U. S. AIR FORCE

PROJECT RAND

RESEARCH MEMORANDUM

ON A DISTRIBUTED COMMAND AND CONTROL SYSTEM CONFIGURATION

Psul Baran

RM-2632

December 31, 1960

Rand
SANTA MONICA, CA. 90406

SUMMARY

Our present concept of command and control systems envisions one or a few centralized command posts, together with centralized computational apparatus, used to process raw data and to transmit detailed instructions over low-redundancy networks to eventual action points.

The advent of the very-low-CEP missile permits the enemy to effect a high probability of disabling the entire system by careful choice of a small number of targets.

The feasibility of a less vulnerable system configuration that utilizes redundancy of command, computation and communications is briefly examined.

CONTENTS

ON A DISTRIBUTED COMMAND AND CONTROL SYSTEM CONFIGURATION

1. INTRODUCTION

Recently J. B. Carne, a staff member of The RAND Corporation, posed three questions concerning command and control communications:

1) Are we serious when we talk about command and control communications vulnerability?

2) If we do not really wish to pay the price for real hardness, then are we not wasting money providing intermediate, sham hardness, worthless during the post-attack period?

3) If so, then might we not provide only the minimum network capability to service our peacetime requirements and eliminate a possibly unfounded dependence upon a capability that will not exist when needed?

The present memorandum starts with the assumption that invulnerability is a necessary requirement for certain command and control systems. Next, the need is shown for redundancy of command structure and redundancy of computational apparatus as well as the need for redundancy of communications. Lastly, the technical feasibility of achieving these three goals is discussed.

2. THE NEED FOR CERTAIN COMMAND AND CONTROL SYSTEMS TO SURVIVE ATTACK

In speaking on thermonuclear war, [1] H. Kahn of the RAND staff has indicated that we need the following command and control capability to withstand the first strike:

1) A fast-response data channel between the effective heads of state of the adversaries. In other words, how can you negotiate, if you cannot communicate your wishes?

2) A capability for the rapid inventory summation of military capability.

3) A capability for the issuance of "nonpredeterministic" military response plans.

4) A capability for maintaining communications between the population and the effective government.

The assumed post-attack environment in this paper will be as described by Kahn. We shall consider the question, "Will the command and control systems being planned at present survive Kahn's war phases 2 through 4?" Specifically, these phases are the following: wartime performance with different pre-attack and attack conditions, the acute fallout problem, and survival and patchup.

3. IMPLICATION OF THE SMALL CEP - HIGH YIELD WEAPON UPON THE COMMAND AND CONTROL SYSTEM

It is clear that we have passed a technological milestone where the CEP of an enemy weapon is approaching a value smaller than the weapon's lethal radius against any conventional electromagnetic communications antenna and buried telephone cable. Heads of state have engaged in an exchange of boasts of the low measured miss distances of warhead delivery systems. The proven development of the small CEP ICBM must evoke a major change of concepts of air defense. Although estimates for such small CEPs have been heard for several years now, it is hoped that measured and advertised values of these tests will convince the command and control system designer of the possibility of the existence of a weapon system capable of a single-shot probability of kill approaching that of the probability of non-abort of the missile itself. The enemy may destroy almost any known point target above ground which it is to his advantage to destroy.

Our communications systems are for the most part conceived upon the model of hierarchical networks. An entire network may be made inoperable at the cost of a few missiles cleverly directed against a few switching center targets. Even if an infinitely hard command center is assumed, the command center remains an ideal target for covert attack. How will our delivered big-L systems respond to such attacks?

Contractors' proposals for the big-L systems have been written in compliance with a short deadline and probably without a full appreciation of the military threat posed by the enemy. Experience shows that the requirement of immunity to enemy action is usually the most difficult system design goal to meet. Thus, we find that the vulnerability requirement is essentially ignored during construction of most large military systems--not by desire, but out of the necessity of meeting most of the other system requirements.

We should appreciate the fact that electronics companies have to operate under many pressures. At times, the military sales department of a contractor will bid on a CPFF contract regardless of how nearly impossible the requirements may be. We should not be surprised if the end system fails to meet the specifications found in General Operational Requirements.

4. THE COMMUNICATIONS VULNERABILITY SOFT SPOT

It is theoretically possible to modify any command post to any desired hardness by digging a deeper and deeper hole in the ground. Eventually a depth is reached below which the shock of the nuclear weapon is so diminished that survival is assured. A more difficult problem is to assure a comparable degree of survival for communications out of the hole.

All electromagnetic systems of communication, with the exception of ELF and VLF, require above-ground antenna structures. ELF and VLF systems are limited to extremely narrow bandwidths and low ranges for reasonable power levels and buried antennas of reasonable size. Seismic communication suffers an even more restricted range and bandwidth capability; buried telephone cable is extremely expensive if placed at the same desired depth as the command post; and shallow burial creates a relatively soft target. Mixtures of highly redundant cable and antenna structures appear to hold somewhat more promise, but whether such a system would stymie a determined enemy is not at all clear.

5. IS COMMAND DECENTRALIZATION POSSIBLE?

1) Why do we find ourselves placing full reliance upon a relatively few critical point targets?

