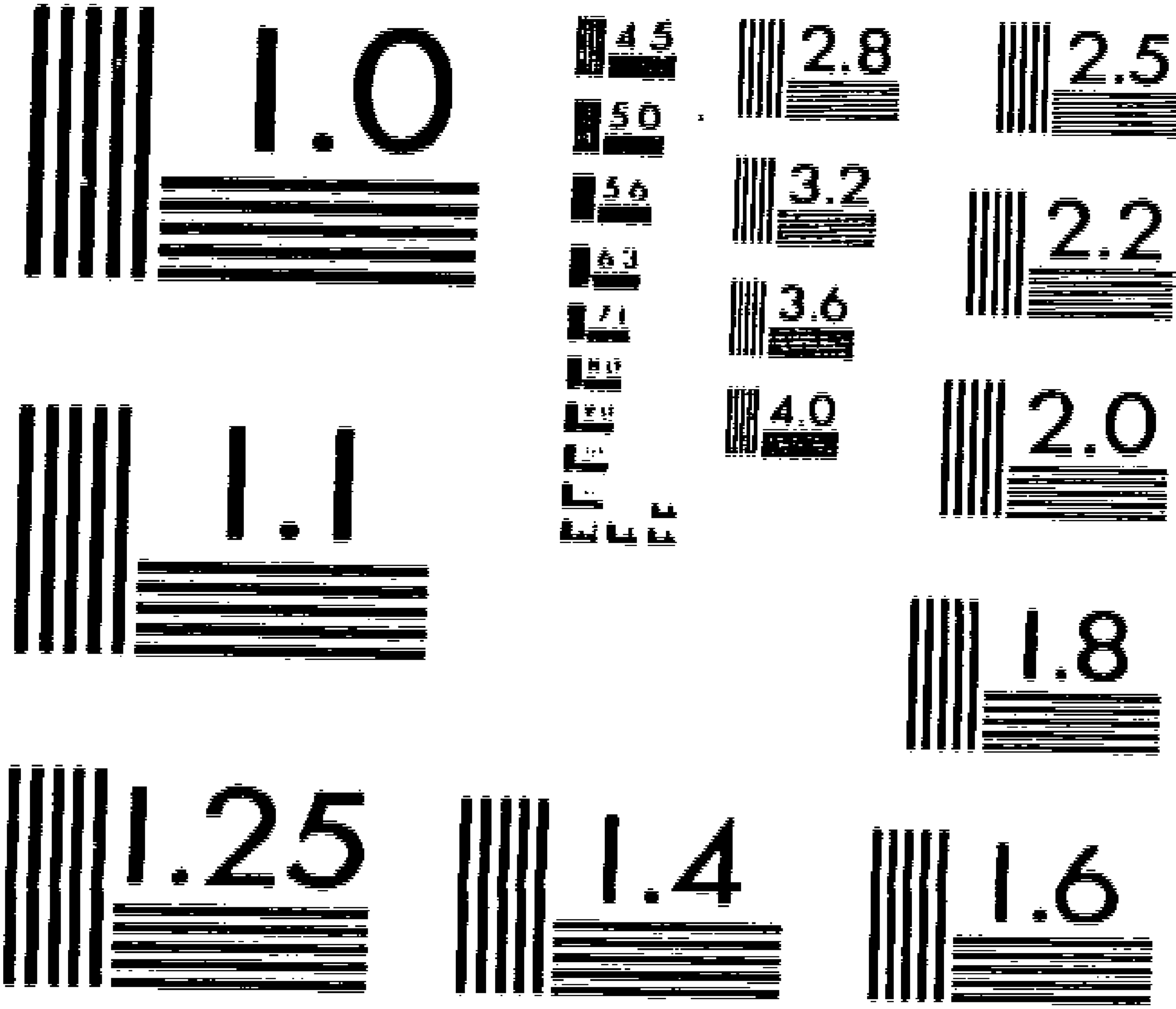
1.0
4.5
5.0
5.6
6.3
7.1
2.8
3.2
3.6
4.0
2.5
2.2
2.0
1.1
1.8
1.25
1.4
1.6

DOCUMENT RESUME

ED 060 861 LI 003 525

AUTHOR	Korfhage, Robert R.; And Others
TITLE	Graph Models for Library Information Networks.
INSTITUTION	Southern Methodist Univ., Dallas, Tex. Computer Science/Operations Research Center.
REPORT NO	TR-CP-710013
PUB DATE	Aug 71
NOTE	33p.; (4 References)
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	Electronic Data Processing; *Information Networks; Library Automation; *Library Networks; *Library Services; *Models; *Public Libraries; Theories
IDENTIFIERS	PLAN; *Public Library Access Network

ABSTRACT

The design and study of library information networks are enhanced by the use of the concepts which have been developed by graph theorists. In this paper this theme is expanded upon, proposing a general network structure which is believed to be a good model for a wide variety of library and other information networks. The basic concepts from graph theory are illustrated with the aid of a hypothetical Public Library Access Network (PLAN). (Author)

GRAPH MODELS FOR LIBRARY INFORMATION NETWORKS

(Technical Report CP-710013)

Robert R. Korfhage
U. Narayan Bhat
Richard E. Nance

Computer Science/Operations Research Center
Institute of Technology
Southern Methodist University
Dallas, Texas

August 1971

ABSTRACT

The design and study of library information networks are enhanced by the use of the concepts which have been developed by graph theorists. In this paper we expand upon this theme, proposing a general network structure which we believe to be a good model for a wide variety of library and other information networks. The basic concepts from graph theory are illustrated with the aid of a hypothetical Public Library Access Network (PLAN).

THE DESIGN OF PLAN.

The word "graph" is used in many different ways, by both mathematicians and non-mathematicians. However, to the graph theorist this word has one precise meaning. It does not refer to bar charts, nor to the curves which trace the values of mathematical functions. Rather, a graph (shown in Figure 1) is a structure which is intuitively associated with figures constructed of lines and points [1]. In particular, we shall be concerned with directed graphs or digraphs [2], whose lines have an orientation from one point into another (Figure 2). Precisely, a digraph consists of a finite set of nodes n_1, n_2, ..., n_k, and a finite set of arcs a_1, a_2, ..., a_h, such that to each arc is associated an ordered pair of nodes. If to arc a_i is associated the node pair (n_s, n_t), we then say that a_i is the arc from n_s into n_t, and indicate this by an arrow on the picture of the digraph (Figure 2).

To illustrate the use of digraphs in the study of information networks, let us consider a hypothetical network called PLAN (Public Library Access Network). The purpose of PLAN is to provide the library users of the forty-eight adjacent states and the District of Columbia with a system which will enable them to access any library anywhere within the territory covered. For purposes of exposition we shall assume that there are only forty-nine libraries in the network, and shall discuss the possible distribution and interconnection of these libraries. In addition to the provision of service, we might also want to impose

