Dynamic Replication Control Strategy for Opportunistic Networks

Salem Sati
Technology of Social Networks
University of Düsseldorf, Germany
sati@cs.uni-duesseldorf.de

Andre Ippisch
Technology of Social Networks
University of Düsseldorf, Germany
ippisch@cs.uni-duesseldorf.de

Kalman Graffi
Technology of Social Networks
University of Düsseldorf, Germany
graffi@cs.uni-duesseldorf.de

***Abstract*—Epidemic routing is proposed as one of the routing protocols for Opportunistic Networks. These kind of networks behave as sparse and/or highly mobile networks in which there may not be a reliable path from source to destination. We study the trade-off between delivery delay/ratio and resource consumption in an Opportunistic Network in which a message has to be spread to each encountered node by epidemic relaying. In addition to the destination, there are several other nodes in the network that can cooperate in relaying the message. We first assume that, at every instant, all the nodes can predict the number of relays storing the message and the number of new message replicators that have received the message. We formulate the problem as a controlled finite discrete Markov chain and derive the optimal closed-loop control (replication policy). However, in practice, the intermittent connectivity in the network implies that the nodes may not have the required perfect knowledge of the system state. To address this greedy issue, we obtain an ordinary differential equation (ODE) approximation for the optimally controlled Markov chain. Finally, we evaluate the performance of the replication control policy over finite networks. Numerical results show that this dynamic replication policy performs close to the optimal closed-loop policy.**

***Index Terms*—Opportunistic Networks, Epidemic Relaying, Markov Chain Modeling, Optimal Replication Control**

I. Introduction

The end-to-end path in delay tolerant networks (DTN) is not guaranteed, therefore, the messages are delivered based on hop-by-hop routing from a source node to the destination node via Store-Carry-Forward fashion. In DTN, a source node or an intermediate node stores messages in its buffer and carries them while moving around. Furthermore, the DTN routing is divided into two main types which are *flooding-based* routing and *utility-based* routing. One of the utility-based routing protocols is PRoPHET [1] which forwards the messages to other nodes based on calculated cost and the messages are delivered to the destination node via multiple hops. The main idea of utility-based routing protocols is to reduce the node resources consumption in terms of storage, bandwidth and power by decreasing the number of message replications. This type of utility-based routing protocols offer a limited message copy such as the *GTMX* [1] forwarding strategy of the PRoPHET routing protocol.

On the other hand, the flooding-based routing protocols replicate the message to every encountered node. The assumption of unlimited node resources is paid against the redundancy of unlimited message copies. One representative protocol of flooding-based routing is the Epidemic [2] routing protocol. As the name implies, the node replicates the message to all encountered neighbor nodes. The epidemic routing protocol is very powerful when the buffer size of nodes is infinite, however, if the buffer size is not sufficient, especially as in reality where the node buffer size is limited, the approach generates unlimited message copies by its flooding behavior which leads to overhead and the routing performance degrades by congestion or buffer overflow. In order to solve the message overhead and node's resources consumption problem of the Epidemic protocol, several quota-based routing schemes have been proposed, such as Spray & Wait [3] and Spray & Focus [4] protocol. In these protocols, the total number of message copies present in a network is limited by a certain number of hops.

In this paper, we propose a replication policy for Epidemic routing, termed *Most Of Storage and Transmission – with Replication Probability Threshold* (*MOST-RPT*) which integrates an optimal delivery probability policy as function of the message hop-count combined with a replication counter and a rule for the optimal number of message copies in the network. The optimality means that *MOST-RPT* maximizes the number of the message copies as inverse of the minimum single message delivery probability based on a particular knowledge of the message carrier about the message copies in network as hop-count. The optimality of *MOST-RPT* relies on the assumptions that (1) inter-contact times are distributed exponentially, or they have at least an exponential tail, (2) nodes move independently of each other as i.i.d. distribution, (3) each node knows the average inter-contact rate of all node pairs in the network.

Our optimal delivery probability policy differs from the existing delivery probability policies in three important points.

1) It is a comprehensive policy which combines the direct delivery probability as replication counter and the global delivery probability from source to destination. In addition, it considers the cumulative delivery probability as hop-count of the message, when the previous node replicates the message to other intermediate nodes.
2) Our optimal delivery probability policy is a dynamic function of two important states of the message copy which are hop-count as storage metric and replication

counter as transmission metric to optimize the delivery of the message copies based on resource-focused thresholds.
3) Our optimal delivery probability policy considers inter-contact time as node mobility pattern in addition to the message information as Time-To-Live (TTL), therefore it is efficient regardless of the nodes' density.

