

# Analysis of Utility-Theoretic Heuristics for Intelligent Adaptive Network Routing

Armin R. Mikler\*, Vasant Honavar and Johnny S.K. Wong

Department of Computer Science

Iowa State University

Ames, Iowa 50011, USA

mikler|honavar|wong@cs.iastate.edu

## Abstract

Utility theory offers an elegant and powerful theoretical framework for design and analysis of autonomous adaptive communication networks. Routing of messages in such networks presents a real-time instance of a multi-criterion optimization problem in a dynamic and uncertain environment. In this paper, we incrementally develop a set of heuristic decision functions that can be used to guide messages along a near-optimal (e.g., minimum delay) path in a large network. We present an analysis of properties of such heuristics under a set of simplifying assumptions about the network topology and load dynamics and identify the conditions under which they are guaranteed to route messages along an optimal path. The paper concludes with a discussion of the relevance of the theoretical results presented in the paper to the design of intelligent autonomous adaptive communication networks and an outline of some directions of future research.

## Introduction

With the unprecedented growth in size and complexity of communication networks, the development of intelligent and adaptive approaches to network management (including such functions as routing, congestion control, etc.) have assumed considerable theoretical as well as practical significance. Knowledge representation and heuristic techniques (Pearl 1984) of artificial intelligence, utility-theoretic methods of decision theory, as well as techniques of adaptive control offer a broad range of powerful tools for the design of intelligent, adaptive, and autonomous communication networks. This paper develops and analyzes some utility-theoretic heuristics for adaptive routing in large communication networks.

Routing (Bertsekas & Gallager 1992) in a communication network refers to the task of propagating a message from its source towards its destination. For each

message received, the routing algorithm at each node must select a neighboring node to which the message is to be sent. Such a routing algorithm may be required to meet a diverse set of often conflicting performance requirements (e.g., average message delay, network utilization, etc.). This makes routing an instance of a multi-criterion optimization problem.

For a network node to be able to make an optimal routing decision, as dictated by the relevant performance criteria, it requires not only up-to-date and complete knowledge of the state of the entire network but also an accurate prediction of the network dynamics during propagation of the message through the network. This, however, is impossible unless the routing algorithm is capable of adapting to network state changes in almost real time.

In practice, routing decisions in large communication networks are based on imprecise and uncertain knowledge of the current network state. This imprecision is a function of the network dynamics, the memory available for storage of network state information at each node, the frequency of, and propagation delay associated with, update of such state information. Thus, the routing decisions have to be based on knowledge of network state over a local neighborhood supplemented by a summary of the network state as viewed from a given node. Motivated by these considerations, a class of adaptive heuristic routing algorithms have been developed over the past few years (Mikler, Wong, & Honavar 1994). Experiments demonstrate that such algorithms have a number of interesting properties including: automatic load balancing and message delay minimization. The work described in this paper is a step toward the development of a theoretical framework for the design and the analysis of such heuristics.

In what follows, we draw upon concepts of utility theory (French 1986) to design and analyze utility-theoretic heuristics for routing in large communication networks. Various heuristics are designed and their properties are precisely analyzed. The paper concludes

---

\*Armin Mikler is currently a post-doctoral fellow at the Scalable Computing Laboratory of the Ames Laboratory (U.S. Department of Energy) at Iowa State University, Ames, IA 50011

with a discussion of the relevance and limitations of the main results and some directions for further research.

## Utility-Theoretic Heuristics for Routing

Routing messages in large communication networks so as to optimize some desired set of performance criteria presents an instance of resource-bounded, multi-criteria, real-time, optimization problem. Our proposed solution to this problem involves the use of *utility-theoretic heuristics*. *Utility* is a measure that quantifies a decision maker's preference for one action over another (relative to some criteria to be maximized) (French 1986). When the result of an action is uncertain, it is convenient to use the *expected* utility of each action to pick actions which maximize the expected utility. The heuristic function enables each node  $n_j$  in the network to select a *best* neighbor in its neighborhood to route a message  $M$  (which it has received or generated) towards its destination.

