0 0    4 0 0 0 0 0    4 5 0 0 0 0    3 0 0 0 0 0    3 5 0 0 0 0
3 4 0 0 0 0    3 4 5 0 0 0

3 0 0 0 0 0    4 0 0 0 0 0    3 4 0 0 0 0    5 0 0 0 0 0    3 5 0 0 0 0
4 5 0 0 0 0    3 4 5 0 0 0

                         LINK NUMBER  1  3, FIRST DIRECTIØN TRAFFIC   0 PER CENT

                         LINK NUMBER  1  4, FIRST DIRECTIØN TRAFFIC 100 PER CENT
3 0 0 0 0 0

                         LINK NUMBER  1  5, FIRST DIRECTIØN TRAFFIC  50 PER CENT
4 0 0 0 0 0    3 0 0 0 0 0    3 4 0 0 0 0
3 0 0 0 0 0    4 0 0 0 0 0    3 4 0 0 0 0

                         LINK NUMBER  2  3, FIRST DIRECTIØN TRAFFIC  50 PER CENT
4 0 0 0 0 0    5 0 0 0 0 0    5 4 0 0 0 0
5 0 0 0 0 0    4 0 0 0 0 0    5 4 0 0 0 0

                         LINK NUMBER  2  4, FIRST DIRECTIØN TRAFFIC 100 PER CENT
5 0 0 0 0 0

                         LINK NUMBER  2  5, FIRST DIRECTIØN TRAFFIC   0 PER CENT

                         LINK NUMBER  3  4, FIRST DIRECTIØN TRAFFIC   0 PER CENT

                         LINK NUMBER  3  5, FIRST DIRECTIØN TRAFFIC 100 PER CENT
4 0 0 0 0 0

                         LINK NUMBER  4  5, FIRST DIRECTIØN TRAFFIC   0 PER CENT
```

Fig. 1 — Sample computer output.

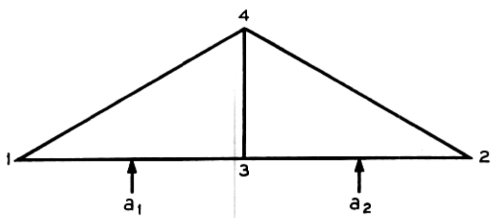
Clearly, the flexible nodal structure upon which the program is based allows certain things to be done by expending extra nodes which are not directly programmed. For example, if it is desired to have two one-way groups in a link, it is necessary merely to assign no trunks to the direct link, and have each of the alternate route patterns contain one node, which is assigned for the purpose. So, a single link would then have 4 nodes and 4 links with trunks assigned, as shown below.



Another possible use is the simulation of progressive graded multiples. Suppose it is desired to simulate the simple multiple shown below.



This is clearly equivalent in terms of loads carried and blocking to the following nodal structure.



In this analogue, a_1 has links 1-4 and 3-4 as an alternate route, while a_2 overflows through 2-4 and 3-4. Thus, if links 1-4 and 2-4 are provided with more trunks than 3-4, they can introduce no blocking and the system corresponds to the graded multiple above where link 3-4 is equivalent to the common group. This sort of flexible structure, then, appears to be useful in many ways, and may in fact come to have application beyond its original intent.

This program was primarily designed as a tool for the evaluation of alternate routing networks and as an aid in formulating principles for their design and administration, although one of the purposes was to assist in the solution of real problems as they arose. Accordingly, studies have been begun with the aim of codifying classes of networks and determining the significant parameters, advantages and disadvantages of each.

III. ANALYSIS OF NETWORKS

As a first step in the study of alternate routing communications networks it is desired to compare the behavior of several alternative configurations under normal and overload conditions. The variables which seem most likely to be significant in determining network performance are:

(1) Number of switching points in the network.
(2) Overall size of the network, perhaps best described as a measure of the average load or number of trunks per link.
(3) Alternate routing procedure used. Thus, a system which allows all traffic to overflow in some specified manner will probably perform differently than one which considers some routes to be "high usage" and from which traffic is alternate routed, and others to be "finals," from which no alternate routing is permitted.
(4) Type of overload encountered. A uniform (system wide) overload, for example, may cause a behavior quite different from an overload which is confined to a particular portion of the network.

In order to estimate the performance of networks when these parameters vary, and yet keep the results simple enough so that they can be easily interpreted, eight different networks were studied, each having two different alternate routing plans. In each case both uniform and nonuniform overloads were considered.

The eight networks studied consisted of two networks with three, two with four, two with five, and two with six nodes. The loads were adjusted so that the average load per link varied from three and one half erlangs per link in the most lightly loaded network to about 28 erlangs per link in the most heavily loaded configuration.

For purposes of convenience, the following terminology will subsequently be used:

(1) A *link* is a connection between two nodes, which may have any number of trunks, including zero.
(2) A *node* is a switching center, characterized by two or more links terminating at it.
(3) If network A is *larger* than network B, it has more nodes.
(4) If network A is *heavier* than network B, it has more offered erlangs per link on the average.

(5) A *hierarchical* alternate route network is one in which at least some of the trunk groups are *high usage*; i.e., traffic which is not carried can be overflowed to other groups, at least some of which are *finals*, which have no alternate routes.
(6) A *symmetrical* alternate route network has only high usage groups.
(7) A *simple* network has only final routes; i.e., no alternate routing is allowed.

The procedure which was followed in all cases was to postulate loads offered to each link in the network. For the sake of generality, these loads are ordinarily unequal, although in some cases equal loads are used in places where it is thought that this will not prejudice the results. Each network was engineered for a loss of 1 per cent on the worst link for simple, symmetrical and hierarchical networks. Loads were then applied corresponding to (a) 25 per cent, 50 per cent, 75 per cent, and 100 per cent uniform overload and (b) 25 per cent, 50 per cent, 75 per cent, and 100 per cent overload on all routes terminating at node 1. The selection of node 1, of course, is quite arbitrary, but this choice appears to be immaterial in the symmetrical case, and is likely to be most typical for hierarchical networks. (The heaviest loads in the hierarchical networks were reserved for the final routes, since this will allow most effective use of the hierarchy and seems to correspond to actual practice.) The simulations ran for 200 holding times for heavy networks and 1000 holding times for light networks, with an additional 20 per cent of this time (i.e., 40 or 200 holding times) discarded at the beginning of each run to remove the initial transient. Results were printed out at five subintervals of the total run, and examined to determine that the initial transient had been removed and the run was long enough to yield results sufficiently accurate for the purposes of this study.

