Online Algorithms for Adaptive Optimization in Heterogeneous Delay Tolerant Networks

Wissam Chahin[1,*], Francesco De Pellegrini[2], Rachid El-Azouzi[1], Amar Prakash Azad[3]

[1]CERI/LIA, University of Avignon, 339, chemin des Meinajaries, Avignon, France
[2]CREATE-NET Via alla Cascata 56/D, Povo, Trento, Italy
[3]SOE, UCSC, USA

Abstract

Delay Tolerant Networks (DTNs) are an emerging type of networks which do not need a predefined infrastructure. In fact, data forwarding in DTNs relies on the contacts among nodes which may possess different features, radio range, battery consumption and radio interfaces. On the other hand, efficient message delivery under limited resources, e.g., battery or storage, requires to optimize forwarding policies. We tackle optimal forwarding control for a DTN composed of nodes of different types, forming a so-called heterogeneous network. Using our model, we characterize the optimal policies and provide a suitable framework to design a new class of multi-dimensional stochastic approximation algorithms working for heterogeneous DTNs. Crucially, our proposed algorithms drive online the source node to the optimal operating point without requiring explicit estimation of network parameters. A thorough analysis of the convergence properties and stability of our algorithms is presented.

1. Introduction

Delay Tolerant Networks (DTNs) are designed to sustain communications in networked systems, where persistent end-to-end connectivity cannot be guaranteed [1–3]. In DTNs, messages are carried from source to destination via relay nodes adopting the so-called "store and carry" forwarding, which leverages on nodes' mobility pattern. The key problem in DTNs is thus to efficiently route messages towards the intended destination. It is worth observing that traditional routing techniques would fail in this context due to frequent disruptions. Furthermore, mobile nodes rarely possess information on the upcoming encounters they are going to experience [4]. A number of schemes have been proposed for efficient message forwarding in DTNs [3, 5, 6].

Disseminating multiple copies of the message in the network is the straightforward routing solution to overcome disconnections. This ensures that at least some of them will reach the destination within some

deadline [7, 8] with high probability. This technique is named epidemic forwarding [9], in analogy to spread of infectious diseases.

In literature, several variants of epidemic forwarding exist, including *spray and wait* [7] and *two hop routing protocol* [10], implementing different trade-offs between delay and number of released copies. We confine our analysis to the *two hop routing protocol* because of two major technical advantages: first, compared to plain epidemic routing it performs natively a better trade-off between the number of released copies and the delivery probability [8]. Second, and most relevant with respect to the algorithmic design we propose later, forwarding control can be fully implemented on board of the source node. Under two hop routing, the source transmits a message copy to mobiles it encounters. A relay forwards the message copy it has to the destination only.

In some literature on DTNs, a common simplifying assumption (adopted for modeling reasons) is that DTN nodes have all similar physical characteristics, e.g., transmission range, mobility patterns, etc. Thus, the network is assumed to be *homogeneous*. Under

*Corresponding author. Email: wissam.chahin@etd.univ-avignon.fr

this assumption, recent work provides insight into the performance of DTNs [8, 10, 11].

However, it is clear that DTN nodes may belong to different categories, e.g., mobiles, laptops, PDAs and thus differ in their transmission range, mobility, etc. A DTN with different types of nodes in turn is classified as *heterogeneous* [12–14]. To this respect, our starting assumption, as in [12] and [13], is that according to their physical characteristics, nodes group into classes homogeneous with respect to routing. More precisely, two nodes belong to the same class if they have *same intermeeting intensities with source and destination nodes*, i.e., the intensities by which they meet source and destination are the same.

In this context, the fundamental question that arises naturally when one models the trade off of network resources for delivery probability (namely, the number of messages or the energy expenditure), is how to exploit diversity of contact patterns to improve performance.

The structure of the paper is as follows. In the next section, we discuss related work and then identify the main contributions of the paper. Section 3 introduces the model. We identify the problem in section 4. In particular, we first aim at deriving the closed-form structure of the optimal forwarding policy for heterogeneous DTNs, both for static and dynamic control policies. Leveraging on the properties of the so determined optimal forwarding policy, in section 5 we design online algorithms which are able to converge to the optimal control policy over time. In particular, we first present the static algorithm then we make a step forward by introducing the dynamic algorithm that uses two-time scale stochastic approximations. It is worth noticing that one of our objectives is to show that online implementation of these algorithms can be made in such a way to depend only on local knowledge at source nodes. We provide extensive numerical results to validate our theoretical derivation in section 6. Section 7 concludes the paper.

In literature, some attempts to address optimal control for heterogeneous DTNs have been performed in [14] and [13]. The algorithmic formulation that we provide here introduces not only a closed form description of the optimal policy, but also an algorithmic distributed implementation suitable for disconnected operations.

2. Related Work and Contribution

With the aim of optimizing network performance, several previous works addressed the control of forwarding schemes in DTNs [4, 8, 10, 11]. The work [10] proposed to control two-hop forwarding and optimized the system performance by choosing the average duration of timers for message discarding.

Authors of [8] considered a homogeneous network and described a general framework for the optimal control of monotone relay policies. The optimal control was proved there to be of dynamic type; the first work claiming the optimality of dynamic policies was [11], limited to epidemic routing. In line with [14], our formulation builds also on multidimensional control, which results in existence of several thresholds. We go a step forward by describing the *closed form structure of the optimal control*. Also, we extend the approach in [15] leveraging stochastic approximation algorithms because they overcome the explicit estimation of network parameters. In fact, such an estimation is per se a difficult task in disconnected systems [16]. Furthermore, we observe that this operation becomes critical in the case of multiple classes of mobiles since a number of such estimates would be required.

In literature, the heterogeneity in mobile ad hoc DTNs is well documented [17, 18]. However, very few papers addressed this aspect from the modeling perspective. One such work is [13]; the authors assume that nodes may migrate from one class to another. In our framework nodes are fixed within one class. Also, in [19], the authors showed that the presence of heterogeneity has controversial effect on the delivery probability. I.e., it cannot be related in a straightforward manner to the performance of the system. In [14] a general setting for the optimality of controls of a DTN was presented. In [12], a routing scheme was proposed in heterogeneous DTNs based on the use of history information to identify the nodes of "highest utility" for routing. In algorithms we present here, source will forward the message according to the node's class and no other a priori information is needed for the source to take the forwarding decision.

Novel Contributions. The main contributions of this work are the following. First, we introduce a new perspective of the optimal forwarding problem in DTNs and characterize its structure. Second, based on this characterization we introduce a class of stochastic approximation algorithms that attain optimality when multiple classes of relays exist and can operate at runtime in spite of the lack of full information on the network state. Moreover, we deeply investigate these algorithms' performance and rigorously prove their convergence to some limit set of Ordinary Differential Equations (ODEs), then we use Lyapunov functions to confirm their stability.

Our work focuses on proposing a general algorithmic approach to optimize forwarding control in a distributive and energy efficient fashion when many classes of nodes co-exist. We believe this approach could be extended to account for protocols other than two hops. In order for this method to be operational in such

scenarios, message replication control could be decentralized in a way that each relay node can control and keep track of the number of message copies it forwards.

3. System Model

Table 1. Glossary of Notations

Symbol	Meaning
$N + 2$	number of nodes
N_i	number of nodes of class i
λ_{si}	pairwise intermeeting intensity for node i with the source
λ_{id}	pairwise intermeeting intensity for node i with the destination
τ	timeout value
Δ	time slot
K	$\lfloor \tau/\Delta \rfloor$
$X_i(n)$	number of infected nodes of class i at time $n\Delta$, (x_i at n=0)
Ψ	energy constraint
Ψ_i	Maximum number of nodes of class i that can be infected
$F_D(t)$	delivery probability at (t)
$\mathbf{U}(.)$	transmission control vector
$U_i(n)$	forwarding probability to class i at time $n\Delta$
$\Pi_I(w)$	projection over I of the value w
h_i	the switch time of a dynamic policy of class i
θ_i	$= \sum_{k=0}^{K-1} U_i(k)$

Consider a network of $N + 2$ mobile nodes (composed of m classes), each equipped with some form of proximity wireless communications. One node, *source node*, has a message to be sent to a destination node. The network is assumed to be sparse, so that, at any time instant, nodes are isolated with high probability. Communication opportunities arise whenever, due to mobility, two nodes get within reciprocal communication range; such events are often named *contacts*. For the ease of reading, we collect all the main symbols used in the paper in Table 1.

The time between contacts of any two nodes, referred to as *intermeeting times*, is assumed to be exponentially distributed [1]. The contact rate is known to converge to a quantity that is independent of number of mobiles N in the network when the N grows large, under the fluid-approximations (see [12, 21, 22]). The validity of such a model has been discussed in [20], and its accuracy has been shown for a number of mobility models (Random

Walker, Random Direction, Random Waypoint) [8, 15]. There exist studies based on traces collected from real-life mobility [1] arguing that intermeeting times may follow a power-law distribution. In [12], it has been shown that the traces and many other exhibit exponential tails after a cutoff point. We choose the exponential intermeeting times model due to the mathematical tractability and the above reasons.

Heterogeneity. Mobiles can have different physical characteristics, such as transmission power, mobility, etc. In our model, nodes' heterogeneity is captured by the distribution of intermeeting times. In particular, intermeeting intensity parameters capture the physical characteristics of nodes [21], dictating the rate at which two nodes meet. More precisely, node i from class k is represented by the tuple $\lambda_i^k = \{\lambda_{si}, \lambda_{id}\}^k$ where λ_{si} (resp. λ_{id}) refers to the inter-meeting intensity of the node i from class k with the source node (resp. with the destination node). We denote the number of mobiles in class k by $N_k \geq 1$, therefore the total number of mobile nodes in m classes are $\sum_{k=1}^{m} N_k = N$. We refer to the m class system as the m-dimensional (mD) system for the sake of brevity.

