Delay and Energy Optimal Two-Hop Relaying in Delay Tolerant Networks

Chandramani Singh, Anurag Kumar, Rajesh Sundaresan

Department of Electrical Communication Engineering

Indian Institute of Science

Bangalore 560012, India

Email: {chandra, anurag, rajeshs}@ece.iisc.ernet.in

Abstract—We study the trade-off between delivery delay and energy consumption in delay tolerant mobile wireless networks that use two-hop relaying. The source may not have perfect knowledge of the delivery status at every instant. We formulate the problem as a stochastic control problem with partial information, and study structural properties of the optimal policy. We also propose a simple suboptimal policy. We then compare the performance of the suboptimal policy against that of the optimal control with perfect information. These are bounds on the performance of the proposed policy with partial information. Several other related open loop policies are also compared with these bounds.

I. Introduction

During the last few years, there has been a growing interest in *delay tolerant networks* (DTNs) [6]. DTNs are sparse wireless ad hoc networks with highly mobile nodes. In these networks, the link between two nodes is up when these are within each other's transmission range, and is down otherwise. In particular, there is no complete route between a source and its destination, most of the time.

In such networks, a common technique to improve packet delivery delay is to disseminate the packet to multiple nodes in the network. In particular, any node that has the packet copies it to another node arriving within its communication range provided the receiving node does not already have the packet. The new node also buffers the packet and acts in the same way. The destination receives the packet when it meets any of the nodes carrying the packet. This "store-carry-forward" paradigm is referred to as *epidemic routing* [14]. Epidemic routing reduces the delivery delay at the cost of inefficient use of network resources such as buffer space, bandwidth and node energy. A variation of epidemic routing that exploits the trade-off between delivery delay and resource consumption is *two-hop relaying* [8]. The source copies the packet to any node that it encounters, and that does not yet have a copy of the packet, but the relays are allowed to transmit the packet only to the destination.

These protocols need to be combined with a so-called "recovery process" that deletes copies of the packet at the nodes carrying it, following its successful delivery to the destination. The packet is deleted to free the buffer and prevent the node from copying it to another node that it meets. On the other hand, a node retains "packet delivered" information in the form of an *anti-packet* that prevents it from accepting another copy of the same packet. Haas and Small [9] suggest the following recovery schemes.

- *immune:* An anti-packet is created at a node only after it meets the destination (and delivers the packet, if the destination did not receive it before).
- *immune_tx:* A node carrying an anti-packet transmits it to another node that is carrying the associated obsolete packet to let that node know of packet delivery.
- *vaccine:* A node carrying an anti-packet forwards it to all other nodes.

In this work we focus on a scenario where a packet needs to be delivered from a source to a destination, over a DTN. There are also other nodes that can work as potential relays. All nodes are mobile. However, relays do not interact with each other. The two-hop relaying protocol is employed. At each meeting with a relay that does not have a copy of the packet, the source has the option of either copying or not copying. When the destination meets the source or a relay that carries the packet, then the packet is delivered. We assume an enhanced version of "immune" recovery scheme in which the source can know of packet delivery via meeting either the destination or a relay with the anti-packet. Copying the packet to relays incurs a transmission cost, but on the other hand increases the number of carriers of the packet, leading to faster delivery. We focus on the problem of optimal control of relaying. Note that copying results in wastage of resources if the packet is already delivered to the destination. However, knowledge of packet delivery is constrained by the same limited connectivity that constrains packet delivery. In such a scenario, the source's decision depends upon its belief about the status of delivery. The source updates its belief continuously with time, and also after each meeting with another node. We formulate the problem as a *partially observable Markov decision process* (POMDP) [4], and characterize the optimal policy. We also propose a simple suboptimal policy.

Related work: Groenevelt et al. [7] model epidemic relaying and two-hop relaying using Markov chains, and derive the average delay and number of copies generated until the time of delivery. Hanbali et al. [10] extend the analysis of two-hop relaying to the case where there is an exponential *time-to-live* timer at each relay. They also consider a limit on the maximum number of transmissions by the source. Zhang et al. [16]

develop a unified framework based on ordinary differential equations to study epidemic routing and its variants. Their models incorporate various recovery schemes as well.

Altman et al. [2] address the optimal relaying problem for a class of *monotone relay strategies* which includes epidemic relaying and two-hop relaying. In particular, they derive *static* and *dynamic* relaying policies. Altman et al. [3] consider optimal discrete-time two-hop relaying. They also employ stochastic approximation to facilitate online estimation of network parameters. In another paper, Altman et al. [1] consider a scenario where active nodes in the network continuously spend energy while *beaconing*. Their paper studies the joint problem of node activation and transmission power control. All these works use fluid approximations to model DTNs and study only open loop controls. Consequently, none of them incorporates the effect of recovery schemes on network performance.

To our knowledge, the only work to have studied closed loop control of optimal relaying in DTNs, is Neglia and Zhang [12]: the optimal policy is a threshold on the number of copies in the network. They demonstrate that open loop policies are indeed inefficient. However, they assume that all the nodes know the number of copies in the network at all times. Moreover, they are informed instantaneously if the destination has received the packet. Thus the performance reported in [12] is a lower bound for the cost in a real system. One simplification due to the assumption of complete information is that the recovery scheme is immaterial.

Our Contributions: We formulate the controlled forwarding problem as a POMDP (Section III), and derive monotonicity results for the value function (Theorem 3.1) and the optimal policy (Theorem 3.2). Next we study a modified control problem that explicitly gives a suboptimal policy for the original problem (Theorem 4.1). Numerical results show that the suboptimal control performs close to optimal control with complete information, and outperforms the open loop control.

II. SYSTEM MODEL

We consider a set of $N + 1$ mobile nodes. These include a source, a destination and $N - 1$ other nodes that act as relays. At $t = 0$, *a single packet* is generated at the source node and is destined to the destination node. Let $\mathcal{N} = \{1, 2, \ldots, N - 1\}$ be the indexed set of relays. The index of the destination node is N. Two nodes may communicate only when they come within transmission range of each other, i.e., at the so called *meeting instants*.

As is customary in this field (see e.g., [9], [16]), the terminology relating to the spread of infectious diseases is used to describe the state of a node. A relay is *susceptible* until it receives a copy of the packet from the source. A node with a copy of the packet is said to be *infected*. Once a relay delivers the packet to the destination or comes to know of the packet delivery by some means, it deletes the packet from its own buffer. Such a relay is said to have *recovered*.

1) Mobility Model: We model the point process of the meeting instants between pairs of nodes as independent Poisson point processes, each with rate λ. Groenevelt et al. [7]

validate this model for common mobility models, e.g., random walker, random direction and random waypoint. In particular, they establish its accuracy under the assumptions of small communication range and sufficiently high speed of nodes.

2) Energy model: We assume that that each transmission incurs a unit energy expenditure at the transmitter.

3) Routing Protocol: As stated before, the two-hop relay protocol is assumed. We also assume an enhanced version of "immune" recovery scheme in which the source can know of packet delivery either via meeting the destination or a relay with the anti-packet. Transmissions between two nodes are assumed to be instantaneous.

III. THE OPTIMAL FORWARDING PROBLEM

At each meeting epoch with a susceptible relay, the source needs to decide whether to copy the packet or not. Copying the packet incurs unit cost, but promotes early delivery of the packet to the destination. Thus, there is a trade-off, and we wish to determine the optimal copying policy of the source. The objective is to minimize

$$\mathbb{E}\{\mathcal{T}_d + \eta \mathcal{E}_c\}, \qquad (1)$$

where $\mathcal{T}_d$ is the delay in delivery of the packet to the destination, $\mathcal{E}_c$ is the total energy consumption due to transmissions of copies of the packet, and η is the parameter that relates energy consumption to delay. Varying η facilitates studying the trade-off between the above two quantities.

A. The MDP Formulation

Let $T_k, k = 1, 2, \ldots$ denote the meeting epochs of the source node with other (destination or relay) nodes; $T_0 := 0$. Thus $T_k, k = 1, 2, \ldots$ are obtained as the superposition of N independent Poisson processes, each of rate λ, and hence are the points of a rate $N\lambda$ Poisson process.

