ABSTRACT

An analytical model of equilibria in a virtual circuit network under the GTE Telenet isolated dynamic routing is presented. In the model the distributions of virtual circuits over the network links are obtained as solutions to an equilibrium programming problem formulated for the network. Average network delay for the packet level is analyzed. The proposed approach is illustrated by simple examples where the results are obtained analytically. General numerical algorithm applicable to multinode networks is given.

I. INTRODUCTION

The purpose of this paper is to present a mathematical model of equilibrium regimes in the GTE Telenet network. Telenet is a packet-switched network operating in a virtual circuit mode [WEIR 82]. In such networks each session utilizes a single route until its termination. The route (virtual circuit) is set up by routing the first packet of every session, the call request packet, from the node of its origin through transit nodes towards the node of its destination. Decisions are made at every node as to which of the eligible links will carry a new virtual circuit. Telenet's routing rule dictates opening the circuit on the link that is currently least saturated. Thus, routing decisions are taken only on the basis of current information available locally in each node; this is isolated adaptive (or dynamic) routing [DAVI 79]. Mathematical modeling of networks under adaptive routing presents considerable difficulties [GERL 81, BERT 81, GEKL 76]. It appears that classical queueing models are too detailed in their description, and lead to virtually intractable models. Ours is a macroscopic model in that it deals with average flows. However, the stochastic nature of the underlying processes constantly manifests itself in the model. In fact, the model owes its validity to the result obtained in [EPGE 84] which states that for large link capacities and proportionally large demands the random numbers of virtual circuits carried by the network links are well represented by the average loads. This result is based on the observation that the process of establishment and termination of virtual circuits is independent of the packet transport process mainly due to enormous difference between comparatively large virtual circuit life span and small packet transport time [GERS 82]. The implication of this is that the description of the equilibrium of the slower virtual circuit level is uninfluenced by the packet level. On the other hand, analysis of the faster packet level is conducted on the premise that the virtual circuit picture is static. Incidentally, all results pertaining to the virtual circuit level are applicable to circuit-switched networks.

To represent an existing routing rather than deduce an optimal routing from some global criterion. As a matter of fact, such global criterion is unlikely to exist for an isolated dynamic routing since in this case we are dealing with many decision making nodes, each pursuing a goal of its own. Not coincidentally, we describe the network equilibria via several interdependent optimization problems each representing a node of the network. Equilibrium programming [GERS 81] is a ready-made tool for our purposes. A four-node example is given where equilibrium flows are found analytically for all possible input flows. In Appendix A we propose a numerical procedure applicable to multinode networks that finds equilibrium distribution of virtual circuits over the network links.

The layout of the paper is as follows: First, we describe the Telenet routing rule. Then routing variables are introduced and the network equilibrium is discussed for the virtual circuit level. We formulate conditions of network equilibrium in relation to the Telenet routing rule. Our next step is to devise an optimization problem for every node of the network. These nodal problems taken together constitute an equilibrium programming problem. We show that the Kuhn-Tucker conditions of optimality for the nodal problems are just another expression for the conditions of network equilibrium. Finally, on the packet level, we discuss the average network delay, one of the major performance characteristics of interest to practitioners.

II. TELNET ROUTING STRATEGY

As was observed, Telenet is a packet-switched network operating in the virtual circuit mode. A fixed end-to-end path is established through the network at session set-up time and all packets associated with that session follow this path. When the session is set up, route selection in the GTE Telenet network is executed on a node by node basis. When a node receives a call request packet, the first packet of every session, it searches its routing tables for a set of eligible outgoing links using the destination address as a search key. Links that belong to the minimum hop paths between the node and the call's destination are selected. An action on the other hand, some of the remaining links are designated as secondary.
The current GTE Telenet routing rule [WEIR 80] prescribes opening of a new virtual circuit on a primary path that is least saturated at the moment. The saturation is evaluated by the ratio of the number of active virtual circuits to the link capacity. If the outgoing links are of equal capacity (which we will assume for simplicity) then the least saturated link is the one that carries the smallest number of virtual circuits. The secondary links are used only if all primary links at the node are fully saturated. If a call request reaches a node at which no primary or secondary links are available the virtual circuit associated with the call is cleared back to the previous node where the route search is resumed with the link of the blocked path excluded. If the blocking occurs at the origin node, the session is assumed to leave the network.

