VULNERABILITY OF COMMUNICATION NETWORKS

by

H. Frank

Memorandum No. ERL-M203

15 March 1967

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

Manuscript submitted: 19 January 1967

The research reported herein was supported in part by the U. S. Army Research Office--Durham under Grant DA-ARO-D-31-124-G776 and in part by the Joint Services Electronics Program (U. S. Army, U. S. Navy and U. S. Air Force) under Grant AF-AFOSR-139-66.

ABSTRACT

A communication system can often be represented as an interconnection of stations and links. We assume that the system is subject to enemy attack aimed at isolating stations from each other. Vulnerability criteria are defined, and asymptotic and recurrence relations are given for computing the vulnerability of several classes of networks with a large number of stations. Based on the analysis procedures, optimum synthesis procedures are suggested. A simple model is analyzed first, then the results of the analysis are extended to include systems with repair, memory, and systems whose interconnections depend on distance. The main objective of this paper is to show under what conditions asymptotic expression for vulnerability can be derived.

I. INTRODUCTION

The problems of determining the "vulnerability" and of designing communication networks which are "invulnerable" to enemy attack is of paramount importance. As yet, there is no complete treatment of either the analysis or synthesis problems in the literature, although a number of partial results are available [1 - 6]. One difficulty which is immediately encountered in vulnerability studies is the lack of a completely suitable "vulnerability criterion." However, the major obstacle to the solution of the problem (given a vulnerability criterion) is the enormous difficulty of obtaining exact, analytic results which are useful for the analysis and synthesis of large systems. This difficulty stems from the fact that vulnerability (or reliability) problems are essentially combinatorial problems, and the number of combinations which must be considered for large systems is usually enormous.

A partial remedy to the above difficulty is to consider networks which possess a high degree of symmetry. Such networks are characterized by relatively few parameters, while the desirability of symmetry can be justified on heurestic grounds. For example, the distributed communication networks discussed by P. Baran [1] appear to possess many of the features of highly invulnerable nets. The results of Baran's Monte Carlo simulations support this conclusion. However,

it is difficult to make quantitative rather than qualitative judgements because detailed analytic information is not available.

In this paper, we develop asymptotic and recursion formulas for computing the vulnerability of classes of networks which are related (in spirit) to Baran's distributed nets. We assume that the nets are subject to an enemy attack aimed at isolating stations from each other. As a criterion of vulnerability (survivability) we use the average fraction of stations which both survive the attack and remain in an arbitrary connected group of stations which can communicate with each other. This criterion is nearly identical to Baran's, except that Baran elects to use the fraction of stations surviving the attack and in the largest connected group.

The scope of this paper is somewhat limited. At all times, we use the simplest formulation of a model that will illustrate a particular point. A number of our results can be generalized to reflect more realistic situations, but our main objective is to establish a methodology with which more complicated models can be attacked.

II. A SIMPLE MODEL

Let N be a communication network with $n+1$ stations v_0, $v_1, \ldots, v_n$ interconnected by a set of B-directed (one-way) links $b_1, b_2, \ldots, b_B$. The topological structure of N may be fixed, as in

the case of a microwave relay system, or it may be time varying, as in a satellite communication network. Alternatively, the structure may be either deterministic or random. In fact, the same network may simultaneously exhibit both qualities; for example, the builders of the network may know the exact location of every link and station while the enemy may know only the probable location of some of the links and stations. In addition, atmospheric conditions, reliability problems, and other random factors may contribute to a state of uncertainty about the exact structure of the net.

The design of an invulnerable net may not be the only consideration of the network designer. Other factors will influence his design; thus, a reasonable objective is to generate a class of networks, with the same survivability features, from which a specific network can be selected. We will consider a simple class of networks generated by a random process. This class is described by the assumptions

A I.a Each station has, on the average, d outwardly directed links. There are no self loops.

A I.b Each link incident at station v_i is also incident at station v_j with probability 1/n (for i, j = 0, . . ., n $i \neq j$).

A I.c All links and all stations are identical.

A specific realization generated by the above process may be highly unsymmetric. However, the essential feature of our model is that before we find a particular realization, all networks in the class

are, in a sense, "symmetric in probability." (The above model has been quite popular with mathematical biophysists, in their studies of random neural nets [7 - 12]. A partial summary of some pertinent results is given by Z. Prihar [2].) We use the following model for the enemy attack strategy:

A II.a Weapons are directed at random into a region of area A. The probability that any given weapon is directed at a region of area Δ ($\Delta \leq A$) is Δ/A.

A II.b The density of weapons is η weapons per unit area.

A II.c All weapons are identical.

In our model, a weapon could represent a ballistic missle. In this case, the assumption that the missiles fall uniformly at random into area A is not totally unrealistic. In other instances, the weapons could be bombs. In this case, the uniformity assumption A II.a is most appropriate when the enemy is uncertain about the physical location of the targets. Such a case could easily occur, for example, in jungle warfare, where the actual target may never be seen or precisely located.

We must specify the interaction between the weapons and the communication net. We will say that a link (or station) is hit if a weapon is directed to within a given distance of that link (or station). Naturally, this distance is a function of the power of the weapon and the structure of the link. The vulnerability of the net is defined by:

A III. a Stations are destroyed if they are hit by at least K_s weapons.

A III. b Links are destroyed if they are hit by at least K_ℓ weapons.

A III. c There is no repair; that is, once a station (or link) is hit, it remains hit.

Under Assumptions A I, A II, and A III, we shall compute the expected fraction of stations which can be reached from a station picked at random after the attack. Let this number be denoted by γ. We are mainly interested in communication nets with a large number of stations and links. Therefore suppose that n is large. Our assumptions describe a class of nets which behave in a uniform manner, both before and after attack. The probability of any station (or link) surviving the attack is identical to the probability of any other station (or link) surviving the attack. Since we have a large net, the Law of Large Numbers [13] is applicable. Consequently, the expected fraction of stations that are destroyed is approximately equal to the probability that any given station is destroyed, and the expected fraction of links that are destroyed is approximately equal to the probability that any given link is destroyed.

Let $f_k(\eta)$ and $g_k(\eta)$ denote the expected fraction of stations and links, respectively, which receive exactly k hits from a raid of density η. From the discussion in the last paragraph, the probability that any given station survives is (approximately)

$$\sum_{k=0}^{K_s-1} f_k(\eta) , \tag{1a}$$

and the probability that any given link survives is

$$\sum_{k=0}^{K_\ell-1} g_k(\eta) . \tag{1b}$$

Let t_1 (or p_1) be the probability that any given weapon hits a given station (or link). This probability depends on the area of vulnerability of the station (link). Given t_1 (or p_1), it is simple to compute $f_k(\eta)$ (or $g_k(\eta)$). In fact, it is well know ([13], p. 150) that $f_k(\eta)$ (or $g_k(\eta)$) is the probability that a Poisson variable with parameter t_1 (or p_1) has k successes. In other words,

$$f_k(\eta) = e^{t_1\eta} \frac{(t_1\eta)^k}{k!} , \tag{2a}$$

and

$$g_k(\eta) = e^{-p_1\eta} \frac{(p_1\eta)^k}{k!} . \tag{2b}$$

For completeness, we will derive Eq. (2a) using a method of A. Biermann [14, 15]. The key feature of Biermann's approach is that it is readily extended to more general situations that do not result in Poisson statistics.

On the average, $(n + 1)\, f_k(\eta)$ stations receive exactly k hits in an attack of density η. Suppose that η is changed by the infinitesmal amount $d\eta$. Then $f_k(\eta)$ will change by the amount $df_k(\eta)$. The change in $f_k(\eta)$ will be caused by stations which have already been hit k - 1 or k times being hit again. The higher order contributions, caused by $j \geq 2$ hits on targets with k - j prior hits can be ignored. Thus

$$df_k(\eta) = -t_1 d\eta\, f_k(\eta) + t_1 d\eta\, f_{k-1}(\eta) \qquad k = 0,\ 1,\ 2,\ \ldots \tag{3}$$

where $t_1 d\eta\, f_j(\eta)$ is the expected fraction of stations which experience j hits from the attack of density η and an additional hit caused by the increase of $d\eta$ $(j = k-1,\ k)$.