2) Is it necessary to organize our command and communications systems so that these "Achilles' heels" exist?

3) Is it possible to design decentralized organizations where over-all control may be safely vested in a large number of places--so large a number as to increase to a prohibitive level the cost of attacks directed at the command centers?

4) What are the requirements for command?

5.1. Separation of Data Networks from the Chain of Command

Wars are generally conducted in an environment where information about the enemy is incomplete. Other things being equal, the best command will emanate from that individual who at that instant possesses most information as to the status of the defense and the status of the offense. The kibitzer can always beat either player in kriegspiel.

Effective military command therefore requires an information transmission and processing capability. In a non-automated management organizational structure, the chain of command and the communications structure are usually regarded as synonymous. It would be helpful if we view the chain-of-command networks and the communications networks as separate entities. The communications network in turn may be factored into two usually coincident networks: a data gathering network and a

command transmission network. Restated, we shall consider three separate networks all having the same topology: a hierarchical chain of delegated authority; an information network to transmit information from the working levels of the organization of the hierarchical chain; and a command network to transmit commands from the point of authority down to the working levels. In a full hierarchical organization we are concerned with the vulnerability of the apex of any of these three coincident networks. The enemy can almost guarantee that the entire organization will be disabled by selective damage of a chosen small chunk of any of these three networks at the modest price of a slight overkill. How, then, should we go about designing these structures so that destruction of one portion is no more disabling than destruction of another portion?

As a starting point let us examine the basic reason for transmitting information up and down the hierarchical chain of command.

The information network provides a maximum of distilled data to the commander. A single human's data input channel is easily saturated by too much data supplied too fast. The command transmission network starts with a highly generalized command. As the single command is relayed down the line it is amplified and appropriate action is taken to implement the command.

One of the proposed functions of a command and control system is automation of the information distillation and

transmission channel. In an automated system the data needed to formulate decisions may be collected in a plurality of geographically remote places merely by duplication of this portion of the communication network.

What is to be clearly noted is that the information network need no longer be tightly coupled to the hierarchical chain of command but may be duplicated at will. Thus, by judicious replication of apparatus, as many "potential" command posts may be created as the pocketbook will allow.

5.2. The Command Generation Subsystem

We are careful in our choice of the leaders in whom we will comfortably vest the responsibility of command. The occupation of command has been assigned to those individuals possessing unusual capabilities. Given a team of 100 men, much time and consideration will be given to seeking out that one individual possessing most of that elusive quality called "leadership." A wide gap is assumed to exist between the effectiveness of the chosen leader and the next ranking alternate. The reason for this is not clear unless it is the fundamental difficulty of providing the alternate leader with the necessary operational experience.* Thus, we must proceed with caution before the statement is made that the command function may be safely performed from a large number of

*W. Ware of the RAND staff has suggested that the new simulation techniques that permit exercising potential leaders might mitigate this problem in the future.

alternative sources. The almost religious regard for the differential acceptance of one individual in the command role versus another is a factor that cannot be lightly dismissed.

How then can we design a system around a concept of potential commanders when the differences of performance are assumed to be so great? Without stating the pros and cons, some conceivable techniques might include:

1) A predetermined "ascendancy to the throne" ordered list.

2) A voting scheme wherein each decision is reviewed by all surviving potential commanders on an equal basis.

3) Voluntary selection of a de facto commander on a per act basis by that individual claiming most knowledge of that act. This technique might seem bizarre, but it occurs very often, and is described in more detail in the next section.

5.3. Virtual Command

It is common practice in military organizations to sign each letter, "Commanding General," although the hand holding the pen may not be that of the Commanding General. In such a circumstance, the writer of the letter may be called the virtual commander.

The question one might now ask is, "What is so sacred about a single commander that he and only he is entrusted with the authority of over-all decision?" Is not authority synonymous with knowledge, and can we not build networks that provide this knowledge to a number of points? This concept is succinctly stated by Warren McCulloch, who describes the shift of point of

control during naval battles: "It is a redundancy of potential command wherein knowledge constitutes authority." He further points out that there is nothing magical about the choice of commander; just as long as we pick the person or persons with most information about a situation, he will probably make good decisions.

To review what has been said:

1) Redundancy of input information is necessary at all possible potential command posts.

2) The amount of redudnancy necessary is that degree that causes attacks directed against the command structure to be no more effective than those directed against other targets of the system.

3) Physical hardness reduces the degree of redundancy necessary, but the limitation of hardness, at least for the present, appears to be that obtainable for the exit communications from extremely hardened command points.

4) With the redundancy of command function and command posts, possible great advantage may be made of secrecy. That is, the defense may play the old shell game; this time with a large number of shells and a sizeable number of peas. The enemy must find <u>all</u> the peas in order to win.

In summary, then: in order to build command and control systems capable of meeting the requirements of the passive air defense concept, physical hardness of command posts and exit communications, redundancy of information networks, redundancy of commanders all appear to be indicated.

6. IS COMMUNICATIONS DECENTRALIZATION POSSIBLE?

The feasibility of synthesizing decentralized communications networks is described in detail in the Appendix. In the Appendix summary there is a spectrum of system mixtures possible ranging from a pure hierarchical network to a non-hierarchical or fully distributed communications network.

