A Threshold Type Policy of a DTN Node Under Fixed Reward Incentive Mechanisms

Tatiana Seregina, Olivier Brun, Balakrishna Prabhu

A Threshold Type Policy of a DTN Node Under Fixed Reward Incentive Mechanisms*

Tatiana Seregina[1,2], Olivier Brun[3], and Balakrishna J. Prabhu[3]

[1]École Nationale de l'Aviation Civile, Toulouse, France
[2]Toulouse Business School, University of Toulouse, France
E-mail: ts.tseregina@gmail.com
[3]LAAS-CNRS, Université de Toulouse, CNRS, Toulouse, France
E-mail: {brun, balakrishna.prabhu}@laas.fr

February 13, 2018

Abstract

The technology of Delay Tolerant Networks (DTN) has been designed to support communication in environments where connectivity is intermittent and communication delays can be very long. We focus on game-theoretic model for DTNs. We consider the model where the source proposes a fixed reward to persuade selfish mobile nodes to participate in relaying messages. The mobile relays can decide to accept or not the packet and then to drop the packet in the future. This game can be modelled as a partially-observable stochastic game. For two relays, we have shown that the optimal policies for the relays relates to the threshold type.

Keywords: delay tolerant networks, reward incentive mechanism, partially-observable stochastic game

1 Introduction

In the last ten years, a substantial research effort has been devoted to Delay Tolerant Networking (DTN) [8,18] for enabling data transfer between mobile devices in environments where there is no communication infrastructure. In such environments, nodes constantly move and can communicate with each other only when they enter each others communication range. Since end-to-end connectivity is only sporadic, DTN implements a store-carry-forward scheme, in which data packets are transiently stored in a network device to be later forwarded to

*The work of T. Seregina was done when she was a PhD candidate at LAAS and appeared in her thesis.

the destination. In other words, DTN divides the end-to-end path into multiple DTN hops, each intermediate node receiving packets and temporarily storing them until an opportunity to send the packet to the destination or to another intermediate node.

The assumption that mobile nodes may serve as relays with a premise that they can store information for a long time before forwarding it reflects a main idea of DTN architecture. Due to random node mobility and uncertainty in connectivity, DTN algorithms commonly use multi-copy routing for message delivery, when the message is delivered if one of the relay nodes with a copy encounters the destination. Replication of the original message by the so-called epidemic routing protocol ensures that at least some copy will reach the destination node with high probability and with a minimum delivery delay. Flooding the network with messages, Epidemic routing leads, however, to a significant resource consumption. To avoid overloading the network with messages while retaining a high delivery performance, the two-hop routing scheme provides simple and more efficient variant of the epidemic-style routing. Under this scheme, forwarding of a message copy is allowed in at most two steps, when a relay received the message from the source can not transmit it to another relay node but only if it encounters the destination.

However, in DTN applications, readiness to participate in forwarding is rather uncommon. DTN nodes are controlled by rational entities, such as people or organizations that can be expected to behave selfishly. When a mobile node needs to conserve its power or due to other individual objectives, it may not be willing to serve as a relay in data transmission, a link may then not be established and the packet will be terminated by the node. Selfish behaviour of DTN nodes and corresponding decentralized nature of their decision making requires mechanisms that should offer appropriate incentives for the nodes to behave in ways that are favourable for the network as a whole.

Several schemes have been proposed in the literature to foster participation of mobile nodes in DTNs: reputation-based schemes [2, 10, 14, 19–21], barter-based schemes [3, 4, 17] or credit-based schemes [5, 6, 6, 7, 11–13, 22–24, 24]. Most of the above works are simulation-based and do not provide explicit expressions of the probability of message delivery. One exception is the work of [1] which proposes a fixed reward to the first relay to deliver the message. Another related work that contains performance analysis of a reward mechanism is [15]. The scheme in there is based upon that of [1]. However, instead of a proposing a fixed reward to each relay, the source proposes a variable reward that depends upon the meeting time with the source and the information given by the source to the relay. In these two schemes, we avoid the use of feedbacks that allow relays to know whether the message has been successfully delivered or not. This is an important technical issue in DTNs since feedback messages may incur large delays. In [15], we assume that the relay proposed by the source to a relay it meets offsets the expected cost of this relay for delivering the message, so that a relay always accepts the message. We obtain explicit expressions of the probability of message delivery under different information settings.

In the present work, we consider the fixed reward mechanism as in [1] but assume that the relays are in competition and do not cooperate. A relay meeting the source is not informed of the existence of other message copies. Assuming a given lifetime for the message and homogeneous relays (that is, all relays have the same probability of meeting the source or the destination at next time step), we considered the discrete-time decision problem faced by a relay. When it meets the source, a relay has to decide whether to accept the message or not, and once the relay has the message it has to choose to retain or to drop it at subsequent decision epochs. Each relay makes its decisions in order to minimize the expected cost it incurs for participating.

We model the interaction between mobile nodes as a stochastic game with partial information. For the single player case, we obtain a necessary condition for the relay to attempt the delivery of the message that reflects a minimal value of the reward and show that the relay's strategy to accept the message from the source is of a threshold type. For two players, we establish that if one of the players follows a threshold type policy then the other one will also use a similar strategy.