The system of equations (3) may be rewritten as

$$df_0(\eta)/d\eta + t_1 f_0(\eta) = 0 \tag{4a}$$

$$df_k(\eta)/d\eta + t_1 f_k(\eta) = t_1 f_{k-1}(\eta) \qquad k = 1,\ 2,\ \ldots \tag{4b}$$

Equation (4a) can be routinely solved to yield

$$f_0(\eta) = e^{-t_1 \eta}. \tag{5a}$$

All of the other $f_k(\eta)$ can be successively computed. For example,

$$df_1(\eta)/d\eta + t_1 f_1(\eta) = t_1 e^{-t_1 \eta}, \tag{6a}$$

or

$$f_1(\eta) = t_1 \eta \, e^{-t_1 \eta} . \tag{6b}$$

And, it is easily shown that, in general,

$$f_k(\eta) = e^{-t_1 \eta} \frac{(t_1 \eta)^k}{k!} . \tag{7}$$

This is the desired expression.

Returning to our network model, we suppose that links b_{i_1}, b_{i_2}, . . ., b_{i_d} are directed away from station v_i. After the attack, suppose that there are on the average, α intact links directed away from v_i, assuming that v_i survives (i = 0, 1, . . ., n). Networks generated by processes satisfying Assumption A I have been studied in detail by several authors [7 - 12]. Solomonoff and Rapaport [8] have obtained the following (approximate) transcendental equation for large n for γ in terms of α, the average number of outwardly directed links at any station:

$$\gamma = 1 - e^{-\alpha\gamma} . \tag{8}$$

Solomonoff [10] has shown that even for small values of n, this equation is extremely accurate. In their model, Solomonoff and Rapaport assumed that the number of stations are fixed and not subject to random disturbances. However, we can generalize their results to our model. In Appendix A we show that

$$\gamma = 1 - \exp\left\{-d\left[\sum_{k=0}^{K_s-1} f_k(\eta)\right]\left[\sum_{k=0}^{K_\ell-1} g_k(\eta)\right]\gamma\right\}. \qquad (9)$$

The proof given in Appendix A is a simple extension of a proof given by Rapaport [9]. We include it for completeness and because both the methods and results of the proof are of interest.

Equation (9) is an equation for the average fraction of surviving stations[1] which can be reached from a station picked at random after the attack in terms of the network parameters. Consider the following network design problem. The designer of the communication system would like to build a system with a guaranteed "invulnerability." That is, he would like to find d, K_s, and K_ℓ the "redundancy levels" of the network so that, on the average at least $\gamma_0 \times 100\%$ of the stations in the net can communicate after the attack. Here we assume that γ_0 is a prescribed number given to the designer. It should be clear that if the designer picks d, K_s, and K_ℓ large enough, he can always guarantee that $\gamma \geq \gamma_0$. Therefore, he would like to find a set of values which guarantee γ_0 with minimum cost.

Suppose that the cost of the system can be expressed as some function of d, K_S, and K_ℓ, say $H(d, K_s, K_\ell)$. The design problem can be stated as:

Find d, K_s, and K_ℓ such that

$$H(d,\ K_s,\ K_\ell) \text{ is minimized} \tag{10a}$$

and

$$\gamma \geq \gamma_0 \,. \tag{10b}$$

Let us solve for the exponent in the right-hand side of Eq. (9). Thus,

$$d\left[\sum_{k=0}^{K_s-1} f_k(\eta)\right]\left[\sum_{k=0}^{K_\ell-1} g_k(\eta)\right] = \frac{-\ln(1-\gamma)}{\gamma} \,. \tag{11}$$

and the constraint given in Eq. (10b) is equivalent to

$$d\left[\sum_{k=0}^{K_s-1} f_k(\eta)\right]\left[\sum_{k=0}^{K_\ell-1} g_k(\eta)\right] \geq \frac{-\ln(1-\gamma_0)}{\gamma_0} \triangleq \Gamma_0 \,, \tag{12}$$

where Γ_0 is a known number.

The design problem is now: Find d, K_s, and K_ℓ such that

$$H(d,\ K_s,\ K_\ell) \text{ is minimized}$$

and

$$d\left[\sum_{k=0}^{K_s-1} \frac{(t_1\eta)^k}{k!}\right]\left[\sum_{k=0}^{K_\ell-1} \frac{(p_1\eta)^k}{k!}\right] \geq e^{(t_1+p_1)\eta}\,\Gamma_0 \,. \tag{13}$$

The solution of the above design problem is routine. K_s and K_ℓ must be integers, and usually will be bounded by some number, beyond which it is not feasible to increase the survivability of a link or station. If

we substitute all pairs of feasible K_s and K_ℓ into Eq. (13) and then solve for the minimum d necessary to satisfy Eq. (13), we can evaluate the cost function at each K_s, K_ℓ and d which satisfy Eq. (13). Any K_s, K_ℓ, and d where the cost function is minimum is an optimal solution.[2]

EXAMPLE 1. We want to design a communication net with a large number of stations so that 90% of the surviving stations can be reached from a station picked at random after the attack. Suppose that the links are invulnerable but the stations are not. Let the cost function $H(d, K_s) = d(\exp 3K_s/2)$. Suppose that the enemy has been successful in locating targets to within one square mile, but that each actual target has a zone of vulnerability of only 0.05 square miles. Then, for a weapon directed at random into the target area, the probability of a hit is $t_1 = 0.05$. Let the density of weapons be 100 weapons per square mile. From Eq. (9),

$$\gamma = 1 - \exp\left\{-d\left[\sum_{k=0}^{K_s-1} \frac{5^k}{k!}\right] e^{-5}\right\}\gamma , \tag{14}$$

and from Eq. (13), our constraint is

$$d\left[\sum_{k=0}^{K_s-1} \frac{5^k}{k!}\right] \geq \frac{e^5(-\ln 0.1)}{0.9} = 374.8 . \tag{15}$$

Suppose that it is not feasible to reinforce stations beyond a survivability of four hits. Then, from Eq. (15), if $K_s = 1$, $d = 375$; if $K_s = 2$, $d = 63$; if $K_s = 3$, $d = 23$; if $K_s = 4$, $d = 10$; and if $K_s = 5$, $d = 5$. Evaluating the cost function, we obtain $H(375, 1) = 1680$; $H(63, 2) = 1265$; $H(23, 3) = 2070$; $H(10, 4) = 4030$, and $H(5, 5) = 5400$. The minimum of the cost function occurs at $d = 63$ and $K_s = 2$. Thus, each station of the system should have 63 outwardly directed links.

III. SOME GENERALIZATIONS OF THE NETWORK MODEL AND THE VULNERABILITY CRITERION

The class of networks generated by the process defined by Assumption A I can be modified in several ways. First we note that a network satisfying A I.b may have parallel links. At station v_i we choose d out of n stations and connect these stations to v_i. Our process of selecting stations is equivalent to sampling a population of n points <u>with replacement</u>; consequently, the same point may be selected more than once. A more reasonable method of selection is to establish links sequentially. The first station is selected equiprobably out of the n possible stations; the second station is selected equiprobably out of the n - 1 remaining stations; · · · ; the <u>d</u>th station is selected equiprobably out of the n - d remaining stations (none of which has been already selected). This scheme is equivalent to sampling a population of n points <u>without replacement</u>; we will call this scheme A I.b'. In other words, A I.b' is:

A I.b' The outward links at station v_i are determined by sampling stations $v_0, v_i, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n$ without replacement a total of d times ($i = 0, 1, 2, \ldots, n$).