There can be multiple source-destination pairs, but we assume that at a given time there is a single message, generated by a tagged source node. The message may eventually have many copies spread in the network.[2] For simplicity, we assume that a message is generated at time $t = 0$ and it remains relevant for some time τ. We do not assume any feedback that allows the source or other mobiles to know whether the message has made it successfully to destination within the allotted time τ.

We adopt a probabilistic *two-hop* routing protocol according to which the source passes the message to mobiles that do not have it with some probability $U \in [u_{min}, u_{max}]$. But, a relay node transmits the message only when it meets the destination node. Such relay policy is *monotone* [8] because the number of copies of the message increases over time. The message is called delivered when the destination receives the first copy of the message.

The problem we address in this paper is to *design online algorithms that drive the source node to an optimal operating point. An optimal operating point is such if the corresponding forwarding policy maximizes the probability to deliver the message by time τ under the constraint on the amount of energy spent by the source.*

---

[1] The *inter-meeting* rate is a function of number of nodes, reciprocal transmission power of nodes, mobility parameters, etc. [20]

[2] Results in the subsequent sections are valid for multiple simultaneous source-destination pairs. But, we should additionally assume that the bandwidth is large enough to ensure that the different forwarding processes are independent, though limited to one message per source. Source s, destination d and the N relays can be accordingly reindexed and the analysis in the subsequent sections remains valid for any such pairs.

The energy spent by the source node relates to the number of message copies transmitted. We assume that the energy spent by relay nodes is not a constraint. In fact, under two hop forwarding relays have to transmit at most once. In particular, we assume that i) each transmission consumes a constant amount of energy, and ii) all other activity requires negligible amount of energy. Under these assumptions the source node spends an energy amount that is proportional to the number of message copies.

A naive way to maximize the delivery probability is to maximize the number of infected nodes. However, under the constraint on the source energy budget we need to tackle an optimal forwarding control able to account for heterogeneity of the relay nodes.

We adopt a discrete time model where the time axis is divided into slots of small duration Δ. Time slot k is the interval $[k\Delta, (k + 1)\Delta]$ and the number of slots is equal to $K = [\tau/\Delta]$. Moreover, the control during $[k\Delta, (k + 1)\Delta]$ is a constant, denoted by $U(k)$.

Forwarding Control. The source node has the possibility to control dynamically the forwarding process to relay nodes: it will forward to nodes in class i with probability U_i. This will slow down the generation of message copies within class i, where $U_i : \{0, \ldots, K - 1\} \to [u_{\min}, u_{\max}], i = 1, \ldots, m$. The control policy the source uses can be expressed by the m-dimensional control vector $\mathbf{U} = \{U_1, \ldots, U_m\}$.

Message Delivery Distribution. The source node aims at maximizing the fluid approximation for the cumulative distribution function (CDF) of the delay[3] $F_D(t) := P(T_d \leq t)$. It is based on a generalization of [8]:

$$F_D(t) = 1 - \exp\left(-\sum_j \lambda_{jd} \int_{s=0}^{t} X_j(s)ds\right). \quad (1)$$

Note that because of monotonicity, maximizing $F_D(t)$ in (1) is equivalent to maximizing $\sum_i \lambda_{is} \int_0^t X_i(s)ds$.

Energy Consumption. Let $\epsilon > 0$ be the energy consumed by the source for transmission of a single copy of the message. As explained in the previous section, energy consumed by the source is significant as compared to the energy spent by other relay nodes under two hop routing. Thus, the total energy consumed by the network to generate message copies during $[0, \tau]$ is

$$\mathcal{E}(\tau) = \epsilon \sum_i [X_i(\tau) - X_i(0)].$$

---

[3] The controlled version reported in (1) derives from the separable differential equation in the form $\frac{d}{dt}F_D(t) = \lim_{h \to 0}\frac{\mathbb{P}[T_d > t + h] - \mathbb{P}[T_d > t]}{h} = [1 - F_D(t)]\sum_j \lambda_{jd}X_j(s)$.

Notice that the total amount of energy spent by the source is proportional to the sum of messages (to all classes) and is the same for each message irrespectively of the mobiles' class.

Optimization Problem. The source's goal is to obtain the *multi class optimal policies* by optimizing over the m-dimensional control vector $\mathbf{U} = \{U_1, \ldots, U_m\}$, where $U_i : \{0, \ldots, K - 1\} \to [u_{\min}, u_{\max}], i = 1, \ldots, m$, which solves

$$\max_{\mathbf{U}} F_D(\tau), \quad \text{subject to} \quad \mathbf{X}(\tau) \cdot \mathbf{1} \leq \Psi, \ \mathbf{X}(0) = \mathbf{x}, \quad (2)$$

where $\mathbf{X}(\tau) \cdot \mathbf{1} = \sum_i X_i(\tau)$. The initial condition $\mathbf{X}(0) = \{X_i(0), i \in \{1, \cdots, m\}\}$, constraint Ψ and total number of copies at time 0, i.e., $\mathbf{x}$ (where $\Psi \geq \mathbf{x}\,\mathbf{1}$) are input for the optimization problem. The control vector is $\mathbf{U}(\cdot) = \{U_i(\cdot) : i \in [1, \cdots, m]\}$ with the m-dimensional support $\mathbf{U}(\cdot) \in [u_{\min}, 1]^m$. Recall that maximizing $F_D(\tau)$ is equivalent to maximizing $\sum_j \lambda_{jd} \int_{s=0}^{\tau} X_j(s)ds$.

The optimal control problem reads

$$\max_{\mathbf{U}} \quad J = \sum_{t=1}^{K-1} \int_{k\Delta}^{(k+1)\Delta} \sum_{i=1}^{m} \lambda_{id}X_i(t)dt, \quad (3)$$

$$\text{subject to} \quad X(\tau) = \sum_{i=1}^{m} X_i(\tau) \leq \psi, \quad (4)$$

where $X_i(n)$ denotes number of mobiles, not including the destination for class i, that have a copy of the message at time $n\Delta$ (i.e. at the beginning of the n-th slot), $X_i(0) = x_i$. Under some standard assumptions, it forms a Markov chain with possible states $1, \ldots, N$ (refer to [15]). It is characterized by

$$X_i(n + 1) = X_i(n) + (N_i - X_i(n))(1 - e^{-\lambda_{si}\Delta U_i(n)}), i = 1, \ldots, m. \quad (5)$$

The objective functional can be rewritten (after integrating (3)) as

$$J = \sum_{i=1}^{m} \tau N_i \lambda_{id} - \frac{\lambda_{id}}{\lambda_{si}} \sum_{n=0}^{K-1} (N_i - X_i(n)) \frac{1 - e^{-\lambda_{si}\Delta U_i(n)}}{U_i(n)}. \quad (6)$$

This model provides a useful framework for the subsequent algorithmic development that we detail later in the paper.

4. Optimal Control

Our goal is to find the optimal transmission policy for the problem stated in (2) for each class of mobiles. In the following, we first obtain the optimal policy from the static class of policies in which the *transmission control* is time-invariant for each class of mobile nodes. To avoid cumbersome notation, we derive the optimal static control for two classes, which can be easily extended for general m classes (we

describe this extension later). Note that static control is suboptimal with respect to dynamic optimal control policies [8]. Yet, it is rather convenient for the sake of implementation simplicity.

In the case of dynamic control policies, we will also obtain the optimal policies, which turn out to be of threshold type. Various methodologies have been developed to establish the threshold structure of optimal transmission policies in DTNs: one based on the Pontryagin maximum principle [8], another based on some sample path comparisons[15], some on stochastic ordering, etc. These approaches, developed in the context of DTNs with one type of population, are not applicable to our problem since the model is no longer scalar. Accordingly, we develop a new approach that establishes the optimality of threshold type policies for each class.

Denote $\overline{X}_i(\tau)$ as the number of message copies the source can transmit to nodes of class i by time τ without any control, i.e. $U_i = 1$. Such transmissions are also referred to as uncontrolled transmission [8]. Indeed, a controlled dynamics simply refers to the slowed down transmission [8, eq. 1] in which the number of message under a control U_i is simply $\overline{X}_i(U_i\tau)$. In our multi-class structure, denoting $\mathbf{U} = \{U_1, \cdots, U_m\}$ as the control vector, we refer the total number of message copies in the network by $\overline{\mathcal{X}}(\{\mathbf{U}\tau\})$ which is the sum of message copies over all classes, i.e. $\overline{\mathcal{X}}(\{U_1\tau, \cdots, U_m\tau\}) = \sum_{i=1}^m \overline{X}_i(U_i\tau)$.

4.1. Optimal Static Control Policy

Our goal is to maximize $F_D(\tau)$ while keeping $\mathcal{E}(\tau)$ low so as to satisfy $\sum_{i=1}^m X_i(\tau) \leq \psi$. In view of (1), this is equivalent to maximizing $\sum_i \lambda_{is} \int_0^t X_i(s)ds$. The control for class i is denoted by U_i under the *static* control policy, a control that is constant over time. In what follows, we derive the optimal static policy for 2D (2 node classes) for the sake of exposition clarity. In the subsequent we illustrate that the method can be directly extended for mD (m node classes).

Theorem 1. (Static Optimal Control-2D): Consider the problem of maximizing $F_D(\tau)$ subject to the energy constraint $\mathcal{E}(\tau) \leq \Psi$.

i. If $\overline{\mathcal{X}}(\{u_{\min}\tau, u_{\min}\tau\}) > \psi$, then there is no feasible solution.

ii. If $\overline{\mathcal{X}}(\{\tau, \tau\}) < \psi$, then the policy $\mathbf{U}$ is optimal if and only if $\mathbf{U} = \{1, 1\}$ for $t \in [0, \tau]$ a.e., otherwise,

iii. If $\overline{\mathcal{X}}(\{\tau, u_{\min}\tau\}) \leq \psi$ then the best static policy is $\mathbf{U} = \{1, v_2\}$, where v_2 is a constant.

iv. If $\overline{\mathcal{X}}(\{\tau, u_{\min}\tau\}) > \psi$, it is sufficient to swap the class indexes in the above statement (for the case iii)).