We see that the routing decisions are taken on the basis of current traffic information available locally in each node. In the taxonomy of routing, this type of routing strategy is classified as dynamic and isolated.

III. MACRO DESCRIPTION

One very important aspect of the Telenet network operation is that the process of establishment and termination of virtual circuits is essentially independent of the packet transport process. The reason for this is twofold. Firstly, it is the count of the virtual circuits on the outgoing links and not the size of the packet queues that determines the routing decisions at the nodes of the network. Secondly, the average duration of a session is much longer than the average time required for the call request packet to travel from its origin to its destination and establish a virtual circuit. In other words, in a slower time frame of the virtual circuit birth-and-death process the effect of the packet transport is negligible. This allows us to consider the virtual circuit level separately and independently of the packet transport level. In fact, the main part of the proposed model is concerned with the virtual circuit level and describes a mathematical model for the equilibrium distribution of the virtual circuits over the links of the network for given inputs. All calculations pertaining to the data packet traffic can be carried out under the assumption that the virtual circuit picture remains static. This assumption is perfectly justified due to the earlier mentioned difference between the comparatively large virtual circuit life span and very fast data packet transport time.

In virtual circuit based networks the links are allowed to carry only a limited number of virtual circuits to avoid congestion in the packet queues. However, the "virtual capacity" of the links is usually large enough to play the role of a large parameter. Asymptotic analysis with respect to this large parameter shows that, for Poisson inputs and exponentially distributed holding time, the network attains an equilibrium under which the number of active virtual circuits carried by each link deviates very little from its equilibrium value; more precisely, the variance (appropriately normalized) tends to zero with the increase in the link virtual capacity [EPGE 84].

In a well designed network the primary routes should be able to accommodate the average peak demand. We assume that this is true for all examples analyzed in this paper. Interestingly, one of the conclusions drawn from the asymptotic analysis mentioned above is that nonprimary routes are used only for service of very rare large fluctuations; in other words, the variability of the underlying processes does not force the network to use the nonprimary routes systematically provided there is enough capacity for the average demand.

Operating experience with the GTE Telenet network where links do have large virtual capacity has shown that "traffic flows in the network are extremely stable and predictable" [WEIR 82].

IV. ROUTING VARIABLES

We will now turn to the mathematical formulation of the model. Networks will be represented by undirected graphs. As an illustration, consider a four-node network in Figure 1.

Let \( Y_{13} \) be the average number of sessions originating at node 1 and destined for node 3. The links (1, 2) and (1, 4) are on the primary (minimum hop) paths for the origin-destination pair (1, 3) and the flow \( Y_{13} \) will be split between them in certain proportions say \( p_{12} \) and \( p_{14} \). \( p_{12} + p_{14} = 1 \). We will call these proportions link routing variables. Clearly, these variables are always non-negative.

In node 3 the flow \( Y_{31} \) destined for node 1 is split between the primary links (3, 2) and (3, 4) in accordance with routing variables \( p_{32} \) and \( p_{34} \). What values these variables are assigned is determined by the routing strategy and about it later. We assume for simplicity that there is no external input at nodes 2 and 4. Viewed from a node, the total flow on an outgoing link typically consists of two components: one shaped at the node and the other shaped elsewhere. For instance, at node 1 the total flow on the link (1, 2) is given by

\[
 f_{12} = p_{12} Y_{13} + p_{32} Y_{31} + f_{12}^* \tag{1}
\]

Here the component \( p_{12} Y_{13} \) is the result of routing decisions made in node 1 while the component \( p_{32} Y_{31} \) is forward on the link (1, 2) via the link (2, 3) by the node 3, and the flow \( f_{12}^* \), although originating in node 1, is not affected by routing there since it is destined for node 2 and has only one primary way to go - the link (1, 2). Thus, in the node 1, \( p_{12} Y_{13} \) is the controlled flow component and \( p_{32} Y_{31} + f_{12}^* \) is the "imposed" component on the link (1, 2). Similarly, total flows on links (1, 4), (2, 3), and (3, 4) are
respectively.