Sketches of the networks are shown in Figs. 2 to 9 along with tables indicating the link loads and trunks assigned for each of the alternate routing doctrines. (Two sketches of each network are provided, one of which can be easily related to a symmetrical alternate routing philosophy while the other suggests a hierarchical doctrine. The dashed lines in the hierarchical sketch denote high usage groups, while the solid lines represent final routes. Since all links are high usage in the symmetrical system this distinction is not needed, and identical solid lines were used throughout.) The simple networks were engineered using the Erlang B tables, while the hierarchical networks were engineered using conventional methods with the sort of hierararchy used in the Bell System, allowing about 0.7 erlangs (25 ccs) on the last trunk in a high usage group. The (hierarchical) configurations were then checked experimentally using the simulator, and adjustments were made where required. The symmetrical

Network Parameters, 3 Nodes — Light

Link Number	Engineered Loads-Erlangs	Engineered Trunks		
		Simple	Symmetrical	Hierarchical
1–2	2	7	6	1
1–3	4	10	8	13
2–3	6	13	11	16
Total	12	30	25	30
Ave/Link	4	10	8.35	10

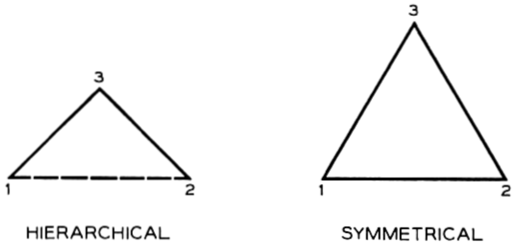


Fig. 2 — Three-node network models with table of link loads and trunk assignments for light loading.

Network Parameters, 3 Nodes — Heavy

Link Number	Engineered Loads-Erlangs	Engineered Trunks		
		Simple	Symmetrical	Hierarchical
1–2	5	11	11	4
1–3	10	18	16	21
2–3	15	24	21	27
Total	30	53	48	52
Ave/Link	10	17.65	16	17.35

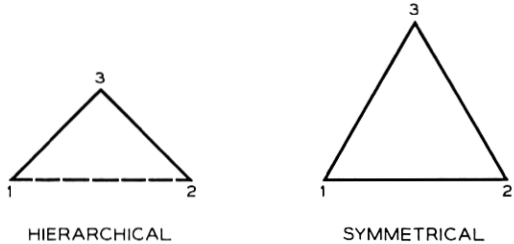


Fig. 3 — Three-node network models with table of link loads and trunk assignments for heavy loading.

Network Parameters, 4 Nodes — Light

Link Number	Engineered Loads-Erlangs	Engineered Trunks		
		Simple	Symmetrical	Hierarchical
1–2	1	5	4	0
1–3	4	10	8	13
1–4	2	7	5	2
2–3	3	8	6	3
2–4	5	11	9	15
3–4	6	13	10	17
Total	21	54	42	50
Ave/Link	3.5	9	7	8.33

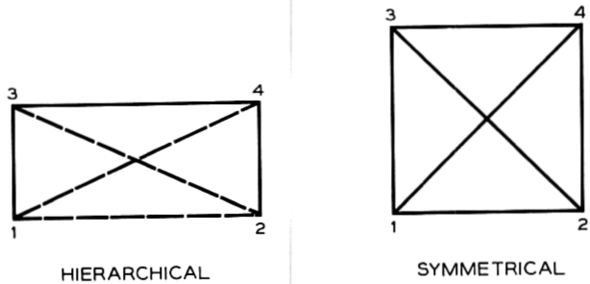


Fig. 4 — Four-node network models with table of link loads and trunk assignments for light loading.

Network Parameters, 4 Nodes — Heavy

Link Number	Engineered Loads-Erlangs	Engineered Trunks		
		Simple	Symmetrical	Hierarchical
1–2	5	11	12	3
1–3	20	30	27	37
1–4	10	18	18	9
2–3	15	24	22	14
2–4	25	36	32	43
3–4	30	42	38	52
Total	105	161	149	158
Ave/Link	17.5	26.8	24.8	26.35

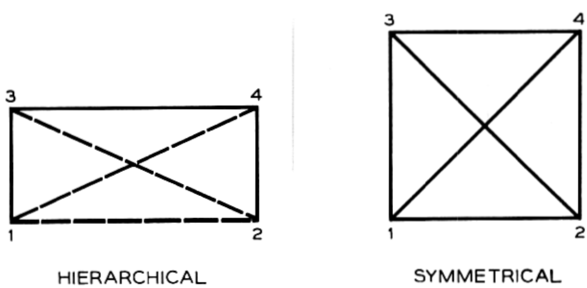


Fig. 5 — Four-node network models with table of link loads and trunk assignments for heavy loading.

NETWORK PARAMETERS, 5 NODES — LIGHT

Link Number	Engineered Loads-Erlangs	Engineered Trunks		
		Simple	Symmetrical	Hierarchical
1–2	1	5	4	0
1–3	5	11	9	17
1–4	2	7	5	1
1–5	2	7	5	1
2–3	3	8	6	2
2–4	4	10	7	3
2–5	6	13	10	19
3–4	6	13	10	21
3–5	4	10	7	4
4–5	7	14	11	23
Total	40	98	74	91
Ave/Link	4	9.8	7.4	9.1

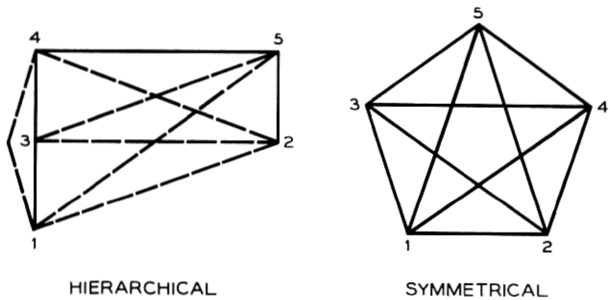


Fig. 6 — Five-node network models with table of link loads and trunk assignments for light loading.

networks were designed to allow each parcel of traffic to overflow through all other nodes in turn and were engineered entirely with the simulator by trial and error. A first estimate of trunk quantities was made using a fixed differential between the load in erlangs and the number of trunks in each link, and corrections were then made as required.