Further, it is possible to build distributed digital data links wherein the point of origin may be authenticated, yet such networks are able to withstand a high degree of link destruction and remain operable provided that at least one link path exists, no matter how complicated a routing path is taken.

6.1. Implication of Distributed Computation

To this point the concept of distributed command and communications has been discussed. We normally think of computers (or future information processors) as highly centralized organs. In order to distribute vulnerability we must also decentralize the computational operation, or else we have gained little. How then shall we decentralize computers in the operational system? Here we must add other problems to the list of essentially unsolved problems facing the computer designer--who may still be thinking in terms of a single highly centralized computer. These new problems include:

1) What is the best way to organize a network of computers to use the fullest capability of whatever portion of the system survives the attack?

2) What is the best method of utilizing information held in the stores of interconnected computers?

3) How can malfunctions and enemy action be isolated to prevent a single defective computer from sending erroneous or fraudulent data to other computers on line?

4) How does one go about trouble-shooting or debugging a complex network of computers? How does one determine who did what to whom, when?

5) What interesting new properties does an interconnected network of computers offer?

6) If it is necessary to duplicate computers at each communications node, how can the unit cost of such replicated computers be kept low enough to permit the system to be built?

7) Is it possible to use the computers themselves to perform much of the switching and housekeeping in relaying digital messages in non-hierarchical communications networks?

8) Is there any way of using the peak load potential of our general purpose digital computer resource, part of which will survive the attack?

9) Is the lead time of new computer design sufficiently great to eliminate its consideration for the generation of command and control systems now reaching the design phase?

The next generation of large-scale digital computers (e.g., the "polymorphic" Ramo-Wooldridge 400, the RCA 601, and the Hughes Modular Computer) is being designed with less centralized internal organization. While these machines may not be wholly applicable to our problems, the development of such machines indicates at least a partial solution of problems 2, 3, 4 and 5.

6.2. Modes of System Death

Assuming the existence of the trio of the three distributed "C's"--Command, Communications, and Computation--we may

ask ourselves, "How will such systems respond during the post-attack period?" Of course we can never be sure, but we suspect that they will exhibit the property of the MacArthurian old soldier: at best they will not die, but merely fade away. If they must die, they should die gracefully.

System performance may be expected to degrade under heavier and heavier stress, but never should the entire system cease operation merely because the enemy has selected reasonable targets, from his point of view, which are critical links for our system.

This, of course, implies that the total system cost will be higher, or that we accept less effective peace time performance, in order to obtain a better post-attack capability. Although a redundancy requirement often brings out the worst in the system designers, it should be remembered that redundancy is fundamental to continued system performance.

7. CONCLUSIONS

The rationale for proceeding with these concepts is that the first duty of the command and control system is to survive. After assured survival we may look at the niceties of how to make the system work. There has been the temptation in the past to merely ignore the vulnerability problem and proceed on the more pleasant aspects of the system design, hoping that these partial results will be immediately useful once the vulnerability problem is somehow solved.

It is therefore suggested that the vulnerability situation be carefully examined earlier rather than later. If the fears stated are supported by fact, then perhaps it would be well to investigate the implications of the need of distributed computation and management structure and examine the feasibility for obtaining such capability in the required time period.

Thus, it may be appropriate to consider a critical examination of the following questions:

1) Is a little post-attack command and control capability worth more than a large capability that requires enemy cooperation to survive?

2) Is a minimum command and control system capability needed on a quick-fix basis?

3) Do "dream" systems that fully satisfy GOR's require quantum jump advancements in the state of the art of:
   a) computers and information processing,
   b) management concepts, and
   c) communications?

4) Should funds now being expended for construction of post-attack systems which will neither survive an attack nor reasonably meet long-term GOR's be invested in more basic development and research?

APPENDIX

Is Communications Decentralization Possible?

Decentralization of communication may mean different things to different people, so a few words may be in order stating the writer's view of the spectrum between centralized and decentralized networks.

Hierarchical and Non-Hierarchical Networks

For ease of visualization the following discussion is presented in pictorial form.

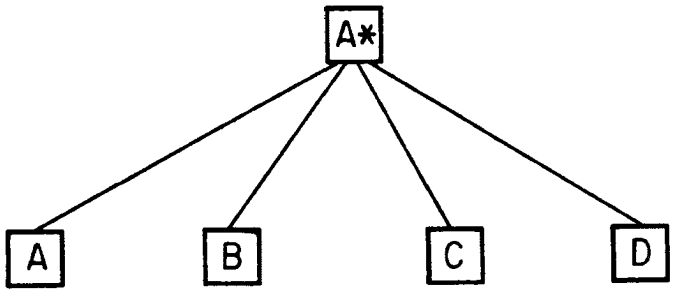

Fig. 1 -- A Small Hierarchical Network

Figure 1 shows what, for obvious reasons, may be called a small hierarchical network. It is called small because there are only four stations, A, B, C, and D. It is called hierarchical because the lines of data flow resemble a hierarchical chain. Thus, stations A, B, C, and D speak only to the upper relay station, A*. Information from station A to station B must always pass through the single relay station, A*.