Let $\mathcal{M}(t)$ be the set of relays that received copies of the packet in $[0, t]$, i.e., are either infected or recovered, and $\mathcal{R}(t)$ be the set of recovered relays at time t. We use $\mathcal{M}_k$ and $\mathcal{R}_k$ to mean $\mathcal{M}(T_k-)$ and $\mathcal{R}(T_k-)$ respectively. We also define $M_k := |\mathcal{M}_k|$, and $R_k := |\mathcal{R}_k|$. Without any loss of generality we assume that $\mathcal{M}_k = \{1, \ldots, M_k\}$ whenever it is nonempty. Then M_k suffices to describe the set $\mathcal{M}_k$. Let I_k and S_k describe respectively, the index and the state of the relay that the source meets at T_k; $I_k \in \{1, \ldots, N\}$, and $S_k \in \mathcal{S} := \{s, i, r, d\}$ where s, i, r stand for susceptible, infected, and recovered relays respectively and d stands for the destination.

The state of the system at meeting epoch T_k is given by the tuple

$$\bar{X}_k := (M_k, \mathcal{R}_k, S_k).$$

Let us also introduce a terminal state τ; the system enters τ when the source meets either a recovered node or the destination. We also define $\bar{\mathcal{N}} := \{0, 1, 2, \ldots, N - 1\}$. The state space then is a subset of $(\bar{\mathcal{N}} \times 2^{\mathcal{N}} \times \mathcal{S}) \cup \{\tau\}$. In particular the system never encounters the set of states $\{(M_k, \mathcal{R}_k, S_k) : S_k \in \{r, d\}\}$, and $\mathcal{R}_k \subset \{1, \ldots, M_k\}$.

The control space is $\mathcal{U} \in \{0, 1\}$, where 1 is for *copy* and 0 is for *do not copy*. Let U_k be the action of the source node at meeting epoch $T_k, k = 1, 2, \ldots$.

Define $\Delta_k := T_k - T_{k-1}$. Thus Δ_k is exponentially distributed with parameter $N\lambda$. To take into account the random meetings between relays and the destination during Δ_k, we define another random variable $Z_k := (D_k, \mathcal{V}_k)$. D_k is the minimum of Δ_k and the packet delivery delay beginning from T_{k-1}, and thus is 0 if $R_{k-1} > 0$. $\mathcal{V}_k$, on the other hand, is the set of relays that recover during $[T_{k-1}, T_k]$. The above embedding convention is shown in Figure 1. We treat the tuple
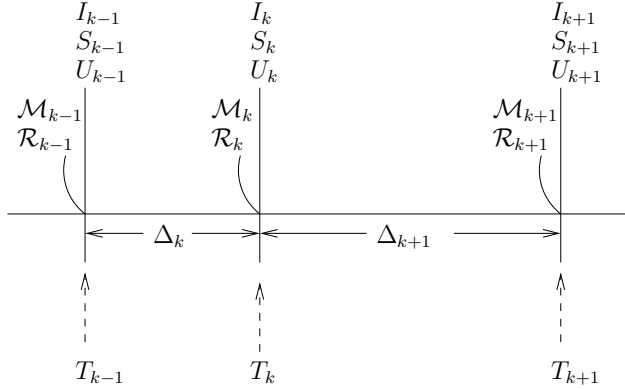

Fig. 1. Diagram showing the embedding convention.

$(\Delta_{k+1}, Z_{k+1}, I_{k+1}, S_{k+1})$ as the random disturbance at epoch T_k; this along with the state $\bar{X}_k$ and the action U_k determines the single stage cost and the next state as given below. While Δ_{k+1} and I_{k+1} are clearly independent of the current state and the action, distributions of Z_{k+1} and S_{k+1} depend on $\bar{X}_k, U_k$ and Δ_{k+1}.

1) Transition structure: From the description of the system model,

$$\bar{X}_{k+1} = \tau, \text{ if } S_{k+1} \in \{r, d\}.$$

Let us define $C_k = U_k 1_{\{S_k = s\}}$, i.e., $C_k = 1$ if and only if a copy is actually made. Then, for $t \in [T_k, T_{k+1})$, $\mathcal{M}(t) = \mathcal{M}_k \cup \{I_k\}$ if $C_k = 1$ and remains unchanged otherwise.[1] In particular,

$$M_{k+1} = M_k + C_k. \tag{2}$$

Also,

$$\mathcal{R}_{k+1} = \mathcal{R}_k \cup \mathcal{V}_{k+1}, \tag{3}$$

and S_{k+1} is a component in the random disturbance. All these show that the next state is indeed a function of the current state, the current action and the current disturbance.

2) Cost Structure: The single stage cost is given by

$$\bar{g}(\bar{X}_k, U_k, \Delta_{k+1}, Z_{k+1}, I_{k+1}, S_{k+1}) = \eta C_k + D_{k+1}.$$

The two terms in the right hand side account for energy consumption and delivery delay respectively. The terminal cost $\bar{g}(\tau) = 0$.

[1] Again a reindexing of nodes makes $\mathcal{M}_k \cup \{I_k\} = \{1, \ldots, M_k + 1\}$.

B. The POMDP Formulation

In reality, knowledge of packet delivery to destination, and hence of $\mathcal{R}(t)$, is constrained by the same limited connectivity that constrains packet delivery. Hence we develop the problem in a *partially observable Markov decision process* (POMDP) framework. At the kth meeting epoch, the source has access to observation (Δ_k, I_k, S_k). Let us denote by H_k the information available to the controller at epoch T_k. We have

$$H_k = (\Delta_1, \ldots, \Delta_k, I_1, \ldots, I_k, S_1, \ldots, S_k, U_1, \ldots, U_{k-1}),$$

where $\Delta_1 = T_1$. Clearly,

$$H_{k+1} = (H_k, U_k, \Delta_{k+1}, I_{k+1}, S_{k+1}), \ k = 0, 1, \ldots$$

We also define the partial history

$$H_{(k+1)-} = (H_k, U_k, \Delta_{k+1}).$$

Considering a new system with state H_k at epoch T_k, we thus get a *completely observable Markov decision process* (COMDP) equivalent of the original POMDP. Following the discussion in [4, Section 5.4], we can make the routine conclusion that $(M_k, P(\mathcal{R}_k|H_k), S_k)$ are sufficient statistics for the POMDP problem, where $P(\mathcal{R}_k|H_k)$ is the probability that the subset $\mathcal{R}_k$ of $\{1, \ldots, M_k\}$ has recovered. Note that M_k is known to the source, S_k is observed, and the only part of state $\bar{X}_k$ that is not observable is $\mathcal{R}_k$. The lack of interaction between relays in our model implies the tremendous simplification that the recovery processes of relays are independent, i.e., if we define vector Ψ_k as $\Psi_{k,j} := 1 - P(\text{the state of relay } j \text{ is } r|H_k)$, then

$$P(\mathcal{R}_k|H_k) = \prod_{j \in \mathcal{R}_k} (1 - \Psi_{k,j}) \prod_{j \notin \mathcal{R}_k} \Psi_{k,j}.$$

Hence we formulate a new COMDP whose state at epoch T_k is $X_k = (M_k, \Psi_k, S_k)$. As before the system enters the terminal state τ when the source meets either a recovered relay or the destination. The state space now is a subset of $(\bar{\mathcal{N}} \times [0, 1]^{\mathcal{N}} \times \mathcal{S}) \cup \{\tau\}$. In particular $\Psi_k \in \Theta(M_k)$ where $\Theta(M) := \{\Psi \in [0, 1]^{\mathcal{N}} : \Psi_j = 1 \text{ for all } j > M\}$, and $S_k \in \{s, i\}$ as before.