In general, routing variables are of macroscopic nature; they are time averaged portions of nodal loads diverted on appropriate links. However, very often our reasoning related to average quantities invokes instantaneous, dynamic considerations. We justify this frivolity by the previously mentioned smallness of the variance. Moreover, we will be calculating the equilibrium virtual circuit loads on the network links by finding appropriate values for routing variables.

V. NETWORK EQUILIBRIUM

As was observed earlier, the Telenet routing rule requires the node to open a new virtual circuit on the least loaded eligible link. Clearly, this tends to minimize the difference between the virtual circuit loads carried by the outgoing links. However, every node tries to even out the load without regard to the consequences of its actions for other nodes; every node pursues its own goal, often being at cross-purposes with its neighbor nodes. The network equilibrium occurs when each node has achieved its goal of best possible load balance on its outgoing links. An equivalent characterization of the network equilibrium is that at each node no flow utilizes a link unless the link's total flow is minimal among the flows on eligible links. This is an adaptation of Wardrop's definition of equilibrium for the case of isolated dynamic routing [WARD 52].

In our four-node example, let us assume that equilibrium flow configuration is such that at node 1 the total flow on link (1, 2) is strictly greater than the total flow on link (1, 4), $f_{12} > f_{14}$. Clearly, this implies that all of the flow on link $Y_{13}$ is directed on the link (1, 4), i.e., $P_{12} = 0$ and $P_{14} = 1$, since otherwise the difference $f_{12} - f_{14}$ would have been only greater, which is contrary to the stated nodal goal. At the other node, any of the three situations are possible $f_{32} = f_{34}$, $f_{32} < f_{34}$ or $f_{32} > f_{34}$.

Assume for definiteness, that $f_{32} > f_{34}$. As in the case of node 1, this implies $P_{32} = 0$ (and hence $P_{34} = 1$) since $P_{32} = 0$ would be contrary to the goal of local routing at node 3 which is to minimize the difference between the total flows on links (3, 2) and (3, 4). Let us now combine the obtained values for the routing variables and the inequalities $f_{12} > f_{14}$ and $f_{32} > f_{34}$. In view of Eqs. (1), this gives

$$f_{14} = P_{14}Y_{13} + P_{34}Y_{31} + f_{14}$$

$$f_{32} = P_{32}Y_{31} + P_{12}Y_{13} + f_{32}$$

$$f_{34} = P_{34}Y_{31} + P_{14}Y_{13} + f_{34}$$

More important however is that the converse is true: inequalities (2) imply the equilibrium flow configuration with the routing $P_{12} = 0$, $P_{14} = 1$ and $P_{32} = 0$, $P_{34} = 1$. In other words, if the inputs $Y_{13}$ and $Y_{31}$ are not large enough to overcome the imbalances $\delta_1 = f_{12} - f_{14}$, and $\delta_3 = f_{32} - f_{34}$ due to the committed flow components, then the equilibrium flows correspond to the routing $P_{12} = 0$, $P_{14} = 1$ at node 1 and $P_{32} = 0$, $P_{34} = 1$ at node 3.

Altogether, there are nine regions in the space of inputs $Y_{13}$ and $Y_{31}$ and nodal imposed imbalances $\delta_1$ and $\delta_3$, each region being associated with a distinct equilibrium flow pattern. These are catalogued in Table 1. Interestingly, in eight out of nine cases, at least one node uses only one of the two available primary routes, i.e., fixed routing sets in at least partially [cf. KLEI 76], p. 346]. Equilibrium with complete balance at both nodes, $f_{12} = f_{14}$ and $f_{32} = f_{34}$, occurs only when the imposed nodal imbalances are equal, $\delta_1 = \delta_3$, condition unlikely to be satisfied. Finally, the presence of splittable inputs at nodes 2 and 4 would not affect the picture presented since the imbalances $\delta_1$ and $\delta_3$ are unaffected by these inputs. The way we analyzed our four-node network quickly becomes impractical for larger networks. The remainder of this paper is devoted to developing an alternate approach.

First, however, we return to routing variables and give a general procedure that will allow us to avoid redundant variables.

