# Exploring and Exploiting Routing Opportunities in Wireless Ad-hoc Networks

A.A. Bhorkar, M. Naghshvar, T. Javidi and B.D. Rao
Department of Electrical Engineering,
University of California San Diego, CA, 92093
{abhorkar, naghshvar, tjavidi, brao}@ucsd.edu

***Abstract*—In this paper, d-AdaptOR, a distributed opportunistic routing scheme for multi-hop wireless ad-hoc networks is proposed. The proposed scheme utilizes a reinforcement learning framework to achieve the optimal performance adaptively even in the absence of reliable knowledge about channel statistics and network model. The scheme extends an earlier proposed scheme [1] which relied on centralized computation. In contrast, d-AdaptOR operates solely based on local information and coordination with other neighboring nodes via network message passing while achieving optimality with respect to an expected average per packet cost criterion.**

## I. Introduction

Opportunistic routing for multi-hop wireless ad-hoc networks has seen recent research interest to overcome deficiencies of conventional routing [2]–[7] as it is applied in a wireless setting. Opportunistic routing decisions are made in an on-line manner, choosing the next relay based on the actual transmission outcomes in addition to a rank ordering of relays. This on-line and sample-path dependent structure of opportunistic schemes improves the performance of routing by exploiting the broadcast nature of wireless transmissions as well as the inherent path and multi-user diversity present in a network. This paper describes d-AdaptOR algorithm which realizes the gains of opportunistic routing in an adaptive and distributed manner. This algorithm extends the centralized algorithm presented in [1].

The authors in [2], [7] provided a theoretical foundation of opportunistic routing as a Markov Decision Problem. In particular, it is shown that the optimal routing decision at any epoch is to select the next relay node based on an index summarizing the expected-cost-to-forward from that node to the destination. This index is shown to be computable in a distributed manner and with low complexity using the probabilistic description of wireless links. The study in [2], [7] provides a unifying framework for almost all versions of opportunistic routing such as SDF [3], GeRaF [4], and EXOR [5]. Here, we note that the variations in [3]–[5] are due to the authors' choices of cost measures to optimize. For instance an optimal route in the context of EXOR is computed so as to minimize the expected number of transmissions (ETX), while GeRaF uses the smallest expected geographical distance from the destination as a criterion for selecting the next-hop.

This work was partially supported by the UC Discovery Grant #com07-10241, Intel Corp., QUALCOMM Inc., Texas Instruments Inc., CWC at UCSD, and NSF CAREER Award CNS-0533035.

The opportunistic algorithms proposed in [2]–[7] implicitly depend on a precise probabilistic model of wireless connections and local topology of the network. In practical setting, however, these probabilistic models have to be "learned" and "maintained". The question of estimation error [8] and learning [1] in the opportunistic routing context has recently received some attention. In this paper, using a reinforcement learning framework, we propose a distributed Adaptive Opportunistic Routing (d-AdaptOR) algorithm which minimizes the expected average per packet cost when zero or erroneous knowledge of transmission success probabilities and network topology is available. This algorithm extends our earlier centralized algorithm [1] to allow for a practical distributed asynchronous implementation with low complexity and overhead costs. The most significant characteristics of the proposed scheme are: 1) it does not assume any initial knowledge of the network, 2) it is distributed, and 3) it is asynchronous, i.e. it only requires a local clock at each node.

The rest of the paper is organized as follows: In Section II, we discuss the system model and formulate the problem. Section III formally introduces our proposed distributed routing algorithm, Adaptive Opportunistic Routing (d-AdaptOR), and its implementation details. We state and prove the optimality of d-AdaptOR algorithm in section IV. We then provide a simulation study of d-AdaptOR in Section V. Finally, we conclude the paper in Section VI.

## II. System Model

We consider the problem of routing packets from the source node $o$ to a destination node $d$ in a wireless ad-hoc network of $d+1$ nodes denoted by the set $\Theta = \{o, 1, 2, \cdots, d\}$. The time is slotted and indexed by $n \geq 0$ (this assumption is not technically critical and is only assumed for ease of exposition). A packet indexed by $m \geq 0$ is generated at the source node $o$ at time $\tau_s^m$ according to an arbitrary distribution with mean $\lambda > 0$.

Given a successful transmission from node $i$ to the set of nodes $S$, the next (possibly randomized) routing decision includes 1) retransmission by node $i$, 2) relaying the packet by a node $j \in S$, or 3) dropping the packet all together. If node $j$ is selected as a relay, then it transmits the packet at the next slot, while other nodes $k \neq j, k \in S$, expunge that packet.

We assume upon a transmission from node $i$, a fixed transmission cost $c_i > 0$ is incurred. Transmission cost $c_i$ can be considered to model the amount of energy used for transmission, the expected time to transmit a given packet, or the hop count when the cost is equal to unity.