Static policy v_2 is given as

$$v_2 = \frac{\overline{X}_1^{-1}(\psi - \overline{\mathcal{X}}_{-1}(u_{\min}\tau))}{\tau}, \text{ and} \qquad (7)$$

$$\overline{\mathcal{X}}_{-1}(u_{\min}\tau) = \overline{\mathcal{X}}(u_{\min}\tau) - \overline{X}_1(U_1\tau). \qquad (8)$$

Proof: Parts (i) and (ii) are obvious. We show part (iii) and (iv) in the following. From (5), the number of nodes with message copies for each class is simply $X_i(t) = N_i - (N_i - X_i(0))e^{-\lambda_{si}U_i t}$, $0 \leq t \leq \tau$. Hence, we can consider ϕ the bijection such that $\phi_i(U_i) = X_i(\tau)$, $i = 1, 2$. From (1), maximizing $F_D(\cdot)$ is equivalent to maximizing the following (a function of $(X_1(\tau), X_2(\tau))$), and it reads in particular

$$J(\cdot) = \sum_{i=1,2} \lambda_{id} \left[N_i\tau - \frac{X_i(\tau) - X_i(0)}{\lambda_{si}U_i(X_i(\tau))} \right], \qquad (9)$$

where $\phi^{-1}(X_i(\tau)) = U_i(X_i(\tau))$. Using (5), we have $U_i(X_i(\tau)) = -\frac{1}{\lambda_{si}\tau} \log\left(\frac{N_i - X_i(\tau)}{N_i - X_i(0)}\right)$. This (bijection) allows the equivalence of maxima with respect to $X_i(\tau)$ to u_i.

Denote $L = \{(X_1(\tau), X_2(\tau)) | \overline{X}_i(u\tau) \leq X_i(\tau) \leq \overline{X}_i(\tau), i = 1, 2$, such that $X_1(\tau) + X_2(\tau) \leq \psi\}$. Indeed, since $\frac{\partial U_i}{\partial X_i(\tau)} > 0$, it is easy to note that $\nabla J \neq 0$ for all points that fall in the interior of L. Moreover, we have

$$\frac{\partial J}{\partial X_i(\tau)} = -\frac{\lambda_{id}\tau}{\lambda_{si}U_i(X_i(\tau))} \left[1 + \frac{X_i(\tau) - X_i(0)}{U_i(X_i(\tau))} \frac{\partial U_i}{\partial X_i(\tau)} \right] < 0,$$

where we used the fact that $X_i(\tau) > X_i(0)$, $U_i(\cdot) > 0$ and $\frac{dX_i(\tau)}{dU_i} > 0$ (from the rule of the derivative of the inverse function) to determine the sign of the bracketed right end product. Therefore, $J(\cdot)$ cannot attain its maximum in the interior of L, so it does in the ∂L (on the boundary). This is also depicted geometrically for the cases ii) and iii) in Fig. 1.

Note, the term $\frac{X_i(\tau) - X_i(0)}{U_i(X_i(\tau))}$ has a negative gradient with respect to $X_i(\tau)$, so that the maximum can only be attained on the intersection of L with the line $X_1(\tau) + X_2(\tau) = \Psi$. In the following, concavity of $J(\cdot)$ ensures unique maxima on one of the boundaries.

The vectorial maximization of $J(\cdot)$ reduces to single variable maximization due to restricting $J(\cdot)$ in the
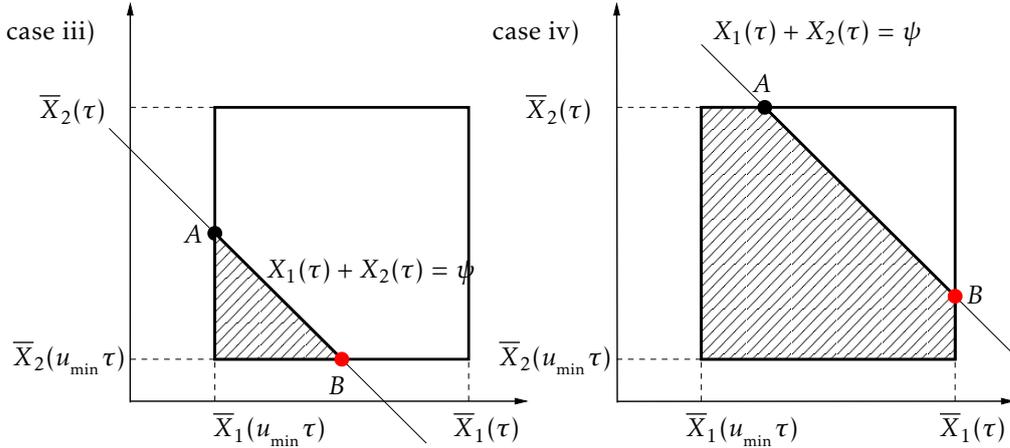
Figure 1. Geometric interpretation of Theorem. 1: case ii) and iii). The feasible region of $X_1(\tau)$ and $X_2(\tau)$ is depicted by horizontal and vertical axis respectively. The shaded area represents the set of points satisfying $X_1(\tau) + X_2(\tau) \leq \Psi$.

region ($X_1(\tau) + X_2(\tau) = \Psi$), e.g., $X_i(\tau)$,

$$J(X_1(\tau)) = \lambda_{1d}\tau \left[N_1 + \frac{X_1(\tau) - X_1(0)}{\log\left(\frac{N_1 - X_1(\tau)}{N_1 - X_1(0)}\right)} \right]$$

$$+ \lambda_{2d}\tau \left[N_2 + \frac{\Psi - X_1(\tau) - X_2(0)}{\log\left(\frac{N_2 - \Psi + X_1(\tau)}{N_2 - X_2(0)}\right)} \right]$$

$$= \lambda_{1d}\tau N_1 \left[1 + \frac{N_1 - X_1(0)}{N_1} f\left(\frac{X_1(\tau) - X_1(0)}{N_1 - X_1(0)}\right) \right]$$

$$+ \lambda_{2d}\tau N_2 \left[1 + \frac{N_2 - X_2(0)}{N_2} f\left(\frac{\Psi - X_1(\tau) - X_2(0)}{N_2 - X_2(0)}\right) \right],$$

where $f(x) = x/\log(1-x)$. To see that $J(X_1(\tau))$ is concave, firstly $f(x) = x/\log(1-x)$ is a concave function in $[0,1]$, since it is differentiable and $\ddot{f}(x) > 0$ in $(0,1)$. Furthermore, $f(Ax + B)$ is still concave for any linear combination of the argument $Ax + B$. Finally, any linear combination of convex functions through nonnegative coefficients is still a concave function.

Now, since $J(X_1(\tau))$ is a concave function in $[\overline{X}_1(u_{\min}\tau), \overline{X}_2(\tau)]$, it attains its maximum at the extrema of the segment, the points corresponding to A and B in Fig. 1. The explicit form for parts (iii) and (iv) is obtained calculating $\overline{X}_1(v_1\tau)$, $\overline{X}_2(v_2\tau)$ under the assumption that $U_2 = u_{\min}$ and $U_1 = 1$ is optimal, respectively. This concludes the proof. ⋄

Remark 1. The explicit closed form of the optimal static policy in parts (iii) and (iv) of theorem 1 can be easily computed as

- The optimal policy for part (iii) is $\mathbf{U} = \{1, v_2\} = \{1, \frac{-1}{\lambda_{s2}\tau} \log(\frac{N - \psi}{N_2 - X_2(0)} - \frac{N_1 - X_1(0)}{N_2 - X_2(0)} e^{-\lambda_{s1}\tau})\}$.

- The optimal policy for part (iv) is $\mathbf{U} = \{v_1, u_{\min}\} = \{\frac{-1}{\lambda_{s1}\tau} \log(\frac{N - \psi}{N_1 - X_1(0)} - \frac{N_2 - X_2(0)}{N_1 - X_1(0)} e^{-\lambda_{s2}\tau u_{\min}}), u_{\min}\}$.

Having illustrated the 2D optimal policy for static control, we note that it can be directly extended to m-dimensions. Without loss of generality, we assume that the indexing of nodes is in the increasing order of the value of λ_i for the following theorem. We denote by $e_k(c)$ for the m-dimensional vector which has 0 on all its components but for the k-th component has a value c.

Theorem 2. (**Static Optimal Control**-mD): Consider the problem of maximizing $\bar{F}_D(\tau)$, in a static policy for the case of m classes (mD case), the following holds

i. If $\overline{\mathcal{X}}(\{\mathbf{1}u_{\min}\tau\}) > \psi$ there is no feasible solution.

ii. If $\overline{\mathcal{X}}(\{\mathbf{1}\tau\}) < \psi$ then the optimal control policy is $\mathbf{U} = \mathbf{1}$, otherwise,

iii. Counting down k from m to 1, for every k :

a. If $\overline{\mathcal{X}}(\{(\mathbf{1} - e_k(1 - u_{\min}))\tau\}) < \psi$, then the optimal control policy is $\mathbf{U} = \{\mathbf{1} - e_k(1 - v_k)\}$, otherwise,

b. If $\overline{\mathcal{X}}(\{(\mathbf{1} - e_{k-1}(1 - u_{\min}) - e_k(1 - u_{\min}))\tau\}) < \psi$, then the optimal control policy is $\mathbf{U} = \{\mathbf{1} - e_{k-1}(1 - v_{k-1}) - e_k(1 - u_{\min})\}$.

iv. If $\overline{\mathcal{X}}(\{(\mathbf{1} - \sum_{k=1}^{m} e_k(1 - u_{\min}))\tau\}) < \psi$, then the optimal control policy is $\mathbf{U} = \{\mathbf{1} - e_1(1 - v_1) - \sum_{k=2}^{m} e_k(1 - u_{\min})\}$.