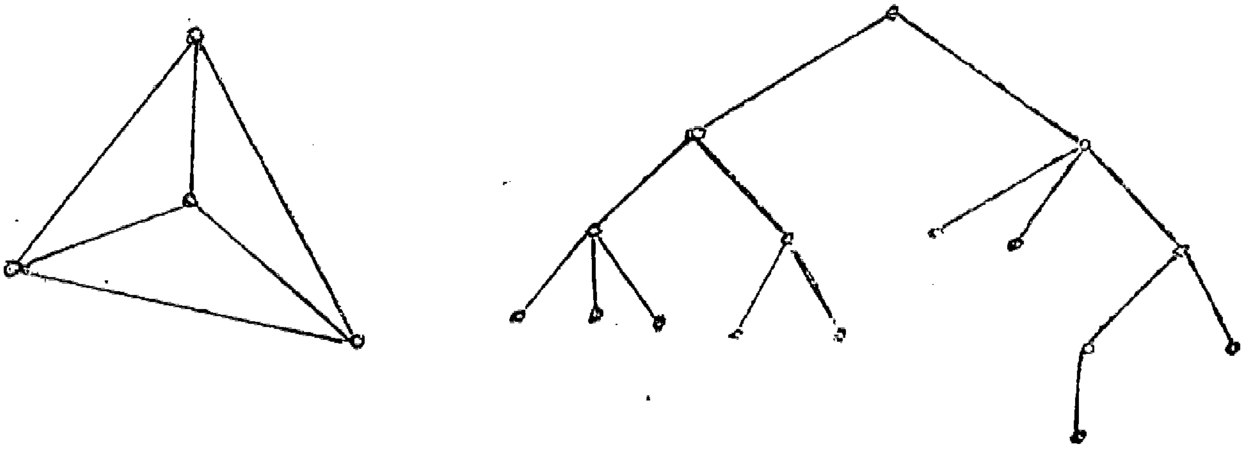

Figure 1. Two examples of graphs

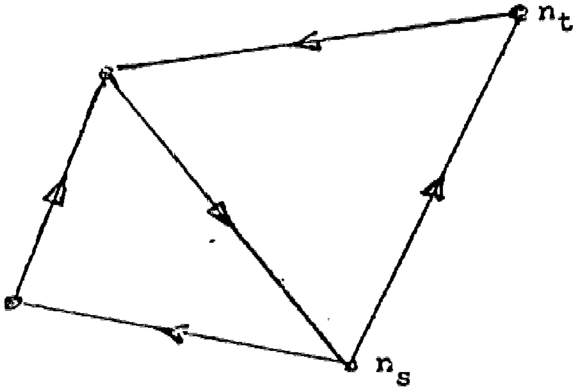

Figure 2. An example of a digraph

other constraints on the system, such as easy access, rapid response, low cost, and the control of information flow within the network.

Let us consider first the distribution of the forty-nine libraries. This distribution can be determined by a number of different criteria. We might choose, for example, to distribute the libraries equally in terms of the area served. This would provide one library for each 61,700 square miles (Figure 3). While such a distribution might have some advantages in terms of cost or other factors, it intuitively does not appear to serve the population equally.

If we think in terms of the state library systems, one natural distribution of the libraries would be to place one in each state capital, and one in Washington (Figure 4). Again, such a distribution would have both advantages and disadvantages.

Since libraries are intended to serve people, a better plan might be to place a library in each of the forty-nine largest cities (Figure 5). Alternatively, one might distribute the libraries, one per 4,080,000 people (Figure 6).

We might suggest many other possible arrangements of the libraries, and even within the arrangements which we have outlined, there is room for variation. Exactly where within a city, state, or area should one place the library for maximum benefit? Each suggested placement of the libraries will have certain advantages and disadvantages in terms of cost, service, and the other constraints imposed on the network. Thus the decision between the proposed placements is not a simple one.

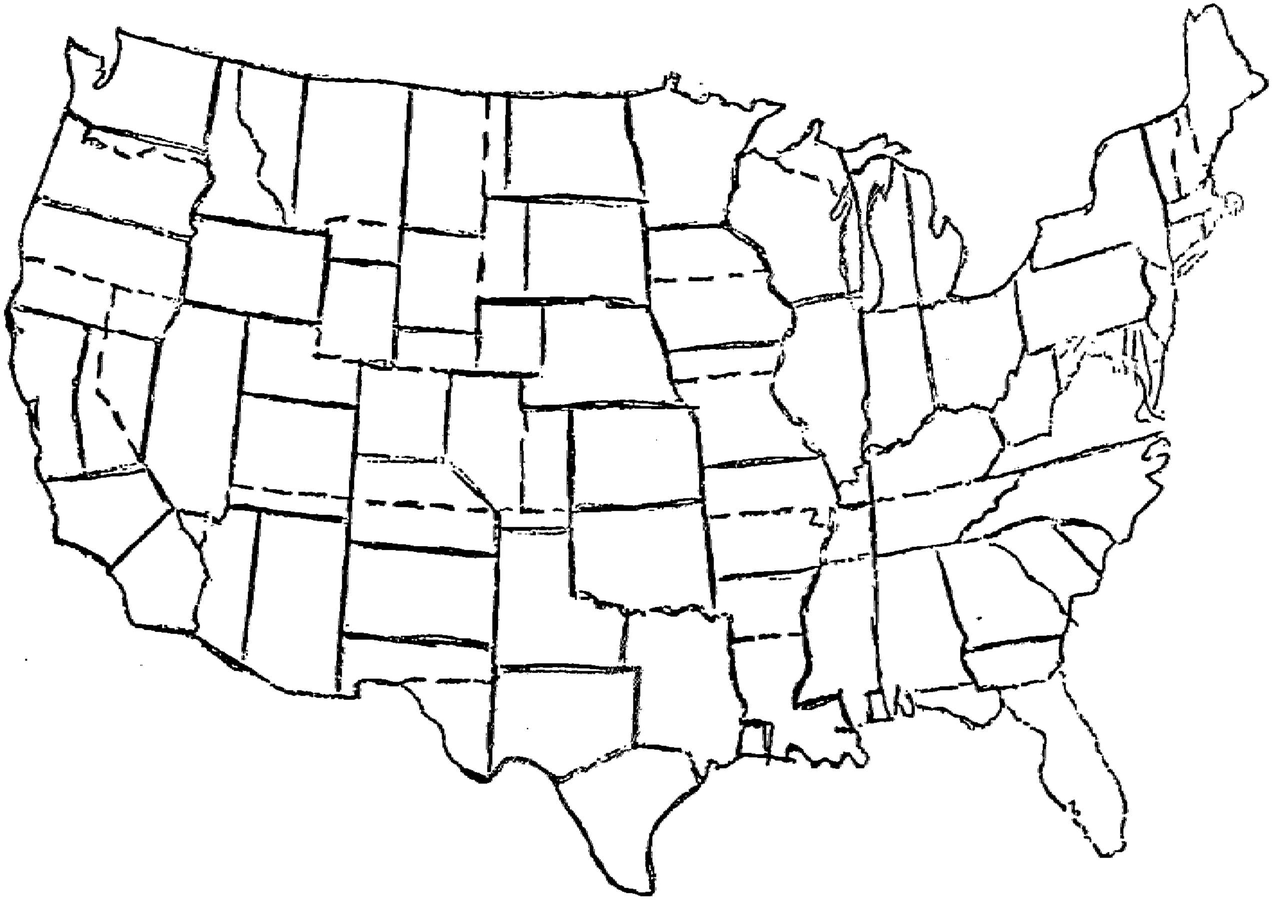

Figure 3. Division of the U. S. into forty-nine approximately equal areas.