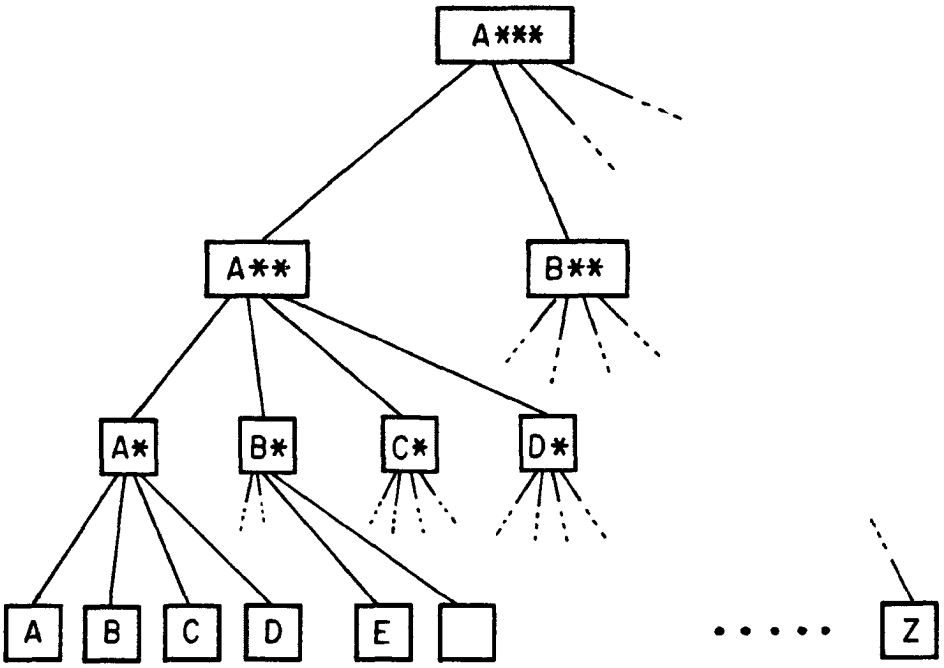

Fig. 2 -- A Large Pure Hierarchical Network

Figure 2 shows a group of such small networks organized into a larger hierarchical network. Note that information from A must always pass up through A* thence to A** thence to B* to be transmitted to station E. This, of course, places a high premium upon the destruction of stations A*, A**, B*, etc. We are interested in other network configuations that minimize this attacker's bonus.

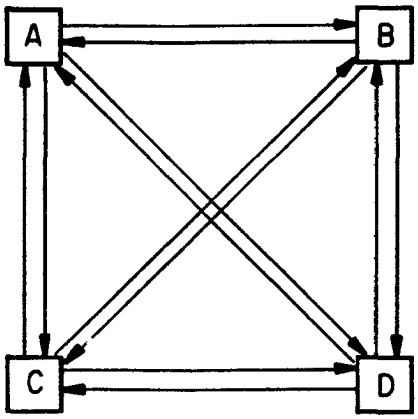

Fig. 3 -- A Small Pure Non-Hierarchical Network

Figure 3 shows a configuration indicating a more desired data flow between stations A, B, C, and D. Here, each station is given a separate channel with which to communicate with all others.

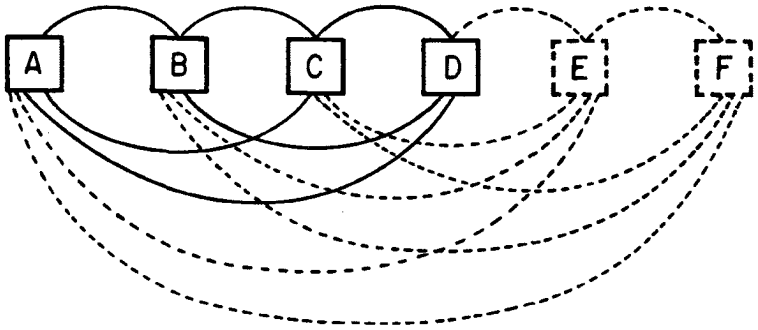

Fig. 4 -- (Figure 3 Redrawn)

Figure 4 merely shows the same network redrawn so that the paths may be better compared with Fig. 1. Two additional stations, E and F are shown in dotted lines in this illustration. The total number of communications links required here is n(n-1), where n is the number of stations. If the number

of stations, n, becomes large, the number of such links becomes exceedingly large. For example, if n equals 100 then the number of links is 9,900. Thus, there is an economic limit to the extent of the use of this form of decentralization of network.

With the realization of the price of the arrangement shown in Fig. 4, we are tempted to examine a pure parallel redundancy hierarchical network.

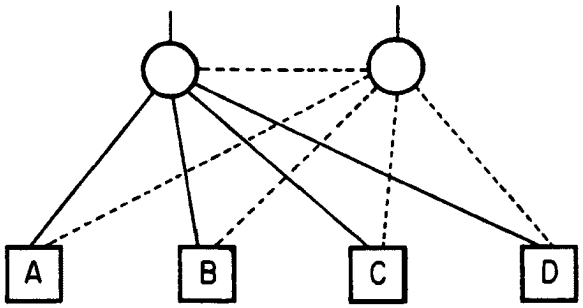

Fig. 5 -- A Parallel Redundant Hierarchical Network

In Fig. 5 the number of links that must be cut to break communications increases linearly with the degree of redundancy.

