Risk sensitive optimal control framework applied to Multi-dimensional delay tolerant networks

Eitan Altman, Veeraruna Kavitha, Francesco De Pellegrini, Vijay Kamble and Vivek Borkar

***Abstract*—Epidemics dynamics can describe the dissemination of information in delay tolerant networks, in peer to peer networks and in content delivery networks. The control of such dynamics has thus gained a central role in all of these areas. However, a major difficulty in this context is that the objective functions to be optimized are often not additive in time but are rather multiplicative. The classical objective function in DTNs, i.e., the successful delivery probability of a message within a given deadline, falls precisely in this category, because it takes often the form of the expectation of the exponent of some integral cost. So far, models involving such costs have been solved by interchanging the order of expectation and the exponential function. While reducing the problem to a standard optimal control problem, this interchange is only tight in the mean field limit obtained as the population tends to infinity.**

In this paper we identify a general framework from optimal control in finance, known as risk sensitive control, which let us handle the original (multiplicative) cost and obtain solutions to several novel control problems in DTNs. In particular, we can derive the structure of state-dependent controls that optimize transmission power at the source node. Further, we can account for the propagation loss factor of the wireless medium while obtaining these controls, and, finally, we address power control at the destination node, resulting in a novel threshold optimal activation policy. Combined optimal power control at source and destination nodes is also obtained. Finally, we demonstrate how risk-sensitive control applies to the case of multi-dimensional DTNs where multiple classes of mobiles exist.

***Index Terms*—Delay Tolerant Networks, Markov Decision Process, Risk Sensitive Control**

I. Introduction

Delay Tolerant Networks (DTNs) gained the interest of the research community in recent past [1], [2]. They have been identified as a promising mean to transport data in intermittently connected networks. DTNs in particular, sustain communications in a networked system where no continuous connectivity guarantee can be assumed [3], [4]. Messages are carried from source to destination via relay nodes adopting store and carry type forwarding protocols; such protocols basically rely on the underlying node mobility pattern. The core problem in DTNs is to efficiently route messages towards the intended destination. We observe that traditional techniques for routing perform very poorly in this context due to frequent disruptions, and furthermore mobile nodes rarely possess information on the upcoming encounters they are going to experience [5], [6]. An intuitive and rather robust solution is to disseminate multiple copies of the message in the network. This is meant to ensure that at least some of them will reach the destination node within some deadline [4], [7].

The above scheme is referred as epidemic-style forwarding [8], which is similar to the spread of infectious diseases. Each time a message-carrying node encounters an uninfected node, it infects this node by passing on the message. Finally, the destination receives the message when it meets an infected node. As in biology, DTNs literature refers to *contacts* as those events when the message can be forwarded.

In this paper, we confine our analysis to the *two hop routing protocol*. This choice is dictated both by efficiency reasons and by the possibility to implement all the forwarding control on board of the source node. In fact, under the two hop routing protocol, the source transmits copies of its message to all mobiles it encounters. A relay, conversely, forwards the message copy it has to the destination only [9].

Optimizing the performance of DTNs requires to maximize the successful delivery probability of a message within a given deadline. However, in order to do so, one has to trade off system resources to increase the success rate. We identify that both source and destination would like to reduce the delivery failure probability. However they have their own individual constraints on their resources (e.g., power) both of which are pivotal in influencing this delivery probability. So, we consider a joint optimization problem with soft constraint on the power of both the source and the destination. We consider two types of power control at source: 1) controlling power per transmission which in turn effects the *transmission range* and hence the contact rate; 2) controlling the number of copies delivered to the contacted mobiles while keeping the power per transmission fixed. The destination only controls the power per transmission. We further consider *the influence of pathloss factor of the underlying wireless medium on the first type of the power control.* Complete characterization of the optimal control is obtained for some of the control problems while for others, interesting analytical properties are derived. Numerical examples are provided to further support our results.

A general framework from optimal control in finance, known as risk sensitive control, is applied to above problems, as it deals with multiplicative costs. Below, we introduce the fundamental structure of this general optimization problem.

Delay Tolerant Networks and Risk Sensitive MDP

Consider a network with a set of I communities. A community i consists of a time varying number of members which we denote by $X_i(t)$. In the context of DTNs, the number X_i usually denotes the number of infected nodes within the i-th community. Now, consider some tagged node which we shall call the "destination". We assume that each member of community i has a Poisson contact process with the destination at a rate of ν_i. Conditioning on the processes $X_i(t)$, these Poisson processes are assumed to be independent, and thus the destination receives messages carried by members of the whole network at a (time varying) rate of $\sum_{i=1}^{I} X_i(t)\nu_i$. The probability that it receives no packets during a time interval $[s, t]$ is thus

$$P(\text{no reception}|X_i(r), r \in [s,t], i = 1, ..., I) = \exp\left(-\int_s^t \sum_i \nu_i X_i(r) dr\right)$$

Unconditioning, we get

$$P_f := P(\text{no reception}) = E\left[\exp\left(-\nu \int_s^t c(r)dr\right)\right];$$
$$c(r) = \sum_i \frac{\nu_i}{|\nu|} X_i(r) \quad (1)$$

where $\nu \neq 0$ is some appropriate constant and E denotes the expectation operator. If $X_i(t)$ are piecewise constant over the intervals $[n\Delta, (n+1)\Delta)$ then the integral in the exponent can be replaced by a summation over time: this is precisely the case considered in the rest of the paper.

Now assume that $X_i(t)$ is the function of some controlled Markov decision process $Y(t), \lambda(t)$ where Y is the state of the process and $\lambda(t)$ is the control. Then the expression P_f is a standard object in the theory of controlled Markov chains: this is the so called "risk sensitive cost criterion", which has had many applications in financial mathematics. The cost function we need to optimize in the paper, will actually be of the form (1) with $\nu > 0$; risk-sensitive control literature in economics classify this case as *risk avert*.[1]

In previous control papers in networking on delay tolerant see [7], [10], one changes the order of expectation and exponent and is thus faced with an optimization of the exponential function of the expected integral (instead of optimizing the expectation of the exponential of the integral). By Jensen's inequality we know that this optimizes a bound over the original function. This bound has been shown to be tight in several contexts as the population size grows (this is the mean field limit) [11].

What is common to the standard interpretation of a risk sensitive cost and to the one we have in our case? *The risk sensitive cost criterion accounts for the time correlation of the integrand. In contrast, the risk neutral bound obtained by optimizing the expectation of the exponent is only a function of the marginal distribution of the random variables.*

There is a wealth of tools to solve control problems with risk sensitive cost. In general one finds the same tools as in standard (additive cost) MDPs but they take instead, a multiplicative form (we shall see this later). Thus a Bellman equation exists as well, and under suitable conditions the actions that optimize this dynamic programming operator at a given state defines the optimal policy there. As in standard MDPs, we can use value iteration (that has in contrast a multiplicative form [12]) to compute the optimal value and policy for a finite horizon cost.

In literature, heterogeneous DTNs with multiple communities, i.e., several classes of nodes, are a current interest of the research ([13]). The tools we introduce throughout this work fit the general case (1). In this paper we mainly consider a single community DTN example. Initial results in a two community example are provided at the end of the paper in section VI.

II. System Model

Consider a large area with N active DTN mobiles and assume that connectivity is not guaranteed all the times. We consider a static source and destination: the communication between the two has to happen only with the help of the active mobiles. Source and destination can transmit/receive only with mobiles that are within their radio range. The radio range depends upon the power used by the transmitter per transmission. The mobiles are wandering freely in the area. A *contact* is said to happen with the source/destination whenever a mobile enters the radio range of the source/destination.

When the area is large and transmission range is small, this contact process is a rare phenomenon and is modeled by a Poisson process, i.e., the time to contact is exponentially distributed [14]. We further assume that each contact duration is sufficient enough to fully transfer the message.

Also, every message generated at the source has to be delivered to the destination within a given deadline T.

The source passes on the copy of the message to some of the contacted mobiles, taking into consideration its own resources. Recall that since we consider the two hop protocol, infected mobiles can only deliver the message to the destination. We say a delivery failed if none of the infected mobiles come in contact with the destination before the deadline T. The source has to spend its power for two purposes: 1) it has to continuously show its presence (done with help of beaconing), 2) it has to deliver the packets to the contacted mobiles using a wireless link, which requires positive transmission power. In any case, these devices are power limited and hence have constraint on the total power that can be used.

The more power the device uses per transmission the bigger is its transmission range and hence larger will be the rate with which the mobiles can contact. This identifies a clear trade off for active time and power in the system. In fact, larger power per transmission causes more mobiles infected per unit time but for a shorter duration, while lower power per transmission causes lesser number of mobiles infected per unit time but for longer duration. Thus the source has to use an optimal power strategy over the period of the deadline T to minimize the probability of delivery failure.