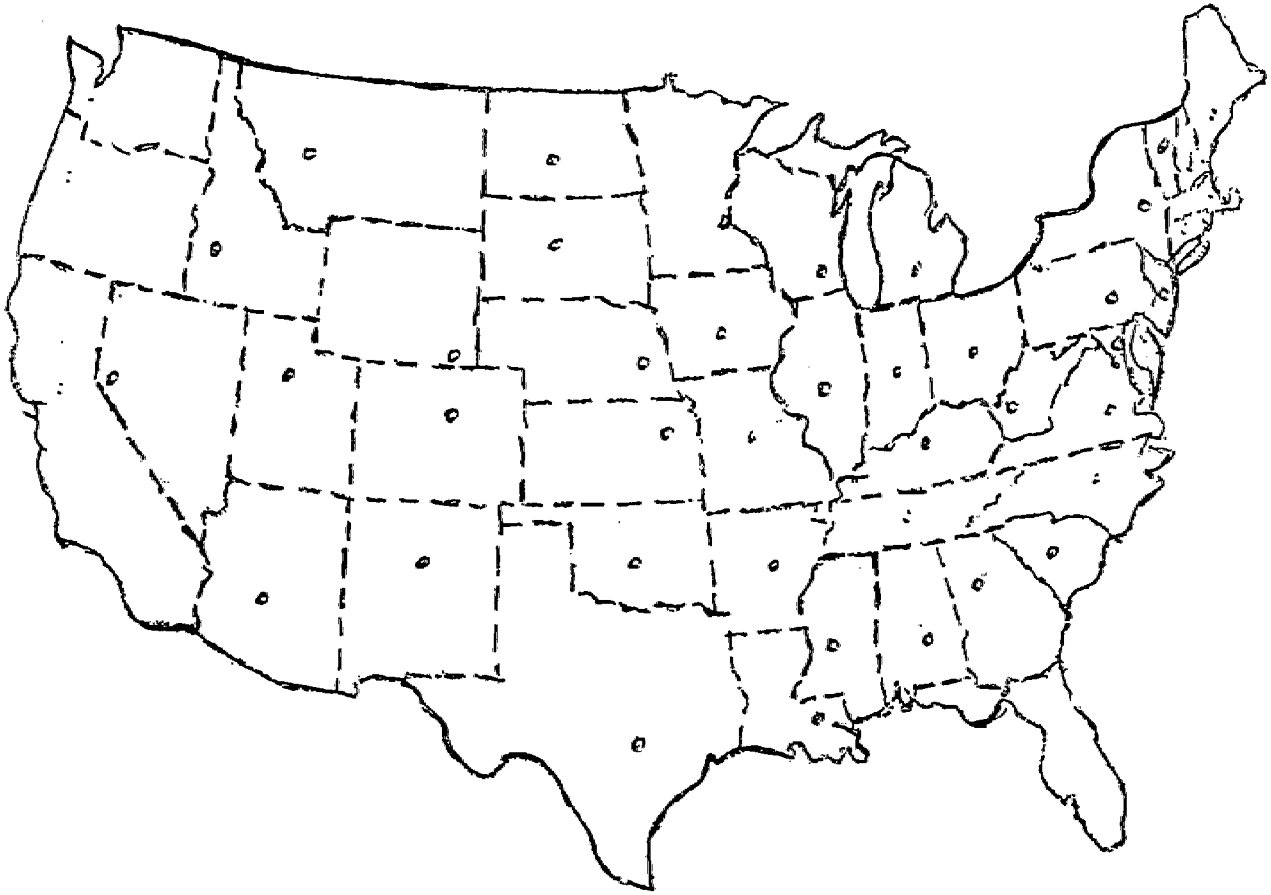

Figure 4. Location of the state capitals

and Washington, D. C.

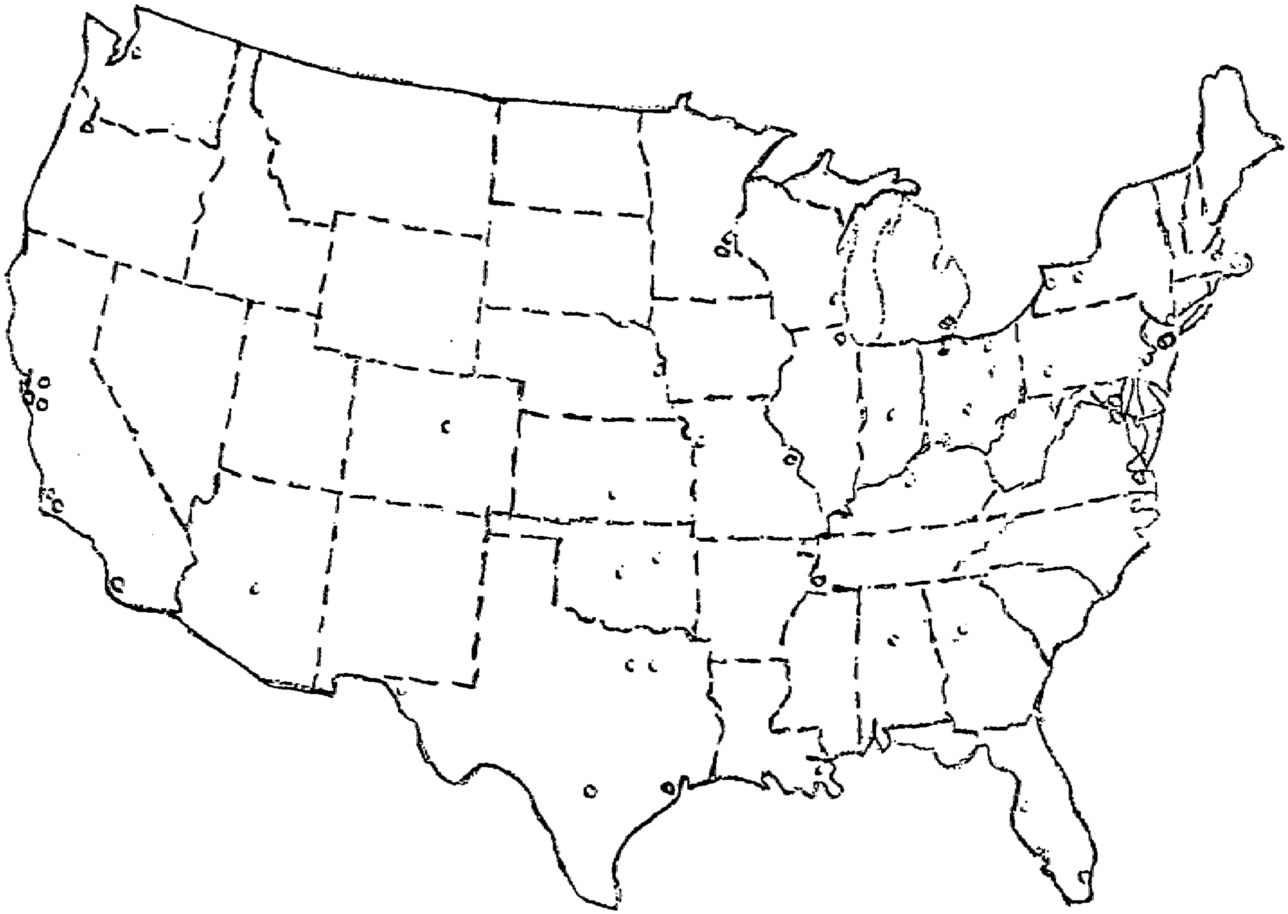

Figure 5. Location of the forty-nine largest cities.