We can include Assumption A I.b' in our model without difficulty. In fact, it is easily shown that Eq. (9) is still valid. This result is not surprising since it is well known that for large populations sampling with and without replacement are virtually identical ([13], p. 57). However, even though both schemes of sampling are equivalent for infinite populations, we can expect some difference for finite populations. Equation (9) is actually a lower bound for γ. This is because a network with no parallel links in the same direction has higher probability of being connected. We can investigate the difference between sampling with and without replacement for finite populations, and simultaneously introduce another important factor into our study of vulnerability.

Most networks have some processing time associated with the links and/or stations. This processing time may be the time necessary to transmit information through the link or the time needed at a station to decode, recode, and retransmit the information. In any event, it is usually desirable to limit the total time a message remains in the network. In many cases, limiting the total time is equivalent to limiting the path lengths of the network routes. Thus, instead of asking for the surviving fraction of stations that can be reached from a given point,

it is reasonable to ask for the fraction of stations that can be reached from a given point by a path of no more than r(r is determined by cost factors, etc.) links.

Let us choose a station at random. Since we have a large number of stations, the probability that any other station is connected to this station is approximately equal to the average fraction of stations that can be reached from this point. Similarly, the probability that any other station is connected to this station by a path of no more than r links is approximately equal to the average fraction of stations at distance r or less from this station. We can compute this number by a recursion formula. For the system satisfying Assumption AI, Rapaport [9] has shown (also see Prihar [1], p. 391) that[3]

$$q(r) = (1 - 1/n)^{X_1}, \tag{16}$$

where

$$X_1 = \alpha(n+1)\,[1 - q(r-1)] \prod_{i=0}^{r-2} q(i)$$

and where q(i) is the probability that a station is more than i links removed and q(0) = n/n+1. If we include Assumptions AII and AIII, we must then solve the recursion formula

$$q(r) = (1 - 1/n)^{X_2}, \tag{17}$$

where

$$X_2 = d\left\{\left[\sum_{k=0}^{K_s-1} f_k(\eta)\right]\left[\sum_{k=0}^{K_\ell-1} g_k(\eta)\right][1 - q(r-1)] \prod_{i=0}^{r-2} q(i)\right\}(n+1)$$

and where

$$q(0) = 1 - \frac{1}{n+1} \sum_{k=0}^{K_s - 1} f_k(\eta) .$$

Then, if p(i) is the probability that a given station is no more than i links removed, $p(i) = 1 - q(i)$ and, in particular, $p(r) = 1 - q(r)$.

A similar recursion formula can be derived for the scheme of sampling withour replacement. (The derivation involves a straight forward modification of the techniques given in Appendix A.) Let $q^*(r)$ represent the probability that any station is more than r links removed from a point chosen at random, when the system is described by Assumptions A I.a, A I.b', A I. c, AII, and AIII. Then, it can be shown that $q^*(r)$ satisfies the recursion formula

$$q^*(r) = \left[1 - \frac{d \sum_{k=0}^{K_s - 1} f_k(\eta) \sum_{k=0}^{K_\ell - 1} g_k(\eta)}{n} \right]^{X_3} \tag{18}$$

where

$$X_3 = (n+1)\left[1 - q(r-1)\right] \prod_{i=0}^{r-2} q(i)$$

and where

$$q^*(0) = 1 - \frac{1}{n+1} \sum_{k=0}^{K_s - 1} f_k(\eta) .$$

Let $P(r)$ and $P^*(r)$ represent the probabilities that any stations is exactly r links removed from a station chosen at random, under Assumptions A I.b and A I.b', respectively. $P(r)$ and $P^*(r)$ are approximately equal to the expected fraction of stations that are connected by at least one path of r links and no path with fewer than r links to the station picked at random. It can be shown that $P(r)$ and $P^*(r)$ are given by

$$P(r) = \left[1 - \sum_{j=0}^{r-1} P(j)\right] [1 - q(r)] \tag{19}$$

and

$$P^*(r) = \left[1 - \sum_{j=0}^{r-1} P^*(j)\right] [1 - q^*(r)] . \tag{20}$$

Furthermore, for large n, both $P(r)$ and $P^*(r)$ satisfy the recurrence relation

$$P(r) = \left[1 - \sum_{j=0}^{r-1} P(j)\right] \left[1 - \exp\left\{- P(r-1)\, d \sum_{k=0}^{K_s-1} f_k(\eta) \sum_{k=0}^{K_\ell-1} g_k(\eta)\right\}\right] \tag{21}$$

with

$$P(0) = P^*(0) = \frac{1}{n+1} \sum_{k=0}^{K_s-1} f_k(\eta) .$$

Now that we know how to compute $P(r)$ (or $P^*(r)$) we can define an alternative vulnerability criterion. As we have already indicated, if the only available path between a pair of stations is "too long," we may consider that the enemy has effectively separated the two stations. The vulnerability index γ does not take this factor into account. In fact,

$$\gamma = \sum_{r=0}^{\infty} P(r) . \tag{22}$$

Define $\gamma(r)$ by

$$\gamma(r) = \sum_{k=0}^{r} P(k) . \tag{23}$$

It should be clear that $\gamma(r)$ is approximately equal to the average number of stations connected by a path of length r or less to a point chosen at random. A reasonable vulnerability constraint is now $\gamma(r) \geq \gamma_0$ and again, the objective is to find d, K_s, and K_ℓ so that the constraint is satisfied with minimum cost. Naturally, the synthesis problem is complicated. In general, the recurrence relations given in Eqs. (19), (20) or (21) must be solved for $P(r)$ (or $P^*(r)$) given values of d, K_s, and K_ℓ. The least cost set is then selected from those which satisfy the constraint.

EXAMPLE 2. A communication network has 100 stations and an average of 20 links per station. Suppose that the links are invulnerable but the stations are not. If the enemy has been successful at locating targets to within one square mile and each target has an area of vulnerability of 0.05 square mile, the probability of a hit is 0.05. Let the density of weapons be 50 weapons per square mile. We would like to find the average surviving fraction of stations that can be reached after the attack from a station chosen at random and the average fraction of stations that can be reached by a path of no more than three links. Assume that K_s is either 2 or 3.

If $K_s = 2$, the probability that a given station survives is

$$\sum_{k=0}^{1} f_k(\eta) = (1 + 2.5)\, e^{-2.5} = 0.287 . \tag{24}$$

Then, from Eq. (9),

$$\gamma = 1 - e^{-20(0.287)\gamma} = 1 - e^{-5.74\gamma} . \tag{25}$$

and γ is very close to unity. However, when we calculate P(r) from Eq. (21), we have that

$$P(0) = (1/100)(0.287) = 0.00284$$

$$P(1) = (1 - 0.00284)(1 - \exp -20(0.284)(0.00284)) = 0.016$$

$$P(2) = (1 - 0.0189)(1 - \exp -20(0.284)(0.016)) = 0.087 \quad (20)$$

$$P(3) = (1 - 0.106)(1 - \exp -20(0.284)(0.086)) = 0.352.$$

Therefore although nearly 100% of the stations can be reached from a station chosen at random,

$$\gamma(3) = 0.00284 + 0.016 + 0.087 + 0.352 = 0.456 \quad (27)$$

and only about 45% of the stations can be reached with paths of three or less links.