The utility  $U_i^d$  of node  $n_i$  (with respect to a destination  $n_d$ ) is computed by a neighboring node,  $n_j$ , as  $n_j$  attempts to route a message  $M$  that it has received, along a desired (e.g., minimum delay) path, to  $M$ 's destination,  $n_d$ . A node  $n_j$  preference-orders its neighbors  $n_i$  according to their respective utilities. We say that the router at  $n_j$  is *indifferent* to the choice between two neighbors  $n_k$  and say  $n_l$  if  $U_k^d = U_l^d$  (where  $n_d$  is the destination of the message  $M$  being routed by  $n_j$ ). We denote the indifference between two nodes as  $n_k \sim n_l$ . We say that a neighboring node  $n_k$  is preferred (by the router at  $n_j$ ) over another neighbor  $n_l$  if  $U_k^d > U_l^d$ . We denote this preference by  $n_k \succ n_l$ .

For the purpose of the analysis that follows, it is assumed that the network is a regular rectangular grid (with adjacent nodes being unit distance of each other). Additional assumptions concerning load and load dynamics are made as necessary. A suitably defined *reward* function provides the directional guidance necessary to route each message towards its destination.

In the regular grid network, let  $D_{i,d}$  denote the Manhattan distance between a node  $n_i$  and  $n_d$ . Other topologies may require the use of other distance measures. We define the *partial reward* for node  $n_i$  as  $R_i^d = f_R(D_{i,d})$ , where  $f_R$  is a *reward function* chosen such that  $\forall i \forall j \ D_{i,d} \leq D_{j,d} \iff f_R(D_{j,d}) \leq f_R(D_{i,d})$ .

There are many possible choices for the reward function  $f_R(\cdot)$ . A particular example of  $f_R(\cdot)$  is given by  $f_R(D_{i,d}) = (m + n) - \frac{D_{i,d}}{m+n}$ , where  $n$  and  $m$  are the dimensions of the grid network. Note that the results that follow are independent of any particular choice of  $f_R(\cdot)$  so long as the reward increases as a message approaches its destination.

We define a *cumulative reward*  $R^P$  obtained by a message  $M$  traveling along a path  $P$  (from its source  $n_s$  to its destination  $n_d$ ) as  $R^P = \sum_{n_i \in P} R_i^d$ .

At each node  $n_i$  along its path  $P$ , the delay encountered by a message  $M$  is modeled by a non-negative, bounded cost  $C_i$ . That is,  $\forall i \ 0 \leq C_i \leq \xi$ . It is further assumed that the penalty  $C_i$  remains constant during the time it takes to make a routing decision for message  $M$  at node  $n_i$ . If cumulative delay is to be minimized, a natural interpretation of  $C_i$  is the delay (on account of load) at  $n_i$ . However, since delays can become unbounded when there is queueing, it may be necessary to discard some messages in order to keep the delay bounded at the expense of message loss. If cumulative load is to be minimized,  $C_i$  is guaranteed to be bounded by the maximum utilization  $\rho \leq 1$ .

The total cost incurred by a message along a path  $P$  is given by  $C^P = \sum_{n_i \in P} C_i$ . We can now define the net *partial payoff*  $Z_i^d$  received by a message  $M$  when it reaches the node  $n_i$  on its way to its destination  $n_d$  as  $Z_i^d = R_i^d - C_i$ . Correspondingly, the total payoff along a path  $P$  is given by  $Z^P = R^P - C^P$ . Let  $\Pi$  be a minimum cost path from a source  $n_s$  to a destination  $n_d$ . The cost  $C^\Pi$  along this path is given by  $C^\Pi = \min_{P \in \mathcal{P}} \{C^P\}$ .

In the discussion that follows, in order to simplify our analysis, we proceed under the assumption that the network is uniformly loaded. This assumption is captured by the following definition:

**Definition 1** If  $\forall i, C_i = \kappa$  ( $0 \leq \kappa \leq \xi$ ), we refer to the network as a *uniform cost network*.

In a uniform cost network, a simple utility function  $U^0$  defined by  $U_i^d = Z_i^d$  is sufficient to route each message along a minimum cost path to its destination. The uniform cost assumption renders the cost component in the payoff function irrelevant to the routing decision. This is no longer true when the network is not a uniform cost network. In what follows, we relax the uniform cost assumption by allowing a single *hotspot* (a node with a high load relative to its neighbors) in an otherwise uniform cost network.

## Routing in presence of a Single Hotspot

**Definition 2** A *hotspot*,  $n_h$ , in an otherwise uniform cost network is a single network node which has a higher load than its neighbors so that a message  $M$  traveling through it incurs a cost  $C_h > \kappa$  (where  $C_i = \kappa \ \forall i \neq h$ ).

