A General Optimization Framework for Stochastic Routing in Wireless Multi-hop Networks

Alejandro Ribeiro[†], Zhi Quan Luo[†], Nikos D. Sidiropoulos[‡], and Georgios. B. Giannakis[†]

Abstract— **We introduce a novel approach to routing based on the so called pairwise packet delivery ratio matrix whose entries represent the probability that a given user decodes the packet transmitted by any other user. We show that this leads naturally to a model in which routing algorithms are described by the evolution of a Markov chain enabling the definition of deliverability criteria in terms of absorbing states. We further introduce optimal routing protocols by selecting the routing matrix from a convex polygon containing all feasible routing matrices. The criteria of optimality include minimization of the packet error probability for a given delay bound and the minimization of the average packet delay. These metrics are correspondingly meaningful in the context of real time transmissions – e.g., voice and/or video – and delay insensitive data – e.g., file transfers.**

Keywords: Routing, Wireless Networks, Markov chains, Optimization

I. Introduction

The rapid decay in the received signal envelope with distance is a unique and challenging feature of wireless networks. Multi-hopping, entailing the division of a longer link into shorter links leading to the destination is the traditional countermeasure. By reducing the average distance between communicating pairs of nodes, multi-hop routing secures, at the very least, significant power savings, when not the feasibility of the communication link itself, providing sufficient motivation for the study of multi-hop routing algorithms for wireless networks.

Multi-hop routing has been widely studied in the context of adhoc networks [5]. These infrastructureless networks rely on peer-to-peer communications. Since the power needed for direct communication from source to destination is usually prohibitively large, multi-hop routing is a must in this context. Although not as widely studied, alternative topologies such as those encountered with collaborative multiple access channels can also benefit from multi-hop routing. Routing algorithms for ad-hoc networks have evolved from the accumulated knowledge about wireline networks. The usual steps are to: i) define a communication radius for each node; ii) draw the corresponding connectivity graph; and iii) invoke shortest path routing to find the optimal route. Most of the differences in multi-hop routing protocols arise in the definition of the associated link metrics. These include path reliability, transmitted power, and mutual interference to name a few [2], [4], [6]–[8].

In this paper, we introduce a general framework for multi-hop routing in wireless networks. Our framework is based on a delivery-ratio (or pairwise packet-error-probability) matrix $\mathbf{R}$ whose $(i, j)th$ entry R_{ij} represents the probability that a packet transmitted by the j^{th} user U_j is correctly received by the i^{th} user U_i. A graph model, can be considered a special case of $\mathbf{R}$ in which the entries $R_{i,j}$ are either 0 or 1. Besides subsuming graph-theoretic models, the delivery ratio matrix offers a more suitable model for the shared, broadcast, and unreliable wireless channel.

The fact that our model offers a more accurate description does not necessarily imply that it is better. The main contribution of the present paper is to show that indeed our stochastic routing protocols (SRP) are advantageous in many circumstances. In particular we will show that

- **(i) Performance improvement.** It has been observed that when modeling the wireless network as a graph the resultant routing matrices tend to use unreliable routes [4] with a consequent reduction on the benefits of optimal routing. Routing matrices obtained by SRP algorithms are based on the measured link reliability (Section II).
- **(ii) Tractability of diverse approaches.** Optimization problems on a graph usually turn out to have combinatorial complexity effectively limiting routing algorithms to formulations that can be solved using variations of shortest path routing. Many optimization problems involving a matrix, however, can be solved in polynomial time using convex optimization techniques [3]. The latter will turn out to be the case with some of the SRP algorithms introduced in this paper (Section II-A).
- **(iii) Relation with graph approaches.** It is possible to stick to graph-network models while taking $\mathbf{R}$ into account; e.g., by defining link metrics as $1/R_{ij}$ [4]. Interestingly, these algorithms appear naturally in our SRP formulation when the optimality criterion is to minimize expected delay (Section III).

In Section IV we present simulations corroborating our analytical claims. We conclude the paper in Section V

II. Stochastic routing protocols (SRP)

Consider a wireless network with $J + 1$ user nodes $\{U_j\}_{j=1}^{J+1}$ in which the first J users $\{U_j\}_{j=1}^{J}$ collaborate to route packets to the destination $D \equiv U_{J+1}$. The physical and multiple access layers are such that if a packet is transmitted by U_j it is correctly *received* by U_i with probability R_{ij} that we arrange in the matrix $\mathbf{R}$. We consider a *per-session* model of routing in which a user node establishing a session is confronted with the routing decisions of its peers that determine the entries R_{ij} of $\mathbf{R}$. Supposing that the probabilities in $\mathbf{R}$ remain invariant over the duration of a session,

* Work in this paper was prepared through collaborative participation in the Communications and Networks Consortium sponsored by the U. S. Army Research Laboratory under the Collaborative Technology Alliance Program, Cooperative Agreement DAAD19-01-2-0011. The U. S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation thereon. The work of N. D. Sidiropoulos was partially supported by a bilateral cooperative research grant of the Greek Secretariat for Research and Technology.