The constant v_k is given by

$$v_k = \frac{\overline{X}_k^{-1}(\psi - \sum_{i=1}^{k-1} \overline{X}_i(\tau) - \sum_{i=k+1}^{m} \overline{X}_i(u_{\min}\tau))}{\tau}.$$

Proof: The proof is a direct extension of the 2D case. ⋄

4.2. Optimal Dynamic Control Policy

In what follows, we consider the problem of maximizing $F_D(\tau)$ under the control $\mathbf{U}(k) \in [u_{\min}, 1]^m$. The solution to the dynamic control problem will be shown to consist of polices involving thresholds. With no loss of generality, we consider the case of two classes (2D) in the following analysis. It can be easily extended to the mD case. This subsection provides the suitable framework for the two-time scale dynamic algorithm introduced later in the paper.

Definition 1 (Threshold policy). A mD *threshold policy* is a control policy $\mathbf{U} : [0, \tau] \rightarrow \{u_{\min}, u_{\max}\}^m$ with the related switching parameter h_i for the i-th class if $U_i(k) = u_{\max}$ for $k \leq h_i$ a.e. and $U_i(k) = u_{\min}$ for $k > h_i$ i.e. component switches from $u_{\max}$ to $u_{\min}$ at most once. We denote h_i the switch time (the *threshold*) of a dynamic policy with respect to the i–th component.

Let ψ_1 be the maximum number of nodes of class 1 that can be infected: then the source can infect ($\psi_2 = \psi - \psi_1$) nodes of class 2. Then, when the constraint is saturated, the optimal threshold policy (U_1^*, U_2^*) satisfies the following relation

$$S_i(\psi_i) \stackrel{\text{def}}{=} \sum_{k=0}^{K-1} U_i^*(k, \psi_i) = -\frac{1}{\lambda_{si}\Delta} \log\left(\frac{N_i - \psi_i}{N_i - X_i(0)}\right).$$

Let $h_i(\psi_i)$ the threshold for the optimal policy U_i^*. Then we have

$$S_i(\psi_i) = h_i(\psi_i) \cdot u_{\max} + (K - 1 - h_i(\psi_i)) \cdot u_{\min} + g(\psi_i), \quad (10)$$

where $g(\psi_i) = -\frac{1}{\lambda_{si}\Delta} \log(\frac{N_i - \psi_i}{N_i - X_i(0)}) - (h_i(\psi_i) \cdot u_{\max} + (K - 1 - h_i(\psi_i)) \cdot u_{\min})$. Notice that (10) defines a bijection, so that $(U_1^*, U_2^*) \sim (\psi_1^*, \psi - \psi_1^*)$: we can now consider the problem from slightly different perspective. The source goal is to find the optimal ψ_1^* that maximizes the delivery probability. The advantage of the new formulation is that the joint constraint is replaced by two separate constraints, one per class. In turn we can express our initial optimization as

$$\max_{\psi_1} \quad \bar{J}(\psi_1) = \max_{\psi_1}(\bar{J}_1(\psi_1) + \bar{J}_2(\psi - \psi_1))$$
$$\text{s.t.} \quad X_1(\tau) \leq \psi_1, X_2(\tau) \leq \psi - \psi_1, \quad (11)$$

where $\bar{J}_i(\cdot)$ is

$$\bar{J}_i(\psi_i) = \tau N_i \lambda_{id} - \frac{\lambda_{id}}{\lambda_{si}} \sum_{k=0}^{K-1} (N_i - X_i(k, \psi_i)) \frac{1 - e^{-\lambda_{si}\Delta \cdot U_i(k, \psi_i)}}{U_i(k, \psi_i)},$$

for $i = 1, 2$. Clearly, the solution of this new problem solves directly the original optimization problem (3).

Theorem 3. There exists an unique optimal value (ψ_1^*) that maximizes the delivery probability in (11).

Proof:

In order to prove the uniqueness, it is sufficient to prove that J is concave in ψ_1. We start by proving that $\bar{J}_1(\psi_1)$ is concave. From equation (10), it follows that $\frac{dg(\psi_1)}{d\psi_1} = \frac{1}{\lambda_{s1}\Delta}$, then the derivative of function $\bar{J}_1$ can be expressed as follows: $\frac{d\bar{J}_1(\psi_1)}{d\psi_1} =$

$$-\frac{\lambda_{1d}}{\lambda_{s1}}(N - X_1^{h_1}(\psi_1) \cdot \frac{\lambda_{s1}\Delta.g(\psi_1)e^{-\lambda_{s1}\Delta.g(\psi_1)} - 1 + e^{-\lambda_{s1}\Delta.g(\psi_1)}}{U_1^{h_1}(\psi_1)^2}$$
$$+\frac{\lambda_{1d}}{\lambda_{s1}} \frac{1 - e^{-\lambda_{si}\Delta.u_{min}}}{u_{min}} \cdot \sum_{k=h_1+1}^{K-1} \frac{dX_1^k(\psi_1)}{d\psi_1} \quad (12)$$

The first term is clearly decreasing in ψ_1, since $g(\psi_1)$ is an increasing function in ψ_1. Let us now determine the derivative of second term as follows:

- For $k = h_1 + 1$, we have

$$X_1^{h_1+1}(\psi_1) = X_1^{h_1} + (N_1 - X_1^{h_1})(1 - e^{-\lambda_{s1}\Delta.g(\psi_1)}),$$

then

$$\frac{dX_1^{h_1+1}(\psi_1)}{d\psi_1} = (N_1 - X_1^{h_1})e^{-\lambda_{s1}\Delta.g(\psi_1)}.$$

- For $k = h_1 + 2$, we have

$$X_1^{h_1+2}(\psi_1) = X_1^{h_1+1}(\psi_1) + (N_1 - X_1^{h_1+1}(\psi_1))(1 - e^{-\lambda_{s1}\Delta.u_{min}}),$$

then

$$\frac{dX_1^{h_1+2}(\psi_1)}{d\psi_1} = \frac{dX_1^{h_1+1}(\psi_1)}{d\psi_1}e^{-\lambda_{s1}\Delta.u_{min}}.$$

- Making all the steps up to $k = K - 1$ we can derive the useful formula of the derivative of the second term in (12):

$$\sum_{k=h_1+1}^{K-1} \frac{dX_1^k(\psi_1)}{d\psi_1} = \frac{dX_1^{h_1+1}(\psi_1)}{d\psi_1}e^{-\lambda_{s1}\Delta.\sum_{k=h_1+2}^{K-1} u_{min}}, \quad (13)$$

Since $g(\psi_1)$ is increasing function in ψ_1, it is easy to check that the derivative of function $X_1^{h_1+1}(\psi_1)$ is a decreasing function, It follows from (13) that the second term is decreasing function in ψ_1. Hence the function $\bar{J}_1$ is a concave function. Using the same steps we can show that $\bar{J}_2(\psi - \psi_1)$ is a concave function in ψ_1, and since the sum of two concave functions is a concave function, $\bar{J}(\psi_1)$ is a concave function, hence proved. ◇

Having characterized the optimal forwarding policies for the m-class DTN, in the following we focus on algorithms that the source can adapt to achieve the optimal performance. In particular, we propose

algorithms in which source node does not require a prior knowledge of system parameters, rather has the ability to estimate online and adjust the forwarding accordingly, we call *blind online algorithms*. We resort on stochastic approximation based approach to design the algorithms.

5. Blind Online Algorithms for Adaptive Optimization

In this section, we propose two online algorithms to attain optimal control of forwarding by the source node: i) Algorithm 1 applies to the static control case , and ii) Algorithm 2 applies to the dynamic control case. Both algorithms are *blind*, do not require a-priori knowledge of network parameters which is intermeeting intensities and number of mobiles in our context. Observe that in the heterogeneous case, each class of nodes has its own (unknown) parameters: intermeeting intensities (λ_{si} and λ_{id}) and the number of nodes (N_i) for each class i. These algorithms will only depend on the source ability to distinguish between classes, e.g., according to node's type (whether it is a throwbox, a smartphone, etc): leveraging on the structure of the optimal solution, this will be enough to find the optimal forwarding control that the source should adopt without explicit estimation of such parameters.

Remark 2. Let $\theta_i = \sum_{k=0}^{K-1} U_i(k), i = 1, 2$, then θ_i identifies both the static and dynamic policies: the static policy is $U_i = \theta_i/K$, while for the dynamic policy threshold writes $h_i = \max\{h \in \mathbb{N} : v(h) = h \cdot u_{max} + (K - h) \cdot u_{min} \leq \theta_i\}$, and $U_i(h) = \theta_i - v(h)$.