1) Transition Structure: Clearly

$$X_{k+1} = \tau, \text{ if } S_{k+1} \in \{r, d\}.$$

If $S_{k+1} \in \{s, i\}$, the controlled Markov process evolves as follows. M_{k+1} is obtained according to (2). In the interval $[T_k, T_{k+1})$, $\mathcal{R}(t)$ is a pure birth continuous time Markov chain, whose evolution is governed by the mobility of infected nodes and the destination. In particular, the conditional distributions of relays' states at $t = T_{(k+1)-}$, denoted by vector $\Psi_{(k+1)-}$, are obtained as

$$\Psi_{(k+1)-,j} = 1 - P(\text{the state of relay } j \text{ is } r|H_{(k+1)-})$$
$$= \begin{cases} \Psi_{k,j} \exp(-\lambda\Delta_{k+1}) & \text{if } j \in \mathcal{M}_{k+1}, \\ 1 & \text{otherwise.} \end{cases}$$

Finally, $\Psi_{(k+1)-}$ is updated immediately after observing I_{k+1} and S_{k+1} as follows. For $I_{k+1} = j, S_{k+1} \in \{s, i\}$,

$$\Psi_{k+1,l} = \begin{cases} 1 & \text{if } l = j, \\ \Psi_{(k+1)-,l} & \text{otherwise.} \end{cases}$$

$$
\begin{aligned}
J(X) \\
= \min\Big\{ & \frac{\prod_{j\in\mathcal{N}}\Psi_j}{(M+N)\lambda} + \mathbb{E}_{\Delta'}\Big\{ P(I'\notin\mathcal{M}|X,0,\Delta')J(M,\Psi^M,s) + \sum_{j\in\mathcal{M}} P(I'=j,S'=i|X,0,\Delta')J(M,\Psi^{M,j},i)\Big\}, \\
& \eta C + \frac{\prod_{j\in\mathcal{N}}\Psi_j}{(M+C+N)\lambda} + \mathbb{E}_{\Delta'}\Big\{ P(I'\notin\mathcal{M}\cup\mathcal{C}|X,1,\Delta')J(M+C,\Psi^{M+C},s) \\
& \qquad\qquad + \sum_{j\in\mathcal{M}\cup\mathcal{C}} P(I'=j,S'=i|X,1,\Delta')J(M+C,\Psi^{M+C,j},i)\Big\}, \\
= \min\Big\{ & \frac{\prod_{j\in\mathcal{N}}\Psi_j}{(M+N)\lambda} + \mathbb{E}_{\Delta'}\Big\{ \frac{N-1-M}{N}J(M,\Psi^M,s) + \sum_{j\in\mathcal{M}} \frac{\Psi_j^M}{N}J(M,\Psi^{M,j},i)\Big\}, \\
& \eta C + \frac{\prod_{j\in\mathcal{N}}\Psi_j}{(M+C+N)\lambda} + \mathbb{E}_{\Delta'}\Big\{ \frac{N-1-M-C}{N}J(M+C,\Psi^{M+C},s) + \sum_{j\in\mathcal{M}\cup\mathcal{C}} \frac{\Psi_j^{M+C}}{N}J(M+C,\Psi^{M+C,j},i)\Big\}\Big\}. \quad (4)
\end{aligned}
$$

We define mappings Γ and Υ that effect these transitions, i.e.,

$$
\begin{aligned}
\Psi_{(k+1)-} &= \Gamma(\Psi_k, M_{k+1}, \Delta_{k+1}), \\
\Psi_{k+1} &= \Upsilon(\Psi_{(k+1)-}, I_{k+1}, S_{k+1}) \text{ if } S_{k+1}\in\{s,i\}.
\end{aligned}
$$

2) Cost Structure: The single stage cost for the above formulated COMDP will be

$$
\begin{aligned}
\mathbb{E}_{\mathcal{R}_k,Z_{k+1}}\{\bar{g}(M_k,\mathcal{R}_k,S_k,U_k,\Delta_{k+1},Z_{k+1},I_{k+1},S_{k+1})|H_k,U_k\} \\
= \mathbb{E}_{\mathcal{R}_k,Z_{k+1}}\{\eta C_k + D_{k+1}|H_k,U_k\},
\end{aligned}
$$

where the subscripts denote the unobserved random variables over which expectation is taken. Taking expectation with respect to the observables $(\Delta_{k+1}, I_{k+1}, S_{k+1})$, the expected single stage cost

$$
\begin{aligned}
g(M_k,\Psi_k,S_k,U_k) \\
= {}& \mathbb{E}_{\Delta_{k+1},I_{k+1},S_{k+1}}\Big\{ \mathbb{E}_{\mathcal{R}_k,Z_{k+1}}\{\eta C_k + D_{k+1}|H_k,U_k\}\Big\} \\
= {}& \eta C_k + P(\mathcal{R}_k=\emptyset|H_k) \\
& \mathbb{E}_{\Delta_{k+1},I_{k+1},S_{k+1}}\Big\{ \mathbb{E}_{Z_{k+1}}\{D_{k+1}|M_k,\mathcal{R}_k=\emptyset,S_k,U_k\}\Big\} \\
= {}& \eta C_k + \Big(\prod_{j\in\mathcal{N}}\Psi_{k,j}\Big)\mathbb{E}_{\Delta_{k+1}}\Big\{ \frac{1-\exp(-(M_k+C_k)\lambda\Delta_{k+1})}{(M_k+C_k)\lambda}\Big\} \\
= {}& \eta C_k + \frac{\prod_{j\in\mathcal{N}}\Psi_{k,j}}{(M_k+C_k+N)\lambda}.
\end{aligned}
$$

This expression has a simple form: there is an energy cost of η if the packet is copied to the relay; with probability $\prod_{j\in\mathcal{N}}\Psi_j$ the source's belief is that the destination has not yet received the packet, and in this case the contribution to the delivery delay in the next stage of the COMDP is the minimum of two independent exponentially distributed random variables, one with rate $(M_k+C_k)\lambda$ and the other with rate $N\lambda$.

3) Policies: A policy π is a sequence of mappings $\{u_k^\pi, k=1,2,\dots\}$, where $u_k^\pi : (\bar{\mathcal{N}}\times[0,1]^{\mathcal{N}}\times\mathcal{S})\cup\{\tau\} \to \mathcal{U}$. The cost function of an admissible policy π for initial state $X=(M,\Psi,S)$ is

$$
J^\pi(X) = \lim_{K\to\infty}\sum_{k=1}^K \mathbb{E}\Big\{ g(X_k, u_k^\pi(X_k))|X_1=X\Big\}.
$$

Let Π be the set of all admissible policies. Then the optimal cost function is defined as

$$
J(X) = \min_{\pi\in\Pi} J^\pi(X).
$$

A policy π is called stationary if u_k^π are identical, say u, for all k. For brevity we refer to it as the stationary policy u. A stationary policy $u^* \equiv \{u^*, u^*, \dots\}$ is optimal if $J^{u^*}(X) = J(X)$ for all states X.

C. Optimal Policy

Since the cost function g takes nonnegative values for all possible values of its arguments, Proposition 1.1 in [5, Chapter 3] implies that, optimal cost function will satisfy the following Bellman equation. For $X=(M,\Psi,S)$,

$$
J(X) = \min_{u\in\{0,1\}}\Big[g(X,u) + \mathbb{E}_{\Delta',I',S'}\Big\{ J(X'|X,u,\Delta',I',S')\Big\}\Big].
$$

Here (Δ',I',S') denote the random disturbance and X' denotes the next state. X' is determined by the transition structure described above. In the following, we also use notations:

$$
\begin{aligned}
\Psi^M &= \Gamma(\Psi,M,\Delta'), \\
\text{and } \Psi^{M,j} &= \Upsilon(\Psi^M,j,i), \text{ for all } M,j,
\end{aligned}
$$

where superscript M shows dependence on the number of nodes carrying the packet. Moreover, $\mathcal{M}\cup\mathcal{C}$ is used to denote the set $\{1,\dots,M+C\}$. Thus the Bellman equation is expanded as in (4) above. The value iteration algorithm can be written as follows

$$
\begin{aligned}
J_0(M,\Psi,S) &= 0, \ S\in\{s,i\}; &(5a) \\
\text{for } k&=1,2,\dots, & \\
J_k(M,\Psi,s) &= \min\{A_k(M,\Psi), \eta + A_k(M+1,\Psi)\}, &(5b) \\
J_k(M,\Psi,i) &= A_k(M,\Psi), &(5c)
\end{aligned}
$$

where,

$$
\begin{aligned}
A_k(M,\Psi) = {}& \frac{\prod_{j\in\mathcal{N}}\Psi_j}{(M+N)\lambda} \\
& + \mathbb{E}_{\Delta'}\Big\{ \frac{N-1-M}{N}J_{k-1}(M,\Psi^M,s) \\
& + \sum_{j\in\mathcal{M}} \frac{\Psi_j^M}{N}J_{k-1}(M,\Psi^{M,j},i)\Big\}. \quad (6)
\end{aligned}
$$