We define the termination event for packet $m$ to be the event that packet $m$ is either received by the destination or is dropped by a relay before reaching the destination. We define termination time $\tau_e^m$ to be a random variable when packet $m$ is terminated. We discriminate amongst the termination events as follows: We assume that upon the termination of a packet at the destination (successful delivery of a packet to the destination), a fixed and given positive reward $R$ is obtained, while if the packet is terminated (dropped) before it reaches the destination, no reward is obtained. Let $r_m$ denote the random reward obtained at the termination time $\tau_e^m$, i.e. it is either zero if the packet is dropped prior to reaching the destination node or $R$ if the packet is received at the destination.

Let $i_{n,m}$ denote the index of the node which transmits packet $m$ at time $n$. The routing scheme can be viewed as selecting a (random) sequence of nodes $\{i_{n,m}\}$ for relaying packets $m = 1, 2, \cdots$. As such, the expected average per packet reward associated with routing packets along a sequence of $\{i_{n,m}\}$ upto time $N$ is:

$$J_N = \mathbf{E}\left[\frac{1}{M_N}\sum_{m=1}^{M_N}\left\{r_m - \sum_{n=\tau_s^m}^{\tau_e^m-1} c_{i_{n,m}}\right\}\right], \qquad (1)$$

where $M_N$ denotes the number of packets terminated upto time $N$ and the expectation is taken over the events of transmission decisions, successful packet receptions, and packet generation times.

**Problem (P)** Choose a sequence of relay nodes $\{i_{n,m}\}$ in the absence of knowledge about the network topology such that $J_N$ is maximized as $N \to \infty$.

In the next section we propose d-AdaptOR algorithm which solves Problem **(P)**. The nature of the algorithm allows for nodes to make routing decisions in distributed, asynchronous, and adaptive manner.

**Remark** The problem of shortest path routing between all source-destination pairs can be effectively decomposed to the problem above where routing from one node to a specific destination is addressed.

## III. Distributed Algorithm

In this section we present an overview and detailed description of the proposed distributed adaptive opportunistic routing (d-AdaptOR) algorithm. In the rest of the paper, we let $\mathcal{N}(i)$ to denote the set of neighbors of node $i$ including node $i$ itself. Let $\mathfrak{S}^i$ be the set of potential reception outcomes due to transmission from node $i \in \Theta$, i.e. $\mathfrak{S}^i = \{S : S \subseteq \mathcal{N}(i), i \in S\}$. For all $S \in \mathfrak{S}^i$, let $A(S)$ denote all the potential routing decisions (actions) for node $i$. $A(S)$ includes the set of nodes $S$ and the termination action $f$, i.e. $A(S) = S \cup \{f\}$. Furthermore, for each node $i$ we define a reward function on states $S \in \mathfrak{S}^i$ and potential decisions $a \in A(S)$ as:

$$g(S,a) = \begin{cases} -c_a & \text{if } a \in S \\ R & \text{if } a = f \text{ and } d \in S \\ 0 & \text{if } a = f \text{ and } d \notin S \end{cases}.$$

### A. Overview of d-AdaptOR

As discussed before, the routing decision at any given time is made based on the successful outcomes and involves retransmission, choice of next relay, or termination. Our proposed scheme makes such decisions in a distributed manner via the following three-way handshake between a node $i$ and its neighbors $\mathcal{N}(i)$.

1) At time $n$ node $i$ transmits a packet.
2) Set of nodes $S_n^i$ that have received the packet successfully from node $i$, transmit acknowledgment packets to node $i$. The acknowledgment packet of node $k \in S_n^i$ includes node's identity and a *message* summarizing its *estimated best score* (EBS) denoted by $\Lambda^k_{max}$.
3) Node $i$ announces node $j \in S_n^i$ as the next transmitter or announces the termination decision $f$.

The routing decision of node $i$ at time $n$ is based on an adaptive (stored) score vector $\Lambda_n(i,\cdot,\cdot)$. The score vector $\Lambda_n(i,\cdot,\cdot)$ lies in space $\mathbb{R}^{v_i}$, where $v_i = \sum_{S\in\mathfrak{S}^i} |A(S)|$, and is updated by node $i$ using the EBS messages $\Lambda^k_{max}$ obtained from neighbors $k \in S_n^i$. Furthermore, node $i$ uses a set of counting variables $\nu_n(i,S,a)$ and $N_n(i,S)$ and a sequence of positive scalers $\{\alpha_n\}_{n=1}^{\infty}$ to update the score vector at time $n$. The counting variable $\nu_n(i,S,a)$ is equal to the number of times set of nodes $S$ have received the packet and decision $a$ has been taken due to transmission from node $i$ upto time $n$, while $N_n(i,S)$ is equal to the number of times set of nodes $S$ have received the packet due to transmission from node $i$ upto time $n$. Lastly, $\{\alpha_n\}_{n=1}^{\infty}$ is a fixed sequence of numbers available at all nodes.