Having established this framework, or procedure for evaluation, a critical question is, What criteria can be used to compare the performance of various types of networks? It is desirable to take account of the efficiency (carried load per dollar of investment) of the network at all times, as well as the grades of service which are provided to each group of customers. Although grade of service here can no longer be interpreted as the small percentage of blocked calls that is ordinarily encountered at normal engineered loads, it is nevertheless incumbent upon the network

NETWORK PARAMETERS, 5 NODES — HEAVY

Link Number	Engineered Loads-Erlangs	Engineered Trunks		
		Simple	Symmetrical	Hierarchical
1–2	5	11	13	3
1–3	35	47	43	57
1–4	15	24	23	15
1–5	10	18	18	9
2–3	20	30	28	19
2–4	25	36	33	25
2–5	40	53	48	64
3–4	45	58	53	75
3–5	30	42	38	31
4–5	50	64	58	81
Total	275	383	355	379
Ave/Link	27.5	38.3	35.5	37.9

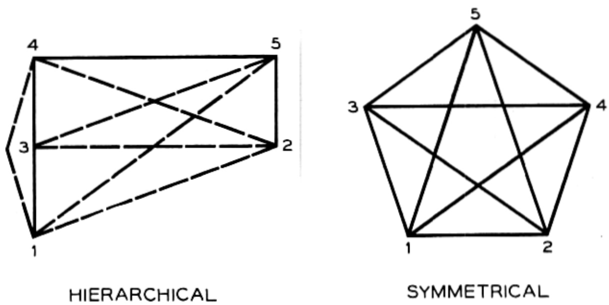


Fig. 7 — Five-node network models with table of link loads and trunk assignments for heavy loading.

designer to consider the extent to which service is degraded on any particular link. Similarly, one would expect the efficiency under overload conditions to be higher than that encountered during normal operation, but the relative efficiencies of networks using various alternate routing doctrines (to carry the same loads) may be rather different. It is clear, for example, that if a call uses a trunk in each of two links, there is a possibility of lower network efficiency being obtained than if it used a trunk in only one link. It is one of the purposes of this study to determine at what overload point such a loss in efficiency takes place, and what if any remedial action can be taken.

Thus, two rather different criteria appear to be important, one of which is essentially an economic variable, and the other a service variable. They are both further complicated by the fact that the first depends on the relative costs of trunks in different links, and the second

NETWORK PARAMETERS, 6 NODES — LIGHT

Link Number	Engineered Loads-Erlangs	Engineered Trunks		
		Simple	Symmetrical	Hierarchical
1–2	1	5	4	0
1–3	5	11	8	16
1–4	1	5	4	1
1–5	3	8	7	3
1–6	2	7	5	2
2–3	2	7	5	2
2–4	6	13	9	18
2–5	3	8	6	3
2–6	4	10	7	4
3–4	4	10	7	4
3–5	6	13	9	21
3–6	4	10	7	4
4–5	5	11	8	5
4–6	7	14	11	23
5–6	7	14	11	24
Total	60	146	108	130
Ave/Link	4	9.75	7.2	8.67

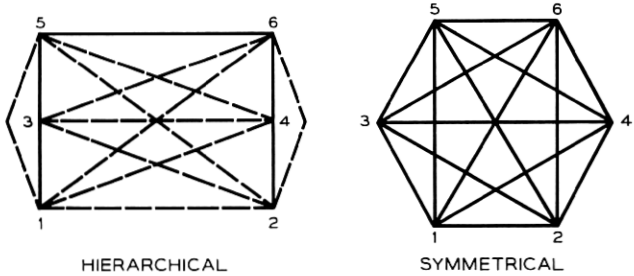


Fig. 8 — Six-node network models with table of link loads and trunk assignments for light loading.

has a different value for every parcel of traffic in the network. In order to simplify these complexities and reduce the number of variables which enter into the measure of performance, only the worst blocking in the network will be considered as the service criterion. This is, of course, conservative, and reflects the difficulties which might occur when alternate routing is canceled and a small parcel of traffic has no trunks in the direct path. The blocking on such a parcel would then be unity, and it would quickly be noticed that a parcel of traffic is isolated.

The problem of assigning costs to trunks is more difficult, of course, since there is no apparent logical worst or best case. Thus the assumptions made here for the relative costs are quite arbitrary and oversimpli-

NETWORK PARAMETERS, 6 NODES — HEAVY

Link Number	Engineered Loads-Erlangs	Engineered Trunks		
		Simple	Symmetrical	Hierarchical
1–2	5	11	15	3
1–3	40	53	48	67
1–4	10	18	19	9
1–5	20	30	28	22
1–6	15	24	24	13
2–3	10	18	19	9
2–4	40	53	48	67
2–5	20	30	28	18
2–6	25	36	33	27
3–4	30	42	38	27
3–5	45	58	53	82
3–6	30	42	38	30
4–5	35	47	43	35
4–6	50	64	58	90
5–6	50	64	58	97
Total	425	590	550	596
Ave/Link	28.3	39.4	36.7	39.7

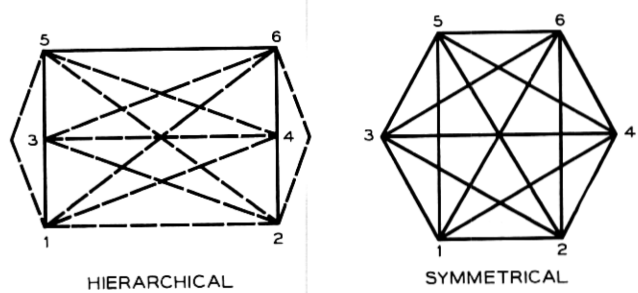


Fig. 9 — Six-node network models with table of link loads and trunk assignments for heavy loading.

fied, but may still be useful in evaluating network performance. Two different assumptions will be made. The first is that all trunks have the same cost. This might be a reasonably realistic assumption in a network where the designers are likely to think of symmetrical alternate routing doctrines. In effect, it states that the distance between any two nodes is not sufficiently different from the distance between any other two nodes to materially affect the cost of trunking facilities between them. Although this may appear to represent an unrealistic geographical situation, it may not be too far in error if the economics of long haul, large cross section trunking facilities are considered. In such systems, the terminal costs make up a large portion of the total trunk cost, and these, of course, are independent of the length of haul. The second assumption is that some routes are half as expensive as the others. For example, in Fig. 4, routes 1–3 and 2–4 are each considered to be half as long as each of the other routes in the network, all of which are assumed to be virtually the same length (all lengths here, of course, refer to costs, which are ordinarily roughly proportional to lengths). This assumption is geographically reasonable, and is, in fact, the kind of layout that is often encountered and which may well have prompted the development of hierarchical alternate routing procedures. Although neither of these weighting schemes may exactly represent an actual case, using each assumption in its logical place may yield more realistic comparative results than would using the same assumption throughout.