If $K_s = 3$, the picture changes drastically. Again it is easy to see that γ is very close to unity. The probability that a station survives is now $(1 + 2.5 + 6.25/2)\, e^{-2.5} = 0.544$. From Eq. (21)

$$P(0) = 0.00533$$

$$P(1) = 0.0555$$

$$P(2) = 0.431 \quad (28)$$

$$P(3) = 0.502 .$$

Therefore, $\gamma(3) = 0.9933$ or in other words, about 99% of the stations in the net can be reached from a station picked at random by a path of no more than three links. Note that in terms of $\gamma \sum_{k=0}^{K_s-1} f_k(\eta)$, if $K_s = 2$, only about 12% of the total stations both survive the attack and

can be contacted from a station chosen at random. However, if $K_s = 3$, over 50% of the total stations will be reachable from the station chosen at random.

In our previous models we considered only systems with directed (one-way) links. Suppose we are given a communication system with undirected (two-way) links, whose stations each have an average of d receives and transmitters. Once a link is established between a pair of stations, both stations can converse with each other. Our vulnerability criterion then becomes the average fraction of stations in an arbitrary connected component[4] of the system. The random process that we will use to generate our system is described by the assumption:

A I.a' Each station has on the average d undirected links. The total number of links is $B = B(n) \sim dn$. There are no self loops or parallel links.

A I.b' There are a total of $\binom{n+1}{2}$ possible links.[5] The network is constructed by choosing the first link equiprobably from among the $\binom{n+1}{2}$ possible links, the second link from among the $\binom{n+1}{2}$ -1 remaining links, . . . the kth link is chosen equiprobably from among the $\binom{n+1}{2}$ -k remaining links.

We would again like to find the average surviving fraction of stations that can be reached from a station chosen at random. We can

attack this analysis problem with the techniques already discussed.[6] However, we will use an alternative approach based upon the work of P. Erdos and A. Renyi [17]. In [17], (an exceptional paper of great scope) Erdos and Renyi establish many asymptotic results describing the growth or "evolution" of graphs generated by random processes specified by A I.a' and A I.b' and other related schemes.

As a first step, we state a theorem (Theorem 6) from [17], which gives the expected number of components of the system.

THEOREM. Let $B(n) \sim dn$ with $d > 1/2$. For large n, the expected number of components is approximately

$$\frac{n}{2d}\left[x(d) - \frac{x^2(d)}{2}\right] \tag{29}$$

where $x(d)$ is the only solution satisfying $0 < x(d) < 1$ of the equation

$$x(d)\, e^{-x(d)} = 2d\, e^{-2d} \cdot \left(x(d) = \sum_{k=1}^{\infty} \frac{k^{k-1}}{k!} \left(2d\, e^{-2d}\right)^k \right).$$

We can use Eq. (29) to find the average fraction of stations in an arbitrary component. (Here, we assume invulnerable stations and vulnerable links, but this limitation is easily removed.) The size of a component is a random variable which is identically and independently distributed for each component. Therefore the average number of

stations in a component is simply n divided by the average number of components. Hence the average fraction of stations in a component is (for large n)

$$\gamma = \frac{2d/n}{x(d) - x^2(d)/2} \tag{30}$$

which for large d ($d \geq 4$) is closely approximated by

$$\gamma = \frac{1}{n} e^{2d} . \tag{31}$$

Equation (31) follows from the fact that for $d \geq 4$, $2de^{-2d}$ is an extremely small number and thus the solution of $x(d) = e^{x(d)}\, 2d\, e^{-2d}$ is nearly exactly $x(d) = 2d\, e^{-2d}$.

Equation (31) shows that for very large n, γ becomes small. This is because almost all stations of the network belong to some small component or to a "giant" component. The size of the "giant" component can be found. According to a theorem of Erdos and Renyi (Theorem 9b):

THEOREM. If $\overline{\gamma}$ denotes the fraction of stations in the largest component of the network and if $B(n) \sim d\, n$ with $d > 1/2$, then for any $\epsilon > 0$,

$$\lim_{n\to\infty} \text{Prob}\left\{\left|\overline{\gamma} - 1 + \frac{x(d)}{d}\right| < \epsilon\right\} = 1 , \tag{32}$$

where

$$x(d) = \sum_{k=1}^{\infty} \frac{k^{k-1}}{k!} \left(2de^{-2d}\right)^k$$

is the solution satisfying $0 < x(d) < 1$ of the equation $x(d)\, e^{-x}(d) = 2de^{-2d}$, and where Prob{ } denotes the probability of the event represented within the brackets.

The last theorem states that the expected fraction of stations in the largest component is $1 - (x(d)/d)$. For large d ($d \geq 4$) this number is closely approximated by

$$\overline{\gamma} \doteq 1 - 2e^{-2d} . \tag{33}$$

Note that these results are related to results obtained by Gilbert [5], Jacobs [4], Austin, Fagan, Penny and Riordan [18], and others [19 - 23]. The effect of an enemy attack can now be included in the model. An attack such as the one described by Assumption A II has the effect of reducing the average degree of each station uniformly. Consequently, d becomes a random variable. Since all the necessary probabilities are known, we could easily find the expected value of $\overline{\gamma}(d)$.

IV. SYSTEMS WITH NONUNIFORM LINKS AND DISTANCE BIAS

In the preceding sections, we assumed that all links were identical (Assumption A I.c). Consequently, each link has the same probability of

surviving a uniform attack. This assumption is essential to our development. On the other hand, most communication nets contain links of different length. Clearly, if two links are of the same construction but of different length, the longer one has a greater probability of being destroyed. This follows from the fact that the longer links are more likely to receive a given number of hits. Our previous development does not include this more general situation. However, there is an alternative approach that we can apply.

Before we discuss this alternative approach, we will introduce an additional generality. We no longer assume that the probability of establishing links from a given station to the other stations of the net is the same for all stations. Instead, we assume that the probability of establishing a link between a pair of stations is dependent on the length of that link. In other words, the probability of establishing a link between a given pair of stations v_i and v_j is a function of the distance between v_i and v_j. Let $\rho(v_i, v_j)$ represent the distance between v_i and v_j. We will assume that our system satisfies

AIV.a The network may be partitioned into regions $Y_1, Y_2, \ldots, Y_m$ of area $\Delta_1, \Delta_2, \ldots, \Delta_m$, respectively, such that within each region there are a large number of stations. The density of the stations is ν stations per unit area and is constant for each region.

AIV.b Within each region Y_i, interconnections are identical, equiprobable, and satisfy Assumption AI.

AIV.c If station v_{s_i} is any sation in region Y_i, for any link b_{s_j} directed away from v_{s_i}, the probability that b_{s_j} contacts a station in the neighborhood of station v_k is $\tau(\rho(v_{s_i}, v_k))\,\Delta$, where $\tau(\cdot)$ is a probability density and Δ is the area of the neighborhood.

AIV.d Every station in Y_i has on the average d outwardly directed links, for $i = 1, 2, \ldots, m$.

For the system described by Assumption AIV, we will derive a recurrence relation for the probability that a station at distance x from a station selected at random is connected to this station by a path consisting of exactly r links and no path consisting of fewer links. As before, this probability is approximately equal to the expected fraction of stations with this property. Our derivation is similar to one given by Rapaport [9].

Let the point selected at random be denoted by v_0. We will use the following procedure to count the number of stations which can be reached from v_0: First we trace all links emanating from v_0. Let the set of stations thus contacted be denoted by S_1. Second, we trace links emanating from stations in S_1 and denote the set of stations thus

contacted by $S_2 \ldots$. At the ith step we trace links emanating from stations in S_{i-1} and denote the new set of stations contacted by S_i. Note that we may contact a given station several times by this procedure. Therefore, at each stage we only trace those links emanating from stations which have never been contacted before (i.e., stations in $S_j - \bigcup_{i=1}^{j-1} S_i$).