The details are presented next.

### B. Detailed description of d-AdaptOR

The operation of d-AdaptOR can be described in terms of initialization and four stages of transmission, reception and acknowledgment, relay, and adaptive computation as shown in Figure 1. For simplicity of presentation we assume a sequential timing for each of the stages. We use $n^+$ to denote some (small) time after the start of $n^{th}$ slot and $(n+1)^-$ to denote some (small) time before the end of $n^{th}$ slot such that $n < n^+ < (n+1)^- < n+1$.

0) **Initialization**:
For all $i \in \Theta, S \in \mathfrak{S}^i, a \in A(S)$, initialize $\Lambda_0(i,S,a) = 0,\ \nu_0(i,S,a) = 0,\ N_0(i,S) = 0,\ \Lambda^f_{max} = -R, \Lambda^i_{max} = 0$.
1) **Transmission Stage**:
Transmission stage occurs at time $n$ in which node $i$ transmits if it has a packet.
2) **Reception and Acknowledgment Stage**:
Let $S_n^i$ denote the (random) set of nodes that have received the packet transmitted by node $i$. In the reception

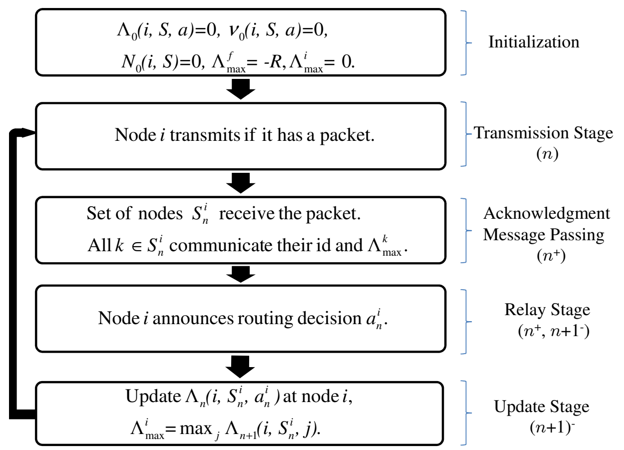



Fig. 1. Flow of the algorithm. The algorithm follows a four-stage procedure: transmission, acknowledgment, relay, and update. In the acknowledgment stage, a node obtains $\Lambda_{max}$ from the neighbors (message passing).

and acknowledgment stage, successful reception of the packet transmitted by node $i$ is acknowledged to it by all the nodes in $S_n^i$. We assume that the delay for the acknowledgment stage is small enough (not more than the duration of the time slot) such that node $i$ infers $S_n^i$ by time $n^+$.
For all nodes $k \in S_n^i$ , the ACK packet of node $k$ to node $i$ includes the EBS message $\Lambda_{max}^k$.
Upon reception and acknowledgment, the counting random variable $N_n$ is incremented as follows:

$$N_n(i,S) = \begin{cases} N_{n-1}(i,S)+1 & \text{if } S = S_n^i \\ N_{n-1}(i,S) & \text{if } S \neq S_n^i \end{cases} .$$

3) **Relay Stage**:
Node $i$ selects a routing action $a_n^i \in A(S_n^i)$ according to the following (randomized) rule:
- with probability $(1 - \epsilon_n(i, S_n^i))$,

$$a_n^i \in \arg\max_{j \in A(S_n^i)} \Lambda_n(i, S_n^i, j)$$

  is selected,[1]
- with probability $\frac{\epsilon_n(i,S_n^i)}{|A(S_n^i)|}$,

$$a_n^i \in A(S_n^i)$$

  is selected randomly,

where

$$\epsilon_n(i,S) = \frac{1}{N_n(i,S)+1}.$$

Node $i$ transmits a control packet which contains information about routing decision $a_n^i$ at some time strictly between $n^+$ and $(n+1)^-$. If $a_n^i \neq f$, then node $a_n^i$ prepares for forwarding in next time slot while nodes $j \in S_n^i, j \neq a_n^i$ expunge the packet. If termination action is chosen, i.e. $a_n^i = f$, all nodes in $S_n^i$ expunge the packet.
Upon selection of routing action, the counting variable $\nu_n$ is updated.

$$\nu_n(i,S,a) = \begin{cases} \nu_{n-1}(i,S,a)+1 & \text{if } (S,a) = (S_n^i, a_n^i) \\ \nu_{n-1}(i,S,a) & \text{if } (S,a) \neq (S_n^i, a_n^i) \end{cases} .$$

4) **Adaptive Computation Stage**:
At time $(n+1)^-$, after being done with transmission and relaying, node $i$ updates score vector $\Lambda_n(i,\cdot,\cdot)$ as follows:
- for $S = S_n^i$, $a = a_n^i$,

$$\Lambda_{n+1}(i,S,a) = \Lambda_n(i,S,a) + \alpha_{\nu_n(i,S,a)} \left( -\Lambda_n(i,S,a) + g(S,a) + \Lambda_{max}^a \right), \quad (2)$$

- otherwise, $\Lambda_{n+1}(i,S,a) = \Lambda_n(i,S,a)$.