<table>
<thead>
<tr>
<th>Equilibrium Flow Configuration</th>
<th>Equilibrium Routing</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_{12} &gt; f_{14}$</td>
<td>$P_{14} = 1$</td>
</tr>
<tr>
<td>$f_{32} &gt; f_{34}$</td>
<td>$P_{34} = 1$</td>
</tr>
<tr>
<td>$f_{12} &lt; f_{14}$</td>
<td>$P_{14} = 0$</td>
</tr>
<tr>
<td>$f_{32} &lt; f_{34}$</td>
<td>$P_{34} = 0$</td>
</tr>
</tbody>
</table>

TABLE 1
VI. SHARING LINKS

Consider now a more general network shown in Figure 2.

At node 3 the flow $y_{31}$ destined for node 1 shares its primary links, (3, 2) and (3, 4), with the flow $y_{37}$ destined for node 7. The latter flow has one additional primary link, (3, 5). The fact that the two flows share some primary links affects the way routing variables are introduced at node 3. If $p_1$ and $p_2$ are the link routing variables for the flow $y_{31}$ and $q_1$, $q_2$, $q_3$ are the link routing variables for the flow $y_{37}$, then there should be a relation between the variables serving different flow components on the same links. Indeed, suppose a call request destined for node 7 is being routed at node 3. Three primary links, (3, 2), (3, 4) and (3, 5), are eligible for the opening of a new virtual circuit. Assume for the moment that link (3, 5), the one that is not available to flow $y_{31}$, is carrying more virtual circuits than either link (3, 2) or link (3, 4). Then, according to the Telenet routing rule, the choice will be made between links (3, 2) and (3, 4), the two least loaded links. This is exactly the choice we have got for the call requests destined for node 1. The outcome should not depend on the destination (whether it is node 1 or whether it is node 7) since only the current loads on the eligible links affect the choice. Hence, we will use the variables $p_1$ and $p_2$ for both the flow $y_{31}$ and the part of the flow $y_{37}$ served by the same links. Schematically, the assignment of routing variables in node 3 is shown in Figure 3. (A general procedure for assignment of routing variables to links has been developed.)

Altogether we now have four routing variables instead of five that we would have if the connection between allocations of the two flows that share links were ignored. By taking into account this connection we made an important step towards "sufficiency" of the model. By that we mean exclusion of infeasible equilibria. In our four-node example the model admits a unique equilibrium except for the symmetric case of $\Delta_1 = \Delta_3$. Symmetry seems to unleash additional degrees of freedom in our model of network equilibrium. This is best illustrated by the two-node example shown in Figure 4.

VII. EQUILIBRIUM PROGRAMMING DESCRIPTION

In this section we will give an equilibrium programming description of the network equilibrium flow configurations.

Equilibrium programming (EP) is a generalization of mathematical programming in which several optimization problems are solved simultaneously; the optimization problems are interdependent in that decision variables of one problem may enter as parameters into the goal function and/or constraints of other problems [GAZA 81]. As an analytical tool EP is innately suitable for modeling systems which comprise several decision-makers, each of whom optimizes a portion of the system.

We will formulate an EP problem whose solutions, called equilibria, satisfy the conditions of equilibrium flow configuration in a network under Telenet-type local routing. Before we do it in general form let us return to the network in Figure 1 for which we have found all equilibrium flow configurations for all inputs. It turns out that equilibrium flows in a network under the Telenet routing rule can be obtained as solutions to a multicomponent optimization problem. We have observed previously that in the Telenet network every node pursues its own goal of achieving the best possible balance of total flows on the outgoing links. Accordingly, every node is assigned a goal function to be optimized over the control.
variables at the disposal of this node. Thus, for node 1 of our four-node example, let

\[ S_1 = -f_{12} \ln f_{12} - f_{14} \ln f_{14} \]  

(3)

where

\[ f_{12} = p_{12} y_{13} + p_{32} y_{31} + f^{+} \]

\[ f_{14} = p_{14} y_{13} + p_{34} y_{31} + f^{+} \]

\[ p_{12} + p_{14} = 1, \quad p_{12} \geq 0, \quad p_{14} > 0 \]