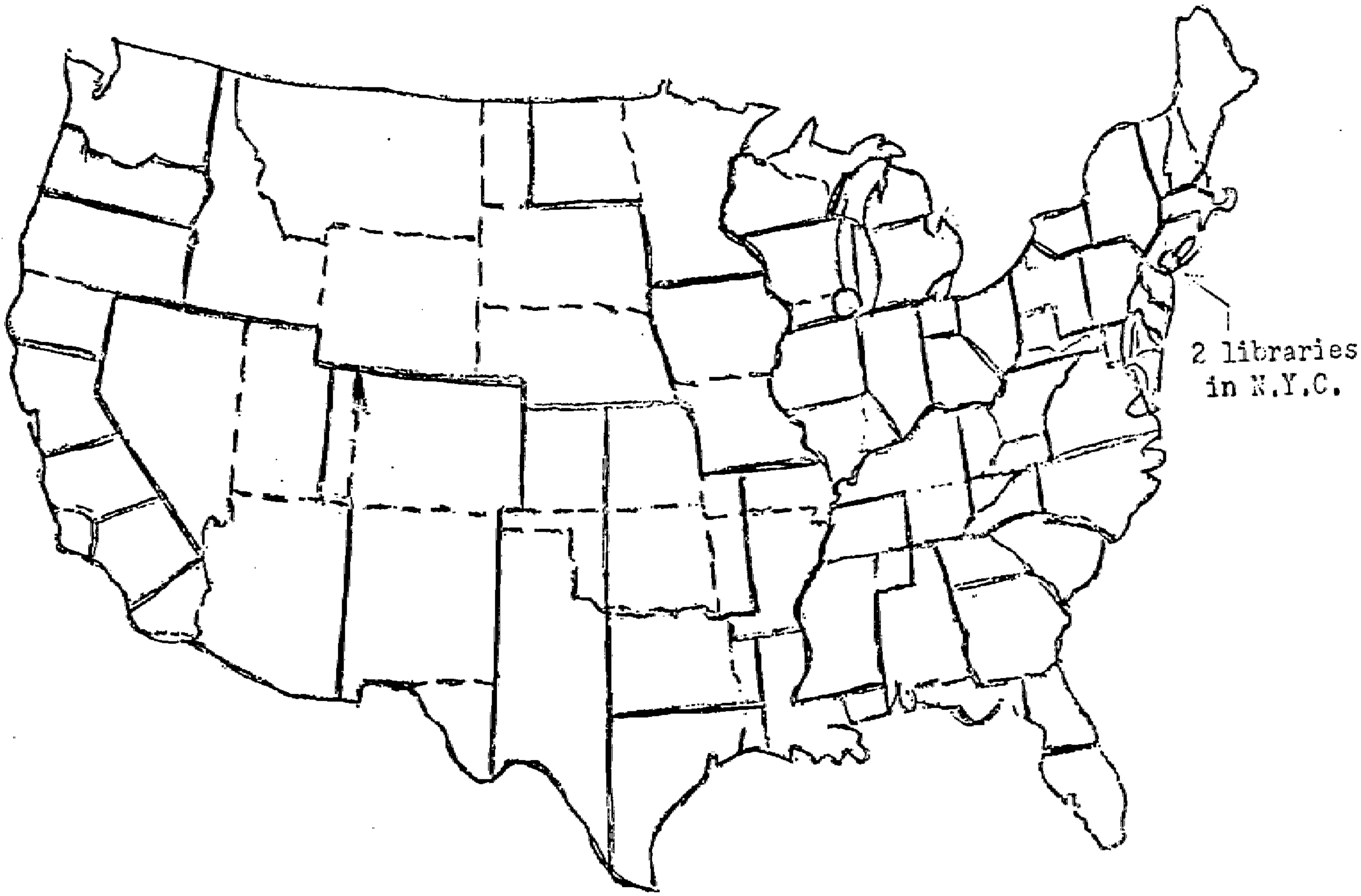

Figure 6. Division of the U. S. into forty-nine areas of approximately equal population.

If our network is to be usable by the public, then there must be more than the forty-nine libraries specified by us. As we include additional libraries, two major problems loom. First, can we design the network in layers, adding a second layer of libraries to the original forty-nine library network, then another layer, and so forth? Or do we need to begin at the beginning every time we enlarge the network? Second, as the libraries become more local in their primary function, they undoubtedly become more specialized and differentiated. We may find not only the public library, but also the university library, specialized libraries relating to particular industries or other groups of users, and probably even computer data banks. It seems reasonable to expect that these differentiated libraries can play quite different roles in the total network.

But let us return to our main network, with forty-nine libraries. At this point we have merely a set of libraries - the nodes in our digraph. To continue the construction of our model we must add the arcs of the graph. These arcs correspond to the information transfer paths of the network. Along some of the paths flow messages requesting information. Other may carry both messages and documents. In any case, a great many possible paths exist for our network. While in actuality these paths may be chosen in a rather complex pattern, we find it convenient to discuss four "pure" types of networks.

In a cyclic network there is precisely one arc leading into each node, and one arc leading from each node. These arcs are so chosen that the entire configuration forms one cycle or loop, with no repeated arcs or nodes (Figures 7). Such a network is relatively inexpensive to install provided that the arcs join nodes which are close together. However, since only one path exists from one node to any other, response time might be rather lengthy. By "installation" we mean the process of bringing the connecting arc structure into existence. In some cases existing information transfer channels (the mail, telephone lines, etc.) can be used; in other situations equipment must be set up. The trade-off between installation cost and response time thus emerges.

A decentralized network provides immediate access from each node to every other one by an arc joining each pair of nodes (Figure 8). This type of network has the maximum number of arcs, and hence installation is quite costly. The PLAN network, for example, would have 2,352 lines, some stretching across the continent. One characteristic shared by both the cyclic and the decentralized networks is the absence of a natural head, or main library. This may be either an advantage or a disadvantage. If one deems it a disadvantage, then a hierarchical network should be considered (Figure 9). This type of network provides, as does the cyclic network, a single path from any one node to any other; hence relatively long response time might be expected. But since it is organized as a branching tree, the hierarchical network provides for relatively simple monitoring and control of the information flow.

Both the cyclic and the hierarchical networks suffer from another defect. Since there is only a single path from any one node to any other, the

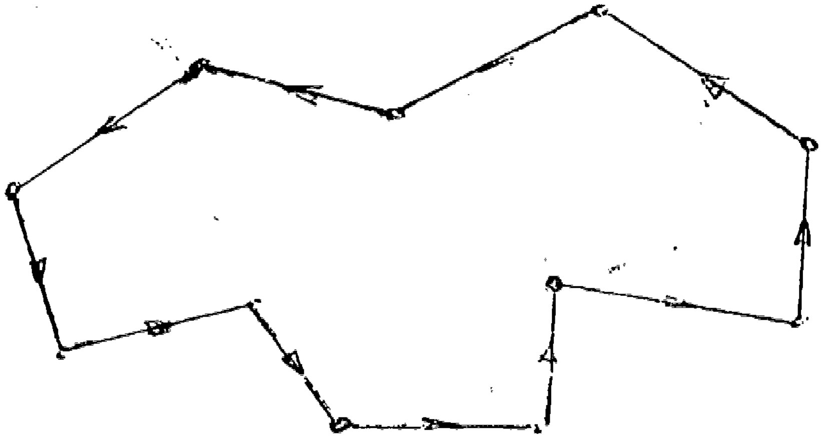

Figure 7. A cyclic network on eleven nodes.

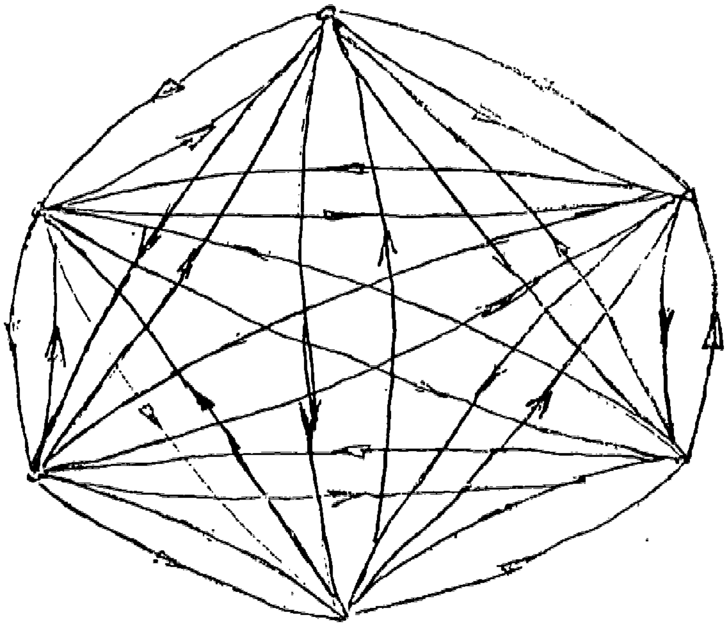

Figure 8. The decentralized network on six nodes.