Again, since the cost function always assumes nonnegative values and the action space is finite, Proposition 1.6 in [5, Chapter 3] implies that $J_k(M, \Phi, S) \to J(M, \Phi, S)$ as $k \to \infty$. Furthermore, a stationary policy u^* is optimal if and only if, for all X, $u^*(X)$ attains minimum in the above Bellman equation (see Proposition 1.3 in [5, Chapter 3]).[2]

Remarks 3.1: Recall that at the first decision instant T_1, X_1 is $X^o := (0, (1, \ldots, 1), s)$ if the source meets any of the relays and τ if it meets the destination. The objective function (1) can then be restated as

$$\min_{\pi \in \Pi} \mathbb{E}^\pi \{\mathcal{T}_d + \eta \mathcal{E}_c\} = \frac{1}{N\lambda} + \frac{N-1}{N} \min_{\pi \in \Pi} J^\pi(X^o) + \eta, \quad (7)$$

where the superscript π shows dependence on the underlying policy. In the right hand side, $\frac{1}{N\lambda}$ is the average delay until the first meeting, and η is the cost of copying to the destination; these have to be borne under any policy.

The constrained problem: Often we aim at optimizing the delivery performance given a constraint on the expected energy consumption (e.g., see[1], [2], [3]). The constrained version of (7) is

$$\min_{\pi \in \Pi} \quad \mathbb{E}^\pi \{\mathcal{T}_d\},$$
$$\text{s. t.} \quad \mathbb{E}^\pi \{\mathcal{E}_c\} \leq \epsilon$$

For any $\eta > 0$, let $\pi(\eta)$ be the optimal solution of (7). Further assume that for a given ϵ in the constrained problem, there is an η_ϵ such that $\mathbb{E}^{\pi(\eta_\epsilon)} \{\mathcal{E}_c\} = \epsilon$. Then $\pi(\eta_\epsilon)$ is optimal for the constrained problem as well.

D. Properties of the Optimal Control

We now describe some properties of the optimal policy. We start with defining the following partial order on $[0, 1]^{\mathcal{N}}$.

Definition 3.1: For $\Psi^1, \Psi^2 \in [0, 1]^{\mathcal{N}}$, $\Psi^1 \geq \Psi^2$ if and only if $\Psi_k^1 \geq \Psi_k^2$ for all $k \in \mathcal{N}$.

In the following the monotonicity properties are meant to be with respect to the above partial order.

Lemma 3.1: 1. If $A_k(M, \Psi) - A_k(M+1, \Psi)$ is increasing in Ψ for all M, so is $J_k(M, \Psi, s) - J_k(M+1, \Psi, s)$.

2. If $A_k(M, \Psi) - A_k(M+1, \Psi)$ is decreasing in M for all Ψ, so is $J_k(M, \Psi, s) - J_k(M+1, \Psi, s)$.

Proof: Observe that

$$J_k(M, \Psi, s) - J_k(M+1, \Psi, s)$$
$$= J_k(M, \Psi, s) - A_k(M+1, \Psi)$$
$$\quad + A_k(M+1, \Psi) - J_k(M+1, \Psi, s)$$
$$= \min\{A_k(M, \Psi) - A_k(M+1, \Psi), \eta\}$$
$$\quad + \max\{0, A_k(M+1, \Psi) - A_k(M+2, \Psi) - \eta\},$$

from which part 1 as well as part 2 directly follow. ∎

Lemma 3.2: 1. For all k and M, the mapping $\Psi \mapsto A_k(M, \Psi) - A_k(M+1, \Psi)$ is monotonically increasing on $\Theta(M)$.

---

²For the modified MDP to be defined soon in the next section, cost functions and action space spaces satisfy hypotheses of Propositions 1.1, 1.3 and 1.6 in [5, Chapter 3]. Hence we use the corresponding results wherever needed.

---

2. For all k and Ψ, the mapping $M \mapsto A_k(M, \Psi) - A_k(M+1, \Psi)$ is monotonically decreasing on $\mathcal{N}$.

Proof: 1. The validity of this claim for $k = 1$ follows from

$$A_1(M, \Psi) - A_1(M+1, \Psi) = \frac{\prod_{j \in \mathcal{N}} \Psi_j}{(M+N)(M+N+1)\lambda}.$$

Now assume that $A_k(M, \Psi) - A_k(M+1, \Psi)$ is increasing in Ψ. Then, from Lemma 3.1, part 1, the same holds true for $J_k(M, \Psi, s) - J_k(M+1, \Psi, s)$. Then, using (5b) and (5c), we get

$$A_{k+1}(M, \Psi) - A_{k+1}(M+1, \Psi)$$
$$= \frac{\prod_{j \in \mathcal{N}} \Psi_j}{(M+N)(M+N+1)\lambda}$$
$$\quad + \mathbb{E}_\Delta \Big\{ \frac{N-2-M}{N} \Big(J_k(M, \Psi^M, s) - J_k(M+1, \Psi^{M+1}, s) \Big)$$
$$\quad + \sum_{j \in \mathcal{M}} \frac{\Psi_j^M}{N} \Big(A_k(M, \Psi^{M,j}) - A_k(M+1, \Psi^{M+1,j}) \Big)$$
$$\quad + \frac{1}{N} \Big(J_k(M, \Psi^M, s) - \Psi_{M+1}^{M+1} A_k(M+1, \Psi^{M+1,M+1}) \Big) \Big\}$$

where we have used the fact that $\Psi_j^{M+1} = \Psi_j^M$ for all $j \in \mathcal{M}$. Using induction on k and Lemma 3.1, part 1, it can be shown that $J_k(M, \Psi^M, s) - J_k(M+1, \Psi^{M+1}, s)$ and $A_k(M, \Psi^{M,j}) - A_k(M+1, \Psi^{M+1,j})$, $j \in \mathcal{M}$ are increasing in Ψ. We omit the details here. The final term can be written as

$$J_k(M, \Psi^M, s) - \Psi_{M+1}^{M+1} A_k(M+1, \Psi^{M+1,M+1})$$
$$= J_k(M, \Psi^M, s) - \exp(-\lambda\Delta) A_k(M+1, \Psi^M)$$
$$= \min\{A_k(M, \Psi^M) - A_k(M+1, \Psi^M), \eta\}$$
$$\quad + (1 - \exp(-\lambda\Delta)) A_k(M+1, \Psi^M).$$

The first equality follows because $\Psi_{M+1}^{M+1} = \exp(-\lambda\Delta)$ and $\Psi^{M+1,M+1} = \Psi^M$ for all $\Psi \in \Theta(M)$. The proof of Theorem 3.1, part 1 below shows that $\Psi \mapsto A_k(M+1, \Psi^M)$ is increasing on $\Theta(M)$. From this, the monotonicity of the left hand side is easily seen. This proves the claim for $k+1$.