Let $P(x, r)$ be the probability that a station at distance x from some other station is exactly r links removed from this station. The expected number of stations to be contacted for the first time from v_0 at the rth stage in the neighborhood of a station at distance x from v_0 is $\nu \Delta P(x, r)$. Therefore, if d links emanate from each station, there is an average of $d\nu \Delta P(x, r)$ new links to be traced from this neighborhood at the $r+1$-st stage of the counting process. Suppose that the expected number of stations are actually contacted at each stage.[7] Then the probability that a station v_j at distance y from v_0 is not contacted by a link emanating from any region is

$$\prod_{i=1}^{m} \left\{ 1 - \tau\left(\rho\left(v_{s_i}, v_j\right)\right)\Delta_i \right\}^{d\nu \Delta_i P(x_i, r)} \tag{34}$$

where v_{s_i} is a station in Y_i and is at distance x_i from v_0. Therefore, the probability that v_j is contacted by at least one link at the $r+1$-st is

$$1 - \prod_{i=1}^{m} \left\{ 1 - \left(\rho(v_i, v_j) \right) \Delta_i \right\}^{d\nu \Delta_i P(x_i, r)} \tag{35}$$

The probability that v_j is not contacted in the first r stages is

$$1 - \sum_{j=0}^{r} P(y, j) ,$$

and hence the probability of <u>first</u> contacting v_j at the r+1-st stage is

$$P(y, r+1) = \left[1 - \sum_{j=0}^{r} P(y, j) \right] \left[1 - \prod_{i=1}^{m} \left\{ 1 - \left(\rho(v_i, v_j) \right) \Delta_i \right\}^{d\nu \Delta_i P(x_i, r)} \right] . \tag{36}$$

Equation (36) is the desired recurrence reaction for P(x, r). Now we can introduce the effect of the enemy attack (Assumptions AII and AIII). The expected number of links emanating from station v_i is d^* given by

$$d^* \triangleq d \left[\sum_{k=0}^{K_s - 1} f_k(\eta) \right] \left[\sum_{j} \left\{ \left[\left(\rho(v_i, v_j) \right) \right] \left[\frac{\rho(v_i, v_j)}{K} \right] \left[\sum_{k=0}^{K_\ell - 1} g_k(\eta) \right] \right\} \right] \tag{37}$$

where $\rho(v_i, v_j)/K$ represents the number of unit areas of vulnerability of the link between v_i and v_j and $g_k(\eta)$ is now the probability that a unit area of link suffers exactly k hits. The probability that there is a link between v_i and v_j is

$$\frac{\rho(v_i, v_j)}{K}\,\tau\Big(\rho(v_i, v_j)\Big)\sum_{k=0}^{K_s-1} f_k(\eta)\sum_{k=0}^{K_\ell-1} g_k(\eta)\ . \tag{38}$$

Therefore, our recurrence relation can be written by substituting Eq. (37) and d^* given by Eq. (37) for d, into Eq. (36). The initial condition for the recursion formula is easily seen to be

$$P(y, 1) = \frac{y}{K}\,p(y)\sum_{k=0}^{K_s-1} f_k(\eta)\sum_{k=0}^{K_\ell-1} g_k(\eta)\ . \tag{39}$$

The recurrence relation given in the last paragraph is, to say the least, complicated; however, this is to be expected since the problem that we posed is inherently difficult. Numerical solutions of the recurrence relation for various values of d, K_s and K_ℓ can be obtained under certain conditions. To solve Eq. (36) we must repeatedly evaluate its right hand side. The computational problem is not difficult. However, at each stage, we must compute a product over all regions $Y_1, Y_2, \ldots, Y_m$. Since there are m regions, the number of computations required to find P(y, r) will be on the order of m^r. Therefore, if r is large, Eq. (36) will no longer be an effective means of finding P(y, r). On the other hand, if both the number of regions and the size of r are limited to numbers which allow the recursive solution of Eq. (36) to be practical, we can use this equation to evaluate the vulnerability criterion $\gamma(r)$ defined by Eq. (23). Here, we must generalize this criterion to include the distance factor. Thus, define $\gamma(y, r)$ as

$$\gamma(y,\ r) = \sum_{j=0}^{r} P(y,\ r)\ .$$

Clearly, $\gamma(y, r)$ represents the average number of stations at distance y from v_0 which are connected to v_0 by at least one path of r links and no paths with fewer links. This vulnerability criterion and the recurrence relation could form the basis of minimum cost design procedures similar to those already discussed.

V. SYSTEMS WITH REPAIR, MEMORY, AND OTHER GENERALIZATIONS

Assumption AIII describes the interaction between the enemy attack and the communication system. A valid objection to this hypothesis is the limitation imposed on the builders of the network by not allowing repair (A III.c). Even though a communication net is severely damaged, it may not remain that way for very long. The situation is familiar where within a few hours after an attack, a network is again operating at nearly full capacity. Repair of the network may thus result in a second, third, or even continuous sequence of attacks on the system.

We noted that the proof of Eq. (2) could be extended to more general situations that did not involve Poisson's statistics. We now replace Assumption A III. c by

A III.c' If a station has experienced k hits at time t, the probability that it will be completely repaired in the time interval $[t, t+dt]$ is $r_k^s(t)\,dt$, for $k = 1, 2, \ldots$ If a link has experienced k hits at time t, the probability that it will be completely repaired in the time interval $[t, t+dt]$ is $r_k^\ell(t)\,dt$, for $k = 1, 2, \ldots$

The functions $r_k^s(t)$ and $r_k^\ell(t)$ are known as repair rate functions [24]. Several possible choices could be $r_k(t) = r_k$ for all t (exponential repair), or $r_k(t) = \lambda_k \beta_k t^{\beta_k - 1}$ $\lambda_k, \beta_k > 0$, $t \geq 0$. (r(t) is the repair rate function for a Weibull distribution.)

For a system with repair, Eq. (2) is no longer applicable. However, consider the system of differential equations which now describes the number of hits per station:

$$\begin{aligned} df_0(\eta, t) &= -t_1\, d\eta\, f_0(\eta, t) + \sum_k r_k^s(t)\, dt\, f_k(\eta, t), \\ df_k(\eta, t) &= -t_1\, d\eta\, f_k(\eta, t) + t_1\, d\eta\, f_{k-1}(\eta, t) - r_k^s(t)\, dt\, f_k(\eta, t) \\ & \qquad k = 1, 2, \ldots \end{aligned} \tag{40}$$

Let $\eta = \eta(t)$. Then this system of equations is equivalent to

$$\frac{df_0(t)}{dt} = -t_1 \frac{d\eta(t)}{dt} f_0(t) + \sum_k r_k^s(t) f_k(t) ,$$

$$\frac{df_k(t)}{dt} = -t_1 \frac{d\eta(t)}{dt} f_k(t) + t_1 \frac{d\eta(t)}{dt} f_{k-1}(t) - r_k^s(t) f_k(t) \qquad k = 1, 2, \ldots \tag{41}$$

This system of differential equations can be placed into the normal form [25] where we see that we have a the time-varying system:

$$\begin{bmatrix} f_0'(t) \\ f_1'(t) \\ f_2'(t) \\ \cdot \\ \cdot \end{bmatrix} = \begin{bmatrix} -t_1\eta'(t) & r_1^s(t) & r_2^s(t) & r_3^s(t) & \cdot & \cdot \\ t_1\eta'(t) & -r_1^s(t) - t_1\eta'(t) & 0 & 0 & \cdot & \cdot \\ 0 & t_1\eta'(t) & -r_2^s(t) - t_1\eta'(t) & 0 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \begin{bmatrix} f_0(t) \\ f_1(t) \\ f_2(t) \\ \cdot \\ \cdot \end{bmatrix} \tag{42}$$

where ' indicates differentiation with respect to t and the initial conditions are $f_0(0^-) = 1$, $f_1(0^-) = f_2(0^-) = \cdots = 0$.

The solution of these equations is of course a numerical problem. The degree of difficulty depends on the functions $\eta'(t)$ and $r_k^s(t)$. However, in some cases, the computational problems are relatively routine. One such case occurs when each $r_k^s(t) = r_k$ (exponential repair) and $\eta(t)$ is a periodic function. Another routine case occurs if $\eta'(t)$ can be approximated by a piecewise constant function, and each $r_k^s(t) = r_k$.