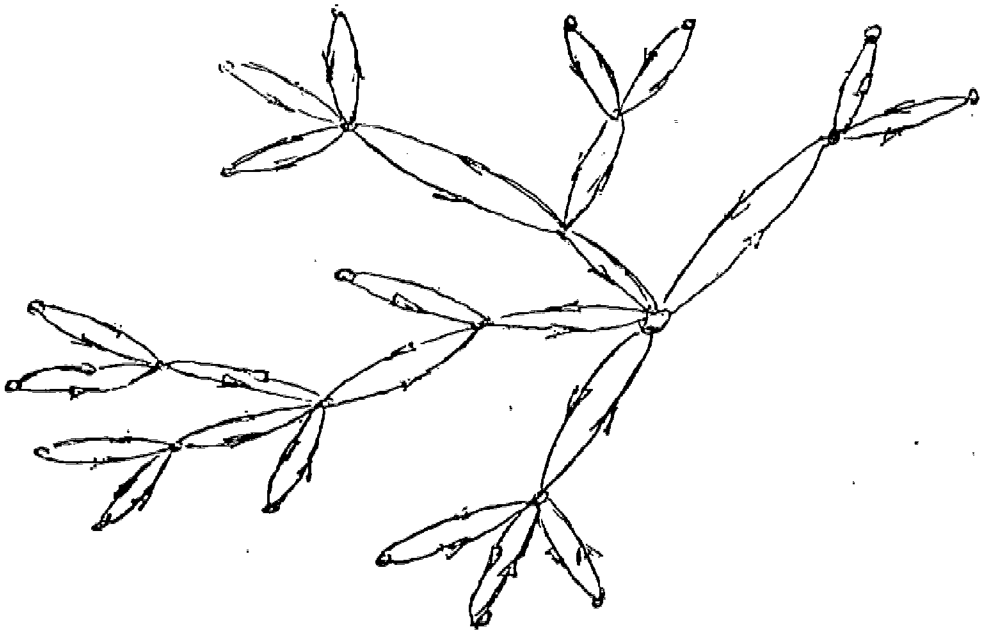

Figure 9. A hierarchical network on twenty-six nodes.

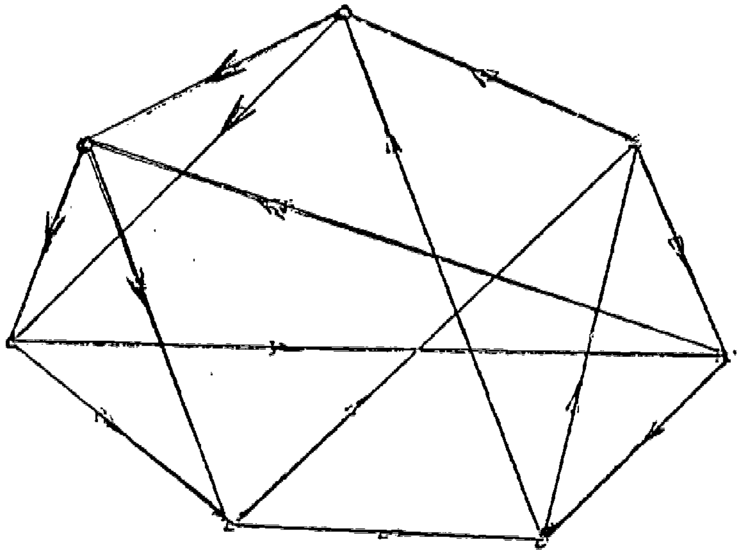

Figure 10. A two-regular network on seven nodes.

breaking of a single arc (information transfer channel) in the network is enough to disrupt communciations. At least one node will not be able to access one or more other nodes. The decentralized network is reliable in the sense that breaking one arc will not disrupt information transfer: there are always other paths to follow. However, as we mention earlier, this type of network involves so many arcs that it is quite expensive. Suppose that we impose the following reliability criterion on the network: from any node to any other node there must be at least two information transfer paths which have no arcs or nodes in common (other than the end nodes). This criterion has been adopted, for example, in the design of the ARPA computer network. What properties does this force on the network? The minimal network which satisfies this criterion is the two-regular network, which has exactly two arcs entering and two arcs leaving each node (Figure 10). Such a network is relatively inexpensive to install, and yet is reliable in the sense that loss of a single information transfer channel does not disrupt the network.

Any realistic network combines the features of each of these "pure" types. For example, one might want PLAN to look like a decentralized network within each of several regions - say, within New England, the southeastern states, and along the Pacific coast. Between these decentralized subnets, one might want PLAN to be two-regular, or cyclic. Then, thinking in terms of state library networks associated with PLAN, one might want the network within each state to be hierarchical. However, at the present stage in modelling information networks we find it exceedingly difficult to properly handle such a complex design. Thus we restrict our attention to the four "pure" types defined above. It should be noted, however, that the general concepts which we define are valid for all networks.

GRAPH CONNECTIVITY.

One criterion which is generally placed on information networks, particularly those involving libraries, is that each user of the network must have access to each information resource within the network. (There may be security classifications or other restrictions which prevent a user from actually reaching a document, but there should be no such restrictions due to the network design.) This concept of <u>total accessibility</u> is related to the graph-theoretic concept of connectedness.

We say that a digraph is <u>weakly connected</u> if, disregarding the directions assigned to the arcs, given any two nodes, n_a and n_b, there is a chain of arcs leading from n_a to n_b. (See Figure 11.) If this is not the case, if there is at least one pair of nodes between which there is no such chain of arcs, we say that the digraph is <u>disconnected</u>. Since we are interested in indicating information flow by the directions assigned to the arcs, weak connectivity is not a sufficient concept for our information network.

We come closer to the desired concept of total accessibility if we require that the digraph be <u>unilaterally connected</u>, that is, that between any two nodes n_a and n_b, there be a chain of arcs which are consistently directed. Thus we should be able to find a chain of directed arcs leading from n_a to n_b, or a chain leading in the opposite.

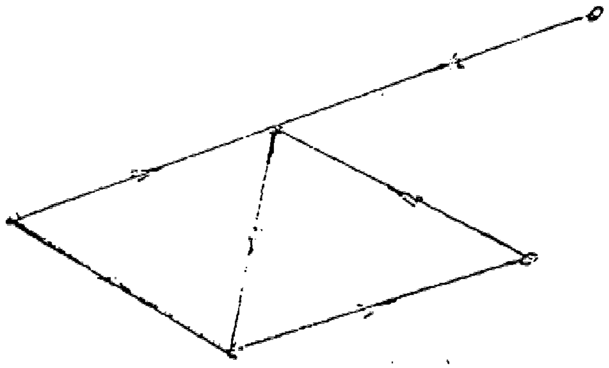

Figure 11. A weakly connected digraph.

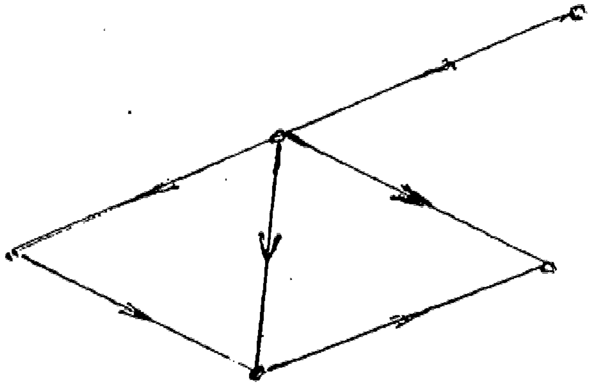

Figure 12. A unilaterally connected digraph.

direction (Figure 12). Unilateral connectivity provides a one-way path. We are able to get the request from the user to the library having the document, or we are able to get the document to the user; but not necessarily both.