Our objective for the *MOST-RPT* policy is that given a resource consumption criteria in form of a certain constraint on the maximum number of replications per single message, the *MOST-RPT* replication rule maximizes the resource allocation based on minimizing the delivery rate of each single message. The basic idea is to model each message replication as an optimal stopping rule problem. In Section II we give therefore an overview on existing work and in Section III we present the *MOST-RPT* policy. In Section IV we present the detailed evaluation of *MOST-RPT* in comparison to current approaches.

II. Related Work

An Opportunistic Network is a subclass of DTN which is often partitioned and only provides a hop-wise end-to-end path between nodes over time, therefore, it uses routing fashion like Store-Carry-Forward. the Direct Delivery [5] is is the simplest DTN routing scheme.In Direct Delivery routing protocol, the source nodes carry their messages to the destination nodes by themselves leading to a low delivery ratio. Another approach is Epidemic Routing which floods the message to all encountered nodes. The Epidemic Routing occupies more system bandwidth, node buffer space and energy consumption for the delivery of the message. The routing protocol Spray & Wait adapts the number of message copies between single, direct copies as in Direct Delivery and an unlimited number of copies. In the first phase, Spray & Wait limits the maximum number of messages copies by a hop count limit L. It then bounds the maximum hop counts of messages to one, i.e. performs direct delivery. The Spray & Wait uses a quota for message delivery as L limits the number of copies between one and infinity.

The analysis and control of unicast DTN routing has been widely studied. Groenevelt et al. [6] modeled epidemic and two-hop routing using Markov chains. They found the relation between the delay and the number of message copies. Zhang et al. [7] derived a framework based on ordinary differential equations (ODEs) to analyze Epidemic Routing. Neglia and Zhang [8] use the optimal control of replication in DTNs of epidemic routing. Their work assumes that all the nodes know the number of nodes carrying the message. The optimal closed-loop control is a threshold-based policy using the number of relays carrying the message. Some DTN routing protocols tried to achieve better performance with different levels of knowledge about the relays. Optimal Probabilistic Forwarding [9] aimed to maximize the delivery ratio of each message based on its hop count and message life. The optimal probabilistic forwarding metric derived by modeling each forwarding task as an optimal stopping rule problem. Jia Xu provides in his paper [10] an Optimal Joint Expected Delay Forwarding (OJEDF) protocol which minimizes the expected delay based on the number of forwarding times per message. Their paper proposes a comprehensive forwarding metric called Joint Expected Delay (JED) which is calculated based on remaining hop-count (or ticket) and message residual lifetime. The aim of this approach is to achieve a near-optimal replication of the message which both provides a maximum possible message delivery ratio, while keeping the overhead for this purpose as small as possible.

III. System Model and Assumptions

In the following, we describe our models for communication, mobility, routing and traffic, as well as the scheduling and dropping policies of buffer management. This framework considers resources during message delivery through node contact duration. Our theoretical framework and its assumptions are described as follows.

Communication Model The system of the mobile DTN environment consists of mobile nodes. Therefore, we assume that there is a finite number (**N**) of nodes in the network. Each of these nodes has a communication interface, such as Wi-Fi or Bluetooth, with finite and limited coverage area, we also assume that all nodes have the same interface. Moreover, we assume that the connection between encountered nodes is established when nodes come in the coverage area of each other. Furthermore, we assume that each node in the system has limited physical storage that can store up to **B** messages and that nodes only have limited energy.

Mobility Model In the DTN context, message flooding occurs only when nodes encounter each other. Thus, the inter-contact time (ICT) between sequential node meetings is the basic delay component of message delivery. The contact duration (CT) of a meeting distribution is a basic property of the message transmission rate, both the values of ICT and CT characterize the mobility model. ICT is related to message delay, so it is the main criteria for end-to-end delay, CT is related to message transmission which is assumed to be small compared to ICT. Therefore, to formulate the optimal replication policy problem, we assume that our mobility model has the following three properties. *First*, contact duration times are distributed exponentially, or they have at least an exponential tail. *Second*, nodes move independently of each other as i.i.d. distribution, where each node moves from one location to an other based on the movement vector of random speed and direction. *Third*, the mobility of the nodes is homogeneous, which means that all nodes have the same meeting rate λ_c.