Let us consider one special case. Suppose each station cannot survive any direct hits but we will repair a station which has suffered exactly one hit in time dt with probability $r_1 dt$. The probability that a station is operating is thus $f_0(t)$ and the differential equation which describes the behavior of $f_0(t)$ is

$$\begin{bmatrix} f_0'(t) \\ f_1'(t) \end{bmatrix} = \begin{bmatrix} -t_1\eta' & r_1 \\ t_1\eta' & -r_1 - t_1\eta' \end{bmatrix} \begin{bmatrix} f_0(t) \\ f_1(t) \end{bmatrix} \tag{43}$$

with initial conditions $f_0(0^-) = 1$ and $f_1(0^-) = 0$. This equation is easily solved numerically or simulated on an analog computer (if $\eta'(t)$ is a "reasonable" function).

EXAMPLE 3. A communication network with a large number of stations is attacked periodically with a density of weapons shown in Fig. 1. Suppose stations cannot survive direct hits but that stations which have experienced exactly one direct hit will be repaired with probability .1dt in time dt (i.e., $r_1 = 0.1$). Assume that links are invulnerability and as before suppose that the probability that a given weapon hits a given station is $t_1 = 0.05$. We would like to compute γ as a function of time.

From Eq. (43), the equation describing the system is:

$$\begin{bmatrix} f_0'(t) \\ f_1'(t) \end{bmatrix} = \begin{bmatrix} -0.05\eta' & 0.1 \\ 0.05\eta' & -0.1 - 0.05\eta' \end{bmatrix} \begin{bmatrix} f_0(t) \\ f_1(t) \end{bmatrix} \tag{44}$$

with $f_0(0^-) = 1$ and $f_1(0^-) = 0$. From time $t = 0$ to time $t = 1$, $\eta'(t) = 20$ and Eq. (44) is

$$\begin{bmatrix} f_0'(t) \\ f_1'(t) \end{bmatrix} = \begin{bmatrix} -1 & 0.1 \\ 1 & -1.1 \end{bmatrix} \begin{bmatrix} f_0 \\ f_1 \end{bmatrix} \tag{45}$$

and on this time interval, the solution is

$$\begin{bmatrix} f_0(t) \\ f_1(t) \end{bmatrix} = \begin{bmatrix} \alpha_0(t) - \alpha_1(t) & 0.1\,\alpha_1(t) \\ \alpha_1(t) & -1.1\,\alpha_1(t) \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{46}$$

where $\alpha_0(t) = 2.14\, e^{-0.73t}$ and $\alpha_1(t) = 1.14\, e^{-1.37t}$. Therefore, at $t = 1$,

$$\begin{bmatrix} f_0(1) \\ f_1(1) \end{bmatrix} = \begin{bmatrix} 0.383 \\ 0.30 \end{bmatrix}. \tag{47}$$

In the time interval $1 \le t \le 4$, the system is described by the equation

$$\begin{bmatrix} f_0'(t) \\ f_1'(t) \end{bmatrix} = \begin{bmatrix} 0 & 0.1 \\ 0 & -0.1 \end{bmatrix} \begin{bmatrix} f_0(t) \\ f_1(t) \end{bmatrix}, \tag{48}$$

and the solution is

$$\begin{bmatrix} f_0(t) \\ f_1(t) \end{bmatrix} = \begin{bmatrix} 1 & 1 - e^{-0.1(t-1)} \\ 0 & e^{-0.1(t-1)} \end{bmatrix} \begin{bmatrix} f_0(1) \\ f_1(1) \end{bmatrix}. \tag{49}$$

Therefore,

$$\begin{bmatrix} f_0(4) \\ f_1(4) \end{bmatrix} = \begin{bmatrix} 0.47 \\ 0.266 \end{bmatrix} . \tag{50}$$

During the time interval $4 \le t \le 5$, Eq. (45) is again applicable. Thus

$$\begin{bmatrix} f_0(5) \\ f_1(5) \end{bmatrix} = \begin{bmatrix} 0.27 \\ 0.25 \end{bmatrix},$$

and on the interval $5 \le t \le 8$ we can use Eq. (48) to show that

$$\begin{bmatrix} f_0(8) \\ f_1(8) \end{bmatrix} = \begin{bmatrix} 0.332 \\ 0.19 \end{bmatrix} .$$

Without repair, an instantaneous attack of density $\eta = 20$ gives

$$f_0(\eta) = e^{-t_1} = 0.368 \tag{51}$$

However, applying the same attack scheme as shown in Fig. 1, without repair (i.e., $r_1 = 0$) we have that

$$f_0(1) = 0.368$$

$$f_0(5) = 0.135$$

$$f_0(8) = 0.135$$

$\gamma(t)$ for given d can now be computed via Eq. (9) since

$$\gamma(t) = 1 - e^{-d f_0(t)\gamma(t)} \tag{52}$$

For example, if $d = 6$, for the system with repair rate $r_1 = 0.1$, $\gamma(8) \approx 0.8$ while for the system without repair, $\gamma(8) \approx 0.03$. Therefore, there is an enormous difference in the behavior of the systems with and without repair.

Once the $f_k(t)$ (and the $g_k(t)$ which are computed in the same manner) are found, these numbers may be substituted into the vulnerability expression, thus synthesis goals have to be modified. Among the unknowns of the system we could include the functions $r_k^s(t)$ and $r_k^\ell(t)$. The cost of repair must then be included in the objective function which is to be minimized. The best synthesis procedure that we can suggest is actually one of repetitive analysis. This is unfortunate but the shortcoming seems to be inherent in the problem.

Another generalization can be introduced if we consider the fact that the kth hit on a target may not be independent of the number k. That is, the probability of a given station (or link) being hit after it already has suffered k hits may vary with k. This situation corresponds to a system with memory. Furthermore, each weapon may be capable of delivering more than a single hit. For example, the weapons may vary in size (power), a hit may cause secondary explosions which have the effect of destroying more of the target, or each weapon may actually be composed of a number of smaller weapons. The latter eventuality occurs if we consider a weapon to be a bomb load, while we compute hits on the basis of the number of impacts of single bombs.

Both of the above generalizations can be included in our model without great difficulty. These extensions are due to Biermann [15] who considered similar problems in a different context. We will briefly formulate the appropriate equations and state the solutions. First, we examine the latter of the two previous generalizations (the possibility of multiple hits per weapon). Let t_i be the probability that a given station receives exactly i hits from a given weapon; let p_i be the probability of this event for the links. Then, if η changes by the amount $d\eta$, $f_k(\eta)$ will change by the amount

$$df_0(\eta) = -(1 - t_0)\, f_0(\eta)$$

$$df_k(\eta) = -(1 - t_0)\, f_k(\eta)\, d\eta + \sum_{j=1}^{k} t_j\, f_{k-j}(\eta)\, d\eta \qquad k = 1,\ 2,\ \ldots \qquad (53)$$

We can write Eq. (53) as the differential equation

$$\frac{df_k(\eta)}{d\eta} = -(1 - t_0)\, f_0(\eta)$$

$$df_k(\eta) = -(1 - t_0)\, f_k(\eta)\, d\eta + \sum_{j=1}^{k} t_j\, f_{k-j}(\eta) \qquad k = 1,\ 2,\ \ldots \qquad (54)$$

Biermann has shown that the solution of this system of equations is given by

$$f_0(\eta) = \left\{\exp -(1 - t_0)\,\eta\right\}$$

$$f_k(\eta) = \exp\left\{-(1 - t_0)\,\eta\right\} \sum_{r=1}^{k} \frac{\eta^r}{r!} \sum_i G_i(k; r) \qquad k = 1, 2, \ldots \quad (55)$$

The term $\sum_i G_i(k; r)$ is found as follows: A partition of k by r ($r \leq k$) is a set of r positive integers whose sum is k. The same r integers arranged in a different order are considered to be a different partition. $G_i(k; r)$ is defined to be the product of the r t_j quantities such that the indices j are a partition of k by r. The expression $\sum_i G_i(k; r)$ represents the sum over all such partitions of k by r.