[1]Given a probability distribution of a nonnegative payoff, an agent is said to be non-sensitive to risk, if the agent has no preference between receiving that payoff or receiving its expectation. This case results in the limit $\nu \to 0$ in (1) (actually logarithm of the cost in (1) divided by ν), which converges to the original (additive cost) optimal control problem. A risk seeking agent, the case with $\nu < 0$, prefers receiving the random variable over its expectation, and a risk avert agent, case with $\nu > 0$, prefers always the expectation.

Also, the transmission range depends upon the propagation coefficient of the wireless medium. Hence the *optimal power strategy depends both upon the total power constraint and the propagation coefficient*.

Unlike the source, the destination receives the message only once and hence it spends most of its power to remain alive (i.e., for beaconing). In some cases it might wish to receive more copies of the messages, due to the possible communication errors. Even in this case, the number of copies received at the destination is much smaller than the number of mobiles infected by the source, as the reception is possible only after the joint occurrence of two rare events. Hence, in all cases, the destination mostly spends its power for remaining alive.

The *source and destination have a common goal, to minimize the delivery failure probability. This they have to do, under their own power constraints.* We obtain optimal policies for both the source and destination which minimize the power spent and maximize the probability of successful delivery. The most important contribution of this paper lies *in using the Risk sensitive MDP tools which facilitate in optimizing the exact cost function unlike many papers of this type.*

The rate of contacts, (λ, γ respectively for source-mobile, destination-mobile contacts), depend upon the speed of the mobiles, propagation characteristics of the wireless medium, the power per transmission at which the source/destination is operating [9], [14]. In Appendix B we relate the power spent to sustain a given contact rate and the attenuation coefficient.

We consider a time slotted system and the goal is to design the time dependent (and or state dependent) contact rates $\{\lambda_k, \gamma_k\}_k$ optimally. Let X_k denote the number of infected mobiles in the beginning of time slot k (which is of unit duration). This represents the state of the system. Then, $X_{k+1} = X_k + I_k$ where with $q_\lambda := e^{-\lambda}$ (probability of no contact in one time slot) for all $n_2 \leq N - n_1$

$$
\begin{aligned}
p^\lambda_{n_1, n_1+n_2} &:= P(I_k = n_2 | X_k = n_1) = p^\lambda_{n_2}(N - n_1) \text{ with} \\
p^\lambda_{n_2}(n_1) &:= \binom{n_1}{n_2} (1 - q_\lambda)^{n_2} q_\lambda^{n_1 - n_2},
\end{aligned}
$$

represents the control dependent transition probabilities. A message is not delivered within the deadline T: if it is "not delivered" by any one of the, X_t, infected mobiles present at the beginning of the slot, during the slot t and if this "no delivery" happens in all the T slots. Conditioned on the state trajectory, the events in various slots are independent and thus the overall delivery failure probability be will be product of delivery failure probability in each slot. The unconditioned delivery failure probability will be the expected value of this product of slot failure probabilities. Thus the probability of the message not reaching the destination within the time deadline T, for a given sequence of contact time parameters $\Lambda := \{\lambda_t\}_{0 \leq t \leq T-1}$, $\Upsilon := \{\nu_t\}_{0 \leq t \leq T-1}$ is given by,

$$
P_f(\Lambda, \Upsilon) = E\left[e^{-\sum_t \nu_t X_t} \right]
$$

Aim is to choose $\Lambda \in \Omega_S^T$, $\Upsilon \in \Omega_D^T$ with $\Omega_S := \{0, \lambda^1, \cdots \lambda^{M_S}\}$ $\Omega_D := \{0, \lambda^1, \cdots \lambda^{M_D}\}$ in a optimal way so as to minimize the failure probability P_f along side minimizing the power consumption at both the source and destination. The power constraint can be introduced in the problem in two ways:

Soft Power Constraint. In some applications, there is no hard constraint on the power spent, but rather one needs to minimize the total power spent. In this case we have two contrasting costs a) the cost for spending the power represented by a function $c(p)$ or equivalently $e^{c(p)}$, b) the cost if the message has not been delivered in terms of P_f. We would like to give fair importance to both the costs, in particular we chose proportional fair treatment to both the costs, by minimizing:

$$
E\left[P_f e^{hc(p)} \right] = E\left[e^{-\sum_t \nu_t X_t + hc(p)} \right],
$$

where h is the weight given to the power cost.

Hard Power Constraint. Some applications have hard constraint on the total power spent: one needs to minimize the failure probability under power constraint (for some $B < \infty$):

$$
\min E\left[e^{-\sum_t \nu_t X_t} \right] \text{ subject to } c(p) \leq B.
$$

We deal with hard constraint in the last section of this paper (section VI) while studying the two community problem while significant parts of this paper work with the soft constraint. We consider two types of power costs $c(p)$ in the following and consider the resulting source and or destination controls. We end this section by introducing the two types of pure source controls (i.e., destination rate ν is constant here). The other (pure destination and combined source-destination) controls are introduced incrementally in the sections after *pure source control.*

Pure source Delivery Control: The source does not change the power per transmission but rather optimizes the total power spent by controlling the number of message deliveries. Thus the transmission range remains constant and hence the (exponential) rate at which contact happens remains constant (say at λ). However when the source comes in contact with a mobile, it delivers the message with a probability $q_t \in [0, 1]$. Thus the effective contact rate is $\lambda_t = \lambda q_t$. This case *is appropriate whenever the source uses some smart techniques to spend minimal power for beaconing while most of it's power is utilized only for delivery of copies of the message. One such example is when the mobiles themselves detect the existence of the source and wake up the otherwise sleeping source.* Thus the power spent by the source depends directly upon the total number of infected mobiles, X_T, at the end of the delivery deadline and thus $c(p) \propto X_T$:

$$
\min_{\Lambda \in \Omega_S^T} E[e^{-\nu \sum_t X_t + hX_T}]. \tag{2}
$$

Pure source Power Control: The source does not control the number of message deliveries but rather controls the power per transmission. The transmission range varies (depending upon the path loss coefficient) with the power per transmission, which in turn varies the mean contact time, $1/\lambda$ (Appendix B).

With this control, the power is also used for maintaining the transmitting range high/low throughout (not just during the delivery of the message) and hence the beaconing capabilities also vary from slot to slot. This case *is appropriate for the situations in which the source has to spend significant power for both beaconing purposes as well as for delivery of message copies.* In this case, the power is inversely proportional to the contact rates (proportionality co-efficient depends on path loss co-efficient) and hence consider the following minimization[2]:

$$\min_{\Lambda \in \Omega_S^T} E\left[e^{-\nu \sum_i X_i + h \sum_i (\lambda_i)^\beta}\right]. \quad (3)$$

We obtain solutions to the above problems using risk sensitive dynamic programming equations (example, [15]).

III. Pure Source Controls

In this section, we obtain complete characterization of the optimal delivery control and some interesting properties of the optimal power control, when destination maintains it's power constant (and hence its contact rate remains constant at ν). We also obtain complete characterization of the optimal power control for the special case of hard controls, i.e., the controls for the case with only two possible actions $\Omega_S = \{0, \lambda_1\}$. These results are obtained using the risk-sensitive dynamic programming (Risk-MDP) equations (presented for example in [15]) for the case with full information. The state of this system is the number of infected customers, X_k, and the source (decision maker) obviously has full state information.