Furthermore, node $i$ updates its EBS message $\Lambda_{max}^i$ for future acknowledgments as:

$$\Lambda_{max}^i = \max_{j \in A(S_n^i)} \Lambda_{n+1}(i, S_n^i, j).$$

## IV. Optimality of d-AdaptOR

We will now state the main result establishing the optimality of the proposed d-AdaptOR algorithm under a time-invariant model of packet reception. Let us characterize the behavior of the wireless channel using a probabilistic *local broadcast model* [7]. The local broadcast model is defined using the transition probability $P(S|i), S \subseteq \Theta, i \in \Theta$, where $P(S|i)$ denotes the probability of successful reception of packet transmitted by node $i$ by all the nodes in $S$. Note that for all $S \neq S'$, successful reception at $S$ and $S'$ are mutually exclusive and $\sum_{S \subseteq \Theta} P(S|i) = 1$. Furthermore, we assume that successful transmissions over different time slots are independent and identically distributed.

**Assumption 1.** *Wireless transmissions of node $i$ at any time can be accurately modeled by $P(S|i)$ independent of time and all other concurrent transmissions.*

Let $S_n^i$ be the set of nodes that have received the packet due to transmission from node $i$ at time $n$, while $a_n^i$ denotes the corresponding routing decision node $i$ takes at time $n$.[2] A distributed routing policy is a collection $\phi = \{\phi^o, \cdots \phi^d\}$ of routing decisions taken at nodes $i \in \Theta$, where $\phi^i$ denotes a sequence of random actions $\phi^i = \{a_0^i, a_1^i, \cdots\}$ for node $i$. The policy $\phi$ is said to be *admissible* if for all nodes $i \in \Theta$, $S \in \mathfrak{S}^i$, $a \in A(S)$, the event $\{a_n^i = a\}$ belongs to the $\sigma$-field $\mathcal{H}_n^i$ generated by $\bigcup_{j \in \mathcal{N}(i)} \{S_0^j, a_0^j, \cdots, S_{n-1,m}^j, a_{n-1,m}^j, S_n^j\}$ . Let $\Phi$ denote the set of such admissible policies.

The following theorem states that d-AdaptOR, denoted by $\phi^*$, maximizes the expected average per packet reward.

[1] In case of ambiguity, node with the smallest index is chosen.

[2] $S_n^i = \emptyset, a_n^i = f$ if node $i$ does not transmit at time $n$.

**Theorem 1.** *Suppose $\sum_{n=0}^{\infty} \alpha_n = \infty$, $\sum_{n=0}^{\infty} \alpha_n^2 < \infty$, and Assumption 1 holds. Then, for all $\phi \in \Phi$,*

$$\lim_{N\to\infty} E^{\phi^*}\left[\frac{1}{M_N}\sum_{m=1}^{M_N}\left\{r_m - \sum_{n=\tau_s^m}^{\tau_e^m-1} c_{i_{n,m}}\right\}\right]$$
$$\geq \limsup_{N\to\infty} E^{\phi}\left[\frac{1}{M_N}\sum_{m=1}^{M_N}\left\{r_m - \sum_{n=\tau_s^m}^{\tau_e^m-1} c_{i_{n,m}}\right\}\right],$$

*where $E^{\phi^*}$ and $E^{\phi}$ are the expectations taken with respect to $\phi^*$ and $\phi$ respectively.*

In the rest of this section, we prove the optimality of d-AdaptOR in two steps. In the first step, we show that $\Lambda_n$ converges almost surely. In the second step we use this convergence result to show that d-AdaptOR is optimal for Problem (**P**).

### A. Convergence of $\Lambda_n$

Let $U : \prod_i \mathbb{R}^{v_i} \to \prod_i \mathbb{R}^{v_i}$ be an operator on vector $\Lambda$ such that,

$$(U\Lambda)(i,S,a) = g(S,a) + \sum_{S'\in\mathfrak{S}^a} P(S'|a) \max_{j\in A(S')} \Lambda(a,S',j).$$

Let $\Lambda^* \in \prod_i \mathbb{R}^{v_i}$ denote the fixed point of operator $U$,[3] i.e.

$$\Lambda^*(i,S,a) = g(S,a) + \sum_{S'\in\mathfrak{S}^a} P(S'|a) \max_{j\in A(S')} \Lambda^*(a,S',j). \quad (3)$$

The following lemma establishes the convergence of recursion (2) to the fixed point of $U$, $\Lambda^*$.