5.1. Blind Online Algorithm for Static Control

Our static algorithm is an extension of [15] to the multi-dimensional case. Each step of the algorithm corresponds to a round of duration τ. For the sake of notation, let $\widehat{X}(\theta_i^k, t), i = 1, 2$ the number of nodes of the ith class that are potentially infected by the source in the current round up to time t, by averaging over several consecutive rounds, using interpolation, the source node is able to obtain an estimate of the average number of copies ($\widehat{X}(\theta_i^k)$) that could potentially be attained using the current set of K policies. $\widehat{X}(\theta_i^k)$ is used to update θ_i according to the formula showed in Algorithm 1, in which, for $I = [\theta_{min}, \theta_{max}]$, the projection function Π_I is defined as follows:

$$\Pi_I(\theta_i) = \begin{cases} \theta_{max} & \text{if} & \theta_i \geq \theta_{max}, \\ \theta_i & \text{if} & \theta_{min} \leq \theta_i \leq \theta_{max}, \\ \theta_{max} & \text{if} & \theta_i \leq \theta_{min}. \end{cases}$$

This stochastic approximation algorithm will implicitly discover the fastest class. This permits to adjust the values of (θ_1, θ_2) for the next round in order to, eventually, estimate the optimal (θ_1, θ_2) in I, which uniquely

---

Algorithm 1 Stochastic approximation of the optimal policy using online estimation (2D static case)

1: **input:** $I = [0, K - 1], \theta_1^0 = \theta_2^0 = K/2, k = 0$
2: **while** $max(|\theta_1^{k+1} - \theta_1^k|, |\theta_2^{k+1} - \theta_2^k|) > \epsilon$ **do**
3: $\quad \widehat{X}_1(\theta_1^k) = \text{interp}(\widehat{X}_1(\theta_1^k), \theta_1^k)$
4: $\quad \widehat{X}_2(\theta_2^k) = \text{interp}(\widehat{X}_2(\theta_2^k), \theta_2^k)$
5: $\quad$ **if** $\widehat{X}_1(\theta_1^k) >= \widehat{X}_2(\theta_2^k)$ **then**
6: $\quad\quad \theta_1^{k+1} = \Pi_I(\theta_1^k + a_k(\Psi - \widehat{X}_1(\theta_1^k)))$
7: $\quad\quad \theta_2^{k+1} = \Pi_I(\theta_2^k + a_k(\Psi - \widehat{X}_1(\theta_1^k)) - \widehat{X}_2(\theta_2^k))$
8: $\quad$ **else**
9: $\quad\quad \theta_1^{k+1} = \Pi_I(\theta_1^k + a_k(\Psi - \widehat{X}_1(\theta_1^k)) - \widehat{X}_2(\theta_2^k))$
10: $\quad\quad \theta_2^{k+1} = \Pi_I(\theta_2^k + a_k(\Psi - \widehat{X}_2(\theta_2^k)))$
11: $\quad$ **end if**
12: $\quad k \leftarrow k + 1$
13: **end while**

---

determine the optimal static policy $\mathbf{U} = (U_1^*, U_2^*)$. In the following theorem we discuss the convergence of our algorithm.

Theorem 4. If the sequence $\{a_k\}$ verifies that $a_k > 0, \forall k, \sum_{k=0}^{+\infty} = +\infty$ and $\sum_{k=0}^{+\infty} a_k^2 < +\infty$, then the sequence (θ_1^k, θ_2^k) converges to the optimal solution (θ_1^*, θ_2^*).

Proof: First, we show that the sequence (θ_1^k, θ_2^k) converges to some limit set of the following Ordinary Differential Equation (ODE)

$$\dot{\theta}^1 = G_1(\theta^1) + z_1 = \Psi - E[X_1 + X_2 | (\theta^1, \theta_{min})] + z_1, \quad (14)$$
$$\dot{\theta}^2 = G_2(\theta^1, \theta^2) + z_2 = \Psi - E[X_1 + X_2 | (\theta^1, \theta^2)] + z_2, \quad (15)$$
$$z = (z_1, z_2) \in -C((\theta^1, \theta^2)),$$

where z_i, $i = 1, 2$, is the projection or constrain term, the minimum force needed to keep the trajectory of the ODEs in $[\theta_{min}, \theta_{max}]$ and the set $C(\vec{\theta})$ is defined as follows [23]: for $\vec{\theta} = (\theta_1, \theta_2) \in (\theta_{min}, \theta_{max})^2$, we have $C(\vec{\theta}) = \{(0, 0)\}$; and for (θ_1, θ_2) in the boundary of $[\theta_{min}, \theta_{max}]^2$, we let $C(\vec{\theta})$ be the infinite convex cone generated by the outer normals at $\vec{\theta}$ of the faces on which $\vec{\theta}$ lies. Put simply, $C(\vec{\theta})$ contains the possible values needed to keep $\vec{\theta}$ in $[\theta_{min}, \theta_{max}]$. For example, if $\theta_1 = \theta_{max}$ and $G_1(\theta_1)$ point out of $[\theta_{min}, \theta_{max}]$ then $z_1(t) = G_1(\theta_1)$. Hence the function $z(\cdot)$ is determined by $[\theta_{min}, \theta_{max}]^2$ and the functions $G_i(\cdot)$, $i = 1, 2$.

Since $G_1(\theta_1)$ (resp. $G_2(\theta_1, \theta_2)$) is decreasing function in θ_1 (resp. θ_1 and θ_2), then the equilibrium is unique. Moreover, it is easy to check that the optimal solution (θ_1^*, θ_2^*) (see theorem 1) is the unique equilibrium of (14)-(15). As discussed in [23], the convergence of such stochastic algorithm is guaranteed when the sequence (a_k) verifies, $a_k > 0, \forall k, \sum_{k=0}^{+\infty} a_k = +\infty$ and $\sum_{k=0}^{+\infty} a_k^2 < +\infty$. We now need to show that
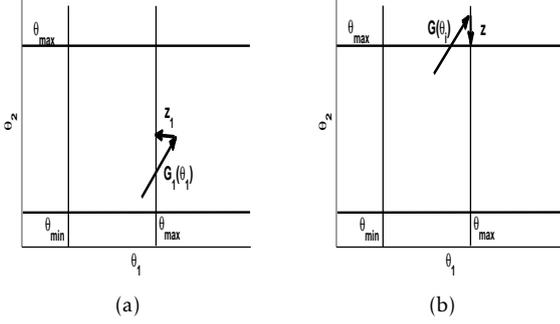
Figure 2. Projection of the value θ_i over I

(θ_1^*, θ_2^*) is globally asymptotically stable of the system (14)-(15). We use the Lyapunov function $V(\theta_1, \theta_2) = (\theta_1 - \theta_1^*)^2 + (\theta_2 - \theta_2^*)^2$. Then we have

$$
\begin{aligned}
\dot{V}(\vec{\theta}) &= 2\dot{\theta}_1(\theta_1 - \theta_1^*) + 2\dot{\theta}_1(\theta_2 - \theta_2^*) \\
&= 2(G_1(\theta_1) + z_1)(\theta_1 - \theta_1^*) \\
&\quad + 2(G_2(\theta_1, \theta_2) + z_2)(\theta_2 - \theta_2^*). \quad (16)
\end{aligned}
$$

Since there is at most one active constraint, we have two cases (see Fig. 2):

- If $\theta_{min} < \theta_i < \theta^{max}$ for $i = 1, 2$, then $z_1 = 0$ and $z_2 = 0$.

- If $\theta_i = \theta_{max}$ (resp. θ_{min}) and $G_i(\theta_{max}) > 0$ (resp. $G_i(\theta_{min}) < 0$) for only one class i, then $z_i = -G_i(\theta_{max})$ (resp. $z_i = -G_i(\theta_{min})$) and $z_{-i} = 0$ for other class.

And because $G_1(\theta_1)$ and $G_2(\theta_1, \theta_2)$ are strictly decreasing functions, it is easy to check that $\dot{V}(\vec{\theta})$ is decreasing. Hence the optimal solution is asymptotically stable.

5.2. Blind Online Algorithm for Dynamic Control

We apply the *two-time-scale stochastic approximation algorithm*, which is a stochastic recursive algorithm. Compared to standard stochastic approximation techniques in literature [23], here some of the components are updated using a step-size much smaller than those of the remaining components. For further insight into the convergence properties of this class of algorithms, the reader is referred to [24].

The algorithm introduced here drives the source to the optimal (ψ_1) that maximizes $J(\psi_1)$, i.e. maximizes the probability of success $F_D(\tau) = 1 - \exp\Big(-\sum_j \lambda_{jd} \int_{s=0}^t X_j(\psi_1, s) ds\Big)$. Since the system parameters (such as N_i and λ_i, $i = 1, 2$) are unknown, the value of $J(\psi_1)$ is also unknown, but a noisy estimate of it is known, namely $f(\psi_1)$ such that $\mathbb{E}[f|\psi_1] = J(\psi_1)$.

The two-time-scale stochastic approximation algorithm is then formulated as follows:

$$
\begin{aligned}
\theta_i^{k+1} &= \Pi_I\Big(\theta_i^k + a_k(\psi_i^k - X_i(\theta_i^k))\Big), \\
&\qquad \psi_2^k = \psi - \psi_1^k, i = 1, 2, \quad (17) \\
\psi_1^{k+1} &= \Pi_H\Big(\psi_1^k + b_k \frac{f(\psi_1^k + c_k) - f(\psi_1^k - c_k)}{c_k}\Big), (18)
\end{aligned}
$$

where $I = [0, K-1]$, $H = [0, \psi]$ and $\{a_k\}, \{b_k\}$ are sequences of non-increasing positive constants satisfying $\sum_{k=0}^{+\infty} a_k = +\infty, \sum_{k=0}^{+\infty} a_k^2 < +\infty, \sum_{k=0}^{+\infty} b_k = +\infty, \sum_{k=0}^{+\infty} b_k^2 < +\infty$, and $\lim_{k \to +\infty} \frac{b_k}{a_k} = 0$. The last condition implies that $b_k \to 0$ at a faster rate than a_k, implying that (18) moves on a slower timescale than (17). An example of such stepsizes are $a_k = \frac{1}{k}$, $b_k = \frac{1}{1+k \log k}$ and so on. Further requirements are imposed on sequence (c_k), which we defer to Thm. 6 for the sake of clearness.

More in detail, at each round k the following steps are executed:

1: Fix (ψ_1) at the value (ψ_1^k) and learn the optimal values of θ_1^k, θ_2^k for $(\psi_1^k + c_k), (\psi_1^k - c_k)$ using the following algorithm:

$$
\begin{aligned}
\theta_1^{k+1} &= \Pi_I\Big(\theta_1^k + a_k(\psi_1^k - X_1(\theta_1^k))\Big), \\
\theta_2^{k+1} &= \Pi_I\Big(\theta_2^k + a_k(\psi_2^k - X_2(\theta_2^k))\Big). \quad (19)
\end{aligned}
$$

2: Measure the noisy estimate of the success probability, $f(\psi_1)$, at $(\psi_1^k + c_k), (\psi_1^k - c_k)$ when θ_1^k, θ_2^k, obtained at the first step, are applied.

3: Use $f(\psi_i^k + c_k), f(\psi_i^k - c_k)$ to update the value of (ψ_1^k) according to Kiefer-Wolfowitz algorithm as shown at step (11) of algorithm 2.