A. Delivery control

The delivery control problem (2) results in the following Risk-MDP equations ([15]), for every n:

$$u^T(n) = e^{-\nu n + hn} \quad (4)$$

$$u^t(n) = \min_{\lambda \in \Omega_S}\left(e^{-\nu n} \sum_{n \le n' \le N} p^\lambda_{n,n'} u^{t+1}(n')\right) \quad (5)$$

$$\lambda^*_t(n) = \operatorname*{argmin}_{\lambda \in \Omega_S}\left(e^{-\nu n} \sum_{n \le n' \le N} p^\lambda_{n,n'} u^{t+1}(n')\right) \quad (6)$$

Here, $\lambda^*_t(n)$ represents the optimal control at time t if the state X_t was at n. The minimum cost and the optimal control for any given initial condition X_0 are obtained by solving the above Risk-MDP equations using backward recursions. The above equations are solved backward recursively (starting with $t = T-1$) to obtain, $u^t(n)$, $\lambda^*_t(n)$ for all possible states $0 \le n \le N$ and all possible times $T \ge t \ge 0$. The minimum cost can now be read as $u^1(X_0)$ while the optimal control given by the sequence, $\lambda^*_1(X_0)$, $\lambda^*_2(X^*_1) \cdots \lambda^*_T(X^*_{T-1})$, where X^*_k is the optimal state trajectory obtained via forward recursion after replacing at each step, λ_t with $\lambda^*_t(X^*_{t-1})$. Solving the MDP equations recursively we obtain (proof in Appendix A):

[2]As explained in App. VII, $\beta = \alpha$ or $\alpha/2$, with α the path loss factor

Lemma 1: Define, $t^*_{Dc} := \inf_{t \le T}\{(T-t)\nu \le h\}$. Then,

$$\lambda^*_t(n) = 0, \quad u^t(n) = e^{-(T-t+1)\nu n + hn} \text{ if } t \ge t^*_{Dc} \text{ and}$$
$$\lambda^*_t(n) = \lambda_M \quad \text{if } t < t^*_{Dc} \quad \text{for all } n. \quad \square$$

Thus the optimal delivery control is a threshold control (deliver with maximum probability till threshold t^*_{Dc} and stop it completely after wards) and the interesting feature here is that the threshold is independent of the state n and the strength of the active population N. It only depends upon h, which represents the weight-age given to the power cost, ν the contact rate for the destination and T the delivery deadline. Larger deadlines (T), larger destination contact rates (ν) increase the threshold t^*_{Dc}. These two factors give more weight-age to the delivery probability cost and hence demand to deliver the message copies for larger durations. On the other hand, with larger h (i.e., larger weight-age to the cost occurred by spending the power) threshold t^*_{Dc} decreases.

B. Power control

The power control (3) will result in the following Risk-MDP equations ([15]) for any state n:

$$u^T(n) = e^{-\nu n}$$

$$u^t(n) = \min_{\lambda \in \Omega_S}\left(e^{-\nu n + h\lambda^\beta} \sum_{n \le n' \le N} p^\lambda_{n,n'} u^{t+1}(n')\right)$$

$$\lambda^*_t(n) = \operatorname*{argmin}_{\lambda \in \Omega_S}\left(e^{-\nu n + h\lambda^\beta} \sum_{n \le n' \le N} p^\lambda_{n,n'} u^{t+1}(n')\right)$$

The optimal control can be computed by backward recursion as in the previous section and as obtained in Appendix A:

$$\sum_{n \le n' \le N} p^\lambda_{n,n'} u^T(n') = e^{-\nu n}(e^{-\nu} + q_\lambda(1 - e^{-\nu}))^{N-n}$$

Clearly at $N = n$, $\lambda^*(T-1, N) = 0$. For $n < N$,

$$u^{T-1}(n) = e^{-\nu(n+N)}\left[\min_{\lambda \in \Omega}\left(e^{Q_n \lambda^\beta}[P_{T-1} q_\lambda + 1]\right)\right]^{N-n}$$

where $Q_n := \frac{h}{N-n}$ and $P_{T-1} := e^\nu - 1$. In this case we could only obtain a partial characterization of the control (proof in Appendix A):

Lemma 2: For all $n \ge n_{off}(\beta, h) := N - h\lambda_1^{\beta-1}$

$$\lambda^*_t(n) = 0 \text{ and } u^t(n) = e^{-(T-t+1)n} \text{ for all } t.$$

For all $n < n_{off}(\beta, h)$,

$$\lambda^*_t(n) \ge \lambda_1 \text{ for all } t. \quad \square$$

The following are consequences of the above lemma:
Switch off population : Immediately after the source infects $n_{off}(\beta, h)$ mobiles (which we refer as *switch off population*), it goes off to sleep mode. As β increases, i.e., as the path loss factor of the wireless medium increases, the switch off population decreases. This is intuitive as with β increasing

more amount of power has to be spent for remaining on and or for delivering the message. Thus the power would be used up more quickly.

Switch off/Sleep time : The above lemma only proves that there again exists a threshold beyond which the source does not deliver messages. In this control, it is done by not using any power for transmission. Hence, here the source actually goes to complete sleep mode after the threshold period. The threshold till which the source delivers is given by:

$$t^*_{Pc} = \inf_{t \le T} \{X^*_t > n_{off}(\beta, h)\},$$

where X^*_t is the optimal state trajectory.

Need not be a threshold policy: For delivery control the optimal threshold was threshold, but here the optimal policy may not be threshold, i.e., it is possible that the optimal control (which is non zero till t^*_{Pc}) can be time varying till t^*_{Pc}.

More power for adverse conditions : Whenever $N \le h\lambda_1^{\beta-1}$, the optimal control is to never switch on. This situation can arise either because the population is not sufficient enough, or because the path loss factor β is high and or because the first non zero contact rate λ_1 is high. This trivial situation can be avoided by reducing the constraint on power cost via the factor h (i.e., by reducing it).

Optimal control spreads over larger time frame and larger states if the first non zero contact rate λ_1 is small enough because as λ_1 decreases the switch off population n_{off} increases to N. That is, when control is possible over finer grid, then the control spreads over larger time frame and for more states. Further, this spread increases with the increase in the path loss factor β (as for $\lambda_1 < 1$, $\lambda_1^{\beta-1} \to 0$ as $\beta \to \infty$).

Finally, the following lemma provides complete characterization of the optimal risk-sensitive control for the case of two level control which is a direct consequence of the Lemma 2.

Lemma 3: When $\Omega_S = \{0, \lambda_1\}$ for all t,

$$\lambda^*_t(n) = \begin{cases} 0 & \text{if} \quad n \ge n_{off}(\beta, h) \\ \lambda_1 & \text{if} \quad n < n_{off}(\beta, h). \end{cases}$$

C. Numerical examples:

In general, since Q_n is a function of the state, the relations that appear in the DP backward recursion at earlier stages do not provide a simple closed form. Numerical results can be obtained for a certain choice of the parameters. In all these results u^* represents the optimal cost. We set $\Omega_S = [0, 1/\lambda_M, \cdots, \lambda_M]$ with $M = 10$. In particular we are interested here in the impact of the coefficient β since it expresses the cost paid by the source due to path loss factor on the optimal control. In Fig. 1 we observe the effect of such parameter for $\beta = 1, 2$. For $\beta = 1$, there exists a thresholding effect both on the state and time. Conversely, for $\beta = 2$, we observe a rather smooth decrease of the optimal control. As given by Lemma 2, for $h > N$ and $\beta = 1$ the control vanishes.

As indicated by the analysis, higher values of the path loss coefficient force the control to become less and less concentrated at small values of the state, compared to the case

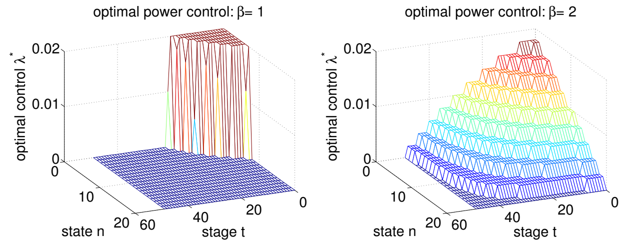

Fig. 1. Power control in the case of $\beta = 1$ (left), $\beta = 2$ (right); $\lambda_M = 2\nu = 2 \cdot 10^{-2}$; $\log(u^*) = -1.53$, $M = 10$, $N = 20$, and $T = 50$.

$\beta = 1$, i.e., to spread the control over larger time scales and larger number of states. It becomes more convenient to forward even when the number of infected nodes is high. However, this is done using a lower probability for smaller values of n. This kind of spread is seen only if the source can also operate at a very small but non zero contact rate (i.e., if λ_1 is small). On the other hand, if the source does not operate at sufficient number of levels of contact rates (as result λ_1 could be large), with larger values of β the control can become trivial. Hence whenever *the wireless medium experiences large propagation losses the source has to operate with higher granularity in the control space, i.e., it will require more smoother controls.*

In Fig. 2 we first depicted the effect of the risk parameter h on the forwarding control: it is apparent that larger penalties (larger h) force the control to switch off at earlier stages. In particular, for $h = N$ the control vanishes, i.e., no forwarding is performed at the source node.

Also, we depicted the impact of the path loss factor on the optimal control given a fixed optimal utility (u^*); we emulated the generation of 1000 optimal control sample paths at the source node and compared the average optimal control corresponding to the same average utility. We maintained u^* more or less same, by varying h with β. We considered two different values of u^* smaller value (higher successful delivery probability) obtained with smaller penalty (smaller h) for power cost and larger one with larger penalty.