Expressions for \( f_{12} \) and \( f_{14} \) were obtained in Sec. IV; constraints on \( P_{12} \) and \( P_{14} \) are straightforward properties of link routing variables. The function \( S_1' \), called nodal entropy, depends on \( P_{12} \) and \( P_{14} \) via \( f_{12} \) and \( f_{14} \). Similarly, the nodal entropy for node 3 is given by

\[ S_3 = -f_{32} \ln f_{32} - f_{34} \ln f_{34} \]  

(4)

where

\[ f_{32} = p_{32} y_{31} + p_{12} y_{11} + f^{+} \]

\[ f_{34} = p_{34} y_{31} + p_{14} y_{11} + f^{+} \]

\[ p_{32} + p_{34} = 1, \quad p_{32} \geq 0, \quad p_{34} \geq 0 \]

Observe that variables \( P_{32} \) and \( P_{34} \) play the role of parameters in the entropy \( S_1 \), while \( P_{12} \) and \( P_{14} \) are parameters in \( S_3 \). Our next step is to write out the Kuhn-Tucker conditions for the nodal entropies. For \( S_1 \), these conditions are

\[ \frac{\partial S_1}{\partial P_{12}} = -y_{13} (\ln f_{12} + 1) = \mu - \lambda_1 \]

\[ \frac{\partial S_1}{\partial P_{14}} = -y_{13} (\ln f_{14} + 1) = \mu - \lambda_2 \]

where \( \lambda_1 \geq 0, \quad \lambda_2 \geq 0, \) and \( \mu \) is real. After some algebra we get

\[ f_{12} = e^{\mu+\lambda_1}, \quad f_{14} = e^{\mu+\lambda_2}. \]  

(5)

These equations are accompanied by the so-called complementarity conditions

\[ \lambda_1 P_{12} = 0, \quad \lambda_2 P_{14} = 0, \quad \lambda_1 \lambda_2 = 0. \]  

(6)

Similarly, the Kuhn-Tucker conditions for a maximum of \( S_3 \) are

\[ f_{32} = e^{\nu+\zeta_1}, \quad f_{34} = e^{\nu+\zeta_2} \]  

(7)

where \( \nu \) is real, \( \zeta_1 \geq 0 \) and \( \zeta_2 \geq 0, \) with the complementarity conditions

\[ \zeta_1 P_{32} = 0, \quad \zeta_2 P_{34} = 0, \quad \zeta_1 \zeta_2 = 0. \]  

(8)

The meaning of the Kuhn-Tucker conditions for \( S_1 \) is as follows: the routing \( P_{12}, P_{14} \) renders \( S_1 \) maximal if and only if there exist real \( \mu \) and non-negative \( \lambda_1 \) and \( \lambda_2 \) such that Eqs. (5) and (6) hold. Assume now that \( P_{12} \) and \( P_{14} \) satisfy these equations and that \( P_{12} > 0. \) Then in view of (6), \( \lambda_1 = 0 \) and thus \( f_{12} \leq f_{14}^+. \) That is, if link \((1, 2)\) is utilized by the flow \( Y_{13} \) then its total flow is less than or equal to the total flow on the other link. Conversely, let \( f_{12} < f_{14}^+ \); then, due to Eqs. (5), \( \lambda_1 < \lambda_2. \) Since \( \lambda_1 \lambda_2 = 0, \) we have \( \lambda_1 = 0 \) and hence \( \lambda_2 > 0. \) Finally, \( \lambda_2 > 0 \) implies via Eqs. (6) that \( P_{12} = 0 \) and thus \( P_{13} = 1. \) That is, all of the flow \( Y_{13} \) is directed onto link \((1, 2). \) To summarize the above argument, the nodal entropy \( S_1 \) is maximized if and only if the routing \( P_{12} \) and \( P_{14} \) is such that the flow \( Y_{13} \) utilizes only the least loaded link (unless the outgoing links carry equal total flows and then both links may be utilized). If the same situation prevails at node 3, then the routings \( P_{12}, P_{14}, P_{32}, P_{34} \) and the corresponding total link flows make up what we previously called equilibrium flow configuration (see Sec. V). Thus network equilibria can be obtained as solutions to simultaneous maximization of the nodal entropies. A proof of this statement for a general network under Telnet-type routing rule follows the argument we used for our four-node example and won't be given here.