2. This can be shown using arguments similar to those in the proof of the first part. We omit the details here. ∎

1) Monotonicity of the Value Function:

Theorem 3.1: The value function has the following properties:

1. For all M and S, $\Psi \mapsto J(M, \Psi, S)$ is monotonically increasing on $\Theta(M)$,

2. For all Ψ and S, $M \mapsto J(M, \Psi, S)$ is monotonically decreasing on $\bar{\mathcal{N}}$,

3. For all M and Ψ, $J(M, \Psi, s) \leq J(M, \Psi, i)$.

Proof: 1. Clearly, for all M, $A_1(M, \Psi)$ is increasing in Ψ. From (5b) and (5c), both $J_1(M, \Psi, s)$ and $J_1(M, \Psi, i)$ are increasing in Ψ. Now assume that $J_k(M, \Psi, s)$ and $J_k(M, \Psi, i)$ are increasing in Ψ. By definition,

$$A_{k+1}(M, \Psi) = \frac{\prod_{j \in \mathcal{N}} \Psi_j}{(M+N)\lambda} + \mathbb{E}_\Delta \Big\{ \frac{N-1-M}{N} J_k(M, \Psi^M, s)$$
$$\quad + \sum_{j \in \mathcal{M}} \frac{\Psi_j^M}{N} J_k(M, \Psi^{M,j}, i) \Big\}.$$

Observe that $\Gamma : [0,1]^{\mathcal{N}} \to [0,1]^{\mathcal{N}}$ and $\Upsilon : [0,1]^{\mathcal{N}} \to [0,1]^{\mathcal{N}}$ are monotone increasing functions. Furthermore, since $\Psi^{M,j}$ does not depend on Ψ_j, $\Psi_j^M J_k(M, \Psi^{M,j}, i)$ is increasing in Ψ_j. Hence, the term inside the expectation is increasing in Ψ for all $\Delta \geq 0$. This shows that $A_{k+1}(M, \Psi)$ is increasing in Ψ. The claim follows by taking limits as $k \to \infty$.

2. It is sufficient to prove that $M \mapsto A_k(M, \Psi)$ is monotonically decreasing on $\bar{\mathcal{N}}$ for all k. First assume $M > 0$, and consider Ψ^o such that $\Psi_j^o = 0$ for all $j \leq M$ and $\Psi_j^o = 1$ otherwise. Then $A_k(M, \Psi^o) - A_k(M+1, \Psi^o) = 0$. Also note that only $\Psi \geq \Psi^o$ are of interest. For all such Ψ, from Lemma 3.2, part 1,

$$A_k(M, \Psi) - A_k(M+1, \Psi) \geq 0, \text{ for all } k.$$

To complete the proof we need to show that $A_k(0, \Psi^1) - A_k(1, \Psi^1) \geq 0$, for all k, where $\Psi^1 := (1, \ldots, 1)$ is the only element of the singleton set $\Theta(0)$. This can be shown using induction on k. We omit the details here.

3. From (5b) and (5c), $J_k(M, \Psi, s) \leq J_k(M, \Psi, i)$ for all k. The claim follows by taking $k \to \infty$. ∎

2) Monotonicity of the Optimal Policy: The following is the main theorem of this section.

Theorem 3.2: The optimal policy $u^* : \bar{\mathcal{N}} \times [0,1]^{\mathcal{N}} \times \{s\} \to \mathcal{U}$ has the following properties.

1. For all M, $\Psi \mapsto u^*(M, \Psi, s)$ is monotonically increasing on $\Theta(M)$.

2. For all Ψ, $M \mapsto u^*(M, \Psi, s)$ is monotonically decreasing on $\bar{\mathcal{N}}$.

Proof: 1. Lemma 3.2, part 1 shows that for all M, $A_k(M, \Psi) - A_k(M+1, \Psi)$ is increasing in Ψ. Taking limit as $k \to \infty$, $A(M, \Psi) - A(M+1, \Psi)$ is increasing in Ψ. Thus for $\Psi^1 \geq \Psi^2$, $A(M, \Psi^2) \geq A(M+1, \Psi^2) + \eta$ implies $A(M, \Psi^1) \geq A(M+1, \Psi^1) + \eta$, i.e., $u^*(M, \Psi^2, s) = 1$ implies $u^*(M, \Psi^1, s) = 1$. This proves the claim.

2. We use Lemma 3.2, part 2 here, and the arguments are similar to those in the proof of the first part. ∎

Hence, for every belief vector Ψ, there exists a threshold on the number of copies: the source stops copying once the number of copies reaches this threshold. Moreover, the threshold increases with Ψ.

Monotonicity with respect to MLR ordering: In POMDP literature, under various constraints, monotonicity of the optimal policy with respect to *monotone likelihood ratio* (MLR) ordering has been shown (see [11], [13]). To define the MLR order over the space of probability mass functions, the underlying state space should be completely ordered. In our case, $2^{\mathcal{N}}$ is not completely ordered. However, we consider a partial order (determined by *set inclusion*) on $2^{\mathcal{N}}$, and introduce the notion of *generalized monotone likelihood ratio* (GMLR), denoted as $\geq_{gr}$, and defined as follows.

Definition 3.2: For $\Phi^1, \Phi^2 \in \mathcal{P}(2^{\mathcal{N}})$, $\Phi^1 \geq_{gr} \Phi^2$ if

$$\Phi^1(\mathcal{R}_1)\Phi^2(\mathcal{R}_2) \geq \Phi^1(\mathcal{R}_2)\Phi^2(\mathcal{R}_1)$$

for all $\mathcal{R}_1 \supseteq \mathcal{R}_2$, $\mathcal{R}_1, \mathcal{R}_2 \in 2^{\mathcal{N}}$.

Now, let $\Psi^1, \Psi^2 \in [0,1]^{\mathcal{N}}$ be as defined in Section III-B. They will induce probability mass functions, $\Phi^1, \Phi^2 \in \mathcal{P}(2^{\mathcal{N}})$, on the sets of recovered relays. Ψ^i and Φ^i are related as

$$\Phi^i(\mathcal{R}) = \prod_{j \in \mathcal{R}} (1 - \Psi_j^i) \prod_{j \notin \mathcal{R}} \Psi_j^i, \; i = 1, 2. \quad (8)$$

The following result is an easy observation.

Proposition 3.1: $\Psi^1 \leq \Psi^2 \iff \Phi^1 \geq_{gr} \Phi^2$.

Proof: See Appendix A. ∎

(8) allows as to map the set of states $\bar{\mathcal{N}} \times [0,1]^{\mathcal{N}} \times \mathcal{S}$ to another set $\bar{\mathcal{N}} \times \mathcal{P}(2^{\mathcal{N}}) \times \mathcal{S}$. The optimal policy $u^*(M, \Psi, S)$ can also be mapped to a policy $\tilde{u}^*(M, \Phi, S)$ defined over the latter set as $\tilde{u}^*(M, \Phi, S) = u^*(M, \Psi, S)$ where Φ and Ψ are related as in (8). Then we have the following result.

Corollary 3.1: $\tilde{u}^*(M, \cdot, s)$ is monotonically decreasing on $(\mathcal{P}(2^{\mathcal{N}}), \geq_{gr})$.

IV. A SUBOPTIMAL CONTROL

In this section we formulate and analyze a suboptimal control problem. The idea is to replace random state transitions with the expected state transition which is similar in spirit to the one proposed by White [15]. Our approximations lead to explicit formula for a simpler policy that depends on Ψ only through the product of its components

$$\phi := \prod_{j \in \mathcal{N}} \Psi_j, \quad (9)$$

which is the probability that the destination has not received the packet. It can be seen that $A_1(M, \Psi)$ depends on Ψ only through ϕ (see (6)). Denote it by

$$\hat{A}_1(M, \phi) = \frac{\phi}{(M+N)\lambda},$$

$$\text{and } \hat{J}_1(M, \phi, s) = \min\{\hat{A}_1(M, \phi), \eta + \hat{A}_1(M+1, \phi)\}.$$

Clearly, $\hat{J}_1(M, \phi, s)$ is concave and increasing in ϕ, and $\hat{A}_1(M, 0) = 0$. It can also be seen that

$$\Psi_j^M A_1(M, \Phi^{M,j}) = \hat{A}_1(M, \phi).$$

Let us replace $A_k(M, \Psi)$ by another function $\hat{A}_k(M, \phi)$ which depends on Ψ only through ϕ and satisfies $A_k(M, \Phi) \leq \hat{A}_k(M, \phi)$. Moreover, $\hat{A}_k(M, \phi)$ is concave, increasing in ϕ, and $\hat{A}_k(M, 0) = 0$. Naturally, given $\hat{A}_k(M, \phi)$, we define,

$$\hat{J}_k(M, \phi, s) = \min\{\hat{A}_k(M, \phi), \eta + \hat{A}_k(M+1, \phi)\}. \quad (10)$$