We observe in Fig. 2 that for small values of the penalty, the behavior of the controls is similar, and resembles a threshold type policy. However, for larger penalties – this has the meaning of increasing the weight of energy expenditure - the control for $\beta = 1$ has still a threshold-like shape, whereas for larger values of the path loss factor, the decrease is gradual. Hence, even in time power spreading occurs, i.e., the source node is forced to reduce the forwarding probability, and to keep on forwarding over time. As in the case of threshold type policies, however, it is still convenient to spend energy as much as possible at the beginning of the interval, which explains the decreasing shape of the control.

IV. Pure Destination Controls

The destination cannot control the number of nodes infected and hence the only meaningful control in this case is power control. This is also meaningful as the destination spends majority of its power for beaconing purposes. In this section,

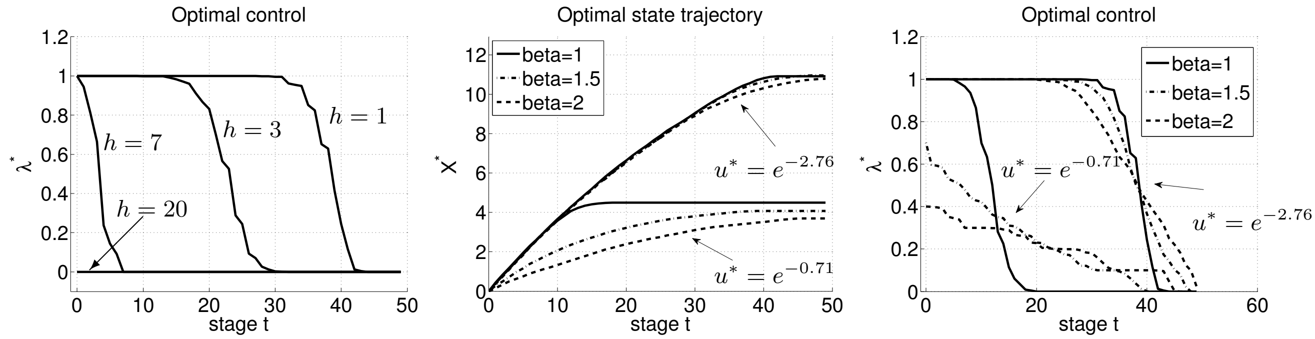

Fig. 2. Effect of h on the optimal control for $\beta = 1$ (left); Effect of β: optimal state trajectory (center) and optimal control (right); settings as in Fig. 1, control values are normalized, reference utilities are superimposed.

we consider the pure destination controls: source maintains its power/probability with which it delivers constant and hence the source contact happens at a constant exponential rate λ. Thus we consider the following:

$$\min_{\Upsilon \in \Omega_D^T} E\left[e^{-\sum_t \nu_t X_t + h \sum_t \nu_t^\beta} \right].$$

The message has to reach the destination before the end of the T^{th} time slot and we assume that only the infected mobiles present at the beginning of the slot can reach destination during the slot. Thus, the destination has to control even in the T^{th} time slot, while the source controls only till slot $T-1$. Thus the corresponding Risk-MDP equations ([15]) for all n are:

$$\begin{aligned} u^{T+1}(n) &= 1 \\ u^t(n) &= \min_{\nu \in \Omega_D} \left(e^{-\nu n + h\nu^\beta} \sum_{n \le n' \le N} p^\lambda_{n,n'} u^{t+1}(n') \right) \\ \nu_t^*(n) &= \operatorname*{argmin}_{\nu \in \Omega_D} \left(e^{-\nu n + h\nu^\beta} \sum_{n \le n' \le N} p^\lambda_{n,n'} u^{t+1}(n') \right) \end{aligned}$$

State ($\{X_t\}$) evolution is independent of control and hence,

$$u^t(n) = \min_{\nu \in \Omega_D} \left(e^{-\nu n + h\nu^\beta} \right) \sum_{n \le n' \le N} p^\lambda_{n,n'} u^{t+1}(n') \, \forall n.$$

Thus, for all t and n, the optimal control

$$\begin{aligned} \nu_t^*(n) &= g_\nu(n,\beta) \text{ where} \\ g_\nu(n,\beta) &:= \arg\min_{\nu \in \Omega_D} (-\nu n + h\nu^\beta). \end{aligned} \qquad (7)$$

With X_t representing the state trajectory, optimal policy, $\nu^*(t) = g_\nu(X_t, \beta)$. It is interesting to observe that the state evolution is independent of the control, the state still influences the system performance and hence the optimal policy depends upon the state value. However the optimal policy could be computed easily, because of the control independent state evolution. *This policy is obtained assuming that the destination also has access to the current state, X_t, the number of infected mobiles. This assumption is satisfied if the source and destination have a very low rate link (back-haul kind), via which few control signals can be exchanged.*

A. State estimation

In situations, where the destination does not have access to the state information, it can estimate the channel state X_t under the following conditions :

Large population limit: For large population, i.e., when N is very large (using strong law of large numbers) for almost all realizations,

$$X_{t+1} = X_t + \sum_{j=1}^{N-X_t} \zeta_j(t) \approx X_t + (N - X_t)(1 - e^{1-\lambda})$$

where $\zeta_j(t)$ is the indicator of the event that mobile j meets the source at time t, whose probability $P(\zeta_j(t) == 1) = 1 - e^{-\lambda}$.

Continuous approximation: As the slot duration tends to zero (note λ by our notations is a product of the actual contact rate and the slot duration and hence as $\lambda \to 0$), one can replace the state progression with that of a solution of the ODE:

$$\dot{X} = \lambda(N - X) \quad X(0) = 0;$$

Thus the optimal policy will approximately be given by:

$$\nu^*(t) = g_\nu(N - Ne^{-\lambda t}, \beta)$$

B. Switch on Threshold

From equation (7) the optimal destination contact, ν^* depend upon time t only via the state at time t, i.e., it depends only upon the number of infected mobiles n. From (7), the optimal control increases with n in the sense: $g_\nu(n,\beta)^{\beta-1} \le n/h$. In fact, there exists a switch on population and a switch on time, reaching which the destination will be switched on:

$$n_{on}(\beta, h) := \inf_n \{ n > h\nu_1^{\beta-1} \}, \ t^*_{Dest} := \inf_t \{ X_t \ge n_{on} \}.$$

The above so because for all $n \ge n_{on}$, $-n\nu_1 + h\nu_1^\beta < 0$ and hence $g_\nu(n,\beta) \ge \nu_1 > 0$. It is interesting to note that *destination has a switch on threshold while the source has a switch off threshold for optimal policies.* We will notice in the coming sections that *even for combined source-destination control, we have same (w.r.t. to the number of infected mobiles) switch on threshold for the destination.*

V. COMBINED SOURCE-DESTINATION CONTROL

The destination always does the power control. The source can either do the power control or the delivery control. Accordingly we have two different the combined controls.

A. Power Control at Source :

The source and destination are far way and static. Hence they can be working in very different geographical locations and hence can experience different path losses. The objective function to be minimized in this case will be :

$$\min_{\Lambda \in \Omega_S^T, \Upsilon \in \Omega_D^T} E\left[e^{-\sum_t \nu_t X_t + h_d \sum_t \nu_t^{\beta_d} + h_s \sum_t \lambda_t^{\beta_s}}\right]$$

This will result in the following risk sensitive Dynamic Programming equations ([15]) for all n (and for all $t \leq T$):

$$u^{T+1}(n) = 1$$

$$u^t(n) = \min_{\nu,\lambda}\left(e^{-\nu n + h_d \nu^{\beta_d} + h_s \lambda^{\beta_d}} \sum_{n \leq n' \leq N} p_{n,n'}^{\lambda} u^{t+1}(n')\right)$$

$$\nu_t^*(n), \lambda_t^*(n)$$

$$= \operatorname*{argmin}_{\nu \in \Omega_D, \lambda \in \Omega_S}\left(e^{-\nu n + h_d \nu^{\beta_d} + h_s \lambda^{\beta_d}} \sum_{n \leq n' \leq N} p_{n,n'}^{\lambda} u^{t+1}(n')\right)$$

It is easy to see that the above optimization can be separated as before,

$$u^t(n) = \min_{\nu \in \Omega_D}\left(e^{-\nu n + h_d \nu^{\beta_d}}\right) \min_{\lambda \in \Omega_S} e^{h_s \lambda^{\beta_s}} \sum_{n \leq n' \leq N} p_{n,n'}^{\lambda} u^{t+1}(n') \, \forall \, n.$$