In the following theorem we proof the convergence of the Kiefer-Wolfowitz part of the algorithm that appears in (18); this serves as an introduction to theorem 6.

Theorem 5. If the sequence (b_k) verifies: $b_k > 0$, $\forall k$, and $\sum_{k=0}^{+\infty} b_k = +\infty, \sum_{k=0}^{+\infty} b_k^2 < +\infty$ and $c_k \to 0$, then (ψ_1^k) converges to the optimal solution (ψ_1^*).

Proof: Consider two sequences $\{b_k, c_k, k \geq 1\}$ satisfying $c_k \to 0$, $\sum_k b_k = \infty$, $\sum_k b_k c_k < \infty$, $\sum_k (b_k/c_k)^2 < \infty$, and the recursive updates of ψ_1:

$$
\psi_1^{k+1} = \psi_1^k + b_k \frac{f(\psi_1^k + c_k) - f(\psi_1^k - c_k)}{c_k}. \quad (20)
$$

This recursive schema converges stochastically to the optimal value (ψ_1^*) that maximizes $J(\psi_1)$ provided that $J(\psi_1)$ satisfies the following conditions:

1. $J(\psi_1)$ is a strictly quasi-concave function.

2. There exists β and B such that $|\psi_1^a - \psi_1^*| + |\psi_1^b - \psi_1^*| < \beta$ implies $|J(\psi_1^a) - J(\psi_1^b)| < B|\psi_1^a - \psi_1^b|$.

3. There exists ρ and R such that $|\psi_1^a - \psi_1^b| < \rho$ implies $|J(\psi_1^a) - J(\psi_1^b)| < R$.

4. For every $\delta > 0$ there exists a positive $\pi(\delta)$ such that $|\psi_1 - \psi_1^*| > \delta$ implies

$$\inf_{0 < \epsilon < \frac{\delta}{2}} \frac{|J(\psi_1 + \epsilon) - J(\psi_1 - \epsilon)|}{\epsilon} > \pi(\delta).$$

In order to prove that $J(\psi_1)$ satisfies these conditions, we only need to prove that $J(\psi_1)$ is concave and has a unique maximum solution ψ_1^* ([25]-lemma 2, theorem 1), which is already proved in theorem 3, hence proved. $\diamond$

Theorem 6. The sequence (θ_i^k, ψ_1^k) defined in the iteration (17) and (18) converges a.s. to $(\theta_i(\psi_1^*), \psi_1^*)$.

Proof: This can be proved directly by the results of ([26], Chapter 6) together with the sure convergence of the two related single-time-scale stochastic approximation algorithms – those defined in (17) and (18), respectively – and that each algorithm has a globally asymptotically stable equilibrium, appearing in the following.

The first algorithm is the aforementioned algorithm (Alg. 1) taking as entries the fixed values of $\widehat{X}_1(\theta_1^k), \widehat{X}_2(\theta_2^k)$ (which are ψ_1^k, ψ_2^k for the current round k). In fact, (17) sees ψ_1 as quasi-static (i.e., 'almost a constant') and it is easy to prove that the sequence (θ_1^k, θ_2^k) converges to some limit set of the following Ordinary Differential Equation (ODE)

$$\begin{aligned}
\dot{\theta}_1 &= G_1(\theta_1) + z_1 \\
&= \psi_1 - E[X_1 | \theta_1] + z_1, \quad z_1 \in -C_1(\theta_1), \quad (21) \\
\dot{\theta}_2 &= G_2(\theta_2) + z_2 \\
&= (\psi - \psi_1) - E[X_2 | \theta_2] + z_2, \quad z_2 \in -C_2(\theta_2), (22)
\end{aligned}$$

where z_i, $i = 1, 2$, is the minimum force needed to keep the solution θ_i in $I = [\theta_{min}, \theta_{max}]$.

If θ_i is in I on some time interval, then $z_i(\cdot)$ is zero on that interval ($C_i(\theta_i)$ contains only the zero element). If θ_i is on the interior of a boundary of I (i.e., θ_i equals either θ_{min} or θ_{max}) and $G_i(\theta_i)$ points out of I, then $z_i(\cdot)$ points backward inside I, i.e. $C_i(\theta_i)$ is the infinite convex cone generated by the outer normals at θ_i of the faces on which θ_i lies. For example, let $\theta_i = \theta_{max}$, with $G_i(\theta_i) > 0$, then, $z_i(t) = -G_i(\theta_i)$. $\diamond$

Following the same reasoning in the proof of theorem 4, it is easy to verify that each of the ODEs (21) and (22) has a globally asymptotically stable equilibrium $\theta_i^*(\psi_1)$.

The second algorithm is the Kiefer-Wolfowitz whose convergence is proved in theorem 5 and its asymptotic behavior is characterized [23] by the ODE

Algorithm 2 Stochastic approximation of the optimal policy using online estimation (2D Dynamic case)

1: **input:** $I = [0, K-1], H = [0, \psi]$, $\theta_1^0 = \theta_2^0 = K/2$, $\psi_1^0 = \psi$, $c_k = \frac{1}{k^{0.001}}$
2: **while** $|\psi_1^{k+1} - \psi_1^k| > \epsilon$ **do**
3:   set $\psi_1^+ = \min(\psi_1^k + c_k, \psi)$, $\psi_2^+ = \psi - \psi_1^+$
4:   $\theta_1^k = \Pi_I\big(\theta_1^k + a_k(\psi_1^+ - \widehat{X}_1(\theta_1^k))\big)$
5:   $\theta_2^k = \Pi_I\big(\theta_2^k + a_k(\psi_2^+ - \widehat{X}_2(\theta_2^k))\big)$
6:   Measure $f(\psi_1^+) = func(\theta_1^k, \theta_2^k)$
7:   set $\psi_1^- = \max(\psi_1^k - c_k, 0)$, $\psi_2^- = \psi - \psi_1^-$
8:   $\theta_1^k = \Pi_I\big(\theta_1^k + a_k(\psi_1^- - \widehat{X}_1(\theta_1^k))\big)$
9:   $\theta_2^k = \Pi_I\big(\theta_2^k + a_k(\psi_2^- - \widehat{X}_2(\theta_2^k))\big)$
10: Measure $f(\psi_1^-) = func(\theta_1^k, \theta_2^k)$
11: $\psi_1^{k+1} = \Pi_H\big(\psi_1^k + b_k(\frac{f(\psi_1^+) - f(\psi_1^-)}{c_k})\big)$
12: $k \leftarrow k + 1$
13: **end while**

$$\begin{aligned}
\dot{\psi}_1 &= G(\theta_1(\psi_1), \theta_2(\psi_1), \psi_1) + z_3 \\
&= \frac{\partial J(\psi_1)}{\partial \psi_1} + z_3, z_3 \in -C_3(\psi_1), \quad (23)
\end{aligned}$$

where z_3, is the minimum force needed to keep the solution ψ_1 in $H = [0, \psi]$. In order to prove that (ψ_1^*) is globally asymptotically stable of the ODE (23), we use the Lyapunov function $V(\psi_1) = (\psi_1 - \psi_1^*)^2$. Then we have

$$\begin{aligned}
\dot{V}(\vec{\psi}_1) &= 2\dot{\psi}_1(\psi_1 - \psi_1^*) \\
&= 2\frac{\partial J(\psi_1)}{\partial \psi_1}(\psi_1 - \psi_1^*) < 0. \quad (24)
\end{aligned}$$

Asymptotic global stability follows from Lyapunov's theorem. $\diamond$

5.3. Discussion on Implementation Issues

Since the proposed algorithms do not require a-priori knowledge of network parameters like intermeeting intensities and number of mobiles of each class, they can be easily implemented in real scenario. In other words, for relay nodes to apply those algorithms, a simple coded version of the proposed algorithms can be loaded on each device's memory. Each time a node needs to transmit a message, i.e., becomes a source, it executes this code. Moreover, as the forwarding protocol employed here is two hop, source can keep track of number of message copies in the network. After few rounds, as we will show in next section, the source will be driven to apply the optimal policies characterized in the previous sections. Certainly timers should be set for algorithm rounds, but those timers will operate distributively on each node.

6. Numerical Investigation

In this section we illustrate numerical experiments that validate our results. Using Matlab scripts, we have studied the impact of the inter-meeting intensity and the energy constraint (ψ) on the 2-D system optimal dynamic policies. We are also interested in the performance of the stochastic approximation algorithms described earlier; we will investigate their ability to drive the two hop forwarding, both in the static as well as dynamic scenario, to the optimal operating point.

In Fig. 3, we reported the optimal control in the case when intermeeting intensities with both the source and the destination are the same ($\lambda_{si} = \lambda_{id}$ for i=1,2), we can temporarily refer to both intensities by mentioning one of them, let us say λ_{s1}.

The first observation on Fig. 3(a) is that when λ_{s1} is small, the control tends to infect both classes over a larger interval. However, as long as $\lambda_{s1} < \lambda_{s2}$, the switch time for class 2 is higher and the source infects more nodes of class 2 (Fig. 3(c)). As λ_{s1} increases, an interesting case is when $\lambda_{s1} = \lambda_{s2}$: the switch times are equal and the source infects the same number of nodes per class.

The situation is flipped after equality is reached: when λ_{s1} exceeds λ_{s2}, the source tends to depend more on class 1 and infects more nodes of that class. This can be explained by the fact that class 1 nodes meet the source and destination with higher frequency ($\lambda_{1d} = \lambda_{s1} > \lambda_{s2}$). Intuitively, the probability of success improves as λ_{s1} increases as in Fig. 3(b).