In general, the equilibrium programming description of the equilibria of a packet-switched virtual circuit network with \( n \) nodes consists of \( n \) interdependent optimization problems. The optimization problems are all similarly structured: for node \( i \) the goal function has the form

\[ S_i = -\sum_j f_j \ln f_j, \]

where \( \sum_j \) goes over the outgoing links utilized by the flows routed at the node; constraints consist of the unity sums for every group of the routing variables (see Appendix A) and expressions for the total link flows \( f_j. \)

Several general methods available for solution of EP problems as well as important existence results can be found in [GAZA 81]. However, the simplicity of the constraints and the special structure of the goal functions in our network EP problems allow us to devise specialized solution procedures. One such procedure is outlined in Appendix B where it is used to find equilibrium virtual circuit configuration for the seven-node network shown in Figure 2.

VIII. PACKET LEVEL: DELAY ANALYSIS

Average network delay is a widely used measure of performance for packet-switched networks. It is the average delay endured by the average packet traversing the network. As has been observed in Sec. III, the packet transport process is evolving much faster than the birth-and-death process of the virtual circuits. The
difference is so significant that a packet traveling through the network almost certainly will encounter no changes in the counts of virtual circuits on the links of its path. In equilibrium the data packets will be pushed through the network in accordance with the route map established for them on the virtual circuit level. This map reflects the equilibrium values of the routing variables that are obtained as a solution to the EP problem for given external inputs. In sum, for each equilibrium virtual circuit configuration, the packet level is modeled as a network under static routing. Each link in this network has two infinite buffer servers, one for each direction. We will assume that the average number of packets per session is large and equal in both directions. (Note here that the largeness of the virtual capacity of the link does not imply smallness of the variances of the data packet queue lengths.) Under this assumption it is not difficult to show that the packet flow is proportional to the virtual circuit flow. Specifically, if \( \lambda_{ij} \) is the packet flow on the link \((i, j)\) then

\[
\lambda_{ij} = \beta f_{ij}
\]

and

\[
\beta = \frac{Nw}{cT}
\]

where

- \( c \) - the link capacity in bits/sec
- \( \tau \) - the session average holding time in sec
- \( N \) - the average number of packets per session
- \( w \) - the average packet length in bits

The formula for the average network delay, \( T \), 

\[(11)\]

\[
T = \frac{1}{ScY_{rs}} \sum_{i,j} \frac{1}{1-\lambda_{ij}}
\]

Formula (11) is valid under the standard assumption of Poissonian interarrival times and the Kleinrock independence assumption of exponentially distributed packet length.

We will rewrite formula (11) in the form which lends itself readily to qualitative analysis:

\[
T = \bar{\tau} \bar{\lambda} \bar{c}
\]

(12)

where

\[
\bar{\tau} = \frac{\sum_{i,j} \lambda_{ij}}{\sum_{i,j} c\lambda_{ij} \sum_{i,j} 1-\lambda_{ij}}
\]

is the average path length (by length we refer to the number of links encountered in the path), and

\[
\bar{c} = \frac{\sum_{i,j} c\lambda_{ij}}{\sum_{i,j} 1-\lambda_{ij}}
\]

is the average (over the network) link delay.

Recall that under the Telenet routing rule, minimum-hop paths are used as primary routes and they carry all of the traffic. Thus, the first factor, \( \bar{\lambda} \), in formula (12) is minimal. As far as the second factor, \( \bar{\tau} \), is concerned it turns out that the Telenet routing rule together with the best load balance achievable locally at the nodes of the network, simultaneously attains minimal average link delay over the same set of control parameters.

IX. CONCLUSIONS

In this paper we have presented a mathematical model of equilibria in a packet-switched network operating in the virtual circuit node under an isolated dynamic routing. The model is comprised of several interdependent maximization problems, one for each node of the network. Solving these problems simultaneously for given external demand gives the equilibrium distribution of virtual circuits over the links of the network. On the basis of this distribution, packet level performance characteristics such as the average network delay are calculated.