Clearly this is less than desirable: there is no sense in getting the request to the proper library if we cannot respond to that request. We need to strengthen the concept of connectivity still further. A digraph is strongly connected if given any two nodes, n_a and n_b, there is a chain of arcs directed from n_a to n_b, and another chain of arcs directed from n_b to n_a. (See Figure 13.) This is the connectivity definition which is needed to realize our concept of total accessibility. One chain of arcs takes the request message from the user at n_a to the library having the document, at n_b. The other chain of arcs returns the document to the user (Figure 14).

Returning briefly to the four types of networks which we have defined, the cyclic, decentralized, hierarchical, and two-regular, we see that each of these is in fact strongly connected. Moreover, if we construct a complex network using these four types as components, and if we insure that the components are at least cyclically connected, then we have a strongly connected network which is suitable for information transfer.

PARAMETERS FOR NETWORKS.

This basic skeleton, the strongly connected digraph, is suitable for network design. However, we must associate with each node and each arc

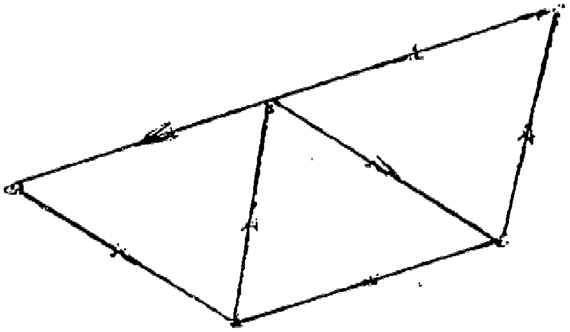

Figure 13. A strongly connected digraph.

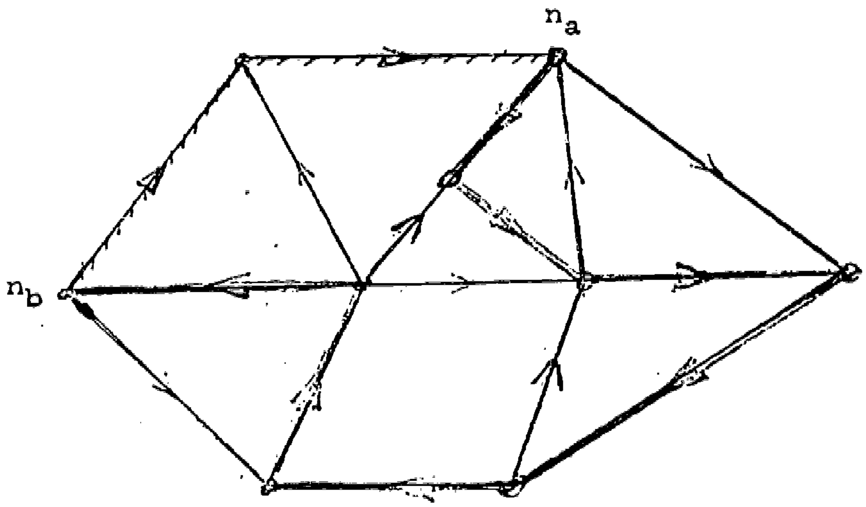

Figure 14. Information transfer paths from n_a to n_b, and from n_b to n_a.

a host of parameters which define the properties applicable to the particular network. While the total behavior of any information network results from the explicit relationship among the nodes and arcs, certain parameters are associated primarily with the nodes (information centers) and others with the arcs (information transfer channels). Thus one may not assume an independence among nodes and arcs; yet one can identify parameters reflecting properties of each without necessarily excluding the other. Parameters associated primarily with the information centers include:

1) the information media,
2) storage requirements,
3) access modes,
4) computer equipment,
5) volume of work load,
6) classification methods,
7) costs associated with the nodes, and
8) probability of servicing a request.

While these parameters influence the operation of the network, we need not be concerned here with the precise manner in which this happens. The problem of adjusting these parameter values is precisely the problem facing designers of libraries and other information centers. We assume that the nodes have certain characteristics, and then in designing the network we use the node characteristics as fixed values. Of course, changing the node characteristics requires redesign of the network.

Of more interest to us are the parameters which are associated primarily with the arcs of the network, that is, with the information transfer channels rather than with the information centers. These parameters include

1) type of transfer,
2) volume of information, in terms of messages and documents,
3) permissible rates of transfer,
4) accuracy of transmission,
5) origin,
6) destination, and
7) cost of information transfer.

The information transfer may involve one of a number of different types of channels. These include messenger service, telephone calls, facsimile transmission, the mail, teletypewriter or other computer terminal transmission, and so forth. Some types of transmission are more suited to the request message than to document transfer; others are equally suited to message and document transfer. One thinks of the typical request as being rather short, hence suitable for expensive high-speed transmission, while a document, being longer, is more suitable for slower, less expensive transfer modes. The distinction between message and document transfer is stated in two previous papers [3], [4].

The volume of transactions to be conducted on the network also influences the design. Here again, one thinks of requests as being more frequent than documents in response to requests. This is partly true since the request may be replicated and sent to many parts of the network, but generally only a single copy of a responding document is required to be transferred.

Already we have alluded to the fact that there are both high-speed and low-speed modes of information transfer. Often one finds that a range of alternatives exists within a network for any single transfer. For example, a document in response to a request might be sent by facsimile, by messenger, or by mail. In actual operation, generally limits are set on either response time, or the cost of information transfer, or both.

The origin and destination of information transfer over the network have a large influence on network design. As we mentioned above, the request message may be replicated and sent to all, or nearly all, information centers in the network. While union catalogs or switching centers may exist in some networks, often the user submitting the request does not know the location of the desired information resource. However, once the document is found, its destination, the user, is known. Thus the transfer paths generated to send this document back to the user are far fewer in number than those necessary to handle the request message. And of course one of the problems persistently plaguing network designers is that of the uneven distribution of users and documents within the network.

Finally, the cost of information transfer is a large factor in network design. Certainly cost is interrelated with many other factors that we have mentioned, and it is entirely possible in an actual network to have parallel information transfer channels at substantially different costs. Hopefully a difference in service accurately reflects the cost difference.

While any comprehensive model must take into account all of the factors which we have discussed, and other similar ones, in the present paper we wish to concentrate on the aspects of the problem which can be modelled by a digraph without including the various parameter values. In particular, we wish to define one measure, and introduce two graphs which are derived from any given network graph, all of which help shed some light on the problems of network design.

NETWORK FLEXIBILITY.

One measure which might be applied in judging a network design is the freedom of choice confronting an information center placing a request on the network. For example, a center in the cyclic network has no choice in the path by which it must route a request, nor in the path by which the response comes. There is only one path around the network. However, a center in the decentralized network has a great variety of paths available for the routing of request messages and document responses. Not only is there a direct connection from one information center to every other one, but also in the event that one connection is broken there are still a number of alternate routes. With this in mind, we define the <u>flexibility</u> of a network to be the quantity

$$F = \frac{Q - N}{N(N - 2)},$$

where Q is the number of arcs in the network, and N is the number of nodes. Note that since a cyclic network has $Q = N$, for such a network the flexibility is $F = F_c = 0$. Similarly, since a decentralized network has $Q = N(N - 1)$ arcs, for such a network the flexibility is $F = F_d = 1$. Continuing in the same vein, a hierarchical network has $Q = 2N - 2$ arcs, hence a flexibility of $F = F_h = 1/N$; a two-regular network has $Q = 2N$ arcs, hence a flexibility of $F = F_t = 1(N - 2)$.