Having reduced the parameters for evaluation to two (worst blocking and load carried per dollar of investment), they can be combined into one by the following argument. Both of these parameters, which will be called B (blocking) and E (efficiency) generally increase with increasing loads (although E may occasionally decrease in a non-simple network). A large value of E is generally desirable, but, of course, a large value of B is not. In fact, quite the reverse is true, and so a high value of $(1 - B)$ is a desirable goal. Furthermore, the two factors will increase under different circumstances. For example, a highly efficient network may readily yield a very high value of E, but will also cause very high blocking. Thus B will be high and $(1 - B)$ low. Conversely, a loosely engineered network is likely to provide good service under overloads, yielding a relatively low B and high $(1 - B)$, but in turn be inefficient, with a low E. In both of these cases, the product $E(1 - B)$, will take on some intermediate value. Accordingly, a figure of merit for networks, called the Performance Measure, will be defined as $M = E(1 - B)$. This number may be dimensioned to lie between zero and one and will pass through a maximum as the load is increased. At engineered levels it will essentially represent the network efficiency, and as the load is increased it will indicate when service or efficiency or both are deteriorating. A high M is clearly a mark of a well performing system, with efficient trunk usage and at least tolerable service, while a low or rapidly decreasing M will show a system which is either being inefficiently used or is providing poor service or both. If $M = 0$ an intolerable situation exists; i.e., either some parcel of traffic is unable to be served or no load is being carried by the network. If a network can be designed to be efficient under normal conditions and to maximize M during moderate or partial overloads, and steps can be taken to prevent too rapid a degradation of this quantity during severe overloads, then it is a reasonable assumption that this design will be satisfactory for its purposes. That is, it will provide efficient communications facilities at normal loads, and will also allow the continuation of at least tolerable service between all points under adverse conditions.

One more modification was made in the results before analysis was undertaken. Since large trunk groups are more efficient than small ones, the group size would naturally affect the values of E and M, both at normal loads and under overload conditions. Accordingly, E and M were then plotted for each network relative to the E and M of the corresponding simple networks. The simple network was chosen as a convenient reference point, since it is easily engineered and requires no complicated switching equipment for implementation. Thus, one of the strong changes in efficiency which is not caused by the alternate routing pattern is largely removed, permitting comparisons among the latter to be more readily made.

IV. RESULTS

In order to investigate the effects of various parameters on network behavior, the network efficiencies, E, and performance measures, M, were calculated for engineered loads and for the various overload conditions which were tested. The relative values of E and M (relative to a simple network designed to carry the same loads) were then plotted versus the degree and kind of overload experienced by the network. Thus Fig. 11 shows the values of M for symmetrical networks, for loads ranging from engineered to 100 per cent overload where the overloads occur uniformly throughout the network. Fig. 13 exhibits the same information for hierarchical networks. Figs. 12 and 14 show the behavior of M with load when the overloads occur only on those links which terminate at node 1 with all other loads remaining at engineered levels. Finally, Figs. 15 through 18 are graphs of efficiency (E) versus load for the same situations as pertain to Figs. 11 through 14. The points from which the (smoothed) curves were plotted are shown in Figs. 12 to 18. They are omitted in Fig. 11 for the sake of clarity.

In order to keep the comparisons between symmetrical and hierarchical networks on a somewhat realistic basis, it is necessary to make some adjustment for the probable differences in geography which are likely to encourage consideration of one or the other type of network. Accordingly, as was mentioned above, certain trunks were considered to cost twice as much as others, which introduced a weighting factor into the values of E and hence into M as well. For example, in the 4 node case shown in Fig. 4, trunks in links 1-3 and 2-4 were considered to be only half as expensive as trunks in the remaining four links in the network. This reduced the cost of the trunk plant to 0.805 times the value which would result if all trunks were assumed to be of equal value in the case of a simple network, and to 0.720 times its former value for the hierarchical case. Thus the relative efficiency of the hierarchical network is increased by a factor of 0.805/0.720 or 1.119. This sort of adjustment was made in all calculations relating to hierarchical networks, while all trunks in symmetrical networks were assumed to be of equal value. The weighting factors obtained for the various networks are tabulated in Fig. 10.

The symmetrical networks which were studied in detail operated in the following fashion. Traffic which was blocked from any link was overflowed to an alternate route consisting of two links in tandem. If the call was blocked on this route it was offered to still another two-link route, and so on until all such routes were exhausted. No call was permitted to use a route which required more than two links in tandem. Some experiments were made on networks which allowed three tandem links to be used, but is was found that they were at best marginally more efficient than the two-link maximum network at engineered loads and deteriorated much more violently under overloads. Therefore, they are not considered further in this paper. The order of selection of alternate routes was arranged to approximately equalize the load overflowed to each route. Although this is probably not the most efficient arrangement, it should be adequate to illustrate the behavior of symmetrical networks.

The hierarchical networks operated in a manner similar to the Bell System toll network, with the difference that whereas in the Bell System the routes are selected link by link, in the simulation the routes are entirely preselected at the originating node. If a network is drawn as shown in the hierarchical sketches in Figs. 2 to 9, the route selection is made by

Overall System Characteristics

Case	Engd. Load/Link	Trunk Value Adjustment for Hierarchy	Average Number of Links/Call					
			Engineered Load		100% Uniform Overload		100% Local Overload	
			Hier.	Symmet.	Hier.	Symmet.	Hier.	Symmet.
3 Nodes, light	4	1.194	1.155	1.014	1.156	1.128	1.203	1.124
4 Nodes, light	3.5	1.119	1.181	1.034	1.185	1.229	1.251	1.184
5 Nodes, light	4	1.181	1.288	1.074	1.305	1.305	1.372	1.210
6 Nodes, light	4	1.194	1.263	1.117	1.328	1.363	1.324	1.245
3 Nodes, heavy	10	1.121	1.076	1.034	1.073	1.164	1.138	1.180
4 Nodes, heavy	17.5	1.064	1.088	1.040	1.121	1.271	1.140	1.202
5 Nodes, heavy	27.5	1.071	1.100	1.062	1.147	1.309	1.151	1.274
6 Nodes, heavy	28.3	1.085	1.134	1.074	1.205	1.364	1.203	1.322

Fig. 10 — Over-all system characteristics.