Thus $\nu_t^*(n) = g_\nu(n, \beta)$ where g_ν is defined in (7). That is, the destination optimal control will be same as that in pure destination control. The source control will however be different, as now, the cost in the new state will depend upon the new state also via the optimal destination control. The optimal destination control only depends upon state n and not on time t directly. Hence, it is possible to define,

$$n_{\nu_{M_D}} := \min_{0 \leq n \leq N} \left\{ g_\nu(n', \beta) = \nu_{M_D} \text{ for all } n' \geq n \right\},$$

which is the minimum number of infected mobiles for which the destination forever switches to the maximum contact rate, ν_{M_D}. The number $n_{\nu_{M_D}}$ can easily be estimated as it is result of a simple optimization defined in (7). In the special case of hard controls, i.e., when $\Omega_D = \{0, \nu_1\}$,

$$n_{\nu_1} = n_{on}(\beta, h) = \min_n \{ n \geq h_d \nu_1^{\beta_d - 1} \}.$$

Once the destination forever switches to its maximum rate, analysis will be similar to that in pure source control and thus,

Lemma 4: Let $n_{off}(\beta, h) := N - h_s \lambda_1^{\beta_s - 1}$. For all t,

$$\lambda_t^*(n) = 0, \;\; u^t(n) = e^{-(T-t+1)n} \text{ if } n \geq \max\{n_{\nu_{M_D}}, n_{off}\}$$

$$\text{and } \lambda_t^*(n) \geq \lambda_1 \text{ for all } t \quad \text{if } n_{\nu_{M_D}} \leq n < n_{off}(\beta, h). \square$$

Numerical Examples: Here we would like to provide insight into the behavior of the joint source-destination control. In Fig. 3 we reported on the the effect of the power control operated at the destination, i.e., based on its dependence on the parameter h_d, on the power control operated at the source node for $\beta = 1$ and $h_s = 1$. In Fig. 3 (left) we observe the optimal policy: the presence of a switch on policy for the destination discourages forwarding at later stages in the case when the number of infected nodes is small. Thus the control in this case is much different from what happens in the case $h_d = 0$, where the threshold seen in over the companion graph in Fig. 1 is decreasing with the state for the same stage lag.

At larger states, there exists the opposite effect due to the cost of forwarding at the source, so that it is not worth to forward at too high rate when a large fraction of nodes have already been infected. Interestingly, the model predicts the presence of states, namely $n = 5$, when the optimal control operates for much longer number of stages at maximum rate.

However, if we increase the cost at the destination, which in turns means that the destination switches on later, there exists a sudden increase of the cost function: as seen in Fig. 3 (center) and (right), once h_s passes from 7 to 8, the optimal control vanishes abruptly, showing a thresholding effect on the cost at the destination.

B. Delivery Control at Source :

The objective function to be minimized in this case is:

$$\min_{\Lambda, \Upsilon} E\left[e^{-\sum_t \nu_t X_t + h_d \sum_t \nu_t^{\beta} + h_s X_T}\right]$$

This will result in the following risk sensitive Dynamic Programming equations ([15]) for all n:

$$u^{T+1}(n) = e^{h_s n}$$

$$u^t(n) = \min_{\nu,\lambda}\left(e^{-\nu n + h_d \nu^{\beta}} \sum_{n \leq n' \leq N} p_{n,n'}^{\lambda} u^{t+1}(n')\right)$$

$$\nu_t^*(n), \lambda_t^*(n)$$

$$= \operatorname*{argmin}_{\nu,\lambda}\left(e^{-\nu n + h_d \nu^{\beta}} \sum_{n \leq n' \leq N} p_{n,n'}^{\lambda} u^{t+1}(n')\right)$$

The above optimization can be separated again and as before, the destination optimal control is same as that in pure destination control case, i.e., $\nu_t^*(n) = g_\nu(n, \beta)$. Further using similar logic we have,

Lemma 5: For all $t \geq t_{Dc}^*$ and for all $n \geq n_{\nu_{M_D}}$,

$$\lambda_t^*(n) = 0, \;\; u^t(n) = e^{-(T-t+1)\nu_{M_D} n + hn},$$

where t_{Dc}^* is defined in Lemma 1 with ν replaced with ν_{M_D}. $\square$

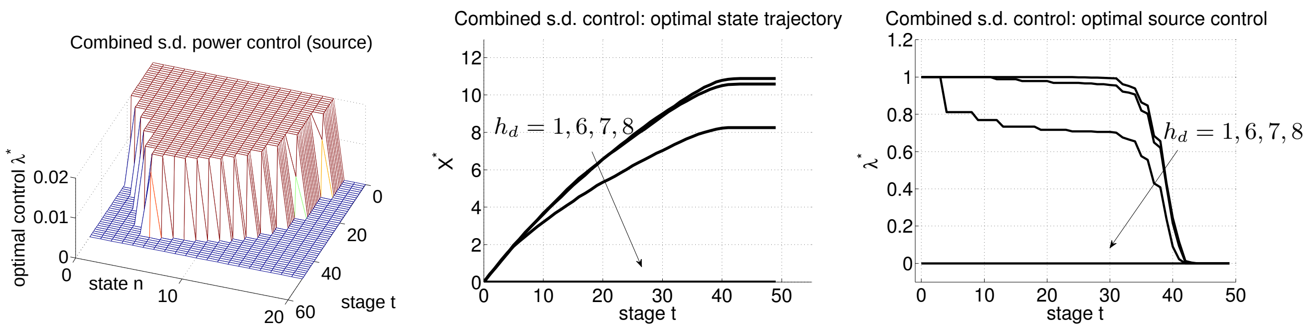

Fig. 3. Effect of h_d on the optimal control for $\beta = 1$ and $h_s = 1$: combined optimal control source destination at the source node (left), average optimal state trajectory (center), optimal source control policy (right); settings as in Fig. 1.

VI. A DTN EXAMPLE WITH TWO COMMUNITIES

In the following section we show how to apply the framework of risk-sensitive control can be extended to a more complicated DTN model. In particular, we consider a multidimensional domain and we refer also to a slightly different cost structure. We also handle *a hard constraint on the power* by appending a function of the power spent up to time t as a part of the current state X_t and then considering a state dependent action space.

Consider N mobile terminals in a two floor building. We consider a discrete time framework, where a fix source in the ground floor wishes to transmit a message to some fix destination in the top floor. There is no direct connectivity between the two floors, and the message can reach the destination only through relaying by mobiles that move from one floor to another. This model is an example of a general situation with two disjoint areas one of them containing the source and the other containing the destination and in which the source/destination stick to their respective areas. The communication between the source and destination is possible because some of the mobiles move across the two areas.

We assume that at each time slot, there is at most one mobile that moves from one floor to the other. The probability of such movement is p_1 for migration from floor 1 to floor 2 and p_2 for migration from floor 2 to floor 1. Both these to and fro migrations are assumed to be independent. If a mobile does move, then the one chosen to move is uniformly distributed among all mobiles on that particular floor.

The contact process that describes the instants at which each mobile at floor 1 (resp. floor 2) meets the source (resp the destination) is a Poisson process with parameter λ (resp. ν). Let $q = 1 - \exp(-\lambda)$ denote the probability that there is at least one encounter between a given mobile at floor 1 and the source. Let $X_i(t)$ be the number of mobiles at floor i at time t with a copy of the packet and let $Y_i(t)$ be the corresponding number without a packet.

The number $\Delta Y_1(t)$ of mobiles that receive a copy of the message in a time slot is thus multinomial with parameter q:

$$p_k(y_1) = P(\Delta Y_1(t) = k | Y_1(t) = y_1) = \binom{y_1}{k} q^k (1-q)^{y_1-k}$$

We assume that the transmission of the message is instantaneous and that a fix amount of energy ε is needed for each transmission.

Assume that there is a fixed source $\overline{s}$ of class s and a destination $\overline{d}$ in some class d.

A. The objective

Assume that a message is available starting from time 0 and that it has some deadline T to arrive at the destination. The number of arrivals of the message at the destination during time $[0, T]$ is a Poisson random variable with parameter $\nu \sum_{t=1}^{T} X_2(t)$. Therefore the probability of having no successful transmission of the message before the expiration of the deadline is given by

$$P_d = E[Z] \quad where \quad Z = \exp\left(-\nu \sum_{t=1}^{T} X_2(t)\right)$$

B. The Trade-off structures and related problems

We assume that the source has a finite battery life and it controls the effective contact rate $\lambda \in \{0, \lambda^M\}$. As considered in the single dimensional case in the previous sections, there are two main causes of energy consumption for the source: 1) The Delivery of the Message to a mobile and 2) Maintaining the transmission radius for beaconing purposes. Again depending on which of these causes is more prominent, we can have two different control problems. But the only difference is that instead of having a smooth trade-off as considered in the single dimensional problems, here we consider that the source has a hard constraint of a finite battery life. We further make the following assumptions:

Assumption A1: The source is aware of the mobility (i.e. which mobile arrives or departs from floor 1). It also recalls to which mobiles it transmits a copy of the message.