These flexibility calculations for the four special networks exhibit properties which are valid for any network. First, since the cyclic network has the minimum number of arcs for a given number of nodes, and the decentralized network has the maximum number of arcs, it follows that for any network $0 \leq F \leq 1$. The values for F_h and F_t exhibit the following general property: if the number of arcs in a network is proportional to the number of nodes, then the flexibility varies inversely with the size of the network. In other words, if we set constant flexibility as a desirable criterion of network design, then as we increase the size of the network the number of arcs must increase more rapidly than the number of nodes - in fact, as the square of the number of nodes. For example, if we were to expand the PLAN network from forty-nine nodes to four hundred ninety nodes, then to maintain the flexibility we would need to increase the number of arcs one hundredfold.

Let us now consider the two graphs derived from a network, which are of assistance in analyzing the network. The first of these defines, in some sense, the "core" of the network; while the second provides an overview. In many networks - for example, the hierarchical networks - we find that in a portion of the network arcs occur in anti-parallel pairs. That is, we find an arc from n_a to n_b, accompanied by one from n_b to n_a. Let us call the graph consisting of all such pairs of arcs the two-skeleton of the given network (Figure 15). In some sense, the two-skeleton represents the "core" or "spine" of the network, along which two-way communication is possible. Note that for the decentralized network and the hierarchical network the two-skeleton is the entire network, while for the cyclic network on three or more nodes the two-skeleton does not exist.

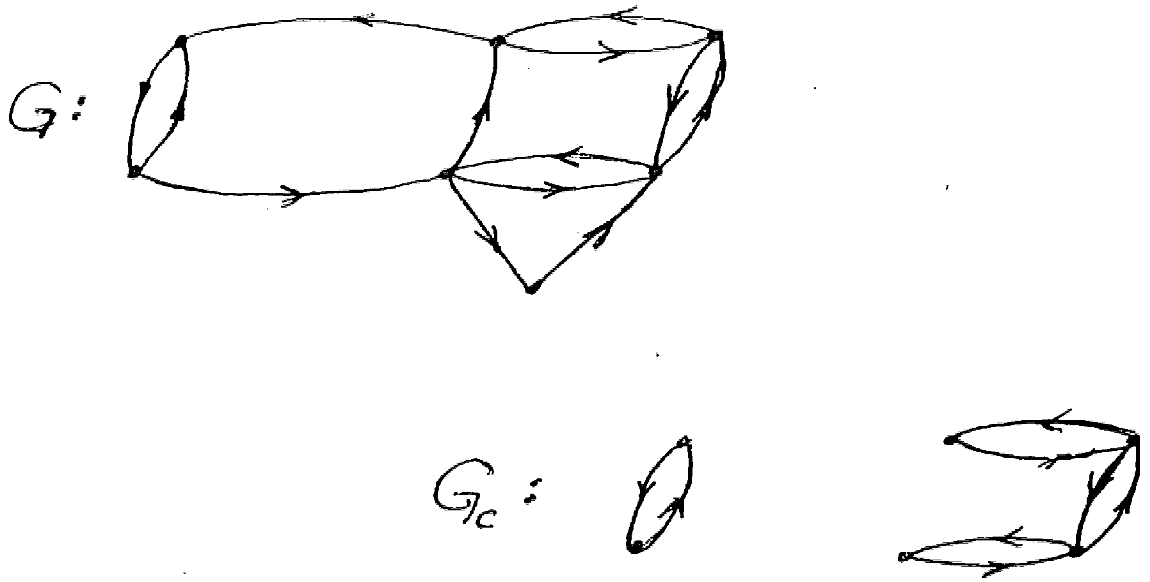

Figure 15. A strongly corrected graph G and its two-skeleton G_c.

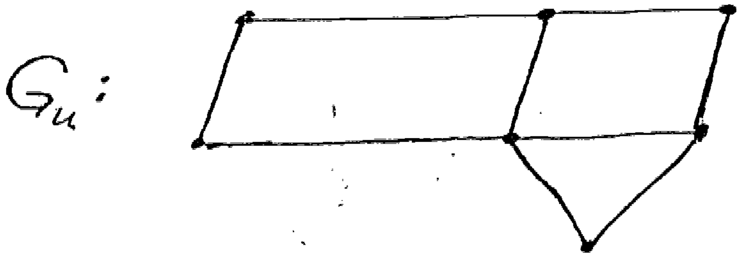

Figure 16. The undirected graph derived from G of Figure 15.

The other derived graph is the undirected graph obtained by removing all directions from the arcs of the network and identifying parallel edges (Figure 16). This graph, showing simply the connections existing throughout the network, without regard to the direction or other properties of these connections, is useful in determing connectivity and several other properties of the network, although by dropping the directionality of the arcs, we do lose some network characteristics.

THE PURPOSE OF MATHEMATICAL MODELS.

We might ask at this point, "Why the need for mathematical rigor?" This question is asked almost as frequently as the comment is made that "these problems cannot be solved by mathematics and models." Why indeed should we go to all the trouble of defining network models thoroughly and carefully? Why not merely assume that everyone knows what a library information network is, and let it go at that? The answer lies in the removal of ambiguity by mathematical definition. One may disagree with the definitions, assumptions, and axioms that are stated; but if the mathematics is correct, one cannot disagree with the conclusions developed from these definitions, assumptions, and axioms. One can, of course, offer different definitions and assumptions, and observe their effect on the resulting conclusions. In such a way one can compare different models, or the same model as applied to different networks. The flexibility measure is a good example of this. One may argue that this measure is not the correct one, and offer alternatives. But assuming that the measure is at least reasonable, one can then draw, as we have, firm conclusions about the characteristics of networks according to the measure.

To the comment on the inadequacy of mathematics, we must agree that mathematics can neither describe nor solve *all* problems associated with information networks. *But neither can any other empirical or analytical technique*. Where applicable, mathematical models should be employed to describe and/or solve network problems amenable to mathematical solution. The modeler must accept the responsibility of stating the premises on which his model is built and conveying the implications of the conclusions that follow. In such a spirit we have conducted this work.

THE GENERAL NETWORK MODEL.

We have discussed a number of different concepts - the "pure" types of networks, the ideas of connectivity, and the parameters associated with a network. We now present a general model which brings together these concepts.

An information network N is a sextuple

$$N = \langle \mathcal{U}, \mathcal{I}, C, A, f, f' \rangle ,$$

where the components of N are defined as below.

$\mathcal{U}$, $\mathcal{I}$, and C are the nodes of the network, representing the *users*, *information resources*, and *information centers* respectively. We require that with each information center $c \in C$ there be associated a non-empty set $U \in \mathcal{U}$ of users, or a non-empty set $I \in \mathcal{I}$ of information resources, or both.

A is the set of directed arcs on $\mathcal{U} \cup \mathcal{I} \cup C$, where an arc (n_a, n_b) denotes that node n_b is directly accessible from node n_a, and where each arc (n_a, n_b) joining nodes of C carries one or both of the labels

m - denoting possible <u>message</u> (request) transfer from
n_a to n_b, or

d - denoting possible <u>document</u> (response) transfer from
n_a to n_b.

Thus we have a typical network model shown in Figure 17. Each user ($u \varepsilon \mathcal{U}$) has direct access to one or more information centers, and each information resource ($i \varepsilon \mathcal{I}$) is accessible to one or more centers. Note that no two users are directly connected: they must communicate via one or more centers. Similarly, no two information resources are directly connected.