hunting up the hierarchy, starting from the terminating office, in the distant region, and then down the hierarchy in the home region. (A region can be thought of as all centers whose final routes ultimately terminate at the same highest level switching center.) Thus, for example, in the four node network shown in Fig. 4, the alternate routes for traffic offered to link 1-2 are:

For half the traffic, 1-4-2, 1-3-2, 1-3-4-2; and for the remainder of the traffic, 1-3-2, 1-4-2, 1-3-4-2.

4.1 *General Observations*

In order to draw conclusions from this study as to the relative merits of various types of alternate routing systems under different load conditions, Figs. 10 to 18 will be examined and the significance of the results discussed.

As a very first step, a cursory examination of all figures reveals the following:

(1) The relative effectiveness of alternate routing networks, whether measured by E or M, tends to decrease with overload, with the decrease occurring more rapidly under uniform than under local overload. Although in some cases the network remains superior to a simple network even for 100 per cent overload, the relative performance at such overloads is almost always poorer than at engineered loads. This is due, of course, to the fact that the average number of links per call increases with overload, causing a decrease in efficiency which may outweigh the gains yielded by the larger effective access provided by the alternate route system. (See Fig. 10.)

(2) Light networks (those with less traffic), gain more from alternate routing than do heavy networks. This seems to occur because systems designed for large parcels of traffic use large efficient groups. Thus providing alternate routes in heavy networks, which increases the effective access somewhat, does not materially increase the efficiency, while the degradation caused by using several links per call is nevertheless present. In lighter networks, the increase in efficiency owing to the larger effective access is substantial, overriding the degradation and causing a considerable gain in effectiveness.

Perhaps to this list should be added:

(3) As mentioned above, symmetrical systems do not appear to benefit from allowing more than two links in tandem to be used by any call. This effect is apparently caused by the decrease in

efficiency which results from using many links per call overriding the gain yielded by increased access. In this situation, of course, the increase in link occupancy may be substantial, while the increase in effective access is likely to be small.

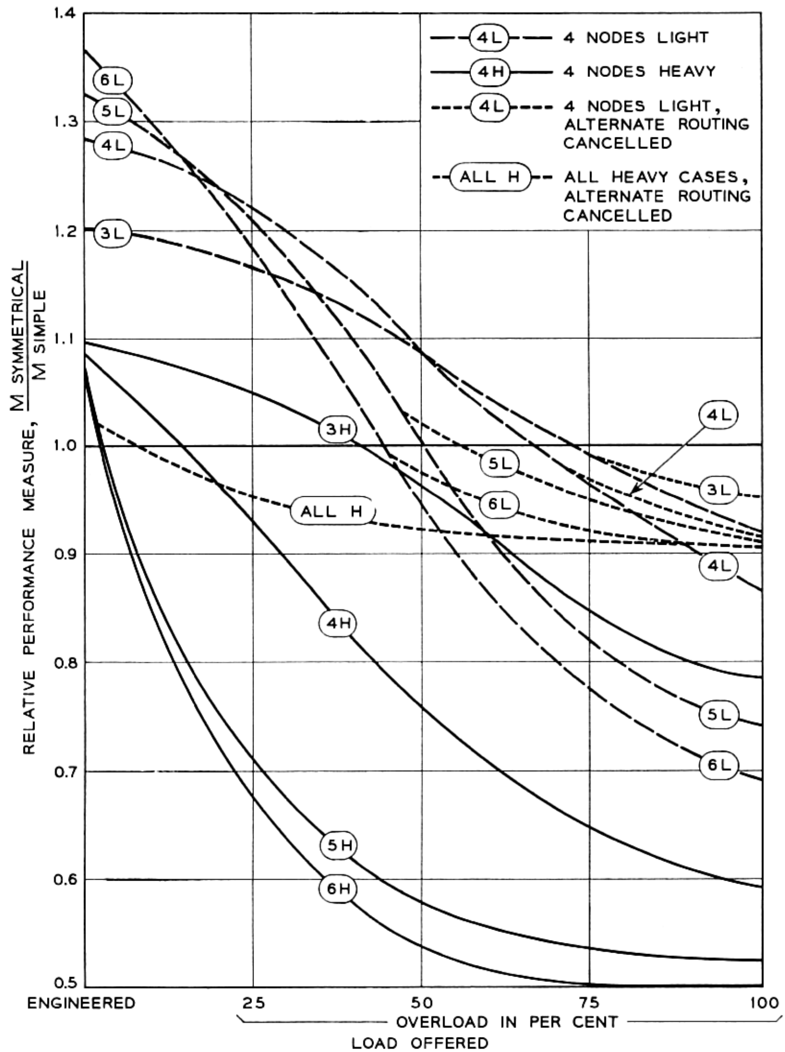


Fig. 11 — Relative performance measure of symmetrical networks under uniform overloads.

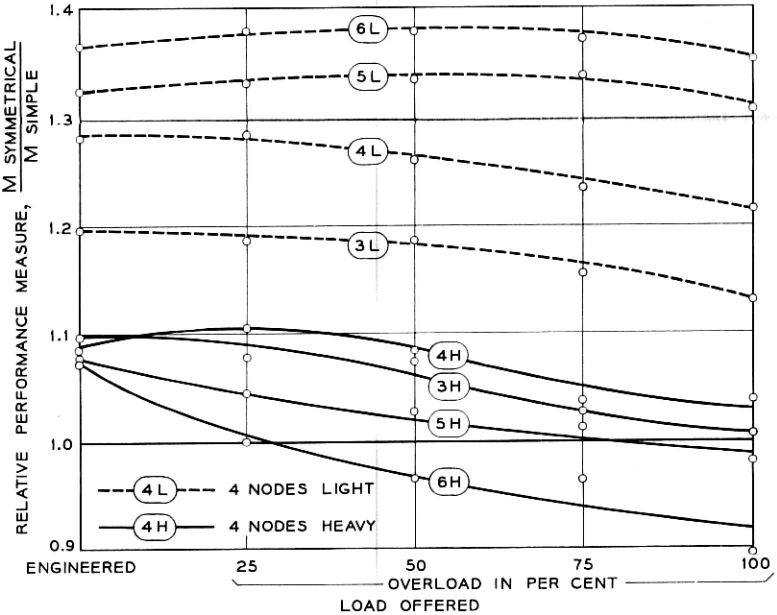


Fig. 12 — Relative performance measure of symmetrical networks under local overloads.