† The authors are with the Dept. of Electrical and Computer Engineering, University of Minnesota, 200 Union Street SE, Minneapolis, MN 55455. Tel/fax: (612)626-7781/625-4583. Email: {aribeiro,luozq,georgios}@ece.umn.edu.

‡ N.D. Sidiropoulos is with the Dept. of Electrical and Computer Engineering, Technical University of Crete, Chania - Crete, Greece 731 00. Tel/fax: +30-28210-37227/+30-8210-37542. Email: nikos@telecom.tuc.gr.

1-4244-0785-0/06/$20.00

our goal is to find a stochastic routing strategy for the session that is optimal in a suitable sense.

Let $e_j(n)$ indicate the binary (0/1) event that the packet is at U_j at time n whose probability we denote by $f_j(n) := \Pr\{e_j(n) = 1\}$. Correspondingly, we define the vectors $\mathbf{e}(n) := [e_1(n), \ldots, e_{J+1}(n)]^T$ and $\mathbf{f}(n) := [f_1(n), \ldots, f_{J+1}(n)]^T$. If the packet is generated at a known source U_s for some $s \in [1, J]$ we have that $f_s(0) = 1$. In general, the packets are generated at a random source with initial distribution $\mathbf{f}(0)$. Regardless of the initial condition what we want from the routing algorithm is for the packet to be eventually delivered to the destination [$(J+1)$-st user]; i.e.,

$$\lim_{n\to\infty} \mathbf{f}(n) = \mathbf{c}_{J+1}, \tag{1}$$

where $\mathbf{c}_{J+1} := [0, \ldots, 0, 1]$ is the $(J+1)$-st vector in the canonical basis of $\mathbb{R}^{J+1}$. Since it is meaningful to focus on routing algorithms which – at least – satisfy (1), we introduce the following definition.

Definition 1 *A routing algorithm ensures deliverability if and only if* (1) *holds for any initial distribution* $\mathbf{f}(0)$.

A routing algorithm is defined by the matrix $\mathbf{T}$ whose ij^{th} entry T_{ij} is the probability that U_j decides to *transmit* (i.e., route) the packet to U_i. Likewise, we define $\mathbf{K}$ with $K_{ij} := \Pr\{e_i(n+1)|e_j(n)\}$ denoting the probability that the packet moves from U_j to U_i between times n and $n+1$. Note that $\mathbf{T}$ and $\mathbf{K}$ are related through $\mathbf{R}$. Indeed, for $i \neq j$ the packet moves from U_j to U_i if and only if it is routed through U_i *and* is correctly decoded; since these two events are independent we have,

$$K_{ij} = T_{ij} R_{ij} \quad \text{for } i \neq j. \tag{2}$$

Since $\mathbf{K}$ and $\mathbf{T}$ are stochastic matrices, columns must sum up to 1 implying that $\mathbf{K}^T\mathbf{1} = \mathbf{1}$ and $\mathbf{T}^T\mathbf{1} = \mathbf{1}$. These two constraints and (2) imply that $\mathbf{K}$ is completely determined by $\mathbf{T}$ (but not viceversa).

Since the $(J+1)$-st user is the destination it will not route the packet, from which we infer $T_{i(J+1)} = 0$, $\forall i \in [1, J]$; and, after taking (2) into account, we arrive at $K_{i(J+1)} = 0$, $\forall i \in [1, J]$. Arguing similarly, it follows that $R_{(J+1)(J+1)} = T_{(J+1)(J+1)} = K_{(J+1)(J+1)} = 1$. Summing up, with properly defined $\mathbf{v}_D \in \mathbb{R}^J$ and $\mathbf{K}_D \in \mathbb{R}^{J\times J}$ we can write

$$\mathbf{K} = \begin{pmatrix} \mathbf{K}_D & \mathbf{0} \\ \mathbf{v}_D^T & 1 \end{pmatrix}. \tag{3}$$

It follows easily by direct substitution that (3) holds if and only if $\mathbf{K}\mathbf{c}_{J+1} = \mathbf{c}_{J+1}$, i.e., if and only if $\mathbf{c}_{J+1}$ is an eigenvector of $\mathbf{K}$ associated with the eigenvalue 1.

For future reference, we define the set of stochastic matrices in $\mathbb{R}^{(J+1)^2}$ as

$$\mathcal{P} = \{\mathbf{T} \in \mathbb{R}^{(J+1)^2} \ : \ \mathbf{T}^T\mathbf{1} = \mathbf{1}, T_{ij} \geq 0, \forall i, j\}. \tag{4}$$

The constraints on $\mathbf{K}$ can be written as $\mathbf{K} \in \mathcal{K}$ with

$$\mathcal{K} = \{\mathbf{K} \in \mathcal{P} : K_{ij} = T_{ij}R_{ij}, \text{for } i \neq j, \mathbf{T} \in \mathcal{P}; \mathbf{K}\mathbf{c}_{J+1} = \mathbf{c}_{J+1}\}. \tag{5}$$

Note that the set $\mathcal{K}$ is a convex polyhedron in $\mathbb{R}^{(J+1)^2}$.