The rest of the paper is organized as follows. We describe the problem in Section 2, and formulate it as a stochastic game with partial information in Section 3. The optimal strategy of a single player is investigated in Section 4, whereas Section 5 is devoted to the game with two relays. Some conclusions are drawn in Section 6.

2 Problem description

Consider a set of nodes in which there is one source, one destination, and N relays. The relays are mobile and meet the source or the destination every once in a while. It shall be assumed that the inter-meeting time between a relay and the source (resp., destination) is a sequence of i.i.d. geometric random variables with distribution function with parameter q (resp. p). Two nodes can exchange data only when they meet. It is assumed that the source and the destination are fixed, and thus cannot communicate directly.

After a message is generated, the source proposes it to every relay that it meets. A relay can choose to either accept the message or reject it. As an incentive, the source offers a fixed reward, say R, to be claimed by the first relay that delivers the message to the destination. We emphasize that only the first relay to deliver the message gets the reward, R. The other relays are not entitled to any share of the proposed reward. A relay that accepts the message incurs certain costs:

1. cost related to the energy spent in receiving the message from the source. This is fixed cost and will be denoted by C_r;

2. energy cost of transmitting the message to the destination in case this relay is the first one to do so. This cost is also fixed, and will be denoted by C_d.

3. and the cost of storing the message while the relay is searching for the destination. We denote by C_s the cost per unit time incurred for storing the message.

Associated with each message is a deadline before which the message remains useful to the destination. Once the deadline has passed, the destination will no longer accept the message from the relays.

The decision problem for a relay, when it meets the source, is whether to accept the message or not. In case it accepts the message, the relay can drop the message at any time if it has not yet delivered it to the destination, and if it is no longer profitable to keep the message. The precise optimization problem for the relays is described next.

3 Stochastic game with partial information

We shall study a discrete-time model of this game. The source generates the message at time instant 0 with a deadline at instant $\tau+1$. It is assumed that the reception of the message from the source and its transmission to the destination each takes one time slot, so that a relay has to meet the destination before time τ in order to get the reward. When a relay meets the source it can decide whether to accept or reject the message (assuming it does not already have it). Once the relay accepts the message it can choose to retain or to drop it in each subsequent time slot until it meets the destination or the deadline of the message expires. Thus, the potential decision epochs for every relay are in the set $\{0, 1, \dots, \tau-1\}$. Each relay has to make decision over multiple stages and its cost depends upon its own actions as well as those of the other relays. The objective of each relay is to minimize expected cost it incurs for participating in the game. This strategic interaction between the relays falls within the framework of stochastic games introduced by [16]. In our model, each relay is aware of its own state but does not know that of the others. Furthermore, it does not know whether the packet has already been delivered to the destination or not. Our game is thus a stochastic game with partial information [9]. We now give some background on this type of games. These games are defined by:

- τ : time horizon (message deadline, in our case)
- $\mathcal{R} = \{1, 2, \dots, N\}$ set of players (relays)
- $\mathcal{E}_j$, $j \in \mathcal{R}$: state space of relay j. We denote by X_n^j the state of player j at time n.
- $\mathcal{A}_j$, $j \in \mathcal{R}$: action space of relay j. We denote by A_n^j the action taken by player j at time n.
- $\mathcal{E} := \bigotimes_{j\in\mathcal{R}} \mathcal{E}_j$.
- $\mathcal{A} := \bigotimes_{j\in\mathcal{R}} \mathcal{A}_j$.

- $\mathcal{B}_j : \mathcal{E}_j \times \{0, 1, \dots, \tau - 1\} \to \mathcal{D}(\mathcal{A}_j)$, where $\mathcal{D}(\mathcal{A})$ is the set of probability measures on $\mathcal{A}$. The set $\mathcal{B}_j(t)$ is the set of mixed strategies available to relay j at every time instants. In other words, an element $\sigma_n^j(x)$ is the probability distribution over the set of actions $\mathcal{A}_j$ used by player j to choose its action when it is in state x at time n.

- P_j, $j \in \mathcal{R}$: transition probability matrix of relay j on the space of its state-action pairs.

- $\mathcal{E}_0$: state space of the packet. This can be 0 or 1 which indicates whether the packet has been delivered or not. We denote by X_n^0 the state of the packet at time n.

- $g_j : \mathcal{E}_j \times \mathcal{A}_j \times \mathcal{E}_0 \to \mathbb{R}$, $j \in \mathcal{R}$: cost function for relay j.

Fix $\sigma := \left(\sigma^j\right)_{j \in \mathcal{R}} \in \bigotimes_{j \in \mathcal{R}} \mathcal{B}_j$. Let $\left\{ \mathbf{Z}_n^\sigma := (X_n^j, A_n^j)_{j \in \mathcal{R}}^{(\sigma)} \right\}_{n=0,\dots,\tau-1}$ be the stochastic process of state-action pairs generated by σ. And, assume that the process X_n^0, $n \geq 1$ is adapted to the natural filtration of $\mathbf{Z}_n^\sigma$. By this we mean that, at every time instant, X_n^0 is measurable with respect to the history of the state-action pairs.