4.2 *Symmetrical Networks*

The curves shown in Figs. 11 through 18, when studied closely, reveal much information regarding the characteristics of the networks considered. In Fig. 11, the high relative performance measure of light symmetrical networks at engineered loads and the rapid decline as the load is increased uniformly is clearly indicated. The heavier networks exhibit a lower relative value of M at engineered loads and also decline rapidly, bringing their performance measure down to very low relative values at high overloads. Such a rapid decrease in M, it would appear, would make it impracticable to install symmetrical systems in many actual applications, were it not for the fact that M can be kept relatively high by canceling alternate routing at some appropriate point. The dotted lines in Fig. 11 indicate the relative performance measure if alternate routing is canceled, and it is clear that this factor can be kept above 0.9, regardless of the size of the network and even for 100 per cent overload. In any event, it does appear that for extremely heavy networks the decline is so precipitous that this method of alternate routing might well prove to be inapplicable. Fig. 12, however, illustrates the real

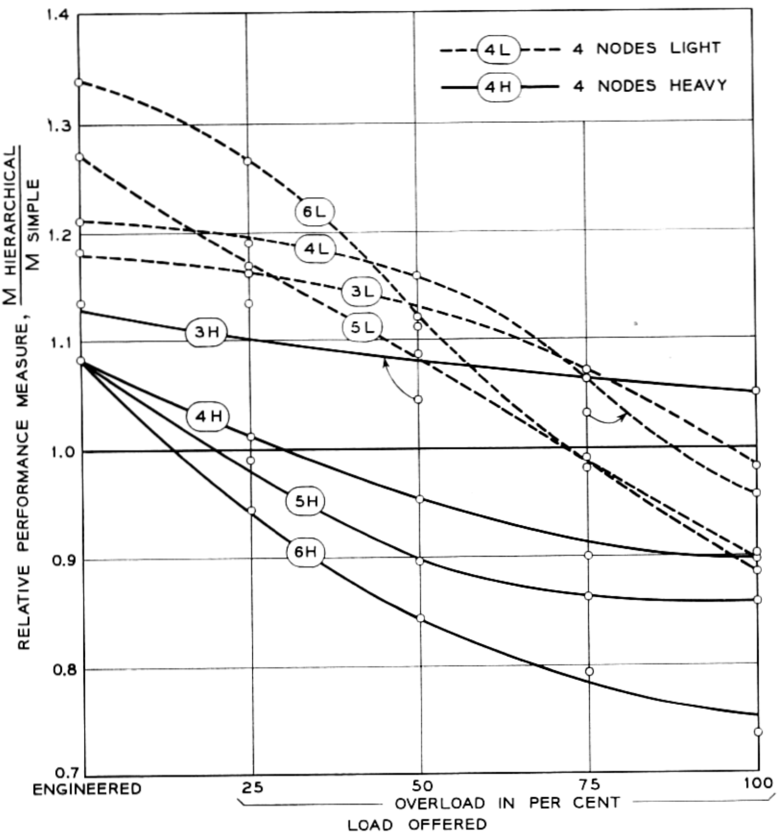


Fig. 13 — Relative performance measure of hierarchical networks under uniform overloads.

strength of the symmetrical routing doctrine. The relative performance measure is shown to be almost constant under local overloads, and remains above unity for all but the largest, heaviest networks. Furthermore, this sort of alternate routing structure is likely to be quite efficient at engineered loads in systems where call setup time and switching delays are no longer negligible, since it generally uses a relatively small number of links per call, as evidenced in Fig. 10.

A symmetrical network structure then can be devised which has the following characteristics:

(1) Performance measure (and thus efficiency) are high at engineered loads.

(2) Local overloads are well tolerated, with the network remaining

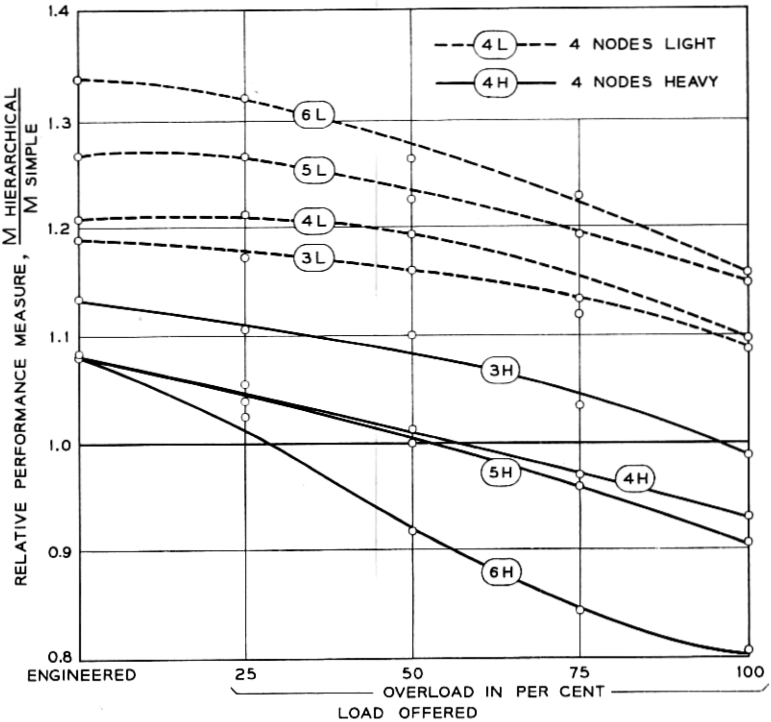


Fig. 14 — Relative performance measure of hierarchical networks under local overloads.

efficient and not allowing any parcel of traffic to suffer excessive blocking.

(3) If alternate routing can be canceled at the appropriate point, then the performance measure can be maintained at a tolerable level even under severe uniform overloads.

(4) The average number of links per call is quite low at engineered loads, increasing rapidly as overloads are applied.

An important practical question in (3), however, is whether a network control can be devised to cancel alternate routing easily, and how the control can determine the degree of overload. Another disadvantage of such networks is the unavailability, at present, of any but the very crudest methods of trunk engineering. However, this type of network is, in principle, capable of satisfying the four points listed above, all of which are desirable and often are difficult to attain concurrently.