The central portion of the network represents the connections between the information centers. It is this portion which really interests us. Within the central portion of the network each arc carries one or both of the labels "m" and "d". The arcs labelled "m" denote channels along which messages or requests for information may be passed. Together, these arcs and their associated nodes form a digraph, G. In the light of our earlier discussion we assume that G be strongly connected, that is, that there be at least one message channel from any given center to any other given center.

Similarly, the arcs labelled "d" denote channels along which documents or responses to requests may be passed. As noted earlier, these channels may not coincide with the m-labelled channels. The digraph G' formed by

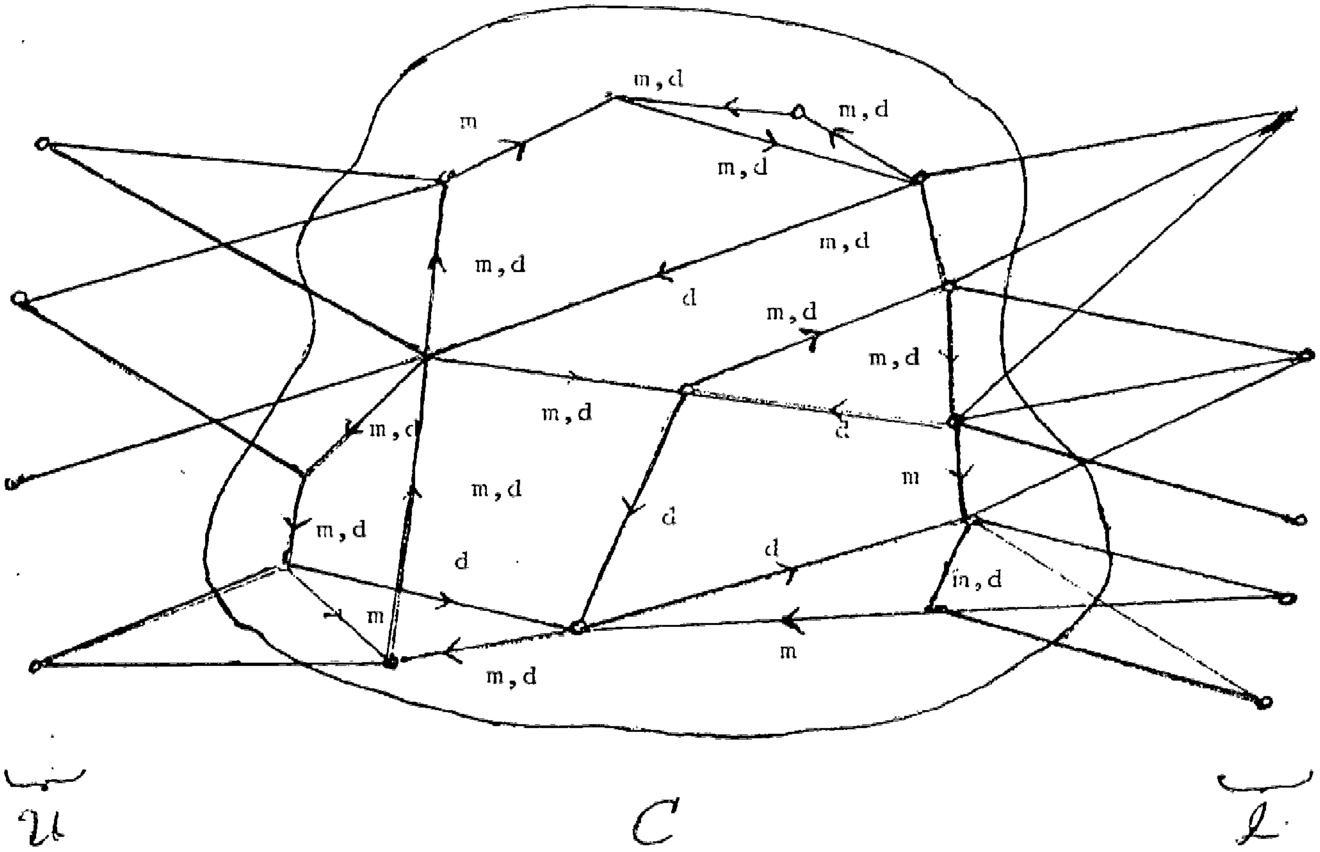

Figure 17. An example of a 𝒰-ℒ-C network

the d-labelled arcs and their associated nodes is also required to be strongly connected. Thus the central portion of our network model, which relates to the information centers, is covered by two strongly connected digraphs, G and G'.

These two digraphs covering $\mathcal{C}$ may be any of the "pure" types which we have discussed, or they may be more complex digraphs. For example, G might be hierarchical, requiring that a request filter gradually up through the network until it can be satisfied. At the same time, G' might be decentralized, so that the document selected in response to a request can be sent directly to the center originating the request.

We have yet to explain the last two components of N, namely f and f'. These mathematical functions define the <u>information transfer structure</u> of the network. Note that since N is strongly connected, any user can access any information resource. But we are interested in <u>how</u> this is accomplished. We wish to be able to compare the costs associated with various modes of access.

Within the digraph G one can define several open paths, that is, consistently directed chains of arcs which do not pass through any one node more than once. Given a user $u \in \mathcal{U}$ attempting to access an information resource $i \in \mathcal{I}$, we are interested in the set of all open paths $\mathcal{P}$ enabling this access. For each path $P \in \mathcal{P}$, the value of f(P) is the set of all ordered pairs (u, i) where $u \in U$ and $i \in I$ are joined by the path $P \in \mathcal{P}$. Thus f(P) defines the paths (including user and information resource) that can be used by a particular user in order to access specific information in the network. Note that all arcs of P are labelled "m".

The function f' is similarly defined for paths P' (consisting of d-labelled arcs) in the digraph G'. Thus f determines the alternatives for message or request flow, while f' determines the alternatives for document or response flow.

This then is our general network model - a model which embodies the concepts of the user, the information center, the information resource; the ideas of message and document flow; and functions defining the information transfer structure for the network. From this basic structure we may proceed in a number of directions. We certainly want to vary the central network structure. In fact, it appears useful to classify networks according to the G-G' structure, as shown in Figure 18.

Furthermore, we wish to attach to the network model some or all of the parameters which we have discussed, and to investigate the relationships among these. For any given assignment of parameter values we find a "topography" for the network - information which is available to the user at various cost levels, $c_1 < c_2 < c_3$ .... Such topographies help us define and study the fine structure of information networks, and refine our concepts and modelling techniques so that we may easily and accurately model the large library networks necessary to handle today's information flow.

G' G	cyclic	decentralized	hierarchical	two-regular	mixed
cyclic					
decentralized					
hierarchical					
two-regular					
mixed					

Figure 18. Schema for the classification of library networks.

REFERENCES

1. Harary, Frank Graph Theory, Addison-Weley, 1969.

2. Harary, Frank, Robert Z. Norman, and Dorwin Cartwright. Structural Models: An Introduction to the Theory of Directed Graphs, John Wiley & Sons, New York, 1965.

3. Nance, Richard E. "An Analytical Model of a Library Netowrk", J. Amer. Soc. for Info. Sci., 21 (1): January-February 1970, pp. 58-66.

4. Nance, Richard E., Robert R. Korfhage, and U. Narayan Bhat. "Information Networks: Definitions and Message Transfer Models", Technical Report CP-710011, Computer Science/Operations Research Center, Southern Methodist University, July 1971.

MPUTER SCIENCE /
ERATIONS RESEARCH CENTER
IU INSTITUTE OF TECHNOLOGY
DALLAS, TEXAS 75222