Routing Model In our theory, we assume that every single message has both an unique id at any time (**t**) and a single destination, i.e. unicast communication, and is assumed to be routed using a replication scheme. Moreover, every single message has a limited overall life time (**TTL**) and a timer (T_{BUF}) of staying in the current node's buffer. During a contact time, the routing protocol should create a list of derivable messages among the ones that are currently in the node buffer. Every time the node transmits a message, both the transmission counter (replication R_c) and message hop count

TABLE I
USED NOTATIONS AND QUANTITIES

No	Parameters	Description
1	N	Number of nodes in the network
2	n	Max relayed nodes in the network
3	ICT	Inter-contact time
4	$P(t)_m$	Probability of single message delivery
5	TTL	Time-To-Live for message
6	$P(t)_N$	Probability of message copies delivery
7	r_t	Number of copies of the message in the entire network at time t
8	T_{BUF}	Queue time in the node buffer
9	nr	Current relayed nodes of the message
10	λ_m	Message replication rate
11	$\grave{R}(t)$	The rate of "infected" nodes carrying the message
12	H_c	Message hop counter
13	R_c	Node message replication counter
14	λ_c	Average meeting rate between two nodes
15	P_t	Probability of message replication
16	B	Max number of messages in the node's buffer
17	K	Number of different applications
18	N_{max}	Total of current message copies in the system
19	$U(m)_{max}$	Utility function for dropping of the message
20	$BS\%$	Free buffer space percentage

(H_c) are increased by one. Thus, different routing protocols may choose different messages. For this paper, we consider the Epidemic Routing protocol as a case study for the following reason. *First*, its simplicity allows us to focus on the problem of resource consumption. *Second*, it can be modeled as greedy algorithm which can be solved by an Ordinary Differential Equation (ODE). *Third*, based on our knowledge, it is the most practical routing protocol implemented by most current empirical DTN middlewares and is recommended by DTNG.

Resource Constraints The resource constraints of our analytical model consider that each mobile node has a physical buffer, which can store up to **B** different messages, which are either self-generated or replicated by other relay nodes. The following two problems can arise each time a new connection is established between encountered nodes. First, the buffer space of the encountered node, because it might be filled with transmitted messages. Second, the order of the transmitting messages, because the aim of scheduling messages is to maximize the overall delivery probability for the entire network. In Table I, we summarize the notations and quantities that we use throughout our model.

The main aim of our framework is to maximize the global delivery ratio in the opportunistic network. We propose a dynamic replication policy for the Epidemic routing protocol, which controls when and to which neighboring nodes to disseminate selected messages. The nodes are considered to have a limited physical buffer size. Our proposed dynamic replication policy determines its scheduling decisions based on the message information and the mobility pattern of the DTN nodes. A utility function computes for each single message two utility values. The proposed dynamic replication policy uses the first utility value of the messages to determine the messages to be transmitted when contacts occur. The second utility value of the message is used to advise which message to drop when the node's buffer is full.

For the computation of the utility values, we use the variables depicted in Table I, all of them are assumed to be locally available to the node. Here, the message delivery probability function ($P_t(N)$), which is essential in a DTN, is the ratio between the number of message copies (r_t) to the number of total nodes (N). The research question remains how to estimate this value. In the context of DTNs it is difficult to find the global number of message copies (r_t) in the network. One possibility is to use the local information a node has and to interpolate these estimations to the whole network. But this information only approximates the number of message copies disseminated at an instant time in the network. In this paper, we propose to model the ratio between the number of the message copies (r_t) and the total number of nodes in the network (N) through a finite-state Markov chain. We can formulate the state diagram of Markov chains of discrete finite state, depicted in Figure 1, as following ordinary differential equation:

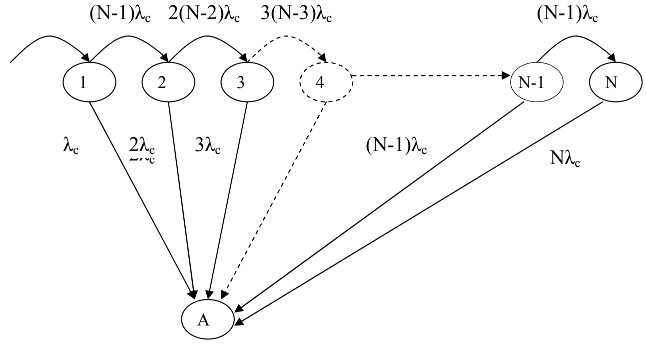


Fig. 1. Transition diagram of Markov chain for finite message dissemination

$$\grave{R}(t) = \frac{\partial r}{\partial t} = \lambda_c r(N - r) \tag{1}$$

Solving equation (1) by integrating we can write the number of message copies at instant time as:

$$r_t = \frac{N}{1 + (N-1)e^{-tN\lambda_c}} \tag{2}$$

In addition we can calculate the probability of message delivery for all messages in the system as follows:

$$P(t)_N = \frac{N}{e^{-tN\lambda_c} + N - 1} \approx 1 - e^{-tN\lambda_c} \tag{3}$$

Now from Eq. (2) we can approximate the function of our system model where the total number of nodes in the network equal to maximum intermediate nodes plus source and destination of the message, therefore, $N = n+2$, furthermore, the message delivery probability function is the ratio of total number of the message copies at instant time t to total number of the node in the network, where the message delivery probability function $\frac{r_t}{N}$ changes from $\frac{1}{N}$ at t=0 to maximum value of 1 at $t = \infty$ where at $t = \infty$ the number of replication r_t equal to maximum number of the nodes in the network N.

$$\frac{r_t}{N} = \frac{1}{1 + ((n+2)-1)e^{-tN\lambda_c}} \tag{4}$$

$$\frac{r_t}{N} = \frac{1}{1 + ne^{-tN\lambda_c} + e^{-tN\lambda_c}} \quad (5)$$

In Eq. (5) the delivery probability function shows exponential behavior. From this Eq. (5), we assume that the replication counter of the message replication at any node is $R_c = e^{-tN\lambda_c}$ and the hop count of the message is $H_c = ne^{-tN\lambda_c}$. As we have always $\lambda_c > \lambda_m$, due to $TTL < t$ in most cases, we therefore calculate the delivery probability function of the message by the following equation:

$$P_t = \frac{r_t}{N} = \frac{1}{1 + H_c + R_c} \quad (6)$$

Our system model is termed *Most Of Storage and Transmission – with Replication Probability Threshold* (**MOST-RPT**).

A. Optimal Dynamic Replication Control Strategy

We derive in this section the optimal replication policy under the assumption that, at any instant time, all the nodes have information about the number of relays carrying the message as hop count and the number of relays that have received the message as replication counter. We aim to determine the optimal replication policy of each message as function of the message delivery probability ($P(t)_N$), i.e., the decision criteria according to which an infected node decides if copying or not the message when it meets a susceptible node. The optimization goal is to minimize the cost based on the idea of dynamic programming which is known as *optimal cost to go* [11]. Analytically, we describe the Markov decision problem which is target to find the policy $\mu = \mu_0, ..., \mu_{T-1}$ that minimizes the following equation.

$$J = P(t)_N = E[\sum_{t=0}^{T-1} g_t(r_t, P_t) + g_T(r_T)] \quad (7)$$

We have the functions $f_0, ..., f_{T-1}$ and we have the cost function of every stage $g_0, ..., g_{T-1}$, in addition we can calculate the terminal cost g_T from Eq. (3) as follows

$$P(t)_{N=1} = P(t)_m = 1 - e^{-\lambda_c TTL} \quad (8)$$

On one hand our model is based on the information of a single message and on the other hand we can a prove that

$$P(t)_m = 1 - e^{-\lambda_c TTL} \approx \lambda_c TTL \quad (9)$$

This is clear from the mathematical fact that

$$\lim_{x \to 0} \frac{e^x - 1}{x} = 1 \quad (10)$$

Note that as $\lambda_c TTL$ approaches to zero the environment will be as sparse network. Now we can find the distributions of independent random variables $r_0, w_0, ..., w_{T-1}$. The whole system can be described as dynamics of the following function

$$r_{t+1} = f_t(r_t, P_t, w_t) \quad (11)$$

From Eq. (8) and Eq. (9) we seek the state feedback policy of

$$P_t = \mu(r_t) \quad (12)$$

Now from Eq. (7) and Eq. (12) we can define the Bellman value function, as *optimal cost-to-go function* [11]

$$V_r^* = min_{\mu_t,...,\mu_{T-1}} E[\sum_{s=t}^{T-1} g_s(r_s, P_s) + g_T(r_T)] \quad (13)$$