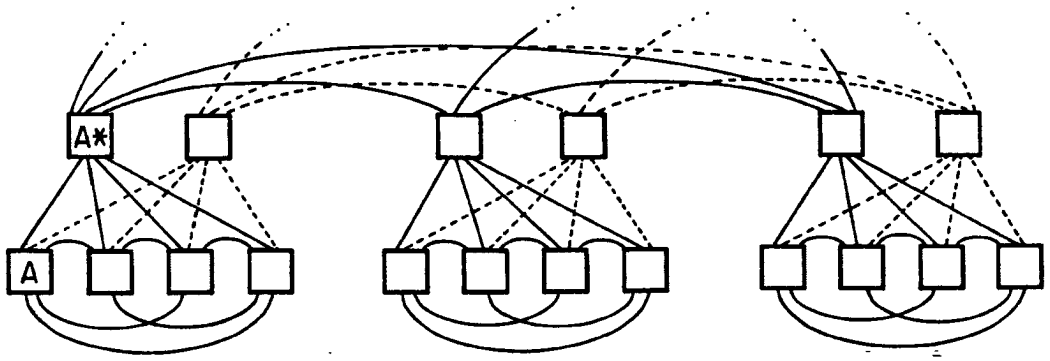

Fig. 6 -- Parallel Redundancy Plus Non-Hierarchical Grouping

Figure 6 shows a mixed system. Here groups of stations are connected in a non-hierarchical network. Remember, the product of n(n-1) is small where n is small. These groups are then connected together in an over-all hierarchical network configuration. Each such level is then connected horizontally with another non-hierarchical network. We can see that a network configuration is beginning to take shape where it is more difficult to break off large pieces of the network from signal contact with the remainder of the network. The limitation of this method is still the weakness of the connections to the top of the pyramid. This network is still a good target even though its destruction leaves relatively large blocks intact but out of contact with one another. Decision actions taken by any one of these large groups might be inconsistent with another's goals.

Let us examine the manner in which communications systems are generally designed; that is, on a geographical basis, where the key optimization criterion is usually minimum cost.

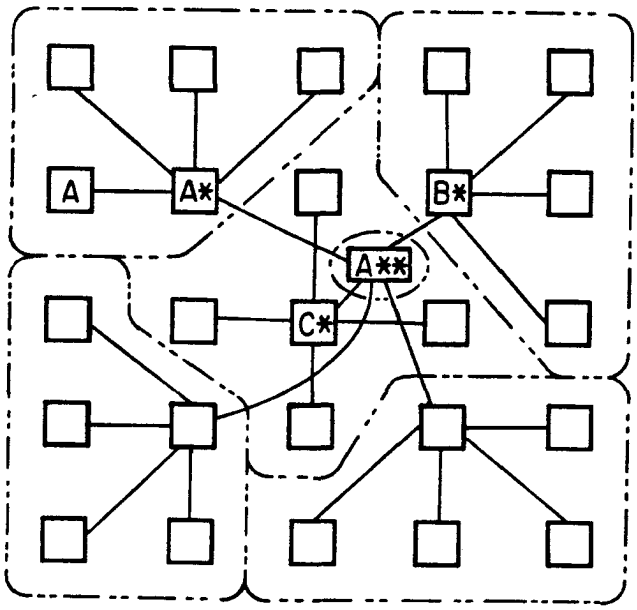

Fig. 7 -- Fig. 2 Redrawn on a Geographical Basis

Each hierarchical group of stations is located near one another geographically and transmits to a central point. Each of these central points then retransmits to a central station. Many national communications systems are organized on this exact basis. For example, SAGE data is transmitted essentially in such a manner.

Thus, another way of viewing Fig. 7 is shown in Fig. 8. Here the stations are shown diagrammatically around a circle while the communications links are shown within the circle.