We can characterize the evolution of $\mathbf{f}(n)$ in terms of $\mathbf{K}$. Indeed, note that due to the law of total probability $f_i(n) = \sum_{j=1}^n \Pr\{e_i(n)|e_j(n-1)\}f_j(n-1) = \sum_{j=1}^n K_{ij}f_j(n-1)$, that we can write in vector-matrix form as,

$$\mathbf{f}(n) = \mathbf{K}\mathbf{f}(n-1) = \mathbf{K}^n\mathbf{f}(0). \tag{6}$$

That is, $\mathbf{f}(n)$ represents the probability evolution of a Markov chain characterized by $\mathbf{K}$ in which the j^{th} state represents the presence of the packet at user U_j. Building on (6), we can find conditions to ensure deliverability of a SRP we describe in the following theorem.

Theorem 1 *The following statements are equivalent:*

(i) *The routing algorithm defined by the matrix* $\mathbf{T}$ *ensures deliverability.*

(ii) *Matrix* $\mathbf{K}$ *describes the probability evolution of an absorbing Markov chain whose unique absorbing state is* $J+1$.

(iii) *The spectral radius of* $\mathbf{K}_D$ *is strictly smaller than one; i.e., with* $\mathrm{eig}(\mathbf{K}_D)$ *denoting the set of eigenvalues of* $\mathbf{K}_D$ *we have* $\rho(\mathbf{K}_D) := \max|\mathrm{eig}(\mathbf{K}_D)| < 1$.

(iv) *The matrix* $\mathbf{K}_D$ *and the vector* $\mathbf{v}_D$ *in* (3) *satisfy* $\mathbf{v}_D^T(\mathbf{I} - \mathbf{K}_D)^{-1} = \mathbf{1}^T$.

Proof: Using induction we can easily show that the n^{th} power of $\mathbf{K}$ can be written as [cf. (3)]

$$\mathbf{K}^n = \begin{pmatrix} \mathbf{K}_D^n & \mathbf{0} \\ \mathbf{v}_D^T \sum_{k=0}^{n-1} \mathbf{K}_D^n & 1 \end{pmatrix}. \tag{7}$$

Upon defining $\mathbf{f}_D(n) := [f_1(n), \ldots, f_J(n)]^T$ we combine (6) and (7) to obtain

$$\mathbf{f}_D(n) = \mathbf{K}_D^n \mathbf{f}_D(0). \tag{8}$$

On the other hand, also note that (1) is true if and only if $\lim_{n\to\infty} \mathbf{f}_D(n) = \mathbf{0}$.

To go from (i) to (ii) note that since for any $\mathbf{K} \in \mathcal{K}$, $\mathbf{K}\mathbf{c}_{J+1} = \mathbf{c}_{J+1}$, $J+1$ is by definition an absorbing state of the Markov chain defined by $\mathbf{K}$. If $j \neq J+1$ is another absorbing state then $\mathbf{K}\mathbf{c}_j = \mathbf{c}_j$ and for $\mathbf{f}(0) = \mathbf{c}_j$ we have that $\mathbf{K}^n\mathbf{f}(0) = \mathbf{c}_j$ for every n; thus $\lim_{n\to\infty} \mathbf{f}(n) = \mathbf{c}_j \neq \mathbf{c}_{J+1}$. This is a contradiction if $\mathbf{T}$ ensures deliverability and consequently $J+1$ is the unique absorbing state.

If (iii) is not true, then $\lim_{n\to\infty} \mathbf{K}_D^n \neq 0$. Hence, there exists a vector $\mathbf{f}(0)$ for which $\lim_{n\to\infty} \mathbf{K}_D^n\mathbf{f}(0) \neq \mathbf{0}$ implying that $J+1$ is not an absorbing state. Thus, (ii) implies (iii).

That (iii) implies (iv) follows after noting that since $\mathbf{1}^T\mathbf{K} = \mathbf{1}^T$, we have that $\mathbf{1}^T\mathbf{K}^n = \mathbf{1}^T$ and asymptotically $\lim_{n\to\infty} \mathbf{1}^T\mathbf{K}^n = \mathbf{1}^T$. But since (iii) also implies that $\lim_{n\to\infty} \mathbf{K}_D^n = 0$, we must have $\lim_{n\to\infty} \mathbf{v}_D^T \sum_{k=0}^{n-1} \mathbf{K}_D^n = \mathbf{1}^T$. To obtain (iv), note that the geometric series is such that $\sum_{k=0}^{\infty} \mathbf{K}_D^n = (\mathbf{I} - \mathbf{K}_D)^{-1}$.

Finally, if (iv) is true then $\lim_{n\to\infty} \mathbf{K}^n = [\mathbf{0}, \ldots, \mathbf{0}, \mathbf{1}]^T$ implying that (i) is true. ■

Theorem 1 gives necessary and sufficient conditions for an SRP to have guaranteed deliverability. None of these conditions is difficult to achieve and, in general, simple routing algorithms; e.g. a random walk through the network with $T_{ij} = 1/J$, will do just fine. A more interesting problem is how to obtain a matrix which guarantees that the limit in (1) is practically achieved with n as small as possible. This motivates different routing algorithms that we can obtain from (6) and analyze next.