Note that since the costs  $C_i$  are bounded by  $\xi$ , it follows that  $C_h \leq \xi$ . Further note that the above definition of a hotspot does not say anything about the relative difference in costs  $C_h$  and  $C_i$ . A more

realistic definition of a hotspot might insist that the cost of routing a message through a hotspot is *significantly larger* than that of routing the same message through a node in the neighborhood of the hotspot. Also, when a network deviates substantially from the uniform cost assumption, it is more useful to focus on the load distribution in the vicinity of a node rather than hotspots. However, to make the analysis mathematically tractable, the discussion that follows focuses on routing in an otherwise uniform cost networks with a single hotspot.

As the uniform cost assumption is relaxed by allowing a single hotspot  $n_h$  with cost  $C_h > C_j \forall j \neq h$  in the network, it is easy to show that relying on partial payoffs alone as utilities for routing messages can result in sub-optimal routes. Consider a grid network with node coordinates increasing as a message  $M$  travels east and south. From the uniform cost assumption, we have  $C_i = C_j = \kappa \forall i, j \neq h$ . Let  $x_s, y_s, x_d, y_d$  be the x and y coordinates of  $M$ 's source and destination, respectively. Let  $x_h$  and  $y_h$  be the x and y coordinates of a hotspot in one of the following configurations:

1.  $x_s \leq x_h \leq x_d \wedge y_s \leq y_h \leq y_d$
2.  $x_s \geq x_h \geq x_d \wedge y_s \geq y_h \geq y_d$

Here, the probability that a shortest path from  $n_s$  to  $n_d$  passes through the hotspot  $n_h$  is non-zero. That is,  $\exists$  a node  $n_i$  in the neighborhood of hotspot  $n_h$  that must decide how to route  $M$  so as to minimize the total cost incurred by  $M$ . As we show below, if this decision is based on a preference ordering induced by the naive utility function  $U^0$  given by  $U_i^d = Z_i^d$ , messages can be routed through the hotspot thereby incurring a higher cost than they would have otherwise.

**Assumption 1** For the discussion below, we assume that the reward functions chosen guarantee that  $\forall n_k \forall n_i$  in the network such that  $|R_i^d - R_k^d| > \xi$  whenever  $D_{i,d} \neq D_{k,d}$ .

This ensures that the cost  $C_i$  of a node  $n_i$ , (and  $n_h$  in particular) does not offset the *guidance* provided through  $R_i^d$  unless two nodes with equal rewards are being compared.

In the following we distinguish 4 canonical cases (see figure 1). We focus in our analysis on configuration 1 above. Similar arguments hold for configuration 2.

#### Case 0

This case combines 4 scenarios of placing nodes  $n_s$ ,  $n_d$ , and  $n_h$  in the grid network, each of which presents a trivial routing problem. In these scenarios, at least two of the nodes  $n_s$ ,  $n_d$ , and  $n_h$  are identical. That is,  $n_s = n_d = n_h$ ,  $n_s = n_d$ ,  $n_s = n_h \neq n_d$ , and

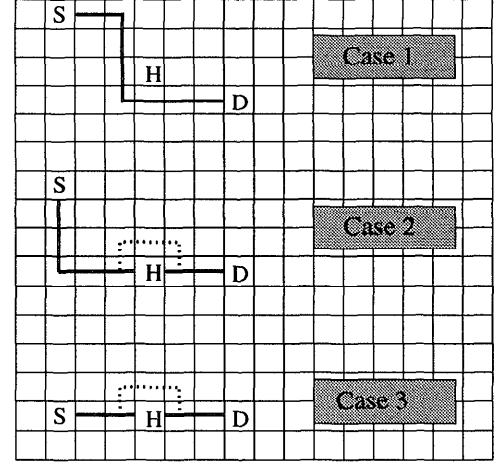


Figure 1: Sample node placement

$n_s \neq n_h = n_d$ . Clearly, in the first two scenarios, no routing decisions are needed as the message source coincides with the destination. Whenever the message source coincides with the hotspot as in the third scenario, the routing algorithm will select a neighbor  $n_k \in H_i$  with the highest utility. Hence, the routing algorithm performs as in the case of a uniform cost network (without hotspots). For the fourth scenario, Assumption 1 assures that  $n_d$  yields the highest partial reward  $R_i^d, \forall i$ , despite the fact that the cost incurred by hotspot conditions reduces its partial payoff. Hence, routing decisions can be made without taking cost  $C_i$  into consideration, as in the case of a network without hotspots.