It is immediate that the above properties for $\hat{A}_k$ hold for $\hat{J}_k(M, \phi, s)$ as well. In particular,

$$J_k(M, \Psi, S) \leq \hat{J}_k(M, \phi, S), \; S \in \{s, i\}. \quad (11)$$

All these facts trivially hold for $k = 1$. We next proceed to define $\hat{A}_k(M, \phi)$ and prove the aforementioned properties inductively. So assume $\hat{A}_k(M, \phi)$ is defined and satisfies all the above properties. Let us revisit (6), and observe that

$$\begin{aligned}
\Psi_j^M A_k(M, \Psi^{M,j}) &\leq \Psi_j^M \hat{A}_k\left(M, \frac{\prod_{l \in \mathcal{N}} \Psi_l^M}{\Psi_j^M}\right) \\
&\leq \hat{A}_k\left(M, \prod_{l \in \mathcal{N}} \Psi_l^M\right)
\end{aligned}$$

where the first inequality follows from the induction hypothesis and the last inequality follows from the hypotheses that $\hat{A}_k(M, \phi)$ is concave in ϕ, and $\hat{A}_k(M, 0) = 0$. We consider the right most expression as an approximation of the left most one, and define, in analogy to (6),

$$\hat{A}_{k+1}(M,\phi) = \frac{\phi}{(M+N)\lambda} + \mathbb{E}_\Delta\Big\{\tfrac{N-1-M}{N}\hat{J}_k(M,\phi^{(M,\Delta)},s)$$
$$+ \tfrac{M}{N}\hat{A}_k(M,\phi^{(M,\Delta)})\Big\}. \quad (12)$$

where $\phi^{(M,\Delta)} := \phi\exp(-M\lambda\Delta)$. The induction hypotheses imply $A_{k+1}(M,\Psi) \le \hat{A}_{k+1}(M,\phi)$. Obviously, $\hat{A}_{k+1}(M,\phi,s)$ is also concave, increasing in ϕ, and $\hat{A}_{k+1}(M,0) = 0$. This completes the induction step.

What we have is another MDP whose state at epoch T_k is $\hat{X}_k = (M_k,\phi_k,S_k)$. The state space is $(\bar{\mathcal{N}} \times [0,1] \times \mathcal{S}) \cup \{\tau\}$ where τ is the terminal state as in Section III-B. The action at epoch T_k is $\hat{U}_k \in \mathcal{U}$ and $C_k = \hat{U}_k 1_{\{S_k=s\}}$ as before. The observation tuple (Δ_{k+1}, S_{k+1}) is the random disturbance at epoch T_k. If $S_{k+1} \in \{r,d\}$ then $X_{k+1} = \tau$. Transitions of M_k and S_k happen as before, and $\phi_{k+1} = \phi_k^{(M_k,\Delta_k)}$. The expected single stage cost is

$$\hat{g}(\hat{X}_k, U_k) = \eta C_k + \frac{\phi_k}{(M_k + C_k + N)\lambda}$$

with $\hat{g}(\tau) = 0$. Policies are also defined as in Section III-B. The value iteration algorithm for this MDP is given by

$$\hat{J}_0(M,\phi,S) = 0, \ S \in \{s,i\};$$
$$\text{for } k = 1,2,\ldots,$$
$$\hat{J}_k(M,\phi,s) = \min\{\hat{A}_k(M,\phi), \eta + \hat{A}_k(M+1,\phi)\},$$
$$\hat{J}_k(M,\phi,i) = \hat{A}_k(M,\phi).$$

where $\hat{A}_k(M,\phi)$ is as defined in (12).

A. An Optimal Policy for the Modified MDP

Let $\hat{J}(M,\phi,S)$ be the optimal value function, and $\hat{u} : (\bar{\mathcal{N}} \times [0,1] \times \mathcal{S}) \cup \{\tau\} \to \mathcal{U}$ be the optimal policy for this MDP. We need the following lemma.

Lemma 4.1: For all k and M, $\hat{A}_k(M,\phi) - \hat{A}_k(M+1,\phi)$ is increasing in ϕ.

Proof: The arguments involved are analogous to those in the proof of Lemma 3.2, part 1. We omit the details here. ∎

Lemma 4.2: $\hat{u}(M,\phi,s)$ is monotonically increasing in ϕ.

Proof: Define,

$$\hat{u}_k(M,\phi,s) = \begin{cases} 0 & \text{if } \hat{A}_k(M,\phi) \le \eta + \hat{A}_k(M+1,\phi), \\ 1 & \text{otherwise.} \end{cases}$$

Then $\hat{u}_k(M,\phi,s)$ is increasing in ϕ follows from Lemma 4.1. A limiting argument as before proves the result. ∎

The following is the main result of this section.

Theorem 4.1: The optimal policy $\hat{u}$ exhibits a threshold behavior:

$$\hat{u}(M,\phi,s) = \begin{cases} 1 & \text{if } \phi > \eta\lambda(M+1)(M+2), \\ 0 & \text{otherwise.} \end{cases}$$

In particular, $\hat{u}(M,\phi,s)$ is monotonically decreasing in M.

Proof: Let us write Bellman equation for the hatted value function.

$$\hat{J}(M,\phi,s)$$
$$= \min\Big\{\hat{A}(M,\phi), \eta + \hat{A}(M+1,\phi)\Big\}$$
$$= \min\Big\{\frac{\phi}{(M+N)\lambda} + \mathbb{E}_\Delta\Big\{\frac{N-1-M}{N}\hat{J}(M,\phi^{(M,\Delta)},s)$$
$$+ \frac{M}{N}\hat{A}(M,\phi^{(M,\Delta)})\Big\},$$
$$\eta + \frac{\phi}{(M+N+1)\lambda}$$
$$+ \mathbb{E}_\Delta\Big\{\frac{N-2-M}{N}\hat{J}(M+1,\phi^{(M+1,\Delta)},s)$$
$$+ \frac{M+1}{N}\hat{A}(M+1,\phi^{(M+1,\Delta)})\Big\}\Big\}.$$

Suppose that $\hat{u}(M,\phi,s) = 0$, i.e., $\hat{J}(M,\phi,s) = \hat{A}(M,\phi)$. In the following we are only interested in the polices which are consistent with Lemma 4.2, namely, $\hat{u}(M,\phi,s) = 0$ implies that $\hat{u}(M,\phi^{(M,\Delta)},s) = 0$ and so $\hat{J}(M,\phi^{(M,\Delta)},s) = \hat{A}(M,\phi^{(M,\Delta)})$, for all $\Delta > 0$. Under this constraint, one can easily show that

$$\hat{J}(M,\phi,s) = \frac{\phi}{(M+1)\lambda}.$$

Next we prove that if $\hat{u}(M+1,\phi,s) = 0$ and $\hat{u}(M,\phi,s) = 1$, then $\hat{u}(M',\phi,s) = 1$ for all $M' < M$. From the Bellman equation

$$\hat{J}(M,\phi,s) = \eta + \hat{A}(M+1,\phi)$$
$$= \eta + \hat{J}(M+1,\phi,s)$$
$$= \eta + \frac{\phi}{(M+2)\lambda}$$
$$< \frac{\phi}{(M+1)\lambda},$$

where the last inequality arises because the right hand side is the cost of $\hat{u}(M,\phi,s) = 0$, which is assumed to be suboptimal. Thus

$$\eta < \frac{\phi}{(M+1)(M+2)\lambda}. \quad (13)$$

Hence,

$$\eta < \frac{\phi}{(M'+1)(M'+2)\lambda}, \ \text{ for all } M' < M.$$

But this implies

$$\hat{J}(M',\phi,s) \le \eta + \hat{A}(M'+1,\phi)$$
$$= \eta + \frac{\phi}{(M'+2)\lambda}$$
$$< \frac{\phi}{(M'+1)\lambda}$$
$$= \hat{A}(M',\phi),$$

which proves the claim that $\hat{u}(M',\phi,s) = 1$. Finally, for a fixed ϕ, $\hat{u}(M,\phi,s) = 1$ if and only if M satisfies (13). ∎

We now make some remarks on the usefulness of this policy $\hat{u}$ for the original POMDP.

Remarks 4.1: 1. The optimal policy for the modified control problem can naturally be mapped to a suboptimal policy for the original control problem.