The system with memory is then easily described by a minor modification of Eq. (54). If the system has memory, the probability t_i depends on the number of hits already taken by the target. Therefore we must consider probabilities of the form $t_{i,j}$. Probability $t_{i,j}$ represents the probability that a given station which has already been hit j times will suffer i additional hits from a single weapon. The differential equations of the system are

$$\frac{df_0(\eta)}{d\eta} + (1 - t_{0,0})\, f_0(\eta) = 0$$

$$\frac{df_k(\eta)}{d\eta} + (1 - t_{0,k})\, f_k(\eta) = \sum_{j=1}^{k} t_{j,k-j}\, f_{k-j}(\eta) \qquad k = 1, 2, \ldots \quad (56)$$

The solution of this system of equations is

$$f_0(\eta) = \exp\left\{-(1 - t_{0,0})\right\}$$

$$f_k(\eta) = \sum_{j=0}^{k} c_{jk}\left\{\exp\ -(1 - t_{0,j})\right\} \qquad k = 1,\ 2,\ \ldots \tag{57}$$

where

$$c_{kk} = -\sum_{j=0}^{k-1} c_{jk} \qquad k = 1,\ 2,\ \ldots$$

and

$$c_{jk} = \left[(1 - t_{0,k}) - (1 - t_{0,j})\right]^{-1} \sum_{i=j}^{k-1} c_{ji}\, t_{k-i,i} \qquad k = 1,\ 2,\ \ldots,$$

$$j = 0,\ 1,\ 2,\ \ldots$$

An alternative approach to using the above equations for finding the $f_k(\eta)$ is to solve Eq. (56) or (54) successively for f_k in terms of $f_{k-1}, \ldots, f_0$. However, this is a routine mathematical exercise and so we will not discuss it further.

VI. CONCLUSIONS AND FURTHER PROBLEMS

To remark that further problems exist is, to say the least, an understatement. In this paper, we have attempted to uncover a methodology with which some vulnerability problems can be studied.

The basis of our study is that if we both generate and destroy systems with "sufficient homogeneity," we can find asymptotic relations and recursion formulas for several vulnerability criterions. Our network model can be modified, as we demonstrated in Sec. III. Further modifications are possible. The description of vulnerability of networks with distance biased links can be used to study more general situations. Its main limitation is the number of regions in the network partitioning and the maximum path length to be considered; both of these numbers must be small.

The author is currently studying some of the computational aspects of the methods suggested in this paper and their generalizations to more sophisticated models. One such model deserves mention here. Suppose a single weapon (of a rather powerful nature) is directed at some point in the net. The actual place where the weapon hits is a random variable with a possibly known probability distribution. Then, if we associate an "energy function" with the weapon, how can we compute the vulnerability index of the net and how can we find minimum energy levels that the stations and links must survive in order to guarantee a given vulnerability index?

REFERENCES

[1] P. Baran, "On distributed communication networks," IEEE Trans on Comm. Tech., Vol. COM-12, pp. 1-9, March 1964.

[2] Z. Prihar, "Communication networks in random behavior," IEEE Trans. on Comm. Tech., Vol. COM-14, No. 4, pp. 389-399, August 1966.

[3] Y. Fu and S. S. Yau, "A note on the reliability of communication networks," J. Soc. Ind. App. Math., Vol. 10, No. 3, pp. 469-474, September 1962.

[4] I. Jacobs, "Connectivity of probabilistic graphs," Tech. Report No. 356, MIT Research Laboratory of Electronics, Cambridge, Mass., September 15, 1959.

[5] E. N. Gilbert, "Random graphs," Ann. Math. Stat., Vol. 30, pp. 1141-1144, December 1959.

[6] H. Frank and S. L. Hakimi, "Probabilistic flows through a communication network," IEEE Trans. Circuit Theory, Vol. CT-12, No. 3, pp. 413-414, September 1965.

[7] H. G. Landau, "On some problems of random nets," Bull. Math. Biophysics, Vol. 14, pp. 203-212, 1952.

[8] R. Solomonoff and A. Rapaport, "Connectivity of random nets," Bull. Math. Biophysics, Vol. 13, pp. 107-117, 1951.

[9] A. Rapoport, "Nets with distance bias," Bull. Math. Biophysics, Vol. 13, pp. 85-91, 1951.

[10] R. Solomonoff, "An exact method for the computation of the connectivity of random nets," Bull. Math. Biophysics, Vol. 14, pp. 153-157, 1952.

[11] A. Rapaport, "Contribution to the theory of random and biased nets," Bull. Math. Biophysics, Vol. 19, pp. 257-277, 1957.

[12] A. Rapaport, "Nets with reciprocity bias," Bull. Math. Biophysics, Vol. 20, pp. 191-201, 1958.

[13] W. Feller, An Introduction to Probability Theory and It's Applications, Vol. 1, New York: John Wiley and Sons, Inc., 1957.

[14] A. Biermann, "A general target theory of radiobiological action," Bull. Math. Biophysics, Vol. 25, pp. 273-296, 1963.

[15] A. Biermann, "A general target theory of radiobiological action:II," Bull. Math. Biophysics, Vol. 25, pp. 367-385, 1963.

[16] G. Hadley, Nonlinear and Dynamic Programming, Reading, Mass.: Addison-Wesley Pub. Co., 1964, Chapter 8.

[17] P. Erdos and A. Renyi, "On the evolution of random graphs," Publications of the Math. Inst. of Hung. Acad. of Sciences, Vol. 5, pp. 17-61, 1960.

[18] T. L. Austin, R. E. Fagen, W. F. Penny, and J. Riordan, "The number of components in random linear graphs," Ann. Math. Stat., Vol. 30, pp. 747-754, 1959.

[19] M. D. Kruskal, "The expected number of components under a random mapping function," Amer. Math. Monthly, Vol. 61, No. 6, pp. 392-397, June - July 1954.

[20] L. Katz, "Probability of indecomposibility of a random mapping function," Ann. Math. Stat., Vol. 26, No. 3, pp. 512-517, September 1955.

[21] B. Harris, "Probability distributions related to random mappings," Ann. Math. Stat., Vol. 31, No. 4, pp. 1045-1062, December 1960.

[22] H. Rubin and R. Sitgreaves, "Probability distributions related to random transformations on a finite set," Tech. Report No. 19A, Applied Math. and Stat. Lab., Stanford University, 1954.

[23] J. E. Folkert, "The distribution of the number of components of a random mapping function," unpublished Ph.D. dissertation, Michigan State University, 1955.

[24] R. E. Barlow and F. Proschan, Mathematical Theory of Reliability, New York: John Wiley & Sons, Inc., 1965, p. 12.

[25] E. A. Coddington and N. Levinson, Theory of Ordinary Differential Equations, New York: McGraw-Hill Book Co., Inc., 1955.

FOOTNOTES

1. $\gamma \sum_{k=0}^{K_s-1} f_k(\eta)$ is the average fraction of stations in the original net that can be reached from a station picked at random after the attack. Since $\gamma \sum_{k=0}^{K_s-1} f_k(\eta)$ is easily found once γ is known, we will use γ as a criterion even though $\gamma \Sigma f_k$ may be more realistic for many cases.

2. The problem can also be formulated as an integer programming problem [16] with <u>linear</u> constraints.

3. Our equation is slightly different than Rapaport's because he allows self loops. Equation (16) is derived in Appendix A.