A. *Fastest convergence rate routing*

The rate of convergence can be either measured on average or for the worst possible initial distribution entailing different criteria for optimal routing. Optimal routing on an average sense is considered in Section III. What we expect from an optimal routing matrix $\mathbf{T}$ is for the convergence rate in (1) to be as fast as possible. The distance – in some sense – between $\mathbf{f}(n)$ and $\mathbf{c}_{J+1}$ can be measured by the norm $\|\mathbf{f}(n) - \mathbf{c}_{J+1}\|_p$ which is to be compared with the original distance $\|\mathbf{f}(0) - \mathbf{c}_{J+1}\|_p$ leading to the following expression for the convergence rate:

$$\xi_p = \sup_{\mathbf{f}(0)\neq\mathbf{c}_{J+1}} \lim_{n\to\infty} \left(\frac{\|\mathbf{f}(n) - \mathbf{c}_{J+1}\|_p}{\|\mathbf{f}(0) - \mathbf{c}_{J+1}\|_p} \right)^{1/n}. \tag{9}$$

This cannot be computed in closed-form for arbitrary p. For $p = 2$, corresponding to the Euclidean norm, the argument in (9) is maximized by the eigenvector associated with the second largest eigenvalue of $\mathbf{K}$. A meaningful routing algorithm is thus to look for the matrix $\mathbf{K} \in \mathcal{K}$ such that

$$\min_{\mathbf{K}\in\mathcal{K}} |\mathrm{eig}_2(\mathbf{K})| = \min_{\mathbf{K}\in\mathcal{K}} \max |\mathrm{eig}(\mathbf{K}_D)| = \min_{\mathbf{K}\in\mathcal{K}} \rho(\mathbf{K}_D), \quad (10)$$

where $\mathrm{eig}_2(\mathbf{K})$ denotes the second largest eigenvalue of $\mathbf{K}$ and $\mathrm{eig}(\mathbf{K}_D)$ the set of eigenvalues of $\mathbf{K}_D$. In establishing the first equality in (10) we used that all the eigenvalues of $\mathbf{K}_D$ are eigenvalues of $\mathbf{K}$ [cf. (3)]; in fact, $\mathrm{eig}(\mathbf{K}) = \mathrm{eig}(\mathbf{K}_D) \cup \{1\}$. The second equality follows from the definition of spectral radius.

Unfortunately, minimizing the spectral radius of a non-symmetric matrix is a notoriously difficult problem, intractable except for small-medium values of J [3]. This motivates an alternative measure of convergence rate based on the vector $\mathbf{f}_D(n) := [f_1(n), \ldots, f_J(n)]^T$ containing the probabilities that the packet is at a certain node other than the destination. The norm of $\mathbf{f}_D(n)$ measures the probability of the packet *not* being delivered at time n. This suggests the metric

$$\zeta_p = \max_{\mathbf{f}_D(n)} \frac{\|\mathbf{f}_D(n+1)\|_p}{\|\mathbf{f}_D(n)\|_p}, \quad (11)$$

which amounts to the worst-case one-step relative reduction of the probability vector $\mathbf{f}_D(n)$ which we want convergent to zero [c.f. (1)]. Similarly to ξ_p, we can define optimal routing in terms of ζ_p. If we further recall that $\mathbf{f}_D(n+1) = \mathbf{K}_D\mathbf{f}_D(n)$ another class of optimal SRPs can be designed to achieve

$$\min_{\mathbf{K}\in\mathcal{K}} \max_{\mathbf{f}_D(n)} \frac{\|\mathbf{K}_D\mathbf{f}_D(n)\|_p}{\|\mathbf{f}_D(n)\|_p} = \min_{\mathbf{K}\in\mathcal{K}} \|\mathbf{K}_D\|_p, \quad (12)$$

where the equality follows from the definition of the p-norm of a matrix. Different from (10), the optimization of (12) is a convex problem for all p since: i) due to the triangle inequality, norms are convex functions of their arguments; and ii) the set $\mathcal{K}$ is a convex polyhedron [c.f. (5)]. For the usual norms, $p = 1, 2, \infty$, solving (12) is either a simple linear program (LP) for $p = 1, \infty$, or a semi-definite program (SDP) for $p = 2$ [3].

In general (10) and (12) are optimized by different matrices $\mathbf{T}$, the pertinent comparisons are discussed in the following remark.

Remark 1 Entailing convex optimization problems – indeed, canonical optimization problems – (12) is tractable for networks with a large number of users J; whereas (10) is only tractable for small-to-medium scale networks. On the other hand, (10) is more meaningful than (12), since the former compares the asymptotic behavior with the initial state while the latter compares two consecutive states. In practice, (12) can be viewed as a tractable approximation to (10).