Let $\mathbf{b}_{-j} \in \mathcal{D}(\mathcal{E}_{-j})$ be the distribution of the initial state of the relays other than j. The expected cost of relay j for σ can then be defined as:

$$V_j(\sigma^j, \sigma^{-j}; x_0^0, x_0^j, \mathbf{b}_{-j}) = \mathbb{E}_{\mathbf{x}_0, \mathbf{b}_{-j}} \sum_{n=0}^{\tau-1} \alpha^n g_j(X_n^j, A_n^j, X_n^0), \tag{1}$$

where α is the discount factor. The terminal cost is assumed to be 0 in every state.

The objective of relay j is to minimize its cost given the strategy of the others. That is,

$$W_j(\sigma^{-j}; x_0^0, x_0^j, \mathbf{b}_{-j}) = \min_{s \in \mathcal{B}_j} V_j(s, \sigma^{-j}; x_0^0, x_0^j, \mathbf{b}_{-j}), \tag{2}$$

and compute

$$\beta_j(\sigma^{-j}; x_0^0, \mathbf{b}_{-j}) = \arg\min_{s \in \mathcal{B}_j} V_j(s, \sigma^{-j}; x_0^0, x_0^j, \mathbf{b}_{-j}), \tag{3}$$

which is the best-response of relay to σ_{-j} given the intial conditions.

This is a partially observable stochastic game (see for example, [9]) since each relay knows only its state but not that of the others. A consequence of the lack of information is that the concept of Markov strategies and Markov equilibrium is not applicable to this setting. The optimal action of a relay in a given state depends on the state of the other relays which is not known to this relay. The probability distribution over the states of the other relays will depend upon the actions they have been taking in the past. This means that a relay will have to keep track of the past actions of the others in order to compute its own action in a given state. The probability of arriving in a given state depends on the actions taken in the past because the action in the current state will depend upon the state of the other relays which is not known.

A policy σ is said to be an equilibrium if

$$\beta_j(\sigma^{-j}; x_0^0, \mathbf{b}_{-j}) = \sigma^j, \ \forall j. \tag{4}$$

The values of different parameters for our model are as follows.

State and action spaces

The state of each relay takes one of the five possible values:

Value	Significance	Action set
0	relay does not have the packet	$\emptyset$
m_s	relay meets the source	$(accept, reject)$
1	relay has the packet	$(drop, keep)$
m_d	relay meets the destination	$\emptyset$
2	relay quits the game	$\emptyset$

In states 0 and 2 the relay does not have a non-trivial action. In state 0 it is waiting to meet the source, while in state 2 it has already quit the game.

Transition matrix

Regarding the contact process that keeps track of the contacts of the relay with the source and the destination, we shall assume i.i.d. contact times. As a consequence, a relay needs to know only the current state of the contact process, and not its entire history to take its decision. In the following, we let p be the probability that a relay meets the destination at the next time step, and q be the probability that it meets the source. The state of each relay evolves according to a time-homogeneous Markov chain whose transition probabilities depend on the action chosen in each state, and is given by:

$$P_j = \begin{array}{c} \\ {\scriptstyle 0} \\ {\scriptstyle m_s} \\ {\scriptstyle 1} \\ {\scriptstyle m_d} \\ {\scriptstyle 2} \end{array} \begin{array}{c} \begin{array}{ccccc} {\scriptstyle 0} & {\scriptstyle m_s} & {\scriptstyle 1} & {\scriptstyle m_d} & {\scriptstyle 2} \end{array} \\ \begin{bmatrix} 1-q & q & 0 & 0 & 0 \\ \mathbb{1}_{reject} & 0 & \mathbb{1}_{accept} & 0 & 0 \\ \mathbb{1}_{drop} & 0 & (1-p)\mathbb{1}_{keep} & p\mathbb{1}_{keep} & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{array}$$

The transition diagram of the Markov chain is shown in Figure 1.

State of the packet

The state of the packet can take two values: 0 (it has not been delivered) or 1 (it has been delivered). The transition probabilities between these two states

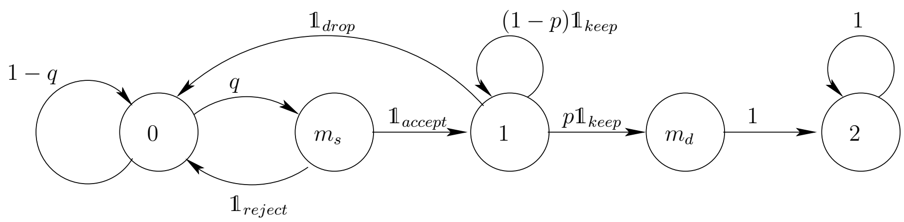

Figure 1: Transition diagram for the Markov chain governing the state of each relay.

depends upon the state of the relays.

$$P(X^0_{n+1} = 1 | X^0_n = 0, \mathbf{X}_n) = P((\cup_{j\in\mathcal{R}}\{X^j_n = 2\} = \emptyset) \cap (\cup_{j\in\mathcal{R}}\{X^j_n = m_d\} \neq \emptyset)),$$
$$P(X^0_{n+1} = 1 | X^0_n = 1, \mathbf{X}_n) = 1.$$

Cost function

The one-step cost incurred by the relay depends on its current state and the action it takes (whether it accepts the packet or not, whether it drops the packet or not). Further, when it meets the destination (that is, in state m_d) the cost incurred depends upon whether any other relay has already delivered the packet or not. Hence

$$\begin{aligned} g(m_s, accept, \cdot) &= C_r, \\ g(1, keep, \cdot) &= C_s, \\ g(m_d, \cdot, 0) &= R - C_d, \end{aligned}$$

and is 0 for all other arguments.