#### Case 1

Let  $P\Delta_{i,j}$  denote the number of minimum hop paths from a node  $n_i$  to node  $n_j$ . This case encompasses all placements of nodes  $n_s, n_h$ , and  $n_d$ , such that

1.  $P\Delta_{s,h} > 1 \wedge P\Delta_{h,d} > 1$  or
2.  $P\Delta_{s,h} = 1 \wedge P\Delta_{h,d} \geq 1$  where  $P\Delta_{s,d} > 1$

Here, the hotspot  $n_h$  does not share either the x or y coordinates of  $n_s$  or  $n_d$ . That is,  $(x_s < x_h < x_d) \wedge (y_s < y_h < y_d)$ . Here, the all partial minimum hop paths from  $n_s$  to  $n_h$  may be part of a minimum cost path from  $n_s$  to  $n_d$  if all nodes  $n_i$  that neighbor the hotspot take action to route  $M$  so as to circumvent  $n_h$ . Thus, the utility function  $U^0$  given by  $U_i^d = Z_i^d$  is guaranteed to route  $M$  on a minimum cost path to its destination  $n_d$ .

**Lemma 1** In a uniform cost network with a single hotspot  $n_h$  located such that  $(x_s < x_h < x_d) \wedge (y_s <$

$y_h < y_d$ ), a routing algorithm which propagates a message  $M$  such that  $U^0$  is maximized at every intermediate step will yield an optimal path  $\Pi$  with cost  $C^\Pi$ .

### Case 2

Here,  $n_s$ ,  $n_d$ , and  $n_h$  are placed such that  $(x_s < x_h < x_d) \wedge (y_s < y_h = y_d)$  or  $(x_s < x_h = x_d) \wedge (y_s < y_h < y_d)$ , i.e.;  $(P\Delta_{s,h} > 1) \wedge (P\Delta_{h,d} = 1)$ .

Assuming the former, there exists a node  $n_i$  with coordinates  $(x_i, y_i)$  with  $(x_s < x_i < x_h) \wedge (y_i = y_h = y_d)$  from which the number of minimum hop routes  $P\Delta_{i,d} = 1$ . Since in a uniform cost network  $n_k \sim n_l$ ,  $\forall k, l \neq h$  the naive utility function  $U^0$  can guide a message  $M$  through  $n_i$ , thereby committing to a path  $P$  with cost  $C^P > C^\Pi$ . Assuming that  $M$  is only routed using utilities to choose among minimum hop routes, the additional cost  $(C^P - C^\Pi)$  is inflicted on  $M$  by  $n_h$ . If  $M$  is permitted to deflect from a minimum hop route, the additional cost  $(C^P - C^\Pi)$  is inflicted by  $n_h$  itself or due to the extended length of  $P$  in circumventing  $n_h$ .

### Case 3

This scenario consists of all placements of  $n_s$ ,  $n_d$ , and  $n_h$  such that  $(x_s = x_h = x_d) \wedge (y_s \leq y_h \leq y_d)$  or  $(x_s \leq x_h \leq x_d) \wedge (y_s = y_h = y_d)$ . Since there is only a single optimal path  $\Pi$  from  $n_s$  to  $n_d$ , i.e.,  $P\Delta_{s,d} = 1$ , message  $M$  must either visit  $n_h$  or deflect from the minimum hop path in order to circumvent  $n_h$ .  $U^0$ , however is not sufficiently informative to guarantee an optimal routing decision. Hence,  $M$  may be routed along a path  $P$  for which  $C^P > C^\Pi$ .

**Assumption 2** In the following we assume that a node  $n_j$  upon receiving a message  $M$  from a neighbor node  $n_i \in H_j$  will refrain from propagating  $M$  back to  $n_i$ .

This is a natural assumption that is meant to avoid the so-called *bouncing of messages* back to a node from which it was routed.

**Lemma 2** In a uniform cost network with a single hotspot  $n_h$ , a routing algorithm based on  $U^0$  will deflect a message  $M$  at most once in order to circumvent  $n_h$  provided bouncing is avoided (via Assumption 2).

