COWLES COMMISSION DISCUSSION PAPER: ECONOMICS NO. 2117

Efficient Communication Networks - The Case of Two-Way Links

Donald Bratton

February 6, 1955

One recalls (CCDP Econ. 2102) that a communication network R for a set V of people provides links of unit length between certain couples of elements of V. In the paper cited, the links were assumed uni-directional; here they will be assumed two-way. Information can be sent from a person x to a person y providing there is a path from x to y in the network. The length of the shortest path is denoted $d_R(x,y)$. (Or simply $d(x,y)$.).

The links being un-oriented, in order that each person can reach each other person, (i.e., that R be "adequate") it is necessary and sufficient that the network be connected. Indeed, the persons x can reach are precisely those belonging to the connected component corresponding to x.

One calls the <u>solution time</u> of R, and denotes by T(R) (or simply T), the number $\max_{x,y} d_R(x,y)$. v denotes the number of people (size of V), and for each network R considered on V, $\ell(R)$ (or simply ℓ) denotes the number of links of R, k(R) (or simply k) the number of connected components of R. In conformity with ordinary usage, the elements of V will be called the vertices (hence the notation).

§1. - We are concerned with the networks which are efficient for the value of T and ℓ. In other words, we study the problem, what is the smallest value of T attainable for a given number v of people using a network of ℓ links?

One first notes that $v \leq \ell + k$, *and that equality holds if and only if the network* R *is without cycles* (i.e., R *is a sum of trees*, i.e., R is a "forest").

This results from elementary homology theory, as $v - \ell = k - b$ where b = basic number of cycles in R (Betti number).

It is clear that R is adequate if and only if it is connected (i.e., $k = 1$). One then has $\ell \geq v - 1$. Now, being given v vertices, one can link them together as a centralized network:

for which $T = 2$ and (as there are no cycles) ℓ attains its lower bound $\ell = v - 1$. It follows that $\ell = v - 1$, $T = 2$, is an efficient point and that the centralized network is efficient. There is thus only one other efficient point -- where $T = 1$, and here necessarily R is the "saturated" network, where $\ell = \binom{v}{2}$. Indeed, it results that

There are (*up to isomorphism*) *just two efficient symmetric networks for each value of* v; *they are the centralized and the saturated networks.*

To verify this result, it is required to show that a network R with $T = 2$ and $\ell = v - 1$ is necessarily centralized. Well, R is a tree where any two vertices are at most two steps apart. There exists a vertex r of R with only one link at it. The relation "the path from x to r passes through y" is a partial ordering $y \leq x$ on R for which r is the bottom element and all the elements are arranged in horizontal layers. Clearly r is the only element in its layer, and there are (as $T = 2$) at most two layers above r. If this top layer is void, then R is centralized with r the center; otherwise the top layer is nonvoid. In the latter case one easily sees that there is just one vertex c in the middle layer and that R is centralized with c as center, Q.E.D.

§2 - One can formulate and settle with the same sort of reasoning more complex problems. V being a finite set (the set of vertices for the networks considered), suppose given a set $S \subset V \times V$, and denote, for each network R on V,

$$(1) \qquad T_S(R) = \max_{(x,y)\in S} d_R(x,y).$$

d_R denotes, of course, the distance between x and y for the network R. $T_S(R)$ is the solution time of R when only certain origins and destinations are relevant.

R is said to <u>achieve</u> S when for each $(x,y)\in S$, there is a path in R from x to y. In order for R to achieve S it is necessary and sufficient that $T_S(R) < \infty$.

R is said to be <u>S-efficient</u> when R is efficient for the parameters $\not{l}$ and T_S.

One sees that in all these questions S appears only via the function T_S. Now, $d_R(x,x) = 0$ and $d_R(x,y) = d_R(y,x)$. Thus, in view of the definition of T_S, one could assume that $(x,x) \in S$ for each $x \in V$, and also that $(y,x) \in S$ whenever $(x,y) \in S$. Then S is itself simply a symmetric network on V! In what follows S is assumed to satisfy these two conditions.

It follows that, in order for R to achieve S, it is necessary and sufficient that connected components of S form a subpartition of those of R.

As a corollary, one deduces that if R achieves S, then $k(R) \leq k(S)$.

Suppose $S \neq \Delta$ (where Δ denotes the network on V where there are zero links). For a network R, in order that $T_S(R) = 1$ it is clearly necessary and sufficient that $R \supset S$ (i.e., R is obtained from S by suitably adding links to S). In particular as $T_S(S) = 1$, S is always S-efficient. Now suppose R is an S-efficient network $\neq S$. As R in particular achieves S, $k(S) \geq k(R) \geq v - \ell(R)$, it follows that $k(S) \geq v - \ell(R)$ i.e., $\ell(R) \geq v - k(S)$.

But for the network R_o obtained by re-forming each connected component of S into a centralised network one has $k(R_o) = k(S)$, $\ell(R_o) = v - k(S)$, and $T_S(R_o) \leq 2$.