Assumption A2: We again consider two hop routing.

- Delivery Control : This case is appropriate when the main source of energy consumption is the delivery of a message. Here the contact rate is held constant at λ^M by maintaining the power of transmission, while the effective contact rate is controlled by choosing a delivery probability $q \in \{0,1\}$. Given the source has a finite battery life, let $c^{max}(< N)$ be the maximum number of transmissions it can afford. Thus the source controls the rate $\lambda \in \{0, \lambda^M\}$ while facing the following optimization problem

$$\min \mathbb{E}\left[\exp\left(-\nu \sum_{t=1}^{T} X_2(t)\right)\right]$$

$$X_1(T) + X_2(T) \leq c^{max}$$

We introduce an extra term, c^s_{left} which counts the number of mobiles infected so far to the state variable (as done in the next subsection) which we refer as $s = \left(Y_1^s, X_1^s, Y_2^s, X_2^s, c^s_{left}\right)$. We then have a state dependent control state, if $c^s_{left} < c^{max}$ then $\Omega_s = \{0, \lambda^M\}$ else $\Omega_s = \{0\}$. It now results in the following Dynamic Programming equations, for all s,

$$\begin{aligned} u^T(s) &= \exp(-\nu X_2^s) \\ u^t(s) &= \min_{\lambda \in \Omega_S}\left(\exp(-\nu X_2^s)\sum_{s' \in S} p^{\lambda}_{ss'} u^{t+1}(s')\right) \\ \lambda_t^*(s) &= \operatorname*{argmin}_{\lambda \in \Omega_S}\left(\exp(-\nu X_2^s)\sum_{s' \in S} p^{\lambda}_{ss'} u^{t+1}(s')\right) \end{aligned}$$

We expect to show that the following will be the optimal control strategy: Deliver always with probability $q = 1$ (hence keeping the contact rate high at λ^M) until the number of transmissions (number of mobiles with the message) reaches c^{max}. We are currently working towards proving this.

- Power Control : This case pertains to the situation where the source spends significant power for beaconing purposes and the contact rate $\lambda \in \{0, \lambda^M\}$ is controlled by varying the power of transmission (choosing a non-zero value of power corresponding to the rate λ^M or zero). Let $\zeta(t) \in \{0,1\}$ be an indicator function indicating whether the effective contact rate chosen at time slot t was high (λ^M). Given the finite battery life of the source, let $\bar{c}^{max}(< T)$ be the maximum number of time slots the source can afford to keep the power of transmission, and hence the contact rate high at λ^M. Thus the source controls the rate $\lambda \in \{0, \lambda^M\}$ while facing the following optimization problem

$$\min \mathbb{E}\left[\exp\left(-\nu \sum_{t=1}^{T} X_2(t)\right)\right]$$

$$\sum_{t=1}^{T} \zeta(t) \leq \bar{c}^{max}$$

This is a dynamic optimization problem with a constraint on the number of times you can take a particular action. In the next section we solve this problem by transforming it into a unconstrained risk sensitive MDP by incorporating this constraint into the state space.

C. Risk Sensitive Dynamic Programming for the power control problem

Let S denote the state space of the Markov Decision Process with each element $s \in S$ of the form

$$s = \left(Y_1^s, X_1^s, Y_2^s, X_2^s, c^s_{left}\right)$$

Here, $c_{left} \in \{0, 1, \cdots, \bar{c}^{max}\}$ denotes the remaining number of time slots the source can keep the contact rate high at λ^M. At each time t, depending on the state and the history, the source chooses the meeting rate $\lambda \in \{0, \lambda^M\}$. Each time the contact rate λ^M is chosen by the source, the c_{left} reduces by 1. When c_{left} reaches 0, the only action available to the source is $\lambda = 0$. Let $p^{\lambda}_{ss'}$ denote the probability of transition from state s to s', $s, s' \in S$ when meeting rate $\lambda \in \{0, \lambda^M\}$ is chosen by the source. These transition probabilities for this model can easily be derived and this is done in Appendix C. Due to Assumptions A1-A2, this again falls into the framework of Risk-MDP with full state information ([15]): 1) in pure source controls, only the source takes decisions, 2) the source can track the exact number of nodes that have and that do not have a copy of the message at each floor.

The following are the Dynamic Programming equations, for all $s \in S$,

$$\begin{aligned} u^T(s) &= \exp(-\nu X_2^s) \\ u^t(s) &= \min_{\lambda \in \Omega_S}\left(\exp(-\nu X_2^s)\sum_{s' \in S} p^{\lambda}_{ss'} u^{t+1}(s')\right) \\ \lambda_t^*(s) &= \operatorname*{argmin}_{\lambda \in \Omega_S}\left(\exp(-\nu X_2^s)\sum_{s' \in S} p^{\lambda}_{ss'} u^{t+1}(s')\right) \end{aligned}$$

As before, $\lambda_t^*(s)$ for $t = 1, \cdots, T$ and $s \in S$ is the optimal Markov Control policy for the process. $u^1(s)$ is the optimized objective function for each starting state $s \in S$.

D. Numerical Examples

We made a few simple observations through our numerical computations which appear to be intuitively obvious:

- At any state $s \in S$ and time slot t if all the mobiles without a copy of the message are at level 1 and if the source has free slots left to transmit, then the optimal action for the source is to keep the contact rate high no matter how much time is left till deadline.
- At any state $s \in S$ and time slot t if none of the mobiles without a copy of the message are at level 1 then the

T_{left}	$\lambda^*_{T_{left}}(A)$	$\lambda^*_{T_{left}}(B)$
1	3	3
2	3	3
3	3	3
4	0	3
5	0	0
6	0	0
7	0	0
8	0	0

TABLE I
OPTIMAL CONTACT RATE AT STATES A AND B

optimal action for the source is to keep the contact rate at 0.

But the optimal policies for the states which lie between these two extreme cases are interesting. In order to illustrate the kind of trade-off faced by the source in this situation and the nature of the control policy, we take up a simple example. Consider $N = 4$ mobiles, $\lambda^M = 3$ and $\nu = 4$. Consider the following states $A = (1, 2, 1, 0, 1)$ and $B = (1, 1, 2, 0, 1)$. In both these states the source is left with only one free slot for keeping the contact rate high. Following is an illustration of the two states.

$$A : \frac{\square}{\square \times \times}$$

$$B : \frac{\square\square}{\square\times}$$

Where $\square$ represents a mobile without a copy of the message while $\times$ represents a mobile bearing a copy of the message. Let the time remaining till deadline T be denoted by T_{left}. For varying values of T_{left} we compute the optimal control values $\lambda^* \in \{0, \lambda^M\}$ at the two states A and B. The results are presented in the table. As suggested by the results, depending on how much time is left, the optimal control of the source at a particular state may vary in the case where the mobiles without a copy of the message are present on both levels. If more time is left till the deadline, the source might want to wait till some of the mobiles without a packet migrate from level 1 to level 2, so that it can infect more mobiles in a single time slot. But as the deadline gets closer, it is better for the source to utilize its free slots quickly rather than waiting.

VII. CONCLUSIONS

In this paper we introduced a novel framework for the control of systems where the objective function has the form of a risk-sensitive cost. We applied such theory to DTNs, where delivery probability presents natively a risk-sensitive form. We provided several variants of the basic control problem where the success probability is traded off for power consumption in presence of propagation loss effects. We specialized the solution to the case when the control is performed at the source, at the destination or jointly. Compared to existing works, a whole set of new state-dependent closed loop policies provide here new insight in the control of DTNs. Finally, we ported our model to the case of multi-dimensional DTNs, where several classes of mobiles exist and the control applied at the source has to account for the type of mobiles that act as relays.

REFERENCES

[1] S. Burleigh, L. Torgerson, K. Fall, V. Cerf, B. Durst, K. Scott, and H. Weiss, "Delay-tolerant networking: an approach to interplanetary Internet," *IEEE Comm. Magazine*, vol. 41, pp. 128–136, June 2003.

[2] L. Pelusi, A. Passarella, and M. Conti, "Opportunistic networking: data forwarding in disconnected mobile ad hoc networks," *IEEE Communications Magazine*, vol. 44, no. 11, pp. 134–141, November 2006.

[3] A. Chaintreau, P. Hui, J. Crowcroft, C. Diot, R. Gass, and J. Scott, "Impact of human mobility on opportunistic forwarding algorithms," *IEEE Transactions on Mobile Computing*, vol. 6, pp. 606–620, 2007.