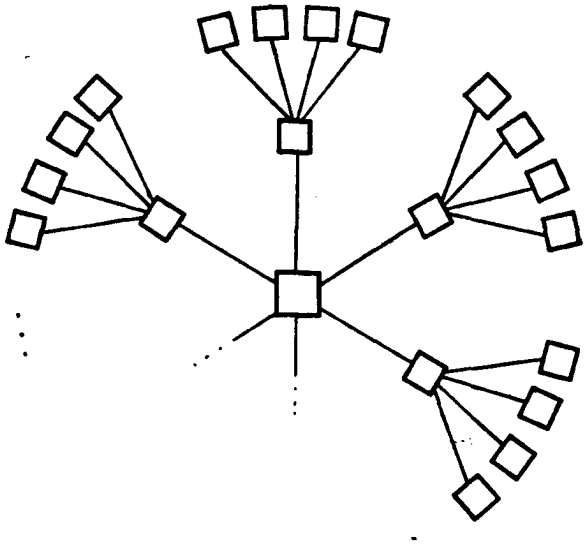

Fig. 8 -- Fig. 2 Again Redrawn

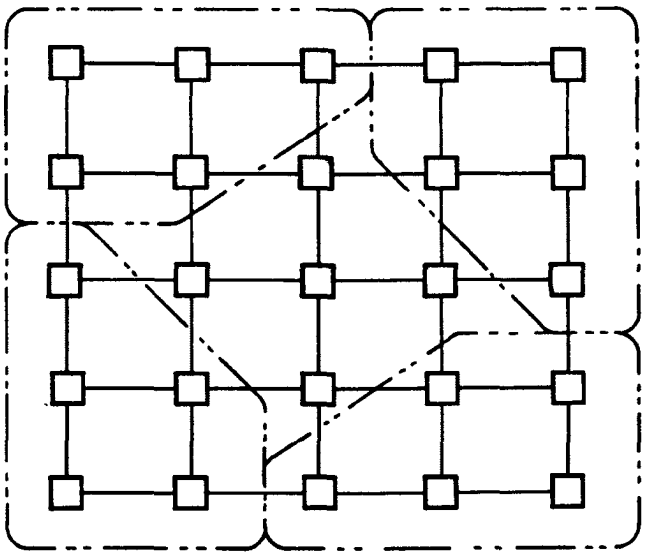

Fig. 9 -- A Non-Hierarchical Network
(See Fig. 7 for comparison)

Figure 9 represents a form of non-hierarchical network wherein each station may relay information from station to station. The geographic form of the interconnection is that of a matrix pattern. Physical link interconnection is shown in Fig. 10.

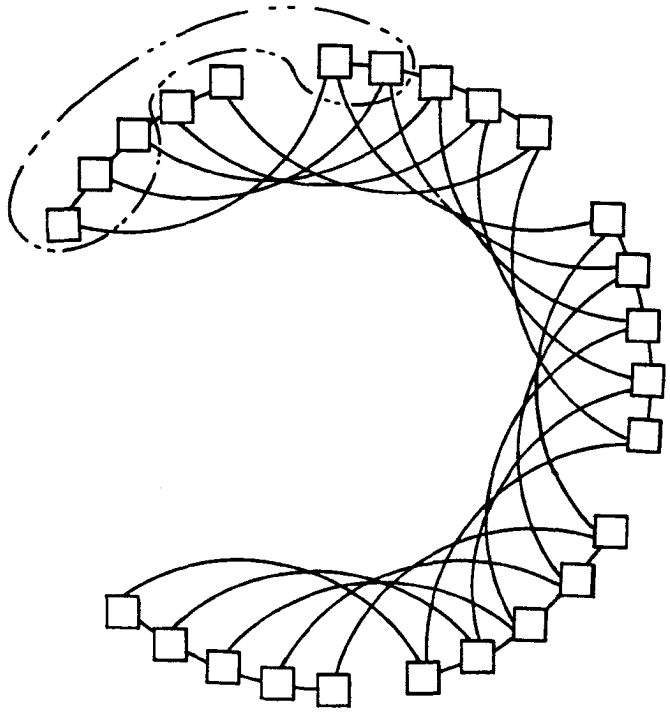

Fig. 10 -- Fig. 9 Redrawn
(See Fig. 8 for comparison)

Here the number of cuts required to isolate any section of the network is a constant and is equal to the number of channels per station. This may be compared to Fig. 8.

Early in this discussion it probably became apparent that we do not normally build pure non-redundant communications systems but rather incorporate some degree of redundancy. Very often there is a touch of non-hierarchical design included.

If the communications systems being examined form closed loops then there is a good probability that it will exhibit the desired property of withstanding slicing better than systems that do not exhibit this topography.