In Fig. 4, we depict the switch times for both classes when each class's nodes meet the source at a rate that is different from the rate at which they meet the destination ($\lambda_{si} \neq \lambda_{id}$ for i=1,2). Fig. 4(a) shows that the switch time of class 1 increases as λ_{1d} increases but still is smaller than the switch time for class 2, as long as ($\lambda_{1d} < \lambda_{2d}$), after this point the source will infect class 1 nodes for longer time because they have higher chance now to meet the destination. The same observation is captured in Fig. 4(b). In Fig. 4(c) we observe that at the beginning when ($\lambda_{1d} < \lambda_{2d}$), the source infects more nodes of class 2 than class 1, then X_1 keeps increasing until it becomes larger than X_2 when ($\lambda_{1d} > \lambda_{2d}$). Intuitively, the probability of success improves as λ_{1d} increases as in Fig. 4(d).

Another important aspect is the impact of the energy constraint on the dynamic policies adopted by the source (see Fig. 5). We notice that the larger the energy constraint is, the higher the switch times are (Fig. 5(a) and 5(c)). This is because increasing ψ allows the source to infect more nodes by augmenting the optimal switch times while satisfying the constraint (Fig. 5(b) and 5(d)). An interesting observation is captured in Fig. 5(b) (Fig.5(d)): when ψ has a small value ($\psi = 10$), the source

(a)   (b)

(c)
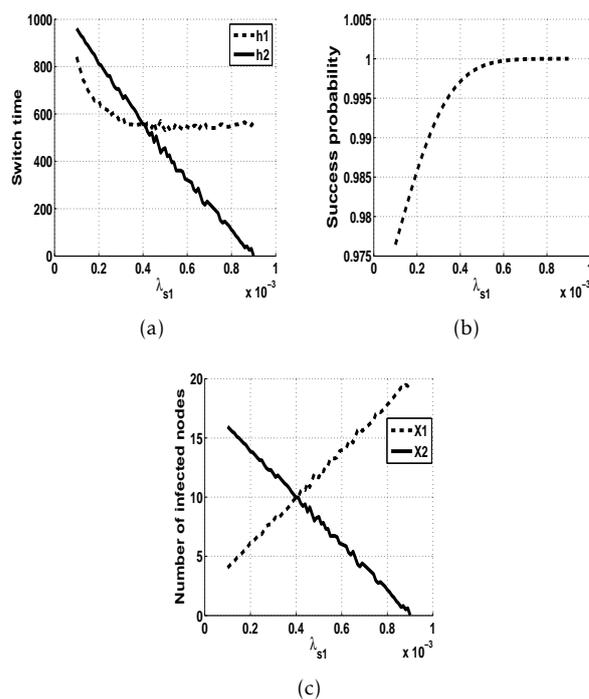
Figure 3. The impact of the inter-meeting intensity ($\lambda_{si} = \lambda_{id}$ for i=1,2) on switching times, delivery probability and number of infected nodes in the 2-D case, where $\psi = 20, N_1 = N_2 = 50, \lambda_{s2} = 0.4 \times 10^{-3}, \tau = 1000$ s.

(a) $\lambda_{s2} = \lambda_{2d} = 0.4 \times$ (b) $\lambda_{s2} = \lambda_{2d} = 0.4 \times$
$10^{-3}, \lambda_{s1} = 0.2 \times 10^{-3}$ $10^{-3}, \lambda_{s1} = 0.6 \times 10^{-3}$

(c) $\lambda_{s2} = \lambda_{2d} = 0.4 \times$ (d) $\lambda_{s2} = \lambda_{2d} = 0.4 \times$
$10^{-3}, \lambda_{s1} = 0.6 \times 10^{-3}$ $10^{-3}, \lambda_{s1} = 0.6 \times 10^{-3}$
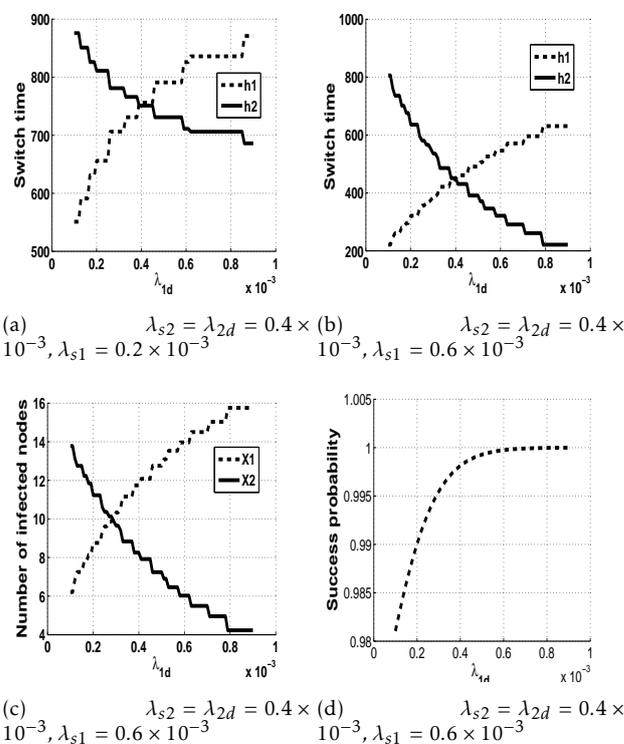
Figure 4. The impact of the inter-meeting intensity ($\lambda_{s1} \neq \lambda_{1d}$) on switching times and number of infected nodes in the 2-D case, where $\psi = 20, N_1 = N_2 = 50, \tau = 1000$ s.

(a) $\lambda_{s1} = 0.6 \times 10^{-3}, \lambda_{s2} = 0.4 \times 10^{-3}$

(b) $\lambda_{s1} = 0.6 \times 10^{-3}, \lambda_{s2} = 0.4 \times 10^{-3}$

(c) $\lambda_{s1} = 0.4 \times 10^{-3}, \lambda_{s2} = 0.6 \times 10^{-3}$

(d) $\lambda_{s1} = 0.4 \times 10^{-3}, \lambda_{s2} = 0.6 \times 10^{-3}$
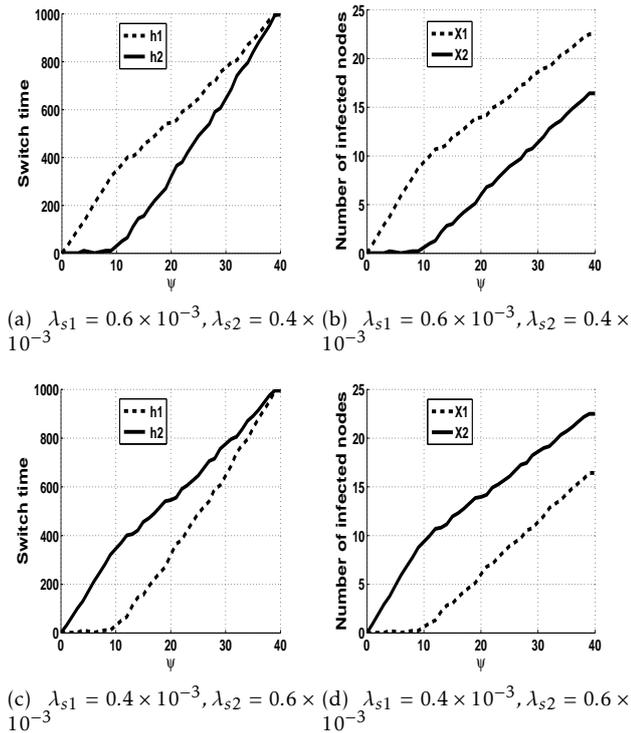
Figure 5. The impact of the energy constraint (ψ) on switching times and number of infected nodes in the 2-D case, where $N_1 = N_2 = 50$.

(a) $\psi = 30$

(b) $\psi = 60$

(c) $\psi = 90$

Figure 6. The stochastic approximation using online estimation ($\lambda_{s1} > \lambda_{s2}$) in the 2-D case (static algorithm 1), where $\tau = 1000, N_1 = N_2 = 1000, \lambda_{s1} = \lambda_{1d} = 0.34 \times 10^{-4}, \lambda_{s2} = \lambda_{2d} = 0.14 \times 10^{-4}$.

forwards the message only to class 1 (class 2) nodes, respectively. This indicates that under a tight energy constraint, the source can rely on one class only (the one with higher λ_s) to have its message delivered. For the case where $\psi = 40$, the source will forward the message all the time to both classes in order to saturate the constraint.

Fig. 6 shows sample paths of our learning algorithm for the problem of static control. In Fig. 6(a) the algorithm is showed to converge to $(U_1^*, U_2^*) = (u_1^*, u_{\min})$ with $0 < U_1^* < 1$. Indeed, since the source may infect at most 30 nodes, from theorem (1, case (iv)) and remark (1), the optimal solution is $(v_1, u_{\min})$ with $v_1 = \frac{-1}{\lambda_{s1}\tau} \log\left(\frac{N-\psi}{N_1-X_1(0)} - \frac{N_2-X_2(0)}{N_1-X_1(0)} e^{-\lambda_{s2}\tau u}\right)$, we can see that U_1 converges to v_1 and U_2 to $u_{\min}$.

In Fig. 6(b) we plot case (iii) of theorem 1 where the energy constraint is 60 and the algorithm converges to $(U_1^*, U_2^*) = (1, v_2)$ with $v_2 = \frac{-1}{\lambda_{s2}\tau} \log\left(\frac{N-\psi}{N_2-X_2(0)} - \frac{N_1-X_1(0)}{N_2-X_2(0)} e^{-\lambda_{s1}\tau}\right)$. In this case the source will infect nodes of both classes in order to saturate the constraint by giving the message with probability 1 to the class with larger intermeeting intensities $\{\lambda_{si}, \lambda_{id}\}$ (class 1) and with some probability v_2 to the other class.

Fig. 6(c) shows the case where the constraint has a large value ($\psi = 90$) and the source can not reach the constraint even when forwarding to both classes with probability 1.

The performance of our learning algorithm in the dynamic case is shown in Fig. 7. In Figs. 7(a) and 7(b) we depict the convergence of ψ_1 (recall that ψ_1 is the maximum number of nodes of class 1 that can be infected) to the optimal value that maximizes the delivery probability under two different values of the energy constraint ($\psi = 20$, $\psi = 50$), and in Figs. 7(c) and 7(d) we show the corresponding switch times for both cases.