From Eq. (8), Eq. (9) and Eq. (13) we can minimize over action w_t and policies $\mu_{t+1}, ..., \mu_{T-1}$ where

$$minP_t = P(t)_m \quad (14)$$

$$w_t = \begin{cases} Copy & P_t > \lambda_c TTL \\ No\ Copy & otherwise \end{cases} \quad (15)$$

From Eq. (11), Eq. (13) and Eq. (15), there is an optimal stationary policy $\mu(t)$ with a threshold behavior as

$$N_{max} = (minP_t)^{-1} = \frac{1}{\lambda_c TTL} \quad (16)$$

The previous equation states that the optimal policy for each node is to copy the message every time it is possible until the total number of infected nodes is equal to N_{max} copies are done, or the message is delivered to the destination. Note that, the threshold value does not depend on the total number of nodes in the network, but simply depends on their pairwise meeting rate λ_c the message life time TTL as the cost function.

B. Buffer Occupancy

We assume that all the nodes of the network have the same traffic parameters, and follow the same replication strategy, We will turn our attention back to the buffer management of the node, in order to investigate its optimality. The number of message copies spread in the network at a random time instant can be counted, assume there are **K** total undelivered messages. We further assume cooperating nodes with no drop events due to delivery or TTL expiration. They assign the available buffer space across the whole network with N nodes, each able to store B message copies among the copies of these messages. Thus, we can write the following equation

$$\sum_{i=0}^{k} r_i - NB \leq 0. \quad (17)$$

Eq. (17) says that the total number of copies (for all messages) must not exceed the available buffer space in all N nodes. Recalling Eq. (3), we can write the expected delay of delivered messages based on the number of message copies as follows.

$$E(t)_N = \frac{1 - e^{-tN\lambda_c}}{\lambda_c} \quad (18)$$

Because the buffer size is equal for all nodes in the network we estimate the buffer space as ratio of the expected delay and the message TTL (Eq. (17) and (18)).

$$BS\% = \frac{E(t)_N}{TTL} \quad (19)$$

With these deduction, we eventually formulate the forwarding and drop policy used in MOST-RPT. As we mentioned in Eq. (11) and Eq. (13), the forwarding policy is based on FIFO with the replication threshold decision calculated as follows:

$$send-queue = \begin{cases} FIFO & P_t > \lambda_c TTL \\ 0 & otherwise \end{cases} \quad (20)$$

To improve the performance of our system model, we consider two cases in which messages are dropped from the buffer. In the first case, the message is dropped when the buffer is full and in the second case, the message's TTL is expired based on Eq. (11), (16) and (19) as follow

$$U(m)_{max} = T_{BUF} P_t^{-1} \quad (21)$$

$$Expired = \begin{cases} P_t <= \lambda_c TTL & buffer < BS\% \\ TTL <= 0 & otherwise \end{cases} \quad (22)$$

With Eq. 20 we define our forwarding policy and with Eq. 21 and 22 the dropping policy of our model.

IV. Experiment and Results

To evaluate our proposed policies *MOST-RPT*, we use following metrics in the comparison with other policies.

1) *Delivery Ratio* is the ratio of the number of delivered messages to the total number of generated messages.
2) *Overhead Ratio* is the average number of intermediate nodes used for one delivered message.
3) *Average Delay* is the average delay of all messages delivered successfully.

To reflect different traffic situations, we consider 3 scenarios. Scenario 1 only considers 80 pedestrians with a buffer size of 5MB, a message creation interval of 25s to 35s and message size ranging from 500KB to 1 MB. In Scenario 2, in addition to those 80 pedestrians from Scenario 1 also 40 cars join in. The scenario reflects a information-centric use case in which small (64KB - 512KB and 512B - 2KB) messages are quickly exchanged. Here the cars have a message creation interval of 1s - 5s. Finally, in Scenario 3 we investigate the effects of larger files and even faster nodes. 80 Pedestrians generate 1MB - 5MB large messages every 60s to 120s. 40 cars generate at an interval of 25s to 35s messages of 64KB to 512KB, while 6 nodes in trains generate every 1s to 5s messages of size 512B to 2KB. During our three different simulation scenarios we change the message TTL from 100 to 500 minutes. Through the variety of scenarios, in combination with the 5 values for the TTL we are able to fully reflect the quality of *MOST-RPT* in comparison to other policies.