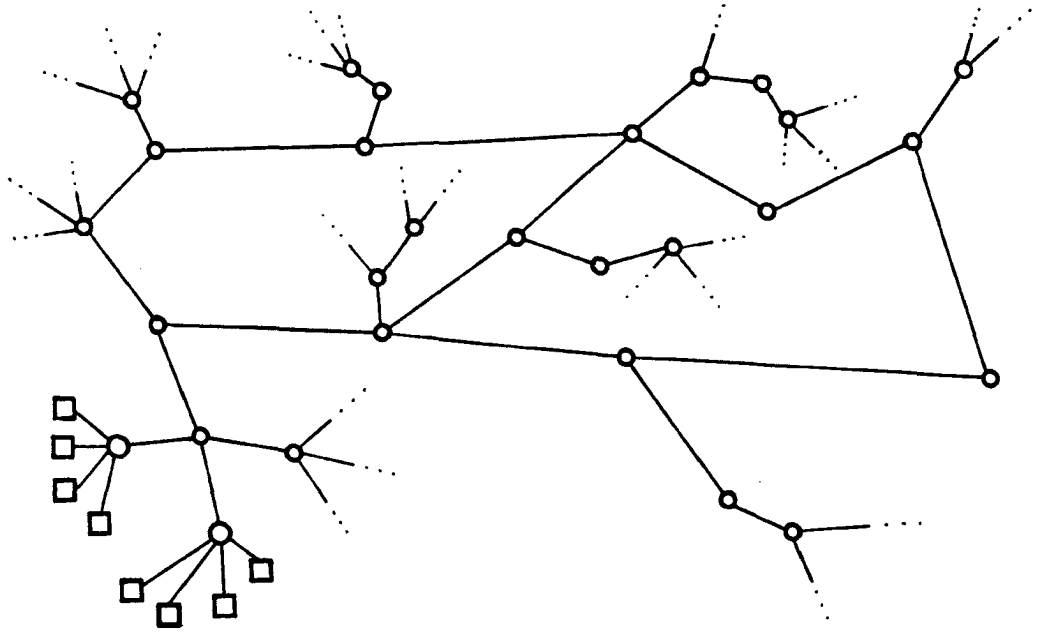

Fig. 11 -- A Mixed System

In summary of this section, there is a spectrum of possible communications systems from a pure hierarchical network to a perfectly distributed system or a non-hierarchical communications network. We are concerned primarily with degree; and it is felt the closer the full looping or non-hierarchical system goal is achieved, generally, the less vulnerable will the system be to a pinpointed destruction of critical key elements.

An Example of a Non-Hierarchical Communications Network

As large pure non-hierarchical communications networks are uncommon it might be well to consider the appearance of such networks and examine their suitability for passive defense.

The simplest conceivable pure non-hierarchical network is the "round robin" or a "bonded copper net," i.e., a non-inductive bilateral network, shown in Fig. 12.

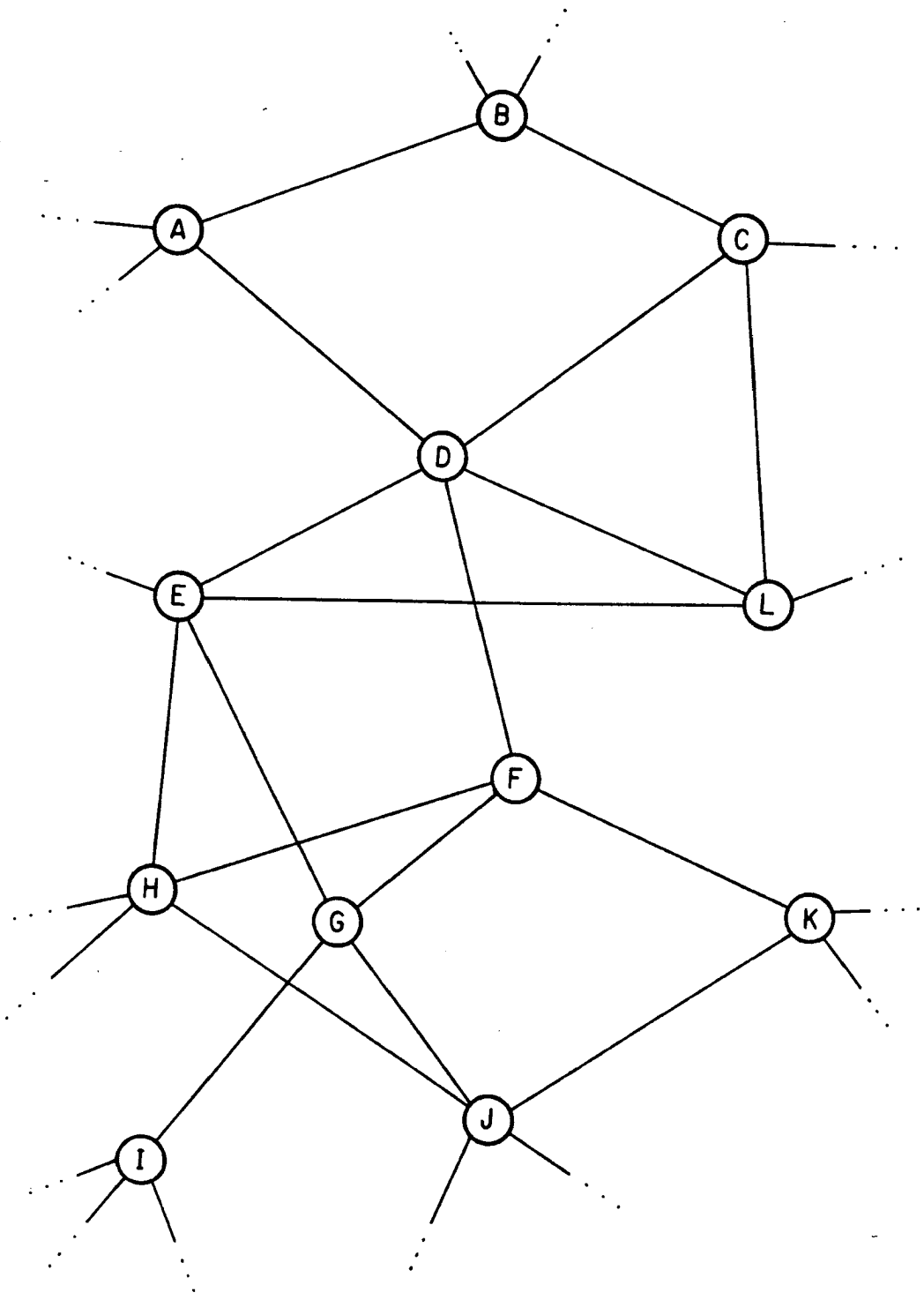

Fig. 12 -- A Non-Hierarchical Network

Imagine for the moment that each line is a telephone line tied together at each lettered station. Each station time-shares the common party telephone line. Jamming of this form of network is child's play. A single equipment malfunction can short out the entire network, and so it is obviously unsuited to the real world.