It follows that, beside the efficient point $\ell = \ell(S)$, $T_S = 1$, there exists at most one other, at $\ell = v - k(S)$, $T_S = 2$. When does the latter point arise? It is clearly necessary for this that $\ell(S) > v - k(S)$, -- i.e., that S is <u>not</u> a forest. This necessary condition is also sufficient. For, S being <u>not</u> a forest, the network $R_o \not\supset S$ (because S possesses a cycle while R_o does not), and hence $T_S(R_o) \neq 1$, i.e., $T_S(R_o) = 2$. We have shown that,

S being an arbitrary symmetric network on V, in order that there exist S-efficient networks $R \neq S$ on V, it is necessary and sufficient that S possesses a cycle (i.e., that S is not a forest). Then such networks R are necessarily forests with the same connected components as S, and $T_S^{(R)} = 2$. (It follows that R has $v - k(S)$ links). And conversely, this efficient point is realized by each network R obtained by arbitrarily reforming the components of S into centralised networks.

§3 - The third problem, more general than each of the preceding, is also settled by the same methods. Let M be a set of messages. By a state of information one means a set $S \subset M \times V$; the relation "$(m,v) \epsilon S$" as interpreted "the message m is known to v". By a network one of course understands a set $R \subset V \times V$ such that $(x,x) \epsilon R$ for each $x \epsilon V$. One calls the compose of S and R, and denotes SR, the set of the $(m,v) \epsilon M \times V$ such that there exists $v \epsilon M$ such that $(m,w) \epsilon S$ and $(w,v) \epsilon R$. Similarly for R_1R_2 where R_i are $\subset V \times V$; the compose of two (symmetric) networks is a (symmetric) network; R^n denotes the compose of n R's together.

One supposes that after an initial state S_o of information, each person (vertex) communicates all the messages he knows to each person linked to him. The resulting state of information is S_oR. If in each unit time interval this communication process is repeated, the state of information when $T = n$ is S_oR^n. (This is a re-phrasing of Shimbel's "fundamental theorem"[1]). Thus, if a desired state of information S_1 is given, in order that S_1 be achieved at $T = n$ it is necessary and sufficient that $S_oR^n \supset S_1$.

The relation $S_oR^n \supset S_1$ can be interpreted in another way. It means that for each $m \epsilon M$ and $v_1 \epsilon V$ such that $(m,v_1) \epsilon S_1$ there exists $v_o \epsilon V$ such that $(m,v_o) \epsilon S_o$ and a path of length $\leq n$ from v_o to v_1

1. A. Shimbel, "Applications of Matrix Algebra to Communication Nets", Bulletin of Mathematical Biophysics 13 (1951) 165-178.

over which the message m can be <u>transported</u>.

One denotes by $T_{S_oS_1}(R)$ the least number n such that $S_oR^n \supset S_1$. R is called <u>(S_o,S_1)-efficient</u> when R is efficient for χ and T_{S_o,S_1}.

S_o and S_1 being supposed given, we seek the (S_o,S_1)-efficient networks.[2]

The more complex function T_{S_o,S_1} can be reverted to the earlier means as follows. Let (R_i) denote the solutions (or equally well, the minimal solutions) of the relation $S_oR \supset S_1$. It results that for each network R,

$$(1) \qquad T_{S_oS_1}(R) = \min_i T_{R_i}(R).$$

To verify this, one observes that the relation "$S_oR^n \supset S_1$" is equivalent to "there exists i such that $R^n \supset R_i$". Hence "$n \geq T_{S_oS_1}(R)$" is equivalent to "there exists i such that $n \geq T_{R_i}(R)$", which is what the equation (1) asserts.

It follows from equation (1) that if R is (S_o,S_1)-efficient, then there exists an i such that R is R_i-efficient and $T_{S_oS_1}(R) = T_{R_i}(R)$. It follows in particular, from §2, that the only possible efficient points are at $T = 1$, $T = 2$. And indeed, we see by §2 that if $T_{S_oS_1}(R) = 1$ then R is one of the R_i and thus is a solution of $S_oR \supset S_1$ with the smallest number of links. On the other hand if $T_{S_oS_1}(R) = 2$ then R is a forest with the same connected components as an R_i; as then $\chi(R) = v - k(R) = v - k(R_i)$, it follows that R_i is a solution of $S_oR_i \supset S_1$ with the largest number of connected components. We have proved the following:

2. Problem posed by R. Radner and A. Tritter in CCDP Econ. No. 2098.

Given initial and terminal states of information S_o, S_1, on the set V, where $S_o \not\supset S_1$, and given ℓ two-way links with which to connect the elements of V into some communication network R, the least time T of achievable communication is given as follows:

$$T = \infty \text{ for } 0 \leq \ell < a,$$

$$T = 2 \text{ for } a \leq \ell < b$$

$$T = 1 \text{ for } b \leq \ell < \infty,$$

where a and b are two numbers determined by S_o and S_1 such that $a \leq b$.

Let the network R = A be a solution of $S_o R \supset S_1$ with the largest number of connected components. R = B a solution with the smallest number of links. One has $a = v - k(A)$, $b = \ell(B)$.

The efficient point at $T = 1$, $\ell = b$, is realized by the network B.

In order that $a = b$, it is necessary and sufficient that there exists a solution of $S_o R \supset S_1$ which is a forest.

If $a \neq b$, the efficient networks at $T = 2$, $\ell = a$, are the forests R with the same connected components as A and such that $T_A(R) = 2$. Each forest of centralized networks where the components are the same as A's then does the trick.

One should note that when $M = V$ and $S_o = \Delta$ then $A = B = S_1$, which shows that these results are consistent with the earlier ones.