The analysis of the performance of a routing algorithm based on  $U^0$  for each of the 4 cases above yields the following theorem:

**Theorem 1** In a uniform cost network with a single hotspot  $n_h$  with  $C_h > \kappa$  (where  $\forall i \neq h, C_i = \kappa$ ), a routing algorithm which propagates a message  $M$  such that  $U^0$  is maximized at every intermediate step is guaranteed to yield a path  $P$  with cost  $C^P$  such that  $C^P - C^\Pi \leq \max((C_h - \kappa), 2\kappa)$ .

The proof of this theorem is given in (Mikler, Honavar, & Wong 1996).

## Eliminating Suboptimality Using A Modified Utility Function

$Z_i^d$ , is determined solely from local information Sub-optimal routing scenarios discussed above arise primarily as a result of a lack of knowledge at  $n_i$  at the time it is routing a message  $M$  to a neighbor  $n_j$ , regarding the likely cost of completing the path from  $n_j$  to the destination of  $M$ , namely,  $n_d$ . As shown in section 2.4, source-hotspot-destination configurations corresponding to scenarios described in Case 2 and Case 3 can result in sub-optimal routes (i.e.,  $C^P > C^\Pi$ ) when routing decisions are based on  $U^0$ .

In what follows, we will modify  $U^0$  to obtain a utility function which is guaranteed to eliminate suboptimal routing decisions that arise in source-hotspot destination placements corresponding to the scenarios in Case 2 and Case 3. We proceed in two steps: First, we define a utility function  $U^1$  that eliminates suboptimal routing decisions that arise in scenarios corresponding to Case 3. We then modify  $U^1$  by introducing a cost estimator function to obtain a utility function  $U^2$  designed to eliminate suboptimal routing decisions that arise in Case 2 scenarios as well.

### Eliminating Sub-Optimality in Case 3

**Definition 3** Let  $U^1$  be a utility function given by:

$$U^1 = \begin{cases} R_j^d & \text{if } \kappa < C_j < 3\kappa \wedge \nexists k (R_j^d = R_k^d) \\ & \wedge (n_j \neq n_d) \\ Z_j^d & \text{otherwise} \end{cases}$$

$U^1$  exploits the fact that messages are to be routed in a uniform cost network with a single hotspot. If routing decisions are based on the preference ordering induced by  $U^1$  in an otherwise uniform cost network with a single hotspot, every message originating in a source  $n_s$  and a destination  $n_d$  that correspond to a source-hotspot destination placement described in Case 3 is guaranteed to be propagated along an optimal path  $\Pi$  between  $n_s$  to  $n_d$ . Using  $U^1$ ,  $n_i$  can decide whether or not to propagate  $M$  through a hotspot  $n_h$  in its neighborhood or to circumvent the hotspot by routing  $M$  through a different neighbor  $n_k \neq n_h$ . In other words, the preference ordering induced by  $U^1$  ensures that at a node neighboring a hotspot in a Case 3 scenario we have:

- $(C_h - C_k) = (C_h - \kappa) > 2\kappa \iff n_k \succ n_h$
- $(C_h - C_k) = (C_h - \kappa) < 2\kappa \iff n_h \succ n_k$ .

Thus all routing decisions based on  $U^1$  in Case 3 scenarios result in optimal (minimum cost) routes. However, it is easy to see that  $U^1$  does not eliminate the possibility of a sub-optimal route in a source-hotspot-destination configurations corresponding to the scenario in Case 2.

**Eliminating Sub-Optimality in Case 2** As shown by the preceding analysis,  $U^1$  can result in a sub-optimal routing decision in a source-hotspot-destination configuration corresponding to the scenario in Case 2. In particular, any routing decision in a configuration corresponding to Case 2 will result in a sub-optimal path  $P$  if it results in the propagation of a message  $M$  to a node  $n_k \in P$  such that  $x_k < x_h < x_d \wedge y_k = y_h = y_d$  or  $x_k = x_h = x_d \wedge y_k < y_h < y_d$ . Routing decisions based on a preference ordering induced by  $U^1$  can lead to such a situation since in a neighborhood  $H_i$  of  $n_i$  such that  $n_h \notin H_i, \forall n_j, n_k \in H_i, n_k \sim n_j$  provided  $R_k^d = R_j^d$ . Note that Case 2 scenarios include all placements of  $n_s, n_h$ , and  $n_d$ , such that  $\forall \{n_i \mid x_i \neq x_d \wedge y_i \neq y_d\} \exists k, l$ , such that  $(n_k \in \Pi) \vee (n_l \in \Pi)$ .