III. Minimum expected delay routing

An alternative approach to optimal routing is to consider the packet delivery time measured in number of hops and look for the matrix $\mathbf{T}$ that minimizes the average packet delay. Packet delay is simply the time n at which the packet is received by $D \equiv U_{J+1}$:

$$\delta = \min\{n \; : \; e_{J+1}(n) = 1\} = \sum_{n=0}^{\infty} [1 - e_{J+1}(n)] \quad (13)$$

where the second equality is true since $1 - e_{J+1}(n) = 1$ if $n < \delta$ and $1 - e_{J+1}(n) = 0$ for $n \geq \delta$; we thus have δ terms equal to 1 in the summation in (13). Starting from (13), the expected delay can be computed as we show in the following theorem.

Algorithm 1 Min. expected delay routing (Dijkstra version)

Require: The packet success probability matrix $\mathbf{R}$
Ensure: The routing matrix $\mathbf{T}$
$\bar{\delta}_j = 1/R_{(J+1)j}$, for $j \in [1, J]$
$\mathcal{U} = \{U_j\}_{j=1}^J$
while $\mathcal{U} \neq \emptyset$ **do**
  $j^* = \arg\min_{j:U_j\in\mathcal{U}} \bar{\delta}_j$
  $\mathcal{U} = \mathcal{U} - \{U_{j^*}\}$
  for all $i : U_i \in \mathcal{U}$ **do**
    if $1/R_{ij^*} + \bar{\delta}_j^* < \bar{\delta}_i$ **then**
      $\bar{\delta}_i = 1/R_{ij^*} + \bar{\delta}_j^*$,
      $T_{ij^*} = 1$; $T_{ij} = 0$ for $j \neq j^*$
    end if
  end for
end while

Theorem 2 *For a routing algorithm ensuring deliverability, the expected delay is given by*

$$\bar{\delta} := \mathrm{E}(\delta) = \mathbf{1}^T(\mathbf{I} - \mathbf{K}_D)^{-1}\mathbf{f}_D(0), \quad (14)$$

where $\mathbf{f}_D(0) = [f_1(0), \ldots, f_J(0)]$ *is the initial distribution for the first* J *users.*

Proof: Taking expected value in (13), using the linearity of the expected value operator and noting that $f_D(n) = \mathrm{E}[e_D(n)]$, we obtain

$$\bar{\delta} = \sum_{n=1}^{\infty}[1 - f_{J+1}(n)] = \sum_{n=1}^{\infty}\sum_{j=1}^{J} f_j(n), \quad (15)$$

where in establishing the second equality we used that $\sum_{j=1}^{J+1} f_j(n) = 1$ and hence $1 - f_{J+1}(n) = \sum_{j=1}^{J} f_j(n)$. But since the innermost summation can be written as $\mathbf{1}^T\mathbf{f}_D(n)$, if we also recall that $\mathbf{f}_D(n) = \mathbf{K}_D^n\mathbf{f}(0)$, we obtain

$$\bar{\delta} = \sum_{n=1}^{\infty} \mathbf{1}^T\mathbf{f}_D(n) = \sum_{n=1}^{\infty} \mathbf{1}^T\mathbf{K}_D^n\mathbf{f}(0) = \mathbf{1}^T\left(\sum_{n=1}^{\infty}\mathbf{K}_D^n\right)\mathbf{f}(0). \quad (16)$$

For routing matrices that ensure deliverability the matrix geometric series in (16) is convergent with $\sum_{n=0}^{\infty}\mathbf{K}_D^n = (\mathbf{I} - \mathbf{K}_D)^{-1}$. Substituting this into (16), (14) follows readily. ■

The expected delay $\bar{\delta}$ is a function of the routing matrix $\mathbf{K}$ *and* the initial distribution $\mathbf{f}(0)$. Using the result in Theorem 2 we can find the matrix that minimizes the expected delay as the argument solving the optimization problem

$$\mathbf{K}_D^*[\mathbf{f}_D(0)] = \arg\min_{\mathbf{K}\in\mathcal{K}} \bar{\delta} = \arg\min_{\mathbf{K}\in\mathcal{K}} \mathbf{1}^T(\mathbf{I} - \mathbf{K}_D)^{-1}\mathbf{f}_D(0) \quad (17)$$

A direct attempt at solving (17) is doomed to failure. Luckily, it turns out that (17) is equivalent to a shortest path routing algorithm as we show in the ensuing theorem.

Theorem 3 *Define the expected delay vector* $\bar{\boldsymbol{\delta}} = [\bar{\delta}_1, \ldots, \bar{\delta}_J] := \mathbf{1}^T(\mathbf{I} - \mathbf{K}_D)^{-1}$ *in which* $\bar{\delta}_j$ *is the expected delay when the packet starts at* U_j*; i.e., when* $\mathbf{f}(0) = \mathbf{c}_j$*; and let* $\bar{\delta}_{J+1} = 0$*. If there exists a matrix* $\mathbf{K}_D$ *ensuring deliverability, the matrix* $\mathbf{K}_D^\dagger \in \mathcal{K}$ *such that*

$$\bar{\delta}_j = \min_i \left\{\frac{1}{R_{ij}} + \bar{\delta}_i\right\}, \quad \bar{\delta}_{J+1} = 0, \quad (18)$$

minimizes the expected delay for any initial distribution $\mathbf{f}(0)$*; i.e.* $\mathbf{K}_D^*[\mathbf{f}_D(0)] = \mathbf{K}_D^\dagger$ *for any* $\mathbf{f}_D(0)$ *and* $\mathbf{K}_D^*[\mathbf{f}_D(0)]$ *as in* (17).