**Lemma 1.** *Let*
(J1) $\Lambda_0(\cdot,\cdot,\cdot) = 0$, $\Lambda^f_{max} = -R$, $\Lambda^i_{max} = 0$ *for all* $i \in \Theta$,
(J2) $\sum_{n=0}^{\infty} \alpha_n = \infty$, $\sum_{n=0}^{\infty} \alpha_n^2 < \infty$.
*Then iterate $\Lambda_n$ obtained by the stochastic recursion (2) converges to $\Lambda^*$ almost surely.*

The proof follows known results on the convergence of a certain recursive supermartingale process presented as Fact **??** in Appendix **??**.

### B. Proof of optimality

Using the convergence of $\Lambda_n$ we show that the expected average per packet reward under d-AdaptOR is equal to the optimal expected average per packet reward obtained for a genie-aided system where the local broadcast model is known perfectly.

In proving the optimality of d-AdaptOR, we take cue from known results associated with a closely related Auxiliary Problem **(AP)**. In the Auxiliary Problem **(AP)**, there exists a centralized controller with full knowledge of the local broadcast model $P(\cdot|\cdot)$ as well as the transmission outcomes across the network [2], [7]. For Auxiliary Problem **(AP)**, a routing policy is a collection $\pi = \{\pi^o, \cdots, \pi^d\}$ of routing decisions taken for nodes $i \in \Theta$ at the centralized controller, where $\pi^i$ denotes a sequence of random actions $\pi^i = \{a_0^i, a_1^i, \cdots\}$ for node $i$. The routing policy $\pi$ is said to be admissible for Auxiliary Problem **(AP)** if the event $\{a_n^i = a\}$ belongs to the product $\sigma$-field $\mathcal{F}_n = \mathcal{P} \times \prod_i \mathcal{H}_n^i$ [10], where $\mathcal{P}$ is the borel $\sigma$-field generated by the random probability measures for the local broadcast model.[4] Let $\Pi$ denote the set of admissible policies for Auxiliary Problem **(AP)**.

The reward associated with policy $\pi \in \Pi$ for routing a single packet $m$ from the source to the destination is then given by

$$J^\pi(\{o\}) := \mathbf{E}^\pi\left[\left\{r_m - \sum_{n=0}^{\tau_e^m-1} c_{i_{n,m}}\right\} |\mathcal{F}_0\right], \quad (4)$$

where $\mathcal{F}_0 = \mathcal{P}$, and the expectation $E^\pi$ is taken with respect to the random events as well as the conditional distributions over action space defined by policy $\pi$. Now, in this setting, we are ready to formulate the following Auxiliary Problem (**AP**) as a classical shortest path Markov Decision Problem (MDP).

**Auxiliary Problem (AP)** Find an optimal policy $\pi^*$ such that for all $\pi \in \Pi$,

$$J^{\pi^*}(\{o\}) = \sup_{\pi\in\Pi} J^\pi(\{o\}). \quad (5)$$

Auxiliary Problem **(AP)** has been extensively studied in [2], [7], [11]. It is shown in [7] that $\pi^*$ exists.

**Fact** 1 (Theorem 2.1 [7]). There exists a function $\pi^* : \bigcup_i \mathfrak{S}^i \to \Theta \cup \{f\}$ such that $a_n^i = \pi^*(S_n^i)$ is an optimal solution for the Auxiliary Problem **(AP)**.[5] Furthermore, $\pi^*$ is such that

$$\pi^*(S) \in \arg\max_{j\in A(S)} V^*(j), \quad (6)$$

where (value) function $V^* : \Theta \cup \{f\} \to \mathbb{R}$ is the unique solution to the following fixed point equation:

$$V^*(d) = R \quad (7)$$
$$V^*(i) = \max(\{-c_i + \sum_{S'} P(S'|i)(\max_{j\in S'} V^*(j))\}, 0) \quad (8)$$
$$V^*(f) = 0. \quad (9)$$

Lastly, $V^*(j)$ is the maximum expected reward for routing a packet from node $j$ to destination $d$:

$$V^*(j) = J^{\pi^*}(\{j\}) = \sup_{\pi\in\Pi} J^\pi(\{j\}).$$

Lemma 2 below states the relationship between the solution of Problem **(P)** and that of the Auxiliary Problem **(AP)**, i.e. it shows that $V^*(o)$ is an outer bound for the solution to Problem **(P)**.

**Lemma 2.** *Consider any admissible policy $\phi \in \Phi$ for Problem* (**P**). *Then*

$$\limsup_{N\to\infty} E^{\phi}\left[\frac{1}{M_N}\sum_{m=1}^{M_N}\left\{r_m - \sum_{n=\tau_s^m}^{\tau_e^m-1} c_{i_{n,m}}\right\}\right] \leq V^*(o).$$

[3] Existence and uniqueness of $\Lambda^*$ is provided in [9].