4 The Single Player Case

In order to get some insights into the structure of the best-response policy of a relay, we shall first consider the case of a single player. In order to simplify notations, we drop the index j of the relay. Since no other relay can deliver the message, the state of the packet is $X^0_n = 0$ until the relay meets the destination, and thus we can further simplify notations by writing $g(x, a)$ instead of $g(x, a, 0)$.

4.1 Dynamic Programming Formulation

Assume that the relay meets the source at instant $t \in [0, \tau]$. For epochs $0, \ldots, t-1$, thus, there are no decisions to be made. For the remaining epochs, the optimal policy can be computed using Dynamic Programming.

Let $V_n(x)$ be the optimal cost-to-go starting in state $x \in \{0, m_s, 1, m_d, 2\}$ at instant n. From the dynamic programming equation,

$$V_n(X_n) = \min_{a \in \mathcal{A}(X_n)} g(X_n, a) + \alpha \mathbb{E} V_{n+1}(X_{n+1}), \tag{5}$$

where α is the discount factor ($0 \leq \alpha < 1$).

At time n, if the relay is in contact with the destination, its terminal cost is

$$V_n(m_d) = C_d - R, \quad n = 1, 2, \ldots, \tau. \tag{6}$$

In particular, we have $V_\tau(m_d) = C_d - R$ at time τ. If at that time the relay has the message and is not in contact with the destination, then it is optimal to drop the message since it is no longer useful, so that $V_\tau(1) = 0$. On the other hand, if the relay does not have the message at instant τ, then it incurs no costs, so that $V_\tau(0) = V_\tau(m_s) = 0$. To summarize, the terminal costs at the instant $n = \tau$ are:

$$V_\tau(m_d) = C_d - R. \tag{7}$$

$$V_\tau(x) = 0, \quad \forall x \neq m_d. \tag{8}$$

The optimal policy at different decision epochs and states can be computed recursively by rolling back (5). If the contact process is history dependent, then the optimal policy is usually computed numerically. However, as we shall see below, the assumption of an i.i.d. contact process enables the derivation of structural properties of the optimal policy.

4.2 To Drop or to Retain

Assume that the relay is in state 1 at instant $\tau - 1$, that is it has the message and it is not in contact with the destination. The relay has to decide whether to drop the message or not. Taking $n = \tau - 1$ in (5), we obtain

$$\begin{aligned} V_{\tau-1}(1) &= \min_{a \in \{keep, drop\}} [g(1, a) + \alpha \mathbb{E} V_\tau(X_\tau)] \\ &= \min\left(0, C_s + \alpha(p V_\tau(m_d) + \overline{p} V_\tau(1))\right), \\ &= \min\left(0, C_s + p\alpha(C_d - R)\right), \end{aligned} \tag{9}$$

where we have used the short-hand notation $\overline{p} = 1 - p$, and the last equality follows from (7)–(8). Thus, if the first term is the minimum, then it is optimal to drop the message at $\tau - 1$, otherwise it is optimal to keep it.

One can recursively develop (9) to compute the optimal policy at step n given that the relay has the message and has not yet encountered the destination. For $n = \tau - 2$, we obtain

$$\begin{aligned} V_{\tau-2}(1) &= \min\left\{0, C_s + \alpha(p V_{\tau-1}(m_d) + \overline{p} V_{\tau-1}(1))\right\}, \\ &= \min\{\ 0,\ C_s + p\alpha(C_d - R), \\ &\qquad C_s(1 + \overline{p}\alpha) + (C_d - R)\alpha(p + \overline{p}p\alpha)\}. \end{aligned} \tag{10}$$

Here, the second and the third terms in the minimum correspond to the cost of retaining the message at instant $\tau - 2$. Thus, if either term is negative, then it is optimal to retain the message. Otherwise, it is optimal to drop the message at instant $\tau - 2$.

More generally, the i^{th} component in the min corresponds to the cost obtained if the action *keep* is played i consecutive times starting from the current decision epoch n, until the relay meets the destination or decides to drop the message. This i^{th} component can be represented as follows,

$$
\begin{aligned}
U_{n,i} &= \sum_{j=1}^{i} (\alpha \overline{p})^{j-1} \left(C_s + (C_d - R)\alpha p\right), && (11)\\
&= (C_s + \alpha p(C_d - R)) \frac{1 - (\overline{p}\alpha)^i}{1 - \overline{p}\alpha}. && (12)
\end{aligned}
$$

The recursion (9) can be developed in terms of $U_{n,i}$ as:

$$V_n(1) = \min(0, U_{n,1}, U_{n,2}, \ldots, U_{n,\tau-n}). \tag{13}$$

The optimal policy at instant n is to retain the message if either of $U_{n,i}$ is negative. Otherwise it is optimal to drop the message at time n. Note from (12) that if $C_s + \alpha p(C_d - R) < 0$, then $U_{n,i} < 0$, $\forall n$ and $\forall i$, and the sequence decreases with i. From (13), one can conlude that if $C_s + \alpha p(C_d - R) < 0$, then the relay will retain the message until it is delivered to the destination or the deadline expires. Otherwise, the relay will drop the message immediately. Thus,

$$R > \frac{C_s}{\alpha p} + C_d, \tag{14}$$

is a necessary condition for the relay to attempt the delivery of the message.