Let us apply a little intelligence at the nodes. Instead of bonding the lines in Fig. 12 as in the first case, let each node examine the messages and relay the messages in a desired manner, the desired manner being:

1) Messages are to travel over the shortest path regardless of the number of paths possible.

2) Messages are to be authenticated as to geographical point of origin (to confound would-be introducers of spurious signals; i.e., jammers) and defective signals rejected.

3) Location of system malfunction is to be pinpointed.

In Fig. 12, messages from D are transmitted to relay points A, C, E, L and F; messages from A are transmitted to B and D, etc.

There are several techniques possible that permit messages to travel effectively over the shortest path between two chosen stations. The method chosen for this illustration is the use of a "handover number." Along with each message is a "handover number" tag that indicates the number of times a message has traveled. Every time a message is re-relayed the message handover number is increased by one. This handover number provides sufficient information to permit a distributed digital

mechanization of an adaption of the Moore algorithm for finding the shortest path through a maze: where two messages of the same origin are received, only that message which bears the lowest handover number of the pair is propagated. See Fig. 13.

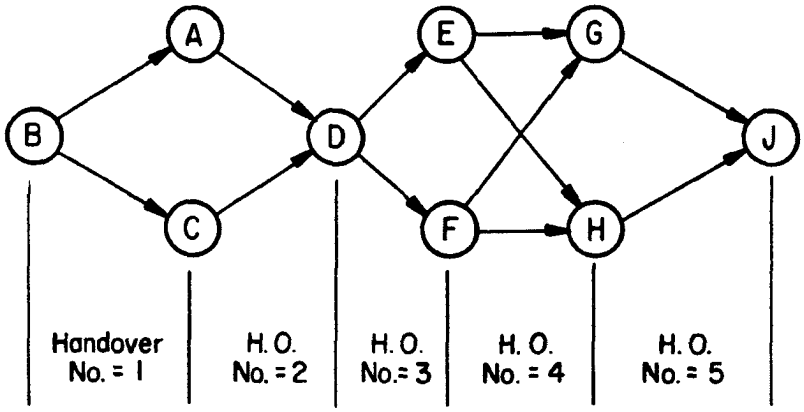

Fig. 13 -- Use of the Handover Number Showing "Shortest Path Messages Between B and J."

Here messages from B are transmitted to A and C, thence to D, thence to E and F, thence to G and H, thence to J. Thus, messages originated at each point are automatically routed over the shortest intact path to all other stations regardless of the number of broken links, provided there is at least one open link. There are, of course, many possible shortest paths or parts of paths between any two stations. Note that many relay points (for example, D, G, H and J) will each receive the same message from B over two separate input channels

simultaneously. Thus, any station incorrectly relaying a message either intentionally or otherwise will create the situation where two different messages will arrive bearing the same handover number. In such a case the proper response of the relay point is not to relay that particular message. The response of the network would be that of cutting out the suspicious links while the correct message travels by other more roundabout, but verified paths.

An amplification of the properties of these distributed networks may be found in available publications. [2, 3, 4]

The problem of seeking the shortest paths through a maze is not a new one. [5, 6] What is to be emphasized is that distributed computation, or totally independent apparatus at each node may be used to provide such routing without reliance upon a vulnerable central computer. Thus, it appears possible to build communications systems wherein no one point is a better target than any other point.

The effects of fraudulent intervenors on the network may be eliminated by utilizing the knowledge of path length. For example, the shortest path from D to B is never less than 2 links; from E or F to B is 3 links, etc. Thus, it is virtually impossible for a clever enemy to "seize" the network or issue fraudulent messages that will be propagated, [2, 3, 4]

REFERENCES

1. Kahn, H., Three Lectures on Thermonuclear War, The RAND Corporation, Paper P-1888, January 20, 1960.

2. Baran, P., Reliable Digital Communications Systems Using Unreliable Network Repeater Nodes, The RAND Corporation, Paper P-1995, May 27, 1960.

3. Baran, P. and R. Hammerly, A Verified Point of Origin Synchronous Digital Data Link Transmission System Using Randomly Surviving Relay Points, The RAND Corporation, Unnumbered Paper, May 20, 1960.

4. Baran, P. and F. Yates, A Non-Synchronous Digital Data Link Transmission System Using Randomly Surviving Relay Points, The RAND Corporation, Unnumbered Paper, May 25, 1960.

5. Dantzig, G. B., On the Shortest Route Through a Network, The RAND Corporation, Paper P-1345, April 29, 1959.

6. Kalaba, R. E., On Some Communications Network Problems, The RAND Corporation, Paper P-1325, April 22, 1958--revised June 3, 1959.