Proof: See appendix. ■

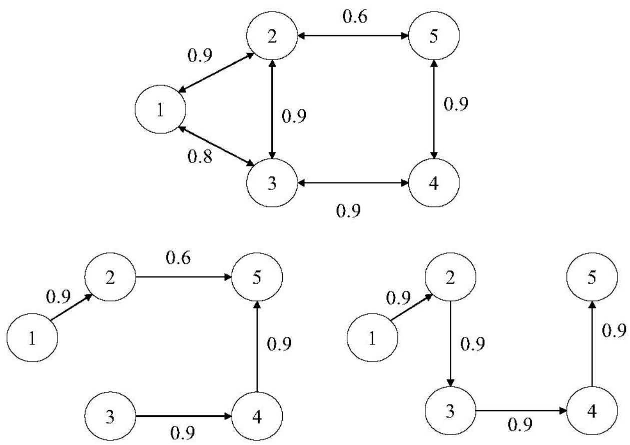

Fig. 1. For a simple connectivity graph (top) the minimum expected delay routing algorithm in (17) tends to select short routes (left); fastest convergence rate routing as per (10) selects longer routes with more reliable hops.

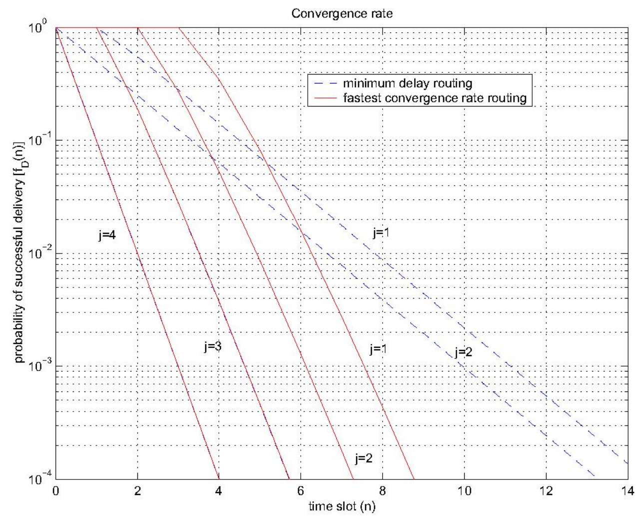

Fig. 2. Convergence rate for the network in Fig. 1. For a fixed time delay fastest convergence rate routing yields a smaller packet error probability.

Characterizing the solution as in (18) indicates that $\mathbf{K}_D^*$ can be found as the shortest path route (SPR) in a fully connected graph with the arc between U_i and U_j having weight $1/R_{ij}$. Indeed, let $\mathbf{i} := (i_1, \ldots, i_k)$ with $k \in [2, J+1]$, $i_1 = j$ and $i_k = J+1$ be an arbitrary sequence starting at U_j and finishing at U_{J+1}. Proceeding recursively, we find that (18) is equivalent to

$$\bar{\delta}_j = \min_{\mathbf{i}} \left\{ \sum_{l=1}^{\#(\mathbf{i})-1} \frac{1}{R_{i_l i_{l+1}}} \right\}, \qquad (19)$$

where $\#(\mathbf{i})$ is the cardinality of $\mathbf{i}$. By definition, (19) is the SPR between j and $J+1$ among all the possible routes $\mathbf{i}$. In fact, the relation in (18) is Bellman's principle of optimality, which we know holds true for the shortest path route [1, Chap.5]. This implies that the solution to minimum expected delay routing can be found in $O(J^2)$ steps using dynamic programming algorithms, e.g. Bellman-Ford, Dijkstra, or Floyd-Warhall [1, Chap.5].

Also important, and contrary to what (17) suggested, minimum expected delay routing does not depend on the initial distribution. The average delays $\bar{\delta}[\mathbf{f}(0)]$ for different initial distributions $\mathbf{f}(0)$ are different, but there exists a matrix that minimizes $\bar{\delta}[\mathbf{f}(0)]$ for *all* $\mathbf{f}(0)$. Such matrix is the solution of the problem

$$\mathbf{K}^* = \arg\min_{K \in \mathcal{K}} \mathbf{1}^T (\mathbf{I} - \mathbf{K}_D)^{-1} \mathbf{1} \qquad (20)$$

that can be obtained using Algorithm 1. Note that for a given $\mathbf{f}(0)$ there might be other solutions to (17), but none will outperform $\mathbf{K}^*$ in (20).

IV. Simulations

The fastest convergence rate SR algorithm in (10) maximizes the packet delivery probability for a given, sufficiently large, time index n. On the other hand, minimum expected delay routing as per (17) minimizes the expected time elapsed until packet delivery. The subtle differences between these two approaches are exemplified in Figs. 1 and 2.