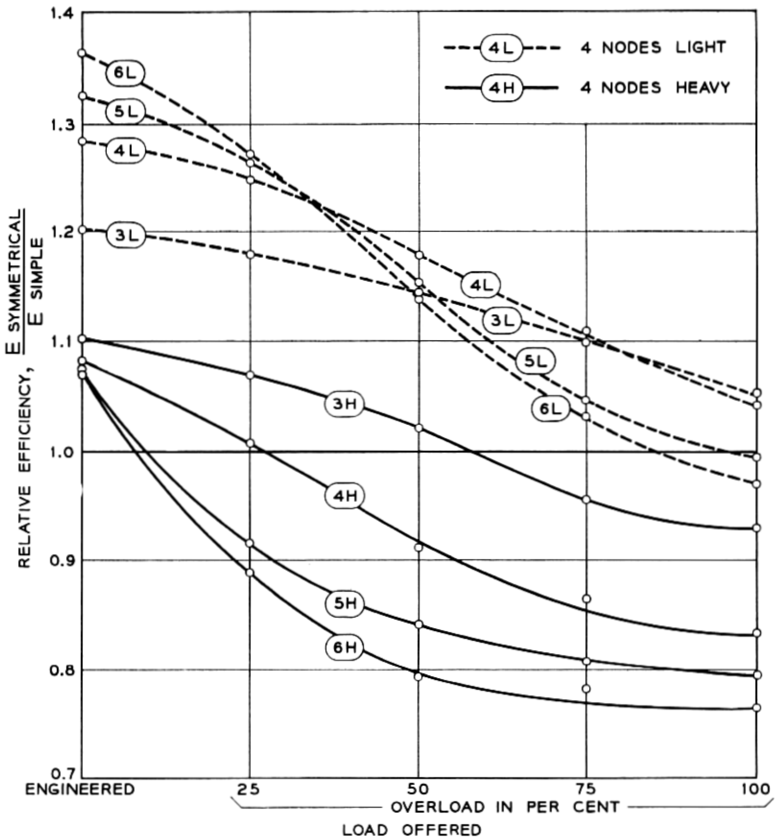


Fig. 15 — Relative efficiency of symmetrical networks under uniform overloads.

4.3 *Hierarchical Networks*

In Fig. 13, the relative performance measure of hierarchical networks under uniform overloads is shown. A comparison with Fig. 11 indicates that the relative M is higher at engineered loads for symmetrical than for hierarchical light networks and not too different for heavy networks, although the decline with uniform overload is more rapid in the former case. In the hierarchical system, however, the relative M cannot be increased by complete cancellation of alternate routing, since this increases the blocking on some parcels of traffic which are offered to high usage groups to a high level. Fig. 13 then shows that, although light

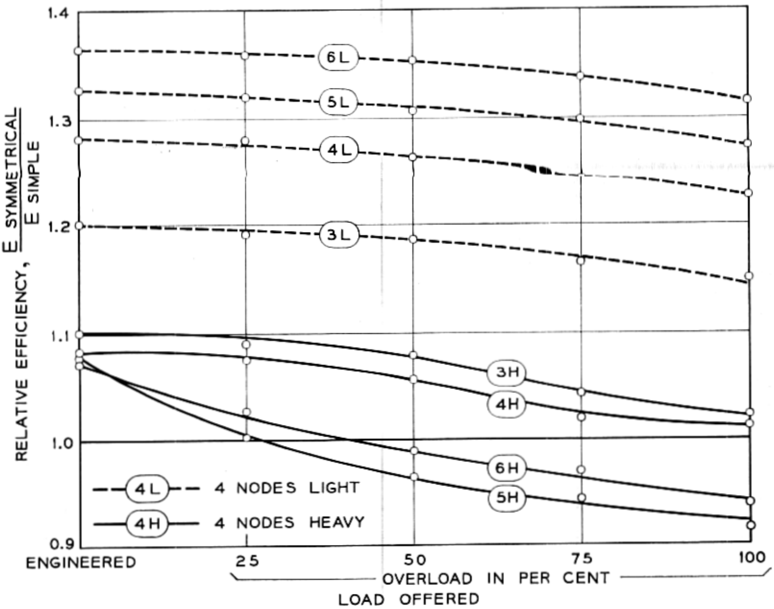


Fig. 16 — Relative efficiency of symmetrical networks under local overloads.

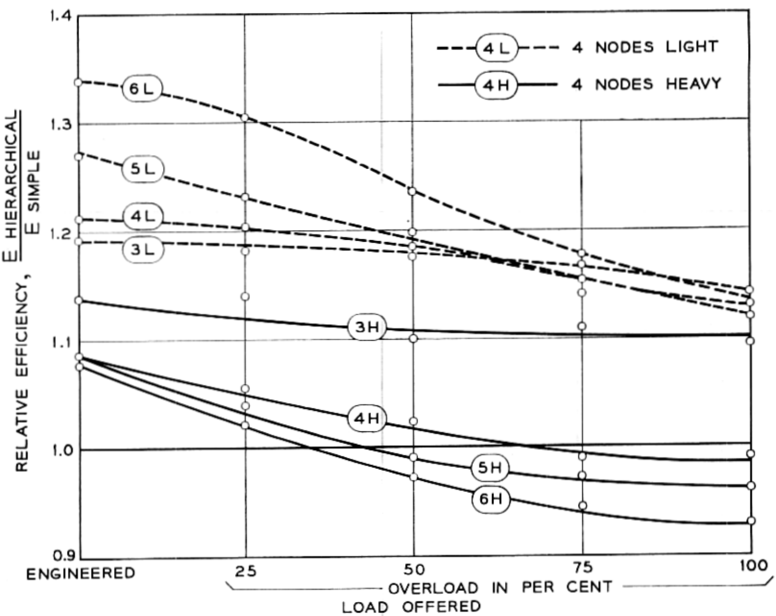


Fig. 17 — Relative efficiency of hierarchical networks under uniform overloads.

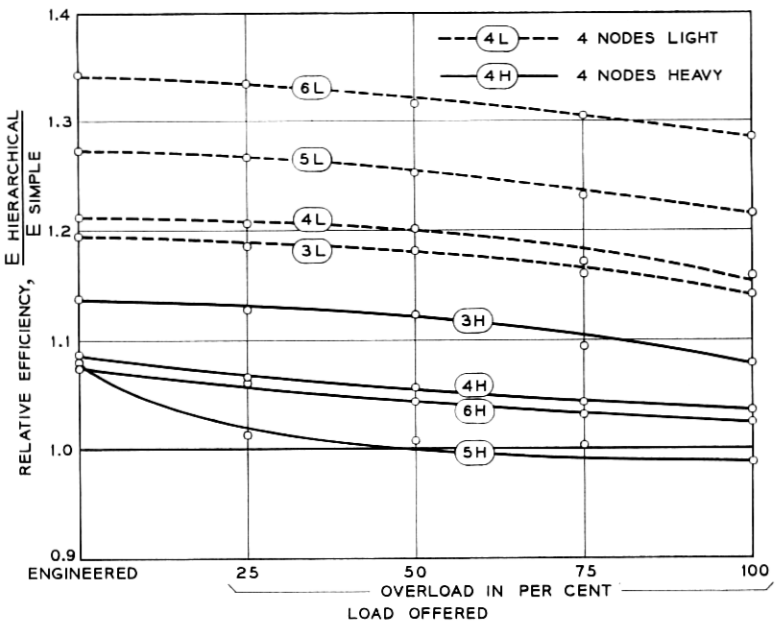


Fig. 18 — Relative efficiency of hierarchical networks under local overloads.