[4] T. Spyropoulos, K. Psounis, and C. Raghavendra, "Efficient routing in intermittently connected mobile networks: the multi-copy case," *ACM/IEEE Transactions on Networking*, vol. 16, pp. 77–90, Feb. 2008.

[5] M. M. B. Tariq, M. Ammar, and E. Zegura, "Message ferry route design for sparse ad hoc networks with mobile nodes," in *Proc. of ACM MobiHoc*, Florence, Italy, May 22–25, 2006, pp. 37–48.

[6] W. Zhao, M. Ammar, and E. Zegura, "Controlling the mobility of multiple data transport ferries in a delay-tolerant network," in *Proc. of IEEE INFOCOM*, Miami USA, March 13–17 2005.

[7] E. Altman, T. Başar, and F. De Pellegrini, "Optimal monotone forwarding policies in delay tolerant mobile ad-hoc networks," in *Proc. of ACM/ICST Inter-Perf*. Athens, Greece: ACM, October 24 2008.

[8] A. Vahdat and D. Becker, "Epidemic routing for partially connected ad hoc networks," Duke University, Tech. Rep. CS-2000-06, 2000.

[9] R. Groenevelt, P. Nain, and G. Koole, "The message delay in mobile ad hoc networks," *Performance Evaluation*, vol. 62, no. 1-4, pp. 210–228, October 2005.

[10] E. Altman, G. Neglia, F. De Pellegrini, and D. Miorandi, "Decentralized stochastic control of delay tolerant networks," in *Proc. of INFOCOxM*, Rio de Janeiro, Brazil, April 19-25 2009.

[11] E. Altman, "Competition and cooperation between nodes in delay tolerant networks with two hop routing," in *Proc. of Netcoop*, Eindhoven, The Netherlands, November 2009.

[12] D. Jacobson, "Optimal stochastic linear systems with exponential criteria and their relation to differential games," *IEEE Trans. Automat. Control*, vol. 18, pp. 124–131, 1973.

[13] A. Chaintreau, J.-Y. L. Boudec, and N. Ristanovic, "The age of gossip: Spatial mean-field regime," in *Proc. of ACM Sigmetrics*, Seattle, June 2009.

[14] R. Groenevelt and P. Nain, "Message delay in MANETs," in *Proc. of Sigmetrics*. Banff, Canada: ACM, June 6 2005, pp. 412–413, see also R. Groenevelt, Stochastic Models for Mobile Ad Hoc Networks. PhD thesis, University of Nice-Sophia Antipolis, April 2005.

[15] S. Coraluppi and S. I. Marcus, "Risk-sensitive queueing," in *Proc. 35th Annual Allerton Conf. on Communication, Control, and Computing*, Urbana, IL, September 28-October 1.

APPENDIX A

Proof of Lemma 1: We start with time $t = T - 1$. For this time, one can easily simplify the λ dependent part of (5) as,

$$\sum_{n \le n' \le N} p^{\lambda}_{n,n'} u^T(n')$$

$$= \sum_{0 \le n' \le N-n} \binom{N-n}{n'} q_\lambda^{N-n-n'} (1-q_\lambda)^{n'} e^{-(\nu-h)(n'+n)}$$

$$= \quad e^{(-\nu+h)n} (q_\lambda + (1-q_\lambda)e^{-\nu+h})^{N-n}$$

Now if $\nu < h$ then, $e^{-\nu+h} > 1$ and because of monotonicity of the function $\lambda \mapsto e^{-\nu+h} + q_\lambda(1 - e^{-\nu+h})$ we have,

$$\lambda^*_{T-1}(n) = 0 \text{ and } u^{T-1}(n) = e^{-2\nu n + hn}$$

If instead $\nu > h$ then $\lambda^*_{T-1}(n) = \lambda_M$ for all n. Assume that at the t-th step $u^t(n) = e^{(-(T-t+1)\nu+h)n}$. We can we write the inductive step

$$\begin{aligned} u^{t+1}(n) &= \min_{\lambda\in\Omega_S}\left(e^{-\nu n}\sum_{n\le n'\le N} p^\lambda_{n,n'} e^{(-(T-t+1)\nu+h)n}\right) \\ &= e^{(-\nu(T-(t+1)-1)+h)n} \\ &\quad \min_{\lambda\in\Omega_S}(q_\lambda + (1-q_\lambda)e^{(-(T-t+1)\nu+h)n}))^{N-n} \end{aligned}$$

from which again the control, as at stage $T-1$, only takes extreme values. Thus the first part of the lemma is proved.

For the second part, let t_1 represent the greatest t less than t^*_{Dc}, i.e., t_1 is the greatest time with $(T-t_1)\nu > h$. Define $\chi := e^{-(T-t_1)\nu+h}$ and note that $\chi < 1$ and that $u^{t_1+1}(n) = \chi^n$. As before, for any n, the following simplification results

$$\sum_{n\le n'\le N} p^\lambda_{n,n'} u^{t_1+1}(n') = \chi^n\left(\chi + q_\lambda(1-\chi)\right)^{N-n}.$$

By monotonicity of the function $\lambda \mapsto \chi + q_\lambda(1-\chi)$ we have, $\lambda^*_{t_1}(n) = \lambda_M$ and

$$\begin{aligned} u^{t_1}(n) &= \zeta^n(\chi + q_{\lambda_M}(1-\chi))^N \text{ for all } n \text{ where} \\ \zeta &:= \frac{e^{-\nu}\chi}{\chi + q_{\lambda_M}(1-\chi)}. \end{aligned}$$

Let $t_2 := t_1 - 1$. Then by simplifying as before,

$$\begin{aligned} &\sum_{n\le n'\le N} p^\lambda_{n,n'} u^{t_1}(n') \\ &\qquad = \zeta^n\left(\chi + q_{\lambda_M}(1-\chi)\right)^N\left(\zeta + q_\lambda(1-\zeta)\right)^{N-n}. \end{aligned}$$

Clearly $\zeta < 1$ and hence for all n, $\lambda^*_{t_2}(n) = \lambda_M$ and

$$\begin{aligned} u^{t_2}(n) &= \left(\chi + q_{\lambda_M}(1-\chi)\right)^N\left(\zeta + q_{\lambda_M}(1-\zeta)\right)^N \\ &\qquad \left(\frac{\zeta e^{-\nu}}{\zeta + q_{\lambda_M}(1-\zeta)}\right)^n. \end{aligned}$$

The last sentence of the lemma is proved by backward induction, using similar logic, all the way till $t = 1$. □

Proof of Lemma 2: Note that $\lambda^*_t(n) = \arg\min_{\lambda\in\Omega_S} f_{T-1}(\lambda)$ with

$$f_{T-1}(\lambda) := e^{Q_n\lambda^\beta - \lambda}[P_{T-1} + e^\lambda] \tag{8}$$

We obtain λ^*_t with the help of continuous space minimizer $\lambda^* := \arg\min_{\lambda\in[0,\lambda_M]} f_{T-1}(\lambda)$ While f_{T-1} is minimized over the interval $[0, \lambda_M]$, we note that it is a product of two functions. The first function is initially decreasing with λ and then starts to increase while the second function is always increasing with λ. The first function at maximum can decrease till $\bar\lambda := \inf_\lambda\{Q_n\lambda^\beta < \lambda\}$, because after this $\bar\lambda$, both the functions are increasing with λ. Thus, $\lambda^* < \bar\lambda$. Thus if the first non zero rate $\lambda_1 > \bar\lambda$, i.e., if

$$Q_n\lambda_1^\beta > \lambda_1, \text{ then } \arg\min_{\lambda\in\Omega_S} f_{T-1}(\lambda) = 0, u^{T-1} = e^{-2\nu n},$$

as in this case $f_{T-1}(0) < f_{T-1}(\lambda_1)$ and f_{T-1} is increasing beyond λ_1. Therefore,

$$\lambda^*_{T-1}(n) = 0 \text{ and } u_{T-1}(n) = e^{-2\nu n} \text{ for all } n \ge n_{off}(\beta, h).$$

For such n, in similar way we can write that,

$$\begin{aligned} u^{T-2}(n) &= e^{-\nu(n+2N)}\left[\min_{\lambda\in\Omega} f_{T-2}(\lambda)\right]^{N-n} \quad \text{with} \\ f_{T-2}(\lambda) &:= \left(e^{Q_n\lambda^\beta-\lambda}[e^{2\nu} - 1 + e^\lambda]\right) \end{aligned}$$

Now using similar logic as before,

$$\lambda^*_{T-2}(n) = 0 \text{ and } u_{T-2}(n) = e^{-3\nu n} \text{ for all } n \ge n_{off}(\beta, h).$$

This continues and we have the first part of the lemma.