The resulting routing matrices for minimum expected delay and fastest convergence rate routing are shown in Fig. 1. We can see that the former algorithm tends to select short routes sometimes containing unreliable hops (left) as exemplified by the link $U_2 \rightarrow U_5$ used to route U_1 and U_2's packets. Whereas, the latter uses longer routes but tends to use more reliable hops (right), as we can see from the use of the $U_2 \rightarrow U_3$ link to route U_1 and U_2's traffic. This is a manifestation of the different optimization criteria. The expected delay for routing U_2's packets is 1.67 for minimum expected delay routing and 3.33 for fastest convergence rate routing. The difference in convergence rate is shown in Fig. 2. To achieve a packet error probability of 10^{-4}, U_2's delay is 7.2 for fastest convergence rate routing and 13.1 for minimum expected delay routing.

Similar conclusions are reached for the more realistic example in Fig. 3 representing a randomly generated network with 20 nodes. In this figure we depict the connectivity graph as well as the result of the minimum expected delay, fastest convergence rate, and minimum 2-norm SRP obtained from (12) with $p = 2$. In this case it is also true that minimum expected delay prefers shorter routes, while fastest convergence rate prefers longer routes containing more reliable hops. Minimum 2-norm routing is the only algorithm considered that yields routing matrices implying non-deterministic routing; i.e. having $T_{ij} \neq 1, 0$ for some i, j.

For real time delay-sensitive applications; e.g. audio and/or video conferencing, fastest convergence routing is a better alternative. This is corroborated by Fig. 4 (top) showing the convergence rate for the network in Fig. 3. For a delay of 14 hops, fastest convergence rate routing yields a packet error probability of 10^{-4} for the least favored user; for the same delay, minimum expected delay routing achieves a packet error probability of 10^{-2}. For delay-tolerant applications; e.g. file transfers, the average delay metric is better suited since to deliver a large number of packets the total number of required hops is significantly smaller. This is illustrated in Fig. 4 (bottom) where we see that for minimum expected delay routing most packets are delivered in a few hops and a few packets take a long time to be delivered. For fastest convergence rate routing none of the packets took more than 8 hops to be delivered but the total number of hops required to deliver all the packets was larger.

V. Conclusions

We introduced a general framework for stochastic routing in wireless multi-hop networks. Deviating from the traditional graph models we considered an approach based on the packet delivery probability matrix and showed that different routing algorithms can be described by the evolution of a properly defined Markov

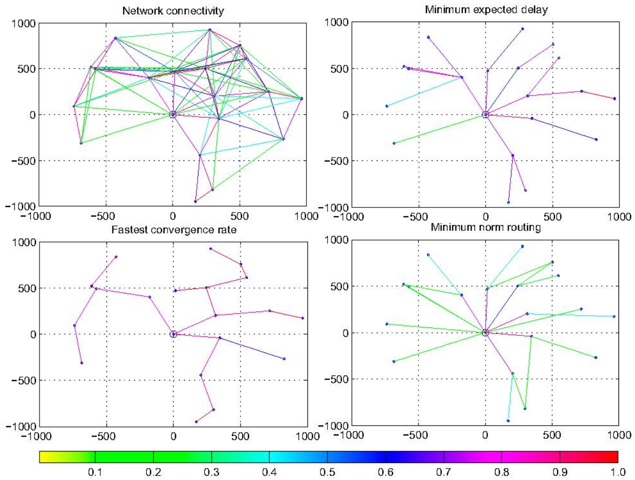

Fig. 3. A randomly generated network with 20 nodes, the color scale represents the elements of the matrix $\mathbf{K}$. Note how fastest convergence rate routing selects routes with large values of K_{ij}.

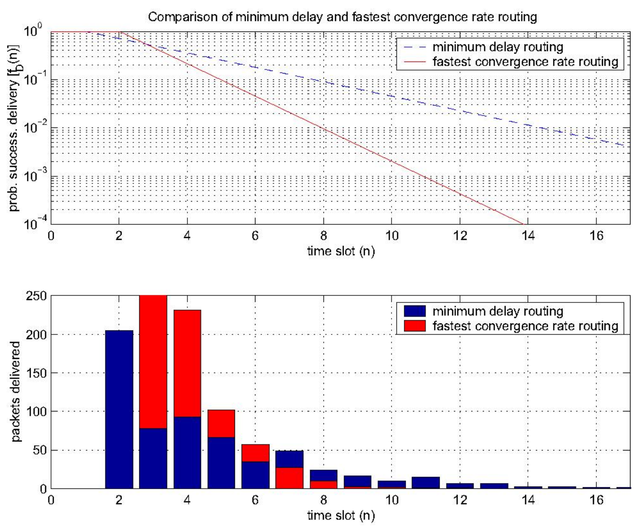

Fig. 4. Convergence rate of the least favored user for the network in Fig. 3 (top) and histogram of packet delivery times for a randomly chosen user (bottom). Fastest convergence rate routing is favored for time sensitive traffic.

chain. This connection permits characterization of a properly defined deliverability condition in terms of absorbing states of the Markov chain and the eigenvalues of the corresponding probability evolution matrix.