[4] $\sigma$-field $\mathcal{P}$ captures the knowledge of underlying local broadcast model and assumes a well-defined prior on these models.

[5] In other words there exists a stationary, deterministic, and Markov optimal policy for Auxiliary Problem **(AP).**

*Proof:* The proof is given in [9]. Intuitively the result holds because the set of admissible policies $\Phi$ in **(P)** is a subset of admissible policies $\Pi$ in **(AP)**. ■

Lemma 3 gives the achievability proof for Problem (P) by showing that for all $\delta > 0$, the expected average per packet reward of d-AdaptOR is no less than $V^*(o) - \delta$.

**Lemma 3.** *For any* $\delta > 0$*,*

$$\liminf_{N\to\infty} E^{\phi^*}\left[\frac{1}{M_N}\sum_{m=1}^{M_N}\left\{r_m - \sum_{n=\tau_s^m}^{\tau_e^m-1} c_{i_{n,m}}\right\}\right] \geq V^*(o) - \delta.$$

*Proof:* The proof is given in Appendix A. ■

Lemmas 2 and 3 imply that

$$\lim_{N\to\infty} E^{\phi^*}\left[\frac{1}{M_N}\sum_{m=1}^{M_N}\left\{r_m - \sum_{n=\tau_s^m}^{\tau_e^m-1} c_{i_{n,m}}\right\}\right]$$

exists and is equal to $V^*(o)$. This together with Lemma 2 establishes the proof of Theorem 1.

## V. Simulations

In this section, we provide simulation results in which the performance of d-AdaptOR is compared against suitably chosen candidates: Stochastic Routing (SR) [2], EXOR [5] and a conventional routing algorithm Ad hoc On-Demand Distance Vector Routing (AODV) [12]. Both SR and EXOR are distributed mechanisms in which the probabilistic structure of the network is used to implement opportunistic routing algorithms. As a result, their performance will be highly dependent on the precision of empirical probability associated with link, $p_{ij}$. To provide a fair comparison, hence, we have considered modified versions of SR and EXOR in which the algorithms adapt $p_{ij}$ to the history of packet reception outcomes, while rely on the updates to make routing decisions (separated scheme of estimation and routing).

Our simulations are performed in QualNet. Simulations consist of a random topology of 16 nodes distributed uniformly over an area of 90mx90m. Each node is equipped with 802.11b radios transmitting at 11 Mbps. The wireless medium is modeled as to include Rician fading and Log-normal shadowing with mean 4dB and the path loss follows the two-ray model in [13] with path exponent of 3.

Packets are generated according to a CBR source with rate 10 packets/sec. They are assumed to be of length 512 bytes equipped with simple CRC error detection. The acknowledgment packets are short packets of length 24 bytes transmitted at rate of 11 Mbps, while FO packets are transmitted at reliable lower rate of 1Mbps. Cost of transmission is assumed to be one unit, while reward for successfully delivering a packet to the destination is assumed to be 20.

Fig. 2(b) plots the expected average per packet reward obtained by the candidate routing algorithms versus network operation time. The optimal algorithm with complete knowledge of link probabilities is also plotted for comparison. We first note that as expected, ADOV performs poorly compared to the opportunistic schemes as it is strictly suboptimal. In particular, Fig. 2(b) shows that the d-AdaptOR algorithm outperforms opportunistic schemes EXOR and SR by at least 5% given sufficient number of packet deliveries. Fig. 2(b) shows that SR performs poorly relative to d-AdaptOR algorithm since it fails to explore possible choices of routes and often results in strictly suboptimal routing policy.

This figure also shows that the randomized routing decisions employed by d-AdaptOR work as a double-edge sword. This is the mechanism through which network opportunities are exhaustively explored until the globally optimal decisions are constructed. At the same time, these randomized decisions lead to a short term performance loss. One should note that due to the exploratory nature of the d-AdaptOR algorithm, during initial startup time EXOR and SR perform better than d-AdaptOR.

## VI. Conclusions

In this paper, we proposed d-AdaptOR, an adaptive routing scheme which maximizes the expected average per packet reward from source to the destination in the absence of any knowledge regarding network topology and link qualities. d-AdaptOR allows for a practical distributed implementation with provably optimal performance. Simulation results show that d-AdaptOR outperforms the existing opportunistic protocols in which statistical link qualities are empirically built and the routing decisions are greedily adapted to the empirical link models.

## References


[1] A.A. Bhorkar and M. Naghshvar and T. Javidi and B.D. Rao, "An Adaptive Opportunistic Routing Scheme for Wireless Ad-hoc Networks," in *ISIT*, 2009.