4.3 To Accept or to Reject

Assume that the relay is in state m_s at instant t, that is it in contact with the source and has not the message. The relay has to decide whether to accept the message or not. The optimal cost at t is:

$$
\begin{aligned}
V_t(m_s) &= \min(0, g(m_s, \mathit{accept}) + \alpha V_{t+1}(1)) \\
&= \min(0, C_r + \alpha V_{t+1}(1)), && (15)
\end{aligned}
$$

where $V_{t+1}(1)$ can be computed from (13). Thus, if at time t the second term is negative, then it is optimal to accept the message from the source. Otherwise, it is optimal to reject it. In particular, if condition (14) is satisfied, $U_{n,i}$ is a decreasing function of i and equation (13) yields

$$V_n(1) = U_{n,\tau-n} = (C_s + \alpha p(C_d - R)) \frac{1 - (\overline{p}\alpha)^{\tau-n}}{1 - \overline{p}\alpha}. \tag{16}$$

We thus obtain that the expected cost for the relay if it accepts the message is

$$g(m_s, accept) + \alpha V_{t+1}(1) = C_r + U_{t+1,\tau-t-1}. \tag{17}$$

We conclude from (15) and (17) that if the relay meets the source at time t, it will accept the message provided that

$$C_r + U_{t+1,\tau-t-1} < 0. \tag{18}$$

Note that (16) implies that $U_{t+1,\tau-t-1}$ increases with t. Since $U_{t+1,\tau-t-1}$ is negative and increases with t, there exists a threshold t^* such that for $t \leq t^*$, the relay will accept the message, and it will reject the message after t^*. The threshold can be easily computed using the above inequality,

$$t^* = \tau - 1 - \frac{\ln\left(1 + \frac{C_r(1-\overline{p}\alpha)}{C_s + \alpha p(C_d - R)}\right)}{\ln(\overline{p}\alpha)}. \tag{19}$$

5 Game with two relays

We now consider the newtork with two relays. We shall restrict our attention to threshold type policies, that is policies such $\sigma_n^j(m_s) = accept$ if $n \leq \theta_1$ and *reject* otherwise, and $\sigma_n^j(1) = drop$ if $n \geq \theta_2$, and *keep* otherwise. The threshold θ_2 could depend on the meeting time with the source. We shall show that if one relay follows a threshold type policy then the best-response of the other relay is also a policy of threshold type.

We shall thus assume that one of the two relays – say relay 2, follows a threshold type policy. That is, there exist θ_1^2 and $\theta_2^2 > \theta_1^2$ such that

$$\sigma_n^2(m_s) = \begin{cases} accept & \text{if } n \leq \theta_1^2, \\ reject & \text{if } n > \theta_1^2, \end{cases} \tag{20}$$

and

$$\sigma_n^2(1) = \begin{cases} keep & \text{if } n \leq \theta_2^2, \\ drop & \text{if } n > \theta_2^2. \end{cases} \tag{21}$$

As in Section 4, we shall use dynamic programming to derive the best-response policy of the first player to the above policy of relay 2. We let $V_n^1(x)$ be the optimal cost-to-go starting in state $x \in \{0, m_s, 1, m_d, 2\}$ at instant n. As we shall see below, the optimal cost-to-go starting in states m_s and 1 can be expressed in terms of the expected costs when the destination is reached.

5.1 Expected costs when the destination is reached

If at time n relay 1 has the message and is in contact with the destination, then its expected cost is

$$V_n^1(m_d) = \frac{1}{2}(C_d - R)\,\mathbb{P}\big(X_n^2 = m_d\big) + (C_d - R)\,\mathbb{P}\big(X_n^0 = 0, X_n^2 \neq m_d\big)\,, \quad (22)$$

for all $n \in \{1, 2, \ldots, \tau\}$, where it is assumed that if both relays meet the destination at the same time, then each one wins the reward with probability $\frac{1}{2}$.

Define $1 - \delta_n$ as the probability that relay 2 delivers the message at a time $t \leq n$, as estimated by relay 1 . Note that $\delta_{n-1} - \delta_n$ is the probability that the second relay meets the destination with the message precisely at time n. The expected cost $V_n^1(m_d)$ can be written as follows

$$\begin{aligned} V_n^1(m_d) &= \frac{1}{2}(C_d - R)\,(\delta_{n-1} - \delta_n) + (C_d - R)\,\delta_n, \\ &= \frac{\delta_{n-1} + \delta_n}{2}\,(C_d - R). \end{aligned} \quad (23)$$

Lemma 1 proves two fundamental properties of the sequence $V_1^1(m_d), V_2^1(m_d), \ldots$ that will be required to establish the structure of the optimal policy of relay 1.

Lemma 1. *The sequence $V_1^1(m_d), V_2^1(m_d), \ldots$ is such that*

(a) it is non-decreasing with n, and

(b) it is constant for all $n \geq \theta_2^2 + 1$.

Proof. To prove assertion (a), observe that since $\delta_{n-1} - \delta_n$ is the probability was defined before, we have $\delta_{n-1} \geq \delta_n$. Hence the sequence $\delta_1, \delta_2, \ldots$ is non-increasing. With (23), it yields $V_{n+1}^1(m_d) - V_n^1(m_d) = \frac{1}{2}(C_d - R)(\delta_{n+1} - \delta_{n-1}) \geq 0$, which concludes the proof.