These observations suggest the possibility of using an estimate of the cost along paths from  $n_k$  to  $n_d$  as a component of a modified utility function  $U^2$  so as to induce a preference ordering between nodes (where no such preference ordering is induced by  $U^1$ ) so as to eliminate suboptimal routing decisions altogether. In other words,  $U^2$  should be able to induce a preference ordering among nodes  $n_k$  and  $n_l$  in the neighborhood of a node  $n_i$  (the node making the routing decision for a message  $M$ ) such that:  $(n_k \in \Pi) \wedge (n_l \notin \Pi) \implies n_k \succ n_l$ . We now proceed to define a cost estimator function  $E_k^d$  as follows:

**Definition 4** A cost estimator function  $E_k^d(\cdot)$  estimates the cost  $E_k^d$  of a minimal cost path to a destination  $n_d$  from a node  $n_k$ .

It would be nice if the cost estimator function defined above helps  $U^2$  to induce the desired preference ordering necessary to guarantee routing along an optimal path in the scenario corresponding to Case 2. We capture this property by defining what are called *admissible* cost estimator functions.

**Definition 5** A cost estimator function is said to be admissible if  $\forall$  nodes  $n_i$  in the network, for all nodes  $n_k, n_l$  in the neighborhood  $H_i$  of  $n_i$ , it is guaranteed that  $(n_k \in \Pi) \wedge (n_l \notin \Pi) \implies E_k^d < E_l^d$

**Definition 6** We define a utility function  $U^2$  as follows:

$$U^2 = \begin{cases} U^1 & \text{if } x_s = x_d \vee y_s = y_d \\ U_j^d = R_j^d - C_j - E_j^d & \text{otherwise} \end{cases}$$

In the discussion that follows, it is assumed that the cost estimator function  $E_k^d$  is admissible.

The estimate returned by  $E_k^d(\cdot)$  must be based, at the very least, on some knowledge of the current cost distribution in the network. More precise estimates would require knowledge of the network dynamics. If costs associated with each node are allowed to change with time, as would be the case in a more realistic routing task, since  $E_k^d$  is computed at the time a message  $M$  is being considered for propagation through  $n_k$ , to a destination  $n_d$ ,  $E_k^d$  has to reflect changes in network load over time. We need to represent at each node, the cost distribution over the network in a form that is independent of specific destination nodes (because the destinations become known only after arrival of the respective messages). Any such representation, in order to be useful in practice in large networks, must not require the storage and update at (or broadcast to) each node, of cost values for all the nodes in the network regions of the network. Ideally, it must adequately summarize the load values in large regions of the network as viewed from a given node.

These considerations (among others) led us to define a *view*,  $V_k$ , which is maintained in every node in the network (Mikler, Wong, & Honavar 1994). In a rectangular grid network, this view consists of four components, one for each of the four directions - north, south, east, and west. Thus we have:  $V_k = [V_k^N, V_k^S, V_k^E, V_k^W]$ .

Each component  $V_k^\delta : (\delta \in \{N, S, E, W\})$  represents a weighted average of costs  $C_i$  along the minimum hop path from  $n_k$  to the border of the grid network in the direction specified by  $\delta$ . Consider two nodes,  $n_i$  and  $n_k$ , located such that  $n_k \in H_i$  and  $n_k$  is to the east of  $n_i$ , i.e.,  $x_i < x_k \wedge y_i = y_k$ . Then  $V_i^E$  is given by:

$$V_i^E = \frac{C_k + V_k^E}{2}$$

$V_i^N, V_i^S$ , and  $V_i^W$  are computed using analogous formulae.

In the discussion that follows, we assume that sufficient time has elapsed for the view computation to stabilize following major load changes in the network before the view is used in the computation of cost estimates using  $E_k^d(\cdot)$ .

In practice, this assumption need not be satisfied exactly so long as the views are adequately precise to ensure the admissibility of the cost estimator function defined below. Assuming that  $n_d$  is located such that  $x_s < x_d \wedge y_s < y_d$ . Let  $D_i^x = |x_i - x_d|$  and  $D_i^y = |y_i - y_d|$  denote the distance from  $n_i$  to  $n_d$  in  $x$  and  $y$

direction, respectively.  $E_i^d(.)$  is given by:

$$E_i^d(.) = \frac{D_i^x V_i^E + D_i^y V_i^S}{D_i^x + D_i^y}$$

It is easy to verify that this estimator (which is one of several alternatives that are possible) is admissible.