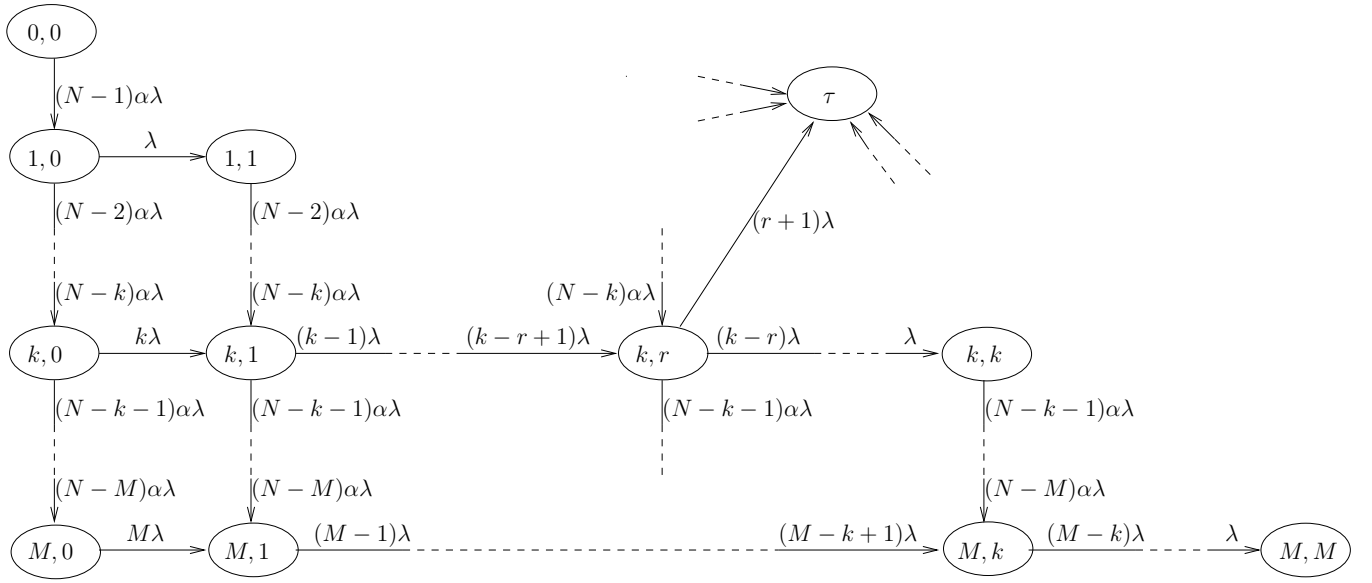$$u(M,\Psi,S) = \hat{u}(M,\phi,S),$$

where ϕ is given by (9).

Fig. 2. Transition rate diagram for the CTMC. Though the diagram does not show all of them, the transition rate from any state (k, r) to τ is $(r+1)\lambda$.

2. Clearly $J(M, \Psi, S) \leq J^u(M, \Psi, S)$. Moreover, $J^u(M, \Psi, S) \leq \hat{J}(M, \phi, S)$, for all M, Ψ and S. This follows because $J_k^\pi(M, \Psi, S) \leq \hat{J}_k^\pi(M, \phi, S)$ for all k and for any policy $\pi \in \Pi$ (See (11)). We thus not only have a suboptimal policy for the original POMDP, but also an upper bound on the cost of this policy.

3. Neglia and Zhang [12] give the following optimal policy for *epidemic* relaying with complete information at all the nodes.

$$\bar{u}(M) = \begin{cases} 1 & \text{if } (M+1)(M+2) < \frac{1}{\eta\lambda}, \\ 0 & \text{otherwise.} \end{cases}$$

The above policy remains optimal even in the setting of two-hop relaying with complete information at the source. This may be viewed as a special case of our setup with $\phi = 1$ at every decision epoch. In this special case, the suboptimal policy derived in this paper, reduces to above policy.

V. NUMERICAL RESULTS

Here we study the performance of the suboptimal closed loop policy proposed in Section IV. First we evaluate the following two open loop schemes that do not exploit either the copying history or the source's beliefs.

1. The source copies to at most M susceptible relays. We optimize over $M \in \{0, \ldots, N-1\}$.

2. The source copies to each susceptible relay it meets with probability α. We optimize over $\alpha \in [0, 1]$.

We call these schemes POOL (partially observed open loop) 1 and POOL 2 respectively. In both schemes the source stops copying if it meets either a recovered relay or the destination. We also compare with the performance of the COMDP (closed loop control with complete information) studied in [12].

To start, let us consider that the source copies with probability α to each of the first M susceptible relay it meets until it has met either a recovered relay or the destination. The state of the network is represented by the tuple $(K(t), R(t))$ where $K(t) \in \{0, \ldots, M\}$ and $R(t) \in \{0, \ldots, K(t)\}$ are the number of infected relays and the number of recovered relays respectively, by time t. As in Section III, when the source meets a recovered relay or the destination, the system enters in a terminal state τ. Under the assumptions made in Section II, $(K(t), R(t))$ is a finite-state, continuous time Markov chain (CTMC) with an absorbing state τ. The transition rate diagram for the CTMC is shown in Figure 2. Though the diagram does not show all transitions to τ, the transition rate from any state (k, r) to τ is $(r+1)\lambda$. In particular, let $p(k, r)$ be the probability that the system visits state (k, r). These probabilities are calculated recursively as follows.

$$p(0, 0) = 1,$$
$$p(k, 0) = \frac{p(k-1, 0)(N-k)\alpha}{\alpha N + (1-\alpha)k}, \quad 1 \leq k \leq M$$
$$p(k, r) = \frac{p(k-1, r)(N-k)\alpha}{\alpha N + (1-\alpha)k} + \frac{p(k, r-1)(k-r+1)}{\alpha N + (1-\alpha)(k+1)},$$
$$1 \leq k < M, 1 \leq r \leq k,$$
$$p(M, r) = \frac{p(M-1, r)(N-M)\alpha}{\alpha N + (1-\alpha)M} + \frac{p(M, r-1)(M-r+1)}{M+1},$$
$$1 \leq r \leq M.$$

Let K be the random number of infected relays before the system enters state τ. Then

$$\begin{aligned} \mathbb{E}\{\mathcal{E}_c\} &= \sum_{k=0}^{M} Pr(K=k)\mathbb{E}\{\mathcal{E}_c | K=k\} \\ &= \sum_{k=0}^{M-1} \sum_{r=0}^{k} \frac{p(k, r)(r+1)}{\alpha N + (1-\alpha)(k+1)}(k+1) \\ &\quad + \sum_{r=0}^{M} \frac{p(M, r)(r+1)}{M+1}(M+1) \end{aligned}$$

Similarly let T_{kr} be the sojourn time in state (k, r) once the
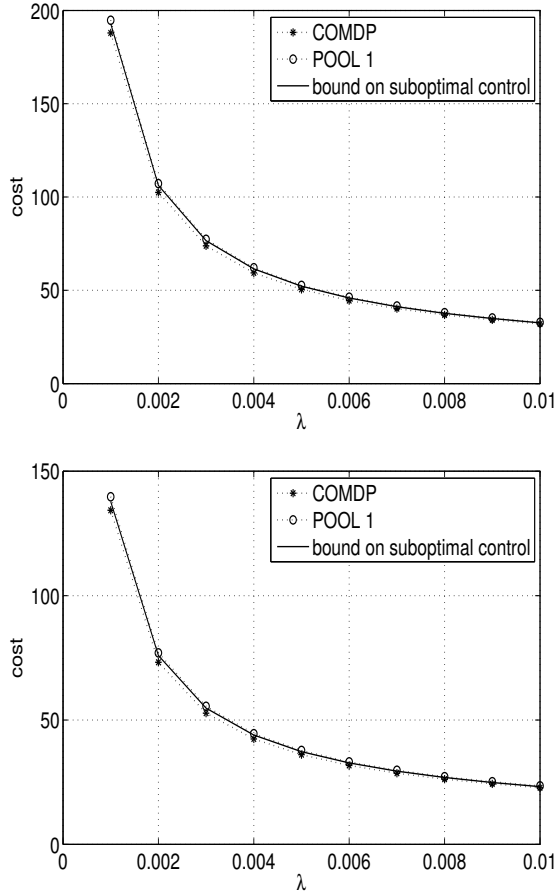
Fig. 3. Cost as a function of λ. The first set of plots is for the number of nodes $N+1 = 51$ and $\eta = 2$ while the second set of plots is for $N+1 = 101$ and $\eta = 1$.