[2] C. Lott and D. Teneketzis, "Stochastic routing in ad hoc wireless networks," *Decision and Control, 2000. Proceedings of the 39th IEEE Conference on*, vol. 3, pp. 2302–2307 vol.3, 2000.

[3] P. Larsson, "Selection Diversity Forwarding in a Multihop Packet Radio Network with Fading channel and Capture," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 2, no. 4, pp. 4754, October 2001.

[4] M. Zorzi and R. R. Rao, "Geographic Random Forwarding (GeRaF) for Ad Hoc and Sensor Networks: Multihop Performance," *IEEE Transactions on Mobile Computing*, vol. 2, no. 4, 2003.

[5] S. Biswas and R. Morris, "ExOR: Opportunistic Multi-hop Routing for Wireless Networks," *ACM SIGCOMM Computer Communication Review*, vol. 35, pp. 3344, October 2005.

[6] S.R. Das S. Jain, "Exploiting path diversity in the link layer in wireless ad hoc networks," *World of Wireless Mobile and Multimedia Networks, 2005. WoWMoM 2005. Sixth IEEE International Symposium on a*, pp. 22–30, June 2005.

[7] C. Lott and D. Teneketzis, "Stochastic routing in ad-hoc networks," *IEEE Transactions on Automatic Control*, vol. 51, pp. 52–72, January 2006.

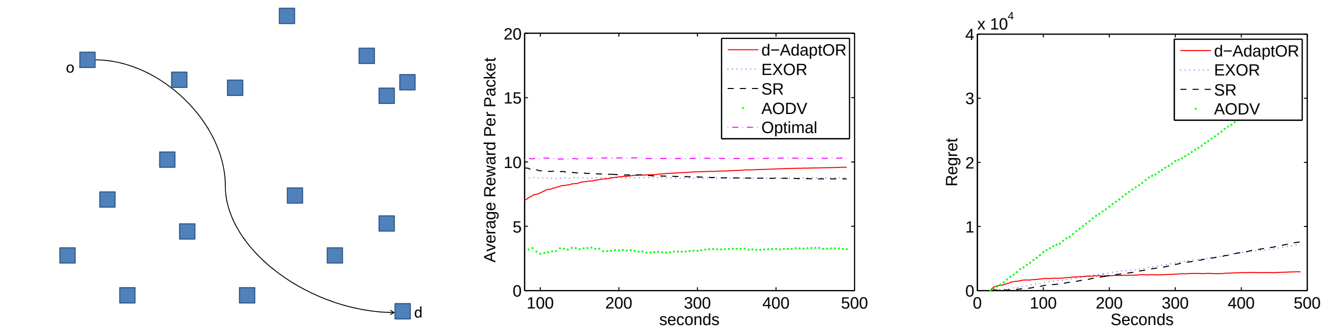
[8] T. Javidi and D. Teneketzis, " Sensitivity Analysis for Optimal Routing in Wireless Ad Hoc Networks in Presence of Error in Channel Quality Estimation," *IEEE Transactions on Automatic Control*, pp. 1303–1316, August 2004.

[9] A.A. Bhorkar, M. Naghshvar, T. Javidi, and B.D. Rao, "A routing policy for wireless ad-hoc networks," Tech. Rep., University of California, San Diego, 2008, Available at http://dsp.ucsd.edu/~bhorkar/files/Adaptor_journal.pdf.

[10] Sidney Resnick, *A Probability Path*, Birkhuser, Boston, 1998.

[11] Dimitri P. Bertsekas and John N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*, Athena Scientific, 1997.

[12] William Stallings, *Wireless Communications and Networks*, Prentice Hall, second edition, 2004.




(a) Node topology : Nodes placed uniformly over an area of 90m × 90m

(b) Expected Average Reward

(c) Regret

Fig. 2. d-AdaptOR vs distributed SR, EXOR, and AODV. $\alpha_n = \frac{1}{n\log(n)}$, $R = 20$, $c_i = 1$ for all $i$.

[13] J. Doble, *Introduction to radio propagation for fixed and mobile communications*, Artech House, Boston, 1996.