To better understand Fig. 7, let us observe Figs. 7(a) and 7(c) where the energy constraint has a small value ($\Psi = 20$), we notice that since class 1 has bigger intermeeting intensities ($\lambda_{s1}, \lambda_{1d}$), the algorithm drives the source to forward the message to class 1 for longer time resulting in more infected nodes in class 1 (19 infected nodes out of 20), which is the value at which the probability of success attains its maximum Fig.7(e). While in Figs. 7(b) and 7(d) where ψ is large ($\psi = 50$), the algorithm converges to ($\psi_1^* = 32$) the value that maximizes the success probability in Fig.7(f). Note that since it meets nodes of class 1 more often ($\lambda_{s1} > \lambda_{s2}$), the source can infect more nodes of class 1 with smaller switching time for this class while it has to forward

(a) $\psi = 20$

(b) $\psi = 50$

(c) $\psi = 20$

(d) $\psi = 50$

(e) $\psi = 20$

(f) $\psi = 50$
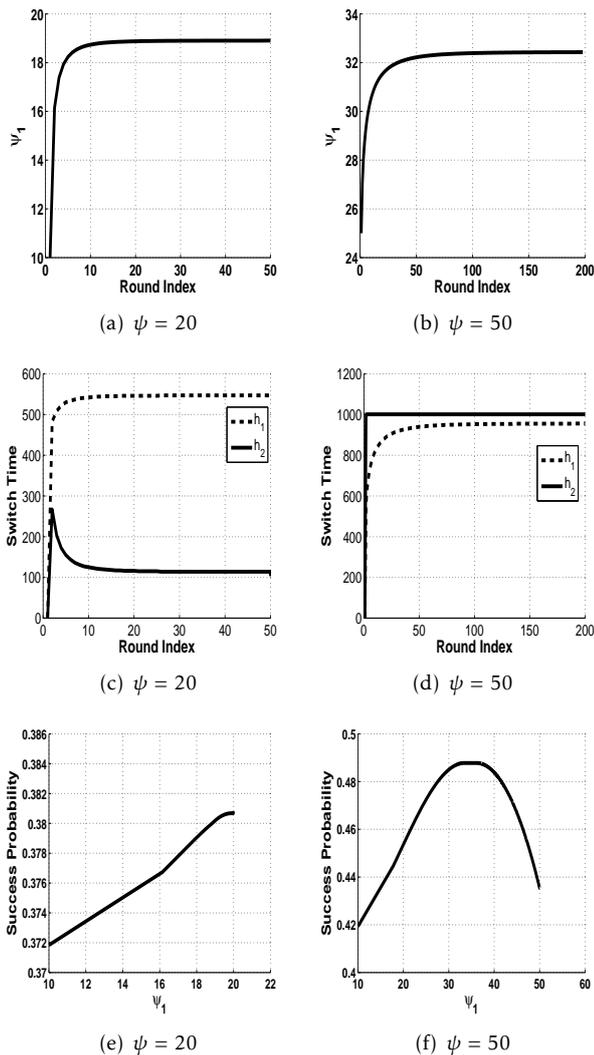
Figure 7. The stochastic approximation using online esti‐mation $(\lambda_{s1} > \lambda_{s2})$ in the 2–D case (Dynamic algorithm 2), where $\tau = 1000, N_1 = N_2 = 1000, \lambda_{s1} = \lambda_{1d} = 0.34 \times 10^{-4}, \lambda_{s2} = \lambda_{2d} = 0.14 \times 10^{-4}$.

all the time to class 2 in order to saturate the energy constraint.

7. Conclusion

In this paper we studied the problem of optimal relaying for DTNs consisting of *heterogeneous* mobile nodes. We have considered controlled two hop forwarding policies for source-destination message delivery using multiple classes of relays. Using our model we characterized the optimal forwarding policy in the family of the multi-dimensional DTNs for both static and dynamic control cases.

Finally, we have designed novel algorithms based on the theory of stochastic approximations. Those algorithms enable nodes in a heterogeneous DTN to

tune up independently and optimally the parameters for both static and dynamic optimal forwarding policies. Thus, nodes autonomously adapt to the operating point of a system comprising multiple classes of nodes. The distinctive feature of our so-called *blind online algorithms* is that the source node does not need to know explicitly the system parameters a priory. Instead, our novel implementation works at runtime permitting online adaptation to the specific existing conditions.

References

[1] CHAINTREAU, A., HUI, P., CROWCROFT, J., DIOT, C., GASS, R. and SCOTT, J. (2007) Impact of human mobility on opportunistic forwarding algorithms. *IEEE Transactions on Mobile Computing* **6**: 606–620.

[2] BURLEIGH, S., TORGERSON, L., FALL, K., CERF, V., DURST, B., SCOTT, K. and WEISS, H. (2003) Delay-tolerant networking: an approach to interplanetary Internet. *IEEE Comm. Magazine* **41**: 128–136.

[3] PELUSI, L., PASSARELLA, A. and CONTI, M. (2006) Opportunistic networking: data forwarding in disconnected mobile ad hoc networks. *IEEE Communications Magazine* **44**(11): 134–141.

[4] ZHAO, W., AMMAR, M. and ZEGURA, E. (2005) Controlling the mobility of multiple data transport ferries in a delay-tolerant network. In *Proc. of IEEE INFOCOM* (Miami USA).

[5] ZHANG, Z. (2006) Routing in intermittently connected mobile ad hoc networks and delay tolerant networks: overview and challenges. *Communications Surveys & Tutorials, IEEE* **8**(1): 24–37.

[6] JONES, E.P. and WARD, P.A. (2006) Routing strategies for delay-tolerant networks. *Submitted to ACM Computer Communication Review (CCR)* .

[7] SPYROPOULOS, T., PSOUNIS, K. and RAGHAVENDRA, C. (2008) Efficient routing in intermittently connected mobile networks: the multi-copy case. *ACM/IEEE Transactions on Networking* **16**: 77–90.

[8] ALTMAN, E., BAŞAR, T. and DE PELLEGRINI, F. (2010) Optimal monotone forwarding policies in delay tolerant mobile ad-hoc networks. *Performance Evaluation* **67**(4): 299 – 317.

[9] VAHDAT, A. and BECKER, D. (2000) *Epidemic Routing for Partially Connected Ad Hoc Networks*. Tech. Rep. CS-2000-06, Duke University.

[10] HANBALI, A.A., NAIN, P. and ALTMAN, E. (2006) Performance of ad hoc networks with two-hop relay routing and limited packet lifetime. In *Proc. of Valuetools* (New York NY USA: ACM): 49.

[11] NEGLIA, G. and ZHANG, X. (2006) Optimal delay-power tradeoff in sparse delay tolerant networks: a preliminary study. In *Proc. of ACM SIGCOMM CHANTS 2006*: 237–244.

[12] SPYROPOULOS, T., TURLETTI, T. and OBRACZKA, K. (2009) Routing in delay-tolerant networks comprising heterogeneous node populations. *Mobile Computing, IEEE Transactions on* **8**(8): 1132 –1147.

[13] CHAINTREAU, A., BOUDEC, J.Y.L. and RISTANOVIC, N. (2009) The age of gossip: Spatial mean-field regime. In

Proc. of ACM SIGMETRICS (Seattle, Washington, USA).

[14] De Pellegrini, F., Altman, E. and Basar, T. (2010) Optimal monotone forwarding policies in delay tolerant mobile ad hoc networks with multiple classes of nodes. In *proc. of WiOpt WDM Workshop* (Avignon, France).

[15] Altman, E., Neglia, G., De Pellegrini, F. and Miorandi, D. (2009) Decentralized stochastic control of delay tolerant networks. In *Proc. of INFOCOM* (Rio de Janeiro, Brazil).

[16] Guerrieri, A., Carreras, I., Pellegrini, F.D., Miorandi, D. and Montresor, A. (2010) Distributed estimation of global parameters in delay-tolerant networks. *Elsevier Comput. Commun.* **33**(13): 1472–1482.

[17] Hui, P., Crowcroft, J. and Yoneki, E. (2008) Bubble rap: social-based forwarding in delay tolerant networks. In *Proc of ACM MobiHoc* (New York, NY, USA: ACM): 241–250. doi:http://doi.acm.org/10.1145/1374618.1374652.

[18] Conan, V., Leguay, J. and Friedman, T. (2007) Characterizing pairwise inter-contact patterns in delay tolerant networks. In *Proc. of ACM Autonomics*: 1–9.

[19] Lee, C.H. and Eunt, D.Y. (2009) Heterogeneity in contact dynamics: helpful or harmful to forwarding algorithms in DTNs? In *Proc. of WiOPT* (Seoul, Korea): 72–81.

[20] Groenevelt, R. and Nain, P. (2005) Message delay in MANETs. In *Proc. of Sigmetrics* (Banff, Canada: ACM): 412–413. See also R. Groenevelt, Stochastic Models for Mobile Ad Hoc Networks. PhD thesis, University of Nice-Sophia Antipolis, April 2005.

[21] Zhang, X., Neglia, G., Kurose, J. and Towsley, D. (2007) Performance modeling of epidemic routing. *Elsevier Computer Networks* **51**: 2867–2891.

[22] Bakhshi, R., Cloth, L., Fokkink, W. and Haverkort, B. (2008) Meanfield analysis for the evaluation of gossip protocols. *Sigmetrics Performance Evaluation Review archive* **36**: 31–39.

[23] Kushner, H.J. and Yin, G.G. (2003) *Stochastic Approximation and Recursive Algorithms and Applications* (Springer, 2nd Edition).

[24] Borkar, V.S. (1997) Stochastic approximation with two time scales (Elsevier).

[25] (2010) *Stochastic approximation to optimize the performance of human operators*.

[26] Borkar, V.S. (2008) *Stochastic Approximation: A Dynamical Systems Viewpoint* (Cambridge University Press).