system enters (k, r). Only T_{k0}s account for the delay. So

$$
\begin{aligned}
\mathbb{E}\{\mathcal{T}_d\} &= \sum_{k=0}^{M} p(k,0)\mathbb{E}\{T_{k0}\} \\
&= \sum_{k=0}^{M-1} \frac{p(k,0)}{(\alpha N + (1-\alpha)(k+1))\lambda} + \frac{p(M,0)}{(M+1)\lambda}
\end{aligned}
$$

Define $V(M, \alpha) := \mathbb{E}\{\mathcal{T}_d + \eta\mathcal{E}_c\}^3$ to be the cost associated with this system. Clearly, POOL 1 chooses M which minimizes $V(M, 1)$ while POOL 2 chooses α which minimizes $V(N-1, \alpha)$. Moreover, to study the problem with complete information one can define another terminal state

$$
\bar{\tau} := \tau \cup \{(k, r) : 1 \le k \le M, 1 \le r \le k\}.
$$

Assume $\bar{K}$ to be the number of infected relays before the system enters state $\bar{\tau}$ and redefine $\mathbb{E}\{\mathcal{E}_c\}$ as before. $\mathbb{E}\{\mathcal{T}_d\}$ remains same. Let $\bar{V}(M, 1)$ be the associated cost. Then the optimal control chooses M that minimizes $\bar{V}(M, 1)$. In Figure 3, we plot the optimal value of the modified problem as

---

[3]In the expressions of $\mathbb{E}\{\mathcal{E}_c\}$ and $\mathbb{E}\{\mathcal{T}_d\}$ the first summations on the right hand sides disappear if $M = 0$.

a function of λ. In view of Remark 4.1, part 2, this is an upper bound on the cost of original problem with the suboptimal policy. We also plot costs of COMDP and POOL 1. The two sets of plots are for $N+1 = 51, \eta = 2$ and $N+1 = 101, \eta = 1$ respectively. It is observed that the three policies perform close to each other and the suboptimal control marginally outperforms the open loop control.

VI. CONCLUSION

We studied optimal two-hop relaying problem in DTNs that employ an enhanced version of "immune" recovery scheme. In particular, we formulated the problem as a POMDP and characterized the optimal policy (Section III). We also derived a suboptimal policy (Section IV) in an explicit form. An optimal policy with complete information at the source gives the least cost among all the policies. On the other hand open loop policies that do not exploit either the copying history or the source's beliefs are likely to perform worse. We compared the performances of these policies and that of the proposed suboptimal policy. The costs of the COMDP and the open loop policy are close to each other implying that the performance gain of the closed loop control is not substantial. Furthermore, the more efficient recovery schemes, e.g., *immune_tx* and *vaccine*, are also not likely to bring substantial performance gains.

APPENDIX A
PROOF OF PROPOSITION 3.1

The only if part: It is sufficient to consider Ψ^1, Ψ^2 which differ in only one, say jth coordinate. Let $\Psi_j^1 < \Psi_j^2$. Now, fix $\mathcal{R}_1 \supset \mathcal{R}_2$.

1) If $j \in \mathcal{R}_2$ or $j \notin \mathcal{R}_1$, $\frac{\Phi^i(\mathcal{R}_1)}{\Phi^i(\mathcal{R}_2)}$ will be same for $i = 1, 2$.
2) If $j \notin \mathcal{R}_2$ but $j \in \mathcal{R}_1$,

$$
\frac{\Phi^1(\mathcal{R}_1)}{\Phi^1(\mathcal{R}_2)} = \frac{1 - \Psi_j^1}{\Psi_j^1} > \frac{1 - \Psi_j^2}{\Psi_j^2} = \frac{\Phi^2(\mathcal{R}_1)}{\Phi^2(\mathcal{R}_2)}.
$$

The if part: We prove this by contraposition. Consider Ψ^1, Ψ^2 and define $\mathcal{R}_1, \mathcal{R}_2$ such that $\Psi_j^1 > \Psi_j^2$ for all $j \in \mathcal{R}_1$ and $\Psi_j^2 > \Psi_j^1$ for all $j \in \mathcal{R}_2$. If $\Psi^1 \not\le \Psi^2$ then $\mathcal{R}_1 \ne \emptyset$. We now argue that $\Phi^1 \not\ge_{gr} \Phi^2$. Indeed, it can be easily verified that

$$
\frac{\Phi^1(\mathcal{R}_1 \cup \mathcal{R}_2)}{\Phi^1(\mathcal{R}_2)} < \frac{\Phi^2(\mathcal{R}_1 \cup \mathcal{R}_2)}{\Phi^2(\mathcal{R}_2)}.
$$

REFERENCES

[1] Eitan Altman, Amar Prakash Azad, Francesco De Pellegrini, and Tamer Basar. Optimal activation and transmission control in delay tolerant networks. In *Proceedings of IEEE INFOCOM*, San Diego, CA, USA, March 2010.

[2] Eitan Altman, Tamer Basar, and Francesco De Pellegrini. Optimal monotone forwarding policies in delay tolerant mobile ad-hoc networks. In *Proceedings of the 3rd International Conference on Performance Evaluation Methodologies and Tools*, Athens, Greece, October 2008.

[3] Eitan Altman, Giovanni Neglia, Francesco De Pellegrini, and Daniele Miorandi. Decentralized stochastic control of delay tolerant networks. In *Proceedings of IEEE INFOCOM*, Rio de Janeiro, Brazil, April 2009.

[4] Dimitri P. Bertsekas. *Dynamic Programming and Optimal Control, Vol. 1*. Athena Scientific, Belmont, Massachusetts, 3rd edition, 2005.

[5] Dimitri P. Bertsekas. *Dynamic Programming and Optimal Control, Vol. 2*. Athena Scientific, Belmont, Massachusetts, 3rd edition, 2007.

[6] Scott Burleigh, Adrian Hooke, Leigh Torgerson, Kevin Fall, Vint Cerf, Bob Durst, Keith Scott, and Howard Weiss. Delay-tolerant networking: An approach to interplanetary internet. *IEEE Communications Magazine*, 41(6):128–136, June 2003.

[7] Robin Groenevelt, Philippe Nain, and Ger Koole. The message delay in mobile ad hoc networks. *Performance Evaluation*, 62(1–4):210 – 228, October 2005.

[8] Matthias Grossglauser and David N. C. Tse. Mobility increases the capacity of ad-hoc wireless networks. *IEEE/ACM Transactions on Networking*, 10(4):477–486, August 2002.

[9] Zygmunt J. Haas and Tara Small. A new networking model for biological applications of ad hoc sensor networks. *IEEE/ACM Transactions on Networking*, 14(1):27–40, February 2006.

[10] Ahmad Al Hanbali, Philippe Nain, and Eitan Altman. Performance of ad hoc networks with two-hop relay routing and limited packet lifetime. *Performance Evaluation*, 65(6–7):463–483, June 2008.

[11] William S. Lovejoy. Some monotonicity results for partially observed markov decision processes. *Operations Research*, 35(5):736–743, September–October 1987.

[12] Giovanni Neglia and Xiaolan Zhang. Optimal delay-power tradeoff in sparse delay tolerant networks: a preliminary study. In *SIGCOMM Workshop on Challenged Networks*, pages 237–244, Pisa, Italy, September 2006.

[13] U. Rieder. Structural results for partially observed control models. *Mathematical Methods of Operations Research*, 35(6):473–490, November 1991.

[14] Amin Vahdat and David Becker. Epidemic routing for partially-connected ad hoc networks. Technical Report CS-2000-06, Duke University, July 2000.

[15] D. J. White. Approximating markov decision processes using expected state transitions. *Journal of Mathematical Analysis and Applications*, 107(1):167–181, April 1985.

[16] Xiaolan Zhang, Giovanni Neglia, Jim Kurose, and Don Towsley. Performance modeling of epidemic routing. *Computer Networks*, 51(10):2867–2891, July 2007.