## APPENDIX

### A. Proof of Lemma 3

*Proof:* From (3), (7)-(9), and (**??**) we obtain the following equality for all $i \in \Theta,\ S \in \mathfrak{S}^i$,

$$\arg\max_{j\in A(S)} V^*(j) = \arg\max_{j\in A(S)} \Lambda^*(i,S,j)\ . \tag{10}$$

Let

$$b = \min_{i\in\Theta}\min_{S\in\mathfrak{S}^i}\min_{\substack{j,k\in A(S)\\ \Lambda^*(i,S,j)\neq\Lambda^*(i,S,k)}} |\Lambda^*(i,S,j)-\Lambda^*(i,S,k)|.$$

Lemma 1 implies that, in an almost sure sense, there exists packet index $m_1 < \infty$ such that for all $n > \tau_s^{m_1}$, $i \in \Theta$, $S \in \mathfrak{S}^i$, $a \in A(S)$,

$$|\Lambda_n(i,S,a) - \Lambda^*(i,S,a)| \le b/2.$$

In other words, from time $\tau_s^{m_1}$ onwards, given any node $i \in \Theta$ and set $S \in \mathfrak{S}^i$, the probability that d-AdaptOR chooses an action $a \in A(S)$ such that $\Lambda^*(i,S,a) \neq \max_{j\in A(S)} \Lambda^*(i,S,j)$ is upper bounded by $\epsilon_n(i,S)$. Furthermore, since $N_n(i,S) \to \infty$ (Lemma **??**), for a given $\gamma > 0$, with probability 1, there exists a packet index $m_2 < \infty$ such that for all $n > \tau_s^{m_2}$, $\max_{i,S}\epsilon_n(i,S) < \gamma$.

Let $m_0 = \max\{m_1, m_2\}$. For all packets with index $m \le m_0$ the overall expected reward is upper-bounded by $m_0 R < \infty$ and lower-bounded by $-\frac{m_0}{\lambda} d\max_i c_i > -\infty$, hence, their presence does not impact the expected average per packet reward. Consequently, we only need to consider the routing decisions of policy $\phi^*$ for packets $m > m_0$.

Consider the $m^{th}$ packet generated at the source. Let $B_k^m$ be an event for which there exist $k$ instances when d-AdaptOR routes packet $m$ differently from the possible set of optimal actions. Mathematically speaking, event $B_k^m$ occurs iff there exists instances $\tau_s^m \le n_1^m \le n_2^m \ldots n_k^m \le \tau_e^m$ such that for all $l = 1, 2, \ldots, k$

$$\Lambda^*(i_{n_l^m,m}, S_{n_l^m}, a_{n_l^m}) \neq \max_{j\in A(S_{n_l^m})} \Lambda^*(i_{n_l^m,m}, S_{n_l^m}, j),$$

where $S_{n_l^m}$ is the set of nodes that have successfully received packet $m$ at time $n_l^m$ due to transmission from node $i_{n_l^m,m}$. We call event $B_k^m$ a mis-routing of order $k$. For $m > m_0$,

$$Prob(B_k^m) \le (\max_{i,S}\epsilon_n(i,S))^k \le \gamma^k.$$

Now for packets $m > m_0$, let us consider the expected differential reward under policies $\pi^*$ and $\phi^*$:

$$\begin{aligned}
&\mathbf{E}^{\pi^*}\left[\left\{r_m - \sum_{n=\tau_s^m}^{\tau_e^m-1} c_{i_{n,m}}|\mathcal{F}_0\right\}\right] - \mathbf{E}^{\phi^*}\left[\left\{r_m - \sum_{n=\tau_s^m}^{\tau_e^m-1} c_{i_{n,m}}\right\}\right]\\
&= V^*(o) - \mathbf{E}^{\phi^*}\left[\left\{r_m - \sum_{n=\tau_s^m}^{\tau_e^m-1} c_{i_{n,m}}\right\}\right]\\
&= \sum_{k=0}^{\infty}\mathbf{E}^{\phi^*}\left[V^*(o) - \left\{r_m - \sum_{n=\tau_s^m}^{\tau_e^m-1} c_{i_{n,m}}\right\}|B_k^m\right]\\
&\quad \times Prob(B_k^m)\\
&\le \sum_{k=0}^{\infty} k\ R\ Prob(B_k^m) \qquad (11)\\
&\le R\sum_{k=1}^{\infty} k\gamma^k \qquad (12)\\
&= \delta, \qquad (13)
\end{aligned}$$

where $\delta = \frac{\gamma R}{(1-\gamma)^2}$. Inequality (11) is obtained by noticing that maximum loss in the reward occurs if algorithm d-AdaptOR decides to drop packet $m$ (no reward) while there exists a node $j$ in the set of potential forwarders such that $V^*(j) \approx R$.

Thus, for all $\delta > 0$ the expected average per packet reward under policy $\phi^*$ is bounded as

$$\begin{aligned}
&\liminf_{N\to\infty} E^{\phi^*}\left[\frac{1}{M_N}\sum_{m=1}^{M_N}\left\{r_m - \sum_{n=\tau_s^m}^{\tau_e^m-1} c_{i_{n,m}}\right\}\right]\\
&\ge \liminf_{N\to\infty} E^{\phi^*}\left[\frac{1}{M_N}\sum_{m=1}^{M_N}(V^*(o)-\delta)\right]\\
&= V^*(o) - \delta.
\end{aligned}$$

■