We then moved on to introduce stochastic routing algorithms that maximize the convergence rate of the Markov chain, entailing a maximization of the packet delivery probability for a fixed, sufficiently large delay n. This routing approach is meaningful in the context of delay sensitive traffic involved in, e.g voice and/or video conferencing. We further found an expression for the average packet delay measured by the number of hops and defined the corresponding optimal routing problem that minimizes it. Interestingly, we proved that the optimum routing matrix can be obtained as the shortest path route in a fully connected graph with the arc between users having a weight inversely proportional to the corresponding delivery ratio.

APPENDIX

A. Proof of Theorem 3

Given $\bar{\delta}_i$ for $i \neq j$ we solve (14) for $\bar{\delta}_j$ obtaining,

$$\bar{\delta}_j = \frac{1 + \sum_{i=1, i\neq j}^{J+1} K_{ij}\bar{\delta}_i}{1 - K_{jj}}. \tag{21}$$

Since $\sum_{i=1}^{J+1} K_{ij} = 1$ we have that $1 - K_{jj} = \sum_{i=1, i\neq j}^{J+1} K_{ij}$; if we also replace $K_{ij} = T_{ij}R_{ij}$ valid for $i \neq j$ we obtain

$$\bar{\delta}_j = \frac{1 + \sum_{i=1, i\neq j}^{J+1} T_{ij}R_{ij}\bar{\delta}_i}{\sum_{i=1, i\neq j}^{J+1} T_{ij}R_{ij}}. \tag{22}$$

Now, replace the 1 in the numerator by $\sum_{i=1, i\neq j}^{J+1} T_{ij} = 1$ and rearrange terms to arrive at

$$\bar{\delta}_j = \frac{\sum_{i=1, i\neq j}^{J+1} \left(1/R_{ij} + \bar{\delta}_i\right) T_{ij}R_{ij}}{\sum_{i=1, i\neq j}^{J+1} T_{ij}R_{ij}}. \tag{23}$$

It also follows by definition that $(1/R_{ij}) + \bar{\delta}_i \geq \min_i \left(1/R_{ij} + \bar{\delta}_i\right)$ which allows us to bound $\bar{\delta}_j$ in (25) by

$$\bar{\delta}_j \geq \min_i \left(\frac{1}{R_{ij}} + \bar{\delta}_i\right) \frac{\sum_{i=1, i\neq j}^{J+1} T_{ij}R_{ij}}{\sum_{i=1, i\neq j}^{J+1} T_{ij}R_{ij}} = \min_i \left(\frac{1}{R_{ij}} + \bar{\delta}_i\right). \tag{24}$$

The matrix satisfying (18) for all j achieves the lower bound in (24) and thus minimizes $\bar{\delta}_j$ for all j. This proves that if a matrix satisfies (18) it minimizes $\bar{\delta}_j$ for all j. That such a matrix exists follows from the construction in Algorithm 1 that yields a matrix satisfying (18) as long as ensuring deliverability is possible.

For an arbitrary initial distribution we have that

$$\bar{\delta}[\mathbf{f}(0)] = \sum_{j=1}^{J} \Pr\{e_j(0)\}\bar{\delta}_j = \mathbf{f}^T(0)\bar{\boldsymbol{\delta}}. \tag{25}$$

But since all components $f_j(0)$ of $\mathbf{f}(0)$ are non-negative $\bar{\delta}[\mathbf{f}(0)]$ is minimized if all components of $\bar{\boldsymbol{\delta}}$ are minimum. The latter is true if (18) is valid for all j [c.f. (24)] completing the proof. ∎

REFERENCES

[1] D. Bertsekas and R. Gallager, *Data Networks*. Prentice Hall, second ed., 1992.

[2] P. Bose, P. Morin, I. Stojmenovic, and J. Urrutia, "Routing with guaranteed delivery in ad hoc wireless networks," *ACM Wireless Networks*, vol. 7, pp. 609–616, Nov. 2001.

[3] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge University Press, 2004.

[4] D. De Couto, D. Aguayo, J. Bicket, and R. Morris, "A high-throughput path metric for multi-hop wireless routing," in *Proc. of the 9th International ACM Conference on Mobile Computing and Networking*, pp. 134–146, San Diego, CA, Sept. 14-19, 2003.

[5] P. Gupta and P. Kumar, "The capacity of wireless networks," *IEEE Transactions on Information Theory*, vol. 46, pp. 388–404, March 2002.

[6] J. Kuruvila, A. Nayak, and I. Stojmenovic, "Hop count optimal position based packet routing algorithms for ad hoc wireless networks with a realistic physical layer," *IEEE Journal on Selected Areas in Communications*, vol. 23, pp. 1267–1275, June 2005.

[7] T. Nadeem and A. Agrawala, "IEEE 802.11 fragmentation-aware energy-efficient ad-hoc routing protocols," in *Proc. of the 1st IEEE International Conference on Mobile Ad Hoc and Sensor Systems*, pp. 90–103, Fort Lauderdale, FL, Oct. 2004.

[8] I. Stojmenovic, A. Nayak, and J. Kuruvila, "Design guidelines for routing protocols in ad hoc and sensor networks with a realistic physical layer," *IEEE Communications Magazine*, vol. 43, pp. 101–106, March 2005.