**Lemma 3** *For all nodes  $n_i$  in the network, for each message  $M$  from a source  $n_s$  to a destination  $n_d$  that reaches a node  $n_i$ , the routing decision at  $n_i$  based on the preference ordering induced by  $U^2$  will route  $M$  along a path  $P$  selected only from the set of minimum hop paths from  $n_i$  to  $n_d$ , unless  $P\Delta_{i,d} = 1$  and  $(n_h \in P) \wedge (n_h \in H_i)$ .*

The preceding discussion sets the stage for Theorem 2 (proved in Mikler, Honavar, & Wong 1996) that establishes a major property of the utility function  $U^2$ , namely, that it eliminates suboptimal routes in an otherwise uniformly loaded grid network with a single hotspot.

**Theorem 2** *In a uniform cost network with a single hotspot  $n_h$  with an associated cost  $C_h > \kappa$  (where  $\forall i \neq h, C_i = \kappa$ ), a routing algorithm which makes routing decisions at each node based on a preference ordering induced by  $U^2$  is guaranteed to propagate each message  $M$  along a minimum cost path  $\Pi$ .*

## Discussion and Summary

In this paper, we have formulated some simple utility-theoretic heuristic decision functions for guiding messages along a near-minimum-delay path in a large network. We have analyzed some of the interesting properties of such heuristics under a set of simplifying assumptions regarding network topology and load dynamics. For a network with a regular grid topology and certain assumption about load dynamics we have identified the precise conditions under which a simple and computationally efficient utility-theoretic heuristic decision function is guaranteed to route a message along a minimum delay path. This analysis was, at least in part, motivated by a desire to understand and explain the results of a wide range of experiments (Mikler, Honavar, & Wong, 1994) using heuristics that are very similar in spirit to  $U^2$  in more precise mathematical terms.

Given the simplifying assumptions used in our analysis, it is natural to question the applicability of the results when the simplifying assumptions may not hold. It is worth pointing out that experiments with heuristics similar to  $U^2$  display automatic *load balancing* in the network. This suggests that the simplifying assumption of *uniform network load* (except at a hot spot) is useful at least as a crude first approximation of

a more realistic scenario. In the presence of hotspots, the routing functions compensates for this change by redistributing traffic away from the hotspots. This suggests that our analytical results are likely to be useful to guide the design utility-theoretic heuristics for a more complex network. Work in progress is aimed at extending our analysis to a range of increasingly complex scenarios such as: irregular grids; non-uniform load distributions; multiple hotspots or contiguous hotspot regions.

The performance of utility-theoretic heuristics, as described in this paper, critically depends on the existence of an adequately precise *estimator* of the relevant performance measure. It would be useful to analyze different estimators and the resulting heuristics - especially since the design of good heuristics for complex problems is commonly based on solution of *simplified* or *relaxed versions* of the original problem (Pearl 1984). Other interesting research directions include the investigation of methods for adaptation or tuning of heuristics in real-time. For this we may draw upon machine learning techniques that modify existing heuristics as a function of measured network behavior or as a function of information gathered through directed experiments initiated by the network during otherwise idle periods.

## References

- Bertsekas, D. and Gallager, R. 1992. *Data Networks*. Englewood Cliffs, NJ.: Prentice-Hall
- French, S. 1986. *Decision Theory: an introduction to the mathematics of rationality*. New York: Halsted Press.
- Mikler, A.R., Wong, J.S.K., and Honavar, V.G. 1994. Quo Vadis - A Framework for Intelligent Routing in Large Communication Networks. Technical Report TR94-24, Dept. of Computer Science, Iowa State University. To appear in: *The Journal of Systems and Software*.
- Mikler, A.R., Honavar, V.G. and Wong, J.S.K. 1996. Utility-Theoretic Heuristics for Intelligent Adaptive Routing in Large Communication Networks. Technical Report TR95-14, Dept. of Computer Science, Iowa State University.
- Pearl, J. 1984. *Heuristics: Intelligent Search Strategies for Computer Problem Solving*. Reading, MA.: Addison-Wesley.