4. The components of an undirected network are the maximum subsets of stations which can communicate with each other.

5. $\binom{k}{j} \triangleq \frac{k!}{j!(k-j)!}$.

6. For example, we can replace each undirected link by two opposite, directed links. The events that these links are destroyed are not statistically independent. However, if we treat these events as independent, the vulnerabilities that we find are upper bounds to the actual vulnerability of the network. The average number of links incident at each station is in the original network now equal to the

average number of outwardly directed links in the directed network. Hence we can now use Eq. (8) to give an upper bound to γ.

7. This assumption is the basis of most of our results.

APPENDIX A

DERIVATION OF EQUATION (9)

Select an arbitrary surviving station v_i at random and let $S_0 = \{v_i\}$. Let S_1 be the set of stations connected to v_i by links directed from v_i, . . ., and let S_i be the set of stations connected to the stations in S_{i-1} by links directed from S_{i-1}, . . . We will use the following procedure to count the number of station in $\bigcup_{i=0}^{\infty} S_i$. First, we trace all links emanating from S_0 (that is, we find the number of stations in S_1), . . . , at the ith stage, we trace all links emanating from S_{i-1} which have not been already traced (that is, we find the number of stations in S_i which have not already been contacted at a previous stage). In the derivation to follow, we will compute the expected number of new stations to be contacted at the r+1st stage, based on the assumption that the expected number of stations were contacted at the rth stage.

Let p(r) be the probability that a given station is contacted at the rth stage and let q(r) = 1 - p(r). The probability that a station is contacted for the first time at the r+1 st stage is

$$p(r) \prod_{i=0}^{r-1} q(i) = [1 - q(r)] \prod_{i=0}^{r-1} q(i)$$

where q(0) is the probability of not selecting a given station at stage 0.

At the r+1st stage, we examine all station in S_{r+1} by tracing links emanating from S_r. However, if a station in S_r is also in

S_i $(i < r)$, we will not retrace its links. The average number of links emanating from a station that has survived is αt, where t is the probability that any given station survives and α is the average number of links incident out of a surviving station given that its terminal stations have survived. Therefore, the expected number of links to be traced is

$$\alpha t(n+1)\,[1 - q(r)] \prod_{i=0}^{r-1} q(i) ,$$

and the probability that any given vertex is <u>not</u> contacted at the <u>r+1st</u> stage is

$$q(r+1) = \left(1 - \frac{1}{n}\right)^X$$

where $$X = \alpha t[n+1]\,[1 - q(r)] \prod_{i=0}^{r-1} q(i) ,$$

which, for large n, may be written as

$$q(r+1) = \exp\left\{- \alpha t\,[1 - q(r)] \prod_{i=0}^{r-1} q(i)\right\}$$

$$= \exp\left\{-\alpha t\left[\prod_{i=0}^{r-1} q(i) - \prod_{i=0}^{r} q(i)\right]\right\} ,$$

Therefore,

$$\prod_{j=1}^{r+1} q(j) = \prod_{j=1}^{r+1} \exp\left\{-\alpha t\left[\prod_{i=0}^{j-1} q(i) - \prod_{i=0}^{j} q(i)\right]\right\}$$

$$= \exp\left\{-\alpha t \sum_{j=1}^{r}\left[\prod_{i=0}^{j-1} q(i) - \prod_{i=0}^{j} q(i)\right]\right\},$$

and since

$$\sum_{j=1}^{r}\left[\prod_{i=0}^{j-1} q(i) - \prod_{i=0}^{j} q(i)\right] = q(0) - \prod_{i=0}^{r} q(i) ,$$

we have that

$$\prod_{j=1}^{r+1} q(j) = \exp\left\{-\alpha t\left[q(0) - \prod_{i=0}^{r} q(i)\right]\right\} .$$

Now, since $q(0) = 1 - (t/n) \to 1$ as $n \to \infty$ and

$$1 - \gamma = \lim_{r\to\infty} \prod_{i=0}^{r} q(i) ,$$

it follows that

$$\lim_{r\to\infty} \prod_{j=1}^{r+1} q(j) = \lim_{r\to\infty} \exp\left\{-\alpha t\left[q(0) - \prod_{i=0}^{r} q(i)\right]\right\}$$

or

$$1 - \gamma = \exp\{-\alpha t[\gamma]\} .$$

Thus $\gamma = 1 - \exp\{-\alpha t \gamma\}$.

But $t = \sum_{k=0}^{K_s - 1} f_k(\eta)$ and $\alpha = d \sum_{k=0}^{K_\ell - 1} g_k(\eta)$.

And Eq. (9) now follows.

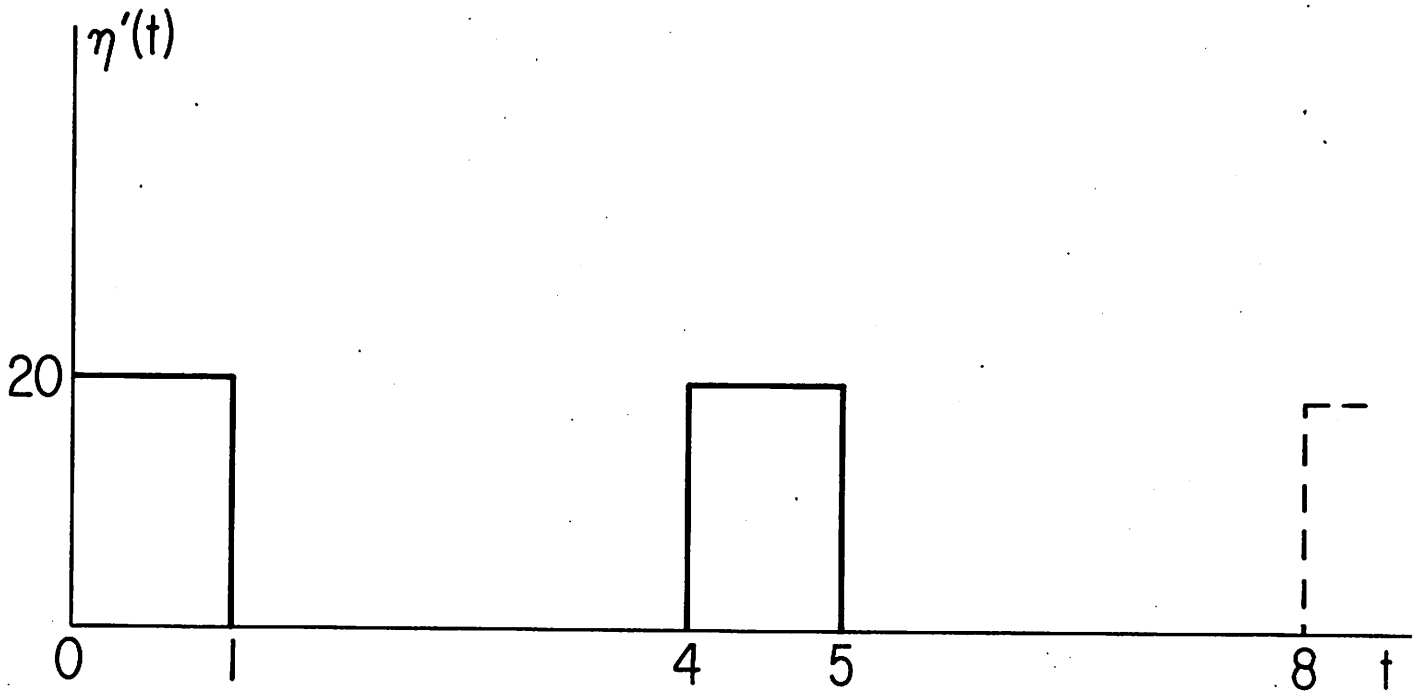

Fig. 1. Instantaneous weapon density for EXAMPLE 3.