Let us now prove assertion (b). Since at time $\theta_2^2 + 1$ the second relay drops the message if it has it, the probability that it delivers the message after that time is 0, implying that $\delta_n = \delta_{\theta_2^2+1}$ for all $n > \theta_2^2$. For $k > \theta_2^2 + 1$, it yields

$$V_k^1(m_d) = \frac{\delta_{k-1} + \delta_k}{2}\,(C_d - R) = \delta_{\theta_2^2+1}\,(C_d - R) = V_{\theta_2^2+1}^1(m_d), \quad (24)$$

which concludes the proof. □

5.2 To drop or to retain

Let us assume that relay 1 is in state 1, that is it has the message but it is not in contact with the destination. It has to decide whether to retain it or to drop it. Proceeding backward in time, we have

$$
\begin{aligned}
V^1_{\tau-1}(1) &= \min_{a\in\{keep,drop\}}[g(1,a)+\alpha\mathbb{E}V^1_\tau(X^1_\tau)], \\
&= \min\left(0, C_s+\alpha p V^1_\tau(m_d)+\alpha\overline{p}V^1_\tau(1)\right), \\
&= \min\left(0, C_s+\alpha p V^1_\tau(m_d)\right), \qquad (25)
\end{aligned}
$$

and

$$
\begin{aligned}
V^1_{\tau-2}(1) &= \min_{a\in\{keep,drop\}}[g(1,a)+\alpha\mathbb{E}V^1_{\tau-1}(X^1_{\tau-1})], \\
&= \min\left(0, C_s+\alpha\left[pV^1_{\tau-1}(m_d)+\overline{p}V^1_{\tau-1}(1)\right]\right), \\
&= \min\left(0, C_s+\alpha p V^1_{\tau-1}(m_d)\right., \\
&\qquad \left. C_s+\alpha p V^1_{\tau-1}(m_d)+\alpha\overline{p}\left[C_s+\alpha p V^1_\tau(m_d)\right]\right). \qquad (26)
\end{aligned}
$$

More generally, we have

$$
V^1_n(1) = \min(0, U_{n,1}, U_{n,2}, \ldots, U_{n,\tau-n}), \qquad (27)
$$

where

$$
U_{n,i} = \sum_{j=1}^{i}(\alpha\overline{p})^{j-1}\left[C_s+\alpha p V^1_{n+j}(m_d)\right]. \qquad (28)
$$

The optimal policy at instant n is to retain the message if $\min\limits_{i=1,\ldots,\tau-n} U_{n,i} < 0$. Otherwise it is optimal to drop the message at n.

We establish below two properties of the $U_{n,i}$.

Lemma 2. *The sequence $\{U_{n,1}\}_{n=1,2,\cdots}$ is a non-decreasing sequence, which is constant starting from $n=\theta_2^2$.*

Proof. We first show that the sequence is non-decreasing. With (28) we have

$$
\begin{aligned}
U_{n+1,1}-U_{n,1} &= C_s+\alpha p V^1_{n+2}(m_d)-C_s-\alpha p V^1_{n+1}(m_d), \\
&= \alpha p\left(V^1_{n+2}(m_d)-V^1_{n+1}(m_d)\right),
\end{aligned}
$$

and with Lemma 1 we can conclude that $U_{n+1,1}\geq U_{n,1}$ that corresponds to the first assertion of the lemma. In order to show that $U_{n,1}=U_{\theta_2^2,1}$ for all $n\geq\theta_2^2$, we use Lemma 1.(b) to obtain

$$
\begin{aligned}
U_{n,1} &= C_s+\alpha p V^1_{n+1}(m_d) \\
&= C_s+\alpha p V^1_{\theta_2^2+1}(m_d) \\
&= U_{\theta_2^2,1}
\end{aligned}
$$

□

Lemma 3. *For all $n \in \{1, 2, \ldots, \tau\}$, if $U_{n,1} \geq 0$, then* $\min_{i=1,\ldots,\tau-n} U_{n,i} = U_{n,1}$.

Proof. Fix $n \in \{1, 2, \ldots, \tau\}$ and assume $U_{n,1} \geq 0$. It is enough to prove that the sequence $U_{n,1}, U_{n,2}, \ldots$ is a non-decreasing sequence. Observing from (28) that $U_{n,i}$ can also be written as follows

$$U_{n,i} = \sum_{j=0}^{i-1} (\alpha \overline{p})^j U_{n+j,1}, \tag{29}$$

we obtain with Lemma 2 that $U_{n,i+1} - U_{n,i} = (\alpha \overline{p})^i U_{n+i,1} \geq (\alpha \overline{p})^i U_{n,1}$. We thus conclude that $U_{n,1} \geq 0$ implies that $U_{n,1}, U_{n,2}, \ldots$ is a non-decreasing sequence, which yields the proof. $\square$

We now show the following result.