When $n < n_{off}$, $Q_n\lambda_1^\beta < \lambda_1$ and so from (8),

$$f_{T-1}(0) > f_{T-1}(\lambda_1) \implies \lambda^*_{T-1}(n) = \lambda_j \text{ for some } j \ge 1.$$

With $o^t_\lambda(n) := e^{-\nu n + h\lambda^\beta}\sum_{n\le n'\ge n} p^\lambda_{n,n'} u^{t+1}(n)$, $u^t(n) = \min_{\lambda\in\Omega_S} o^t_\lambda(n)$. The result is proved by backward mathematical induction: if it holds for some $t+1$ with $\lambda^*_{t+1}(n) = \lambda_j$ for some $j \ge 1$, then it holds at t, as for this n

$$\begin{aligned} o^t_0(n) &= e^{-\nu n} u^{t+1}(n) \\ &= e^{-\nu n} e^{-\nu n + h\lambda_j^\beta}\sum_{n\le n'\le N} p^{\lambda_j}_{n,n'} u^{t+2}(n') \\ &= e^{-\nu n + h\lambda_j^\beta}\sum_{n\le n'\le N} p^{\lambda_j}_{n,n'} o^{t+1}_{\lambda_j}(n') \\ &\ge e^{-\nu n + h\lambda_j^\beta}\sum_{n\le n'\le N} p^{\lambda_j}_{n,n'} u^{t+1}(n') = o^t_{\lambda_j}(n), \end{aligned}$$

and so, $\lambda^*_t(n) \ne 0$ and hence $\lambda^*_t(n) \ge \lambda_1$. ◇

APPENDIX B

Contact times and propagation coefficient:

Let v be the speed of a tagged mobile with range R. Assume that most of the power is used to send short beacons spaced by some time interval dt. Case (i): velocities and ranges are such that all the mobiles in its range are likely to be different than those in the previous beacon transmission. Then the contact rate is proportional to the area covered at each transmission of a beacon, i.e. to R^2. The range needed for a given contact rate is thus of the order of square root of the rate.

Case (ii): If in contrast, the velocities are not so large, then the rate of contacts is proportional to $\pi R v dt$.

Assume we need power p to be able to receive a beacon at a range R. If the path loss is α then we need to transmit at a power of pR^α. Within case 2, the contact rate is linear in R so the transmission power needed behaves like $p\lambda^\alpha$ (where λ is the contact rate). If the price is proportional to the power then we conclude that the price for a given contact rate is of order of λ^α. Within case 1, since the range is of the order of $\sqrt{\lambda}$, then the power needed to obtain λ is of the order of $\lambda^{\alpha/2}$. With cost linear in the power used, this is also the order of the cost for a rate of λ.

Appendix C : Transition probabilities for Two Floor model

Let us define the following vectors which model the migrations of the mobiles in a particular time slot from one level to another:

$$t_1 = (-1, 0, 1, 0, 0)$$

$$t_2 = (0, -1, 0, 1, 0)$$

$$t_3 = (1, 0, -1, 0, 0)$$

$$t_4 = (0, 1, 0, -1, 0)$$

$$t_5 = (-1, 1, 1, -1, 0)$$

$$t_6 = (1, -1, -1, 1, 0)$$

For each state $s \in S$ define the vectors $m_s^{\lambda^M,k} = s + (-k, k, 0, 0, -1)$ and $m_s^{0,k} = s$ for $k = \{0, \cdots, Y_1^s\}$ which model the message transmission by the source at level 1. Then the transition probabilities $p_{ss'}^{\lambda}$ for $s' \in S$ are given as follows:

- Case 1: $m_s^{\lambda,k} = s'$
  If $m_s^{\lambda,k}(3) + m_s^{\lambda,k}(4) = 0$ then
  $$p_{ss'}^{\lambda} = p_k(Y_1^s)(1 - p_1)$$
  Otherwise if $m_s^{\lambda,k}(1) + m_s^{\lambda,k}(2) = 0$ then
  $$p_{ss'}^{\lambda} = p_k(Y_1^s)(1 - p_2))$$
  For remaining cases
  $$p_{ss'}^{\lambda} = p_k(Y_1^s)(1 - p_2)(1 - p_1)$$
  $$+p_k(Y_1^s)p_1 p_2 \frac{m_s^{\lambda,k}(1)}{m_s^{\lambda,k}(1) + m_s^{\lambda,k}(2)} \frac{m_s^{\lambda,k}(3)}{m_s^{\lambda,k}(3) + m_s^{\lambda,k}(4)}$$
  $$+p_k(Y_1^s)p_1 p_2 \frac{m_s^{\lambda,k}(2)}{m_s^{\lambda,k}(1) + m_s^{\lambda,k}(2)} \frac{m_s^{\lambda,k}(4)}{m_s^{\lambda,k}(3) + m_s^{\lambda,k}(4)}$$
- Case 2: $m_s^{\lambda,k} + t_1 = s'$
  If $m_s^{\lambda,k}(3) + m_s^{\lambda,k}(4) = 0$ then
  $$p_{ss'}^{\lambda} = p_k(Y_1^s)p_1 \frac{m_s^{\lambda,k}(1)}{m_s^{\lambda,k}(1) + m_s^{\lambda,k}(2)}$$
  Otherwise
  $$p_{ss'}^{\lambda} = p_k(Y_1^s)p_1(1 - p_2) \frac{m_s^{\lambda,k}(1)}{m_s^{\lambda,k}(1) + m_s^{\lambda,k}(2)}$$
- Case 3: $m_s^{\lambda,k} + t_2 = s'$
  If $m_s^{\lambda,k}(3) + m_s^{\lambda,k}(4) = 0$ then
  $$p_{ss'}^{\lambda} = p_k(Y_1^s)p_1 \frac{m_s^{\lambda,k}(2)}{m_s^{\lambda,k}(1) + m_s^{\lambda,k}(2)}$$
  Otherwise
  $$p_{ss'}^{\lambda} = p_k(Y_1^s)p_1(1 - p_2) \frac{m_s^{\lambda,k}(2)}{m_s^{\lambda,k}(1) + m_s^{\lambda,k}(2)}$$
- Case 4: $m_s^{\lambda,k} + t_3 = s'$
  If $m_s^{\lambda,k}(1) + m_s^{\lambda,k}(2) = 0$ then
  $$p_{ss'}^{\lambda} = p_k(Y_1^s)p_2 \frac{m_s^{\lambda,k}(3)}{m_s^{\lambda,k}(3) + m_s^{\lambda,k}(4)}$$
  Otherwise
  $$p_{ss'}^{\lambda} = p_k(Y_1^s)p_2(1 - p_1) \frac{m_s^{\lambda,k}(3)}{m_s^{\lambda,k}(3) + m_s^{\lambda,k}(4)}$$
- Case 5: $m_s^{\lambda,k} + t_4 = s'$
  If $m_s^{\lambda,k}(1) + m_s^{\lambda,k}(2) = 0$
  $$p_{ss'}^{\lambda} = p_k(Y_1^s)p_2 \frac{m_s^{\lambda,k}(4)}{m_s^{\lambda,k}(3) + m_s^{\lambda,k}(4)}$$
  Otherwise
  $$p_{ss'}^{\lambda} = p_k(Y_1^s)p_2(1 - p_1) \frac{m_s^{\lambda,k}(4)}{m_s^{\lambda,k}(3) + m_s^{\lambda,k}(4)}$$
- Case 6: $m_s^{\lambda,k} + t_5 = s'$
  $$p_{ss'}^{\lambda} = p_k(Y_1^s)p_1 p_2 \frac{m_s^{\lambda,k}(1)}{m_s^{\lambda,k}(1) + m_s^{\lambda,k}(2)} \frac{m_s^{\lambda,k}(4)}{m_s^{\lambda,k}(3) + m_s^{\lambda,k}(4)}$$
- Case 7: $m_s^{\lambda,k} + t_6 = s'$
  $$p_{ss'}^{\lambda} = p_k(Y_1^s)p_1 p_2 \frac{m_s^{\lambda,k}(2)}{m_s^{\lambda,k}(1) + m_s^{\lambda,k}(2)} \frac{m_s^{\lambda,k}(3)}{m_s^{\lambda,k}(3) + m_s^{\lambda,k}(4)}$$
- Also for all $s \in S$ such that $c_{left}^s = 0$ (no more free slots left) $p_{ss'}^{\lambda^M} = p_{ss'}^0$ for all $s' \in S$

where $k = \{0, \cdots, Y_1^s\}$.