Generally speaking, we model a dynamic isolated routing by an equivalent static routing. In fact, the network in equilibrium behaves as if the static routing were employed. Moreover, experience as well as analytical and numerical analysis show that a single path is utilized by the sessions between most of the origin-destination pairs in the network. The temptation therefore is to use the cheaper static routing. One should keep in mind however that this would require a mechanism for routing adjustment in response to changing external demand. Transient processes should also be taken into account.

APPENDIX A

We will use the network in Figure 2 to outline an algorithm of gradient type that finds an equilibrium flow configuration by solving the EP problem formulated for the network.

In node 3 where two inputs are routed towards their respective destinations over two overlapping groups of primary links, a maximum of the nodal entropy

\[
S_3 = f_{32} \ln f_{32} + f_{34} \ln f_{34} + f_{35} \ln f_{35}
\]

is sought subject to

\[
p_1 + p_2 = 1, \quad p_1 > 0, \quad p_2 > 0
\]

and

\[
q_1 + q_2 = 1, \quad q_1 > 0, \quad q_2 > 0
\]

Here the routing variables \( p_1 \), \( p_2 \) and \( q_1 \), \( q_2 \) are as in Figure 3 and \( f_{ij} \) are the total flows on the links with the corresponding indices. The expressions for the flows are

\[
f_{35} = f_{35}' + q_2 \gamma_{37}
\]

\[
f_{34} = f_{34}' + p_2 (q_1 \gamma_{37} + \gamma_{31})
\]

\[
f_{32} = f_{32}' + p_1 (q_1 \gamma_{37} + \gamma_{31})
\]

where \( f_{ij}' \) are the components of the total flows \( f_{ij} \) that are not controlled in node 3. For instance, for link (3, 4), the component \( f_{34}' \) may include the committed flow as well as the flow directed via link (1, 4) at node 1 and via link...
The expressions for nodal entropies $S_1$ and $S_3$ are constructed similarly. To describe the iterative procedure that finds equilibrium values of the routing variables we need the function

$$H(x) = \begin{cases} 1 & \text{if } x > 1 \\ x & \text{if } 0 \leq x \leq 1 \\ 0 & \text{if } x < 0 \end{cases}$$

that will help us to allow for the fact that a routing variable is never greater than one or less than zero.

The procedure is initialized by distributing the external inputs at the network nodes in accordance with some arbitrary set of values for the routing variables. The $k+1$ step of the procedure is given by

$$q_1^{k+1} = H[q_1^k + \Delta (f_{35}^k - \min(f_{32}^k, f_{34}^k))]$$

$$p_1^{k+1} = H[p_1^k + \Delta (f_{34}^k - f_{32}^k)]$$

$$r_1^{k+1} = H[r_1^k + \Delta (f_{14}^k - f_{12}^k)]$$

$$s_1^{k+1} = H[s_1^k + \Delta (f_{76}^k - f_{71}^k)]$$

where $r_1$ and $s_1$ are routine variables for links (1, 2) and (7, 1), respectively, $\Delta$ is the step size, and the rest are as defined above. Strictly speaking, for the procedure to be called gradient it ought to include the derivatives of the optimized function with respect to the control variables. In our formulae instead we have the total link flows of which the derivatives are monotonic functions.

The intent of the procedure is at each iteration to move the flows closer to satisfying the Kuhn-Tucker conditions in all nodes of the network, i.e., to move them closer to the best balance locally achievable. The procedure simultaneously maximizes the nodal entropies $S_1$, $S_3$ and $S_7$ subject to the flow conservation constraints.

The simplicity of this example is in the fact that every group of routing variables consists of two variables that sum up to unity; hence only one of the two need to be traced.

In Figure 5-a are shown the inputs and committed link flows and in Figure 5-b the equilibrium flows obtained by the above procedure.

In general case when variable groups may include more than two variables the following "cautious" procedure has been tested: at every iteration the redistribution is effected only between the activated link(s) carrying the largest total load and the link(s) carrying the smallest total load. Activated means that the routing variable of the link is greater than zero, i.e., the link carries the load that is controllable in the current node.

REFERENCES