The performance of proposed *MOST-RPT* is conducted with 5 different dropping policy of Epidemic routing algorithm. *MOST-RPT* uses the send queue based on the Eq. (20) and the dropping policy which is applied based on the Eq. (21) and Eq. (22). For comparison we select the dropping policies of *FIFO* which is selects the message with the minimum arrival time, *SHLI* which is based on minimum TTL, *MOFO* which selects the message with the maximum replication counter using the higher hop count as tie breaker. In addition, we look at *MaxHop*, the policy which selects the message with the maximum hop count and *MaxRep*, which considers the maximum replication. We compare those 5 different dropping policies with the proposed *MOST-RPT* policy.

We run all six drop policies using the same three scenarios with the listed parameters in Table II and compare their performance with regards to the three above mentioned metrics under different message TTL values and different traffic patterns. The different three traffic patterns were simulated with the default settings of the ONE Simulator [12], [13], using the Epidemic routing protocol with the FIFO replication strategy (send queue) as shown in Table II. From the settings of the scenarios listed in Table II we can calculate the average of TTL as 300 minutes. In addition we run the scenarios to measure the inter-contact time (ICT) between encountered nodes in the scenario. This ICT, in all scenarios, was approximately 5880 seconds long; from the measured ICT we can calculate the meeting rate in the system, this rate λ_c was equal to 0.00017.

From Eq. (9) we can calculate the minimum probability of delivery for a single message which is equal to 0.05. We use it as stopping rule as explained in Eq. (12) to Eq. (14).

For all applied scenarios, based on the idea of the *MOST-RPT* implementation, which considers the replication criteria as resource concentrate, we apply the send-queue based on FIFO sorting which is based on arrival time (oldest buffered message), as every message will be transmitted as replicated message R_c and stored in the buffer H_c as the arrival or buffering time stays greater than zero. Clearly the buffering time will be equal to zero when the message is not yet received or the message is already dropped from the buffer by the node itself, but for our *MOST-RPT* we use FIFO with replication stopping rule as explained in Eq. (15) and Eq. (20). From the condition of the replication stopping rule we can find that the maximum number of replication based on the criteria of minimum delivery probability, we formulate the minimum probability which is based on message information and node mobility by using Eq. (6), Eq. (14) and Eq. (16). From those equations we can calculate that N_{max} is equal to 20 copies.

For the evaluation of the dropping policies we apply six different drop policies to analyze the performance of Epidemic Routing when using FIFO as replication policy for send-queue. In addition we apply our proposed model of *MOST-RPT* but with different drop polices as shown in Eq. (21) and Eq. (22). We consider the buffer occupancy to improve the performance of proposed *MOST-RPT* as shown in Eq. (18) and Eq. (19), we can see that the criteria of the stopping rule should be applied when the node buffer space will be less than 25%.

1) Delivery Ratio: The delivery ratio of the proposed *MOST-RPT* is compared with other scenarios which use scheduling based on FIFO for send queue and different

TABLE II
SIMULATION SETTINGS

No	Settings	Map of downtown Helsinki, Finland
1	Simulation time	12 h
2	Number of devices (n)	126
3	Group Type with Speed	80 Pedestrians (0.5-1.5 km/h) 40 Cars (10-80 km/h) 6 Trains (10-80 km/h)
4	Simulation area	Helsinki, Finland Map
5	Routing protocols	Epidemic
6	Interface type	Simple Broadcast
7	Transmission range	250 m
8	Bandwidth	250 KBps
9	Drop policies used	FIFO, MOFO, SHLI, MaxHop, MaxRep
10	Message size ranges	Scenario 1: 0.5-1 MB Scenario 2: 64-512 KB & 0.5-2 KB Scenario 3: 1-5 MB & 64-512 KB & 0.5-2 KB
11	Message creation interval	Scenario 1: 25-35 s Scenario 2: 25-35 s & 1-5 s Scenario 3: 60-120 s & 25-35 s & 1-5 s
12	Time-to-live (TTL)	100, 200, 300, 400, 500 min
13	Default buffer size	Pedestrians: 5 MB Cars, Trains: 50 MB