networks retain their effectiveness up to 100 per cent overload, large heavy networks show a decline with uniform overloads to a quite low value of relative M.

The behavior of such networks under local overloads is shown in Fig. 14. In these circumstances the relative performance measure declines slowly from the value at engineered load as the local overload is increased. The decline is sufficiently gradual to enable the lighter networks to retain an M greater than unity for all overloads considered. The heavier networks, however, are unable to do this, and the value of relative M for the worst network declines almost to 0.8 for the greatest local overload.

The essential operating characteristics of networks of this basic design then appear to be as follows:

(1) The performance measure (and thus efficiency) tend to be high at engineered loads (if the variation in trunk costs is taken into account).

(2) The performance measure declines at a moderate rate under uniform overload, reaching rather low values for large, heavy

networks. No simple corrective measures are available to improve the situation but no special measures are needed to prevent catastrophic performance degradation. (Certain more complicated corrective measures, such as selective cancellation of alternate routing, might prove effective, but this sort of procedure was not studied.)

(3) Local overloads are moderately well tolerated, with the relative performance measure showing a gradual decline with increasing load, and dipping below unity for some cases.

(4) A relatively large number of links are used per call at engineered loads, and this number increases gradually with overloads.

This is then essentially a moderately well behaving network, providing neither superlative nor intolerable service at any level of load. It requires no complex controls to keep operating reasonably well, and is relatively simple to implement without the need for sophisticated switching equipment at the tandem points. In a real system, with switching delays and appreciable call setup time, however, this type of network may behave badly under overloads, since some calls use many links in tandem, and therefore can tie up a great deal of equipment when processing a call, even though the call is not completed. In fact, the large number of links used per call in hierarchical systems even at engineered loads is a source of inefficiency in such systems.

An apparent peculiarity in all the curves is the superiority of large light networks over small light networks at engineered loads, with the situation reversing as the load increases, so that at 100 per cent overload the small light networks are generally superior. A qualitative explanation of this would again involve the average number of links per call, which increases more rapidly in large networks than in small ones. The heavy networks do not exhibit this effect at all, and the larger heavy networks always appear to perform less well than the smaller ones. Since the larger (heavy) networks were designed to be more heavily loaded than the smaller ones, however, (see Fig. 10), this effect is more likely to be a result of network load than size.

4.4 *Efficiency Curves*

Figs. 15 to 18 portray the network efficiencies in the several cases studied. In general, these curves display a somewhat shallower slope than the corresponding curves for M. This implies that as the load is increased, not only does the relative network efficiency decrease, but the blocking encountered by the most poorly served group of customers also increases more rapidly when alternate routing is in effect than when it is not. The only exception to this is the symmetrical system under local overload (Figs. 12 and 16). In this case the relative efficiency and the relative performance measure decline at about the same rate, which implies that in this system the blocking remains at essentially the same level whether a symmetrical or simple system is in use. This is an important consideration in favor of symmetrical networks, particularly since both efficiency and performance measure remain reasonably high for all types of overloads considered.

V. CONCLUSIONS

The foregoing discussion of various types of alternate routing networks may be of use in determining whether alternate routing structures should be incorporated into particular switching systems and, if so, of what sort they should be. Many of these factors have long been known and used by network designers, and the present study should provide additional documentation. In the case of factors not previously considered, this study may provide justification for their incorporation into future designs. Some of these are as follows:

(1) If the overload capability of the system is not important, some sort of alternate routing system is almost certainly justified on economic grounds.

(2) If *local* overload capability is important, then strong consideration should be given to a symmetrical alternate routing network, since this configuration allows the blocking to be kept to a minimum under local overloads while retaining a high network efficiency.

(3) If *uniform* overload capability is an important consideration, then alternate routing structures should be contemplated with caution, but can still be used if the average load per link is small and appropriate action, such as cancellation of alternate routing (either uniformly or selectively) can be taken as required.

(4) If the average load per link is small, alternate routing almost always is advantageous, while if it is large, the advantage is sometimes questionable.

(5) If the initial efficiency is an important criterion, then the selection of the type of alternate routing may well depend upon the geography of the particular system. Thus, in certain situations, where small towns communicate primarily with nearby cities, a hierarchical structure may be preferable, while if there is a large group of

approximately equal-sized cities spread over the country, then a symmetrical system could prove to be superior.

(6) If switching equipment is expensive or call setup time is long, symmetrical networks may prove to be superior to hierarchical structures at engineered loads regardless of the geography. This would come about because of the large number of links per call, and hence the large amount of switching equipment used by hierarchical networks. Clearly, long setup time in this case would lead to inefficient trunk usage, since trunks in one link would be held while the call progressed along a multi-link path.

(7) Although not shown specifically in these results, a multi-alternate route structure provides service protection, which a simple layout does not, and a well connected symmetrical network is likely to be less vulnerable to damage than a hierarchical system.

Most actual systems, of course, must be designed to be efficient at engineered loads, and yet must also be able to accept either uniform or local overloads without excessive degradation of service. Furthermore, real networks usually serve many small towns communicating primarily with larger cities, which in turn communicate with each other on a roughly equal basis. Therefore, the network designer must decide which of these often conflicting criteria are most important, and develop a system which satisfies these as closely as it can within the limitations imposed by the switching and signaling equipment and the available methods of trunk engineering. It is quite likely that the best system in most situations is some combination of symmetrical and hierarchical networks, not necessarily of the particular kinds studied here. Furthermore, the advent of electronic switching systems and high speed signaling devices has made alternate routing doctrines which are dependent on the state of the system feasible, and these may well prove to be superior than any system with a completely prespecified alternate routing structure. However, an analysis of the basic characteristics of simpler networks is likely to be useful in predicting the behavior and influencing the design of specific, more complex systems. It was this potential application which motivated the studies described in this paper.