Proposition 1. *At time n, $V_n^1(1) < 0$ if and only if $U_{n,1} < 0$.*

Proof. From (27), it is obvious that $U_{n,1} < 0$ implies that $V_n^1(1) < 0$. By contraposition, in order to show that the converse is true, it is enough to show that $U_{n,1} \geq 0$ implies that $V_n^1(1) \geq 0$, which is a direct consequence of Lemma 3. $\square$

According to Lemma 2, the $U_{n,1}$ are non-decreasing with n. Thus, Proposition 1 implies that relay 1 will retain the message as long as $U_{n,1} < 0$, and will drop it once $U_{n,1}$ becomes positive. We are now in position to show that once relay 1 has the message, it uses a threshold type strategy to decide whether to retain it or to drop it.

Proposition 2. *If $U_{\theta_2^2,1} \geq 0$ then there exists threshold $\theta_2^1 \leq \theta_2^2$ such that relay 1 retains the message until θ_2^1 and drops it at time $\theta_2^1 + 1$. Otherwise, if $U_{\theta_2^2,1} < 0$, relay 1 retains the message until it meets the destination or the deadline expires.*

Proof. Let us first consider the case $U_{\theta_2^2,1} \geq 0$. Let t be the time at which relay 1 accepts the message from the source. Since

$$V_t^1(m_s) = min\left(0, C_r + V_{t+1}^1(1)\right),$$

has to be negative for relay 1 to accept the message, this implies that $V_{t+1}^1(1) < -C_r$. According to Proposition 1, $V_{t+1}^1(1) < 0$ in turn implies that $U_{t+1,1} < 0$. Since from Lemma 2 the sequence $U_{1,1}, U_{2,1}, \ldots$ is non-decreasing, $U_{t+1,1} < 0$ and $U_{\theta_2^2,1} \geq 0$ imply that there exists $\theta_2^1 \in [t+1, \theta_2^2]$ such that $U_{n,1} < 0$ for all $n \leq \theta_2^1$ and $U_{\theta_2^1+1,1} \geq 0$. We thus conclude that $V_n^1(1) < 0$ for all $n \leq \theta_2^1$ and $V_{\theta_2^1+1}^1(1) \geq 0$. In other words, relay 1 retains the message until time θ_2^1, and drops it at time $\theta_2^1 + 1$.

Let us now consider the case $U_{\theta_2^2,1} < 0$. According to Lemma 2, the sequence $U_{1,1}, U_{2,1}, \ldots$ is non-decreasing and constant starting from $n = \theta_2^2$. We thus conclude that $U_{n,1} < 0$ for all $n \in \{1, 2, \ldots, \tau\}$. With Proposition 1, it yields $V_n^1(1) < 0$ for all $n \in \{1, 2, \ldots, \tau\}$, implying that the optimal strategy for relay 1 is to retain the message until it meet the destination or the deadline expires. $\square$

According to Proposition 2, the best-response policy of player 1 to the strategy of player 2 is therefore as follows:

$$\sigma_n^1(1) = \begin{cases} keep & \text{if } n \leq \theta_2^1, \\ drop & \text{if } n > \theta_2^1, \end{cases} \tag{30}$$

where the threshold θ_2^1 can be greater than τ.

5.3 To Accept or to Reject

Let t be the time at which relay 1 meets the source. The optimal expected cost at t is:

$$\begin{aligned} V_t^1(m_s) &= \min(0, g(m_s, accept) + \alpha V_{t+1}^1(1)), \\ &= \min(0, C_r + \alpha V_{t+1}^1(1)), \end{aligned} \tag{31}$$

where $V_{t+1}^1(1)$ can be computed from (27). Thus, if at time t the second term is negative, then it is optimal to accept the message from the source. Otherwise, it is optimal to reject it.

Proposition 3. *There exists θ_1^1 such that relay 1 rejects the message if it meets the source at a time $n > \theta_1^1$.*

Proof. Observe that (31) can be written as follows

$$V_t^1(m_s) = \min(0, C_r + \min_{i=1,\ldots,\tau-t-1} U_{t+1,i}).$$

Since Lemma 2 implies that $\min_{i=1,\ldots,\tau-t-1} U_{t+1,i}$ increases with t, we can assert that if at time θ_1^1 the relay rejects the message, i.e., if $\min_{i=1,\ldots,\tau-\theta_1^1} U_{\theta_1^1,i} \geq 0$, then it will also reject it at all subsequent contact times $k > \theta_1^1$ with the source. $\square$

We note that the threshold θ_1^1 can be larger than τ, in which case relay 1 always accepts the message when it meets the source. Similarly, the threshold θ_1^1 can be smaller than 1, in which case relay 1 never accepts the message when it meets the source.

6 Conclusion

We studied the selfish behaviour of DTN nodes incentivised by a reward for participating in message forwarding. The reward is proposed by the source to every relay it meets, but is paid only to the first one that delivers the message. A relay meeting the source is not informed of the existence of other message copies. Assuming a given lifetime for the message, we considered the (discrete-time) decision problem faced by a relay. When it meets the source, a relay has to decide whether to accept the message or not, and once the relay has the message it has to choose to retain or to drop it at subsequent decision epochs. Each relay makes its decisions in order to minimize the expected cost it incurs for participating. We modelled the interaction between mobile nodes as a stochastic game with partial information.

For the single player case, we first obtained a necessary condition for the relay to attempt the delivery of the message that reflects a minimal value of the reward. In fact it implies the minimal reward sufficient to ensure that the player will not drop the message. We then saw that the relay's strategy to accept the message from the source is of a threshold type.