dropping policies, namely FIFO, SHLI, MOFO, MaxHop and MaxRep as shown in Figure 2. From Scenario 1 the delivery ratio of *MOST-RPT* is close to the delivery ratio of the MOFO dropping policy, as the traffic pattern of Scenario 1 is low as shown as in Table II. The high delivery ratio of *MOST-RPT* can also be seen in Scenarios 2 and 3, which have a high traffic rate. The delivery ratio of *MOST-RPT* has improved as well in comparison to the other dropping policies. From Figure 2, we can see that the delivery ratio efficiency is low, when using MaxRep or FIFO dropping policies, because mostly as the messages stay longer in the buffer there is a chance to have a higher replication counter.

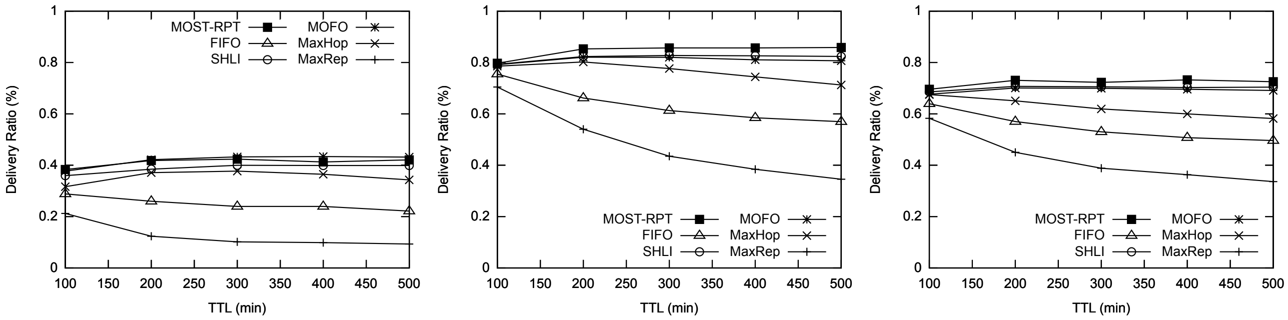
2) Overhead Ratio: The main performance factor to compare *MOST-RPT* with other different buffer managements is the Overhead Ratio, as this factor is related to the resources of both a single node and the network. Due to the unlimited replication of the messages, the Epidemic Router is suffering from the consumption of those resources. Therefore, *MOST-RPT* considers the resource consumption in terms of storage and transmission. Figure 3 shows that *MOST-RPT* has the lowest overhead when compared to the others, as the send queue policy (FIFO) of *MOST-RPT* is controlled by the replication stopping rule based on the probability of the message delivery. This message delivery probability is calculated as a function of hop count (as storage metric) and replication counter (as transmission metric) of the message. Furthermore from Scenario 1 of Figure 3 we notice that *MOST-RPT* has a stable and minimum overhead ratio. This stability derives from the mathematically proved optimality of both the applied stopping rule and the dropping policy. The main idea of replication and dropping decisions should be considered as a trade-off to achieve the desired message delivery. While the message replication will improve the delivery ratio, at the same time we should mention that the dropping rate of the message will increase when using a limited node buffer. This limitation of the buffer space leads to higher overhead and resource consumption. Even when increasing the message creation in Scenario 2 and 3 of Figure 3 we can see that *MOST-RPT* still has a lower overhead ratio and at the same time higher delivery ratio when compared to the others.

3) Delay: In this section we look at the end-to-end average delay, this metric is used as a performance metric for different DTN applications and scenarios. Scenario 1 of Figure 4 shows that *MOST-RPT* has a lower delay compared to the others, because we consider the buffer time as a dropping policy when the buffer is highly occupied with many messages. The message TTL is the sum of all buffer times and transmission times, the buffer times have the highest impact on the message TTL. A message that is held in the buffer will have a higher chance of replication, therefore, we consider the message buffer time as main criteria of the storage and transmission metrics. These functions of replication stopping rule and dropping policy will improve the performance of Epidemic Routing. The performance of Epidemic Routing is optimized as greedy replication problem by *MOST-RPT*. Moreover, we can see that when we increase the rate of message creation as shown in Scenario 2 and 3 of Figure 4 respectively, *MOST-RPT* has lower delay when compared to the other policies.

V. CONCLUSION AND FUTURE WORK