Extending the model to the case of two players, we established that if one of the players follows a threshold type policy then the other one will also use a similar strategy. We thereby have come to the question whether such threshold strategies are an equilibrium of the game. A positive answer to these question is not obvious, however if so it gives strong research impetus and opens up a possibility to fine-tune our reward mechanism.

References

[1] E. Altman. Competition and cooperation between nodes in delay tolerant networks with two hop routing. *In proceeding of NET-COOP*, pages 264–278, 2009.

[2] S. Buchegger and J. L. Boudec. Performance analysis of the confidant protocol: Cooperation of nodes-fairness in distributed ad-hoc networks. In *Proceedings of ACM MobiHoc*, 2002.

[3] L. Buttyan, L. Dora, M. Felegyhazi, and I. Vajda. Barter-based cooperation in delay-tolerant personal wireless networks. In *IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks*, 2007.

[4] L. Buttyan, L. Dora, M. Felegyhazi, and I. Vajda. Barter trade improves message delivery in opportunistic networks. *Ad Hoc Networks*, 8(1):1–14, January 2010.

[5] W. Chahin, H. B. Sidi, R. El-Azouzi, F. De Pellegrini, and J. Walrand. Incentive mechanisms based on minority games in heterogeneous DTNs. In *Proceedings of 25th ITC conference*, Sanghai, China, Sept. 10-12 2013.

[6] B. B. Chen and M. C. Chan. Mobicent: a credit-based incentive system for disruption tolerant network. In *INFOCOM, 2010 Proceedings IEEE*, pages 1–9. IEEE, 2010.

[7] T. Chen and S. Zhong. INPAC: An enforceable incentive scheme for wireless networks using network coding. In *Proceedings of IEEE INFOCOM*, 2010.

[8] Kevin Fall. A delay-tolerant network architecture for challenged internets. In *Proc. ACM SIGCOMM, Karlsruhe, Germany*, pages 27–34, 2003.

[9] M. K. Goush, D. McDonald, and S. Sinha. Zero-sum stochastic games with partial information. *Journal of Optimization Theory and Applications*, 121(1):99–118, 2004.

[10] Q. He, D. Wu, and P. Khosla. SORI: A secure and objective reputation-based incentive scheme for ad hoc networks. In *Proceedings of IEEE WCNC*, 2004.

[11] F. Li and J. Wu. FRAME: An innovative incentive scheme in vehicular networks. In *Proceedings of IEEE International Conference on Communications*, 2009.

[12] X. Y. Li, Y. Wu, P. Xu, G. Chen, and M. Li. Hidden information and actions in multi-hop wireless ad hoc networks. In *Proceedings of ACM MobiHoc*, 2008.

[13] M. E. Mahmoud and X. Shen. PIS: A practical incentive system for multi-hop wireless networks. *IEEE Transactions on Vehicular Technology*, pages 4012–4025, 2010.

[14] S. Marti, T. J. Giuli, K. Lai, and M. Baker. Mitigating routing misbehavior in mobile ad hoc networks. *In Proc. of MobiCom*, pages 255–265, 2000.

[15] T. Seregina, O. Brun, R. Elazouzi, and B. Prabhu. On the design of a reward-based incentive mechanism for delay tolerant networks. *IEEE Transactions on Mobile Computing*, 16(2):453–465, 2017.

[16] Lloyd S. Shapley. *Stochastic Games*, volume 39. 1953.

[17] U. Shevade, H.H. Song, L. Qiu, and Y. Zhang. Incentive-aware routing in DTNs. In *IEEE International Conference on Network Protocols (ICNP)*, pages 238–247, 2008.

[18] Monika S. Shirbhate, Abhishek A. Gulhane, and Smeet D. Thakur. A survey on delay tolerant network. *International Journal of Advance Research in Computer Science and Management Studies*, 2, Issue 12, Desember 2014.

[19] M. Y. S. Uddinand, B. Godfrey, and T. Abdelzaher. RELICS: In-network realization of incentives to combat selfishness in DTNs. In *Proceedings of IEEE International Conference on Network Protocols (ICNP)*, pages 203–212, 2010.

[20] L. Wei, Z. Cao, and H. Zhu. MobiGame: A user-centric reputation based incentive protocol for delay/disruption tolerant networks. In *Proc. IEEE Global Telecommunications Conference (GLOBECOM)*, pages 1–5, 2011.

[21] X. Zhang, X.Wang, A. Liu, Q. Zhang, and C. Tang. Reputation-based schemes for delay tolerant networks. In *Proceedings of 2011 International Conference on Computer Science and Network Technology*, 2011.

[22] S. Zhong, J. Chen, and Y. R. Yang. Sprite, a simple, cheat-proof, credit-based system for mobile ad-hoc networks. In *Proceedings of INFOCOM 2003*, pages 1987–1997, San Francisco, CA, USA, April 2003.

[23] S. Zhong, L. Li, Y. G. Liu, , and Y. R. Yang. On designing incentive-compatible routing and forwarding protocols in wireless ad-hoc networks - an integrated approach using game theoretical and cryptographic techniques. In *Proceedings of ACM MobiCom.*, 2005.

[24] H. Zhu, X. Lin, R. Lu, Y. Fan, and X. Shen. SMART: A secure multilayer credit-based incentive scheme for delay-tolerant networks. *IEEE Transactions on Vehicular Technology*, 58(8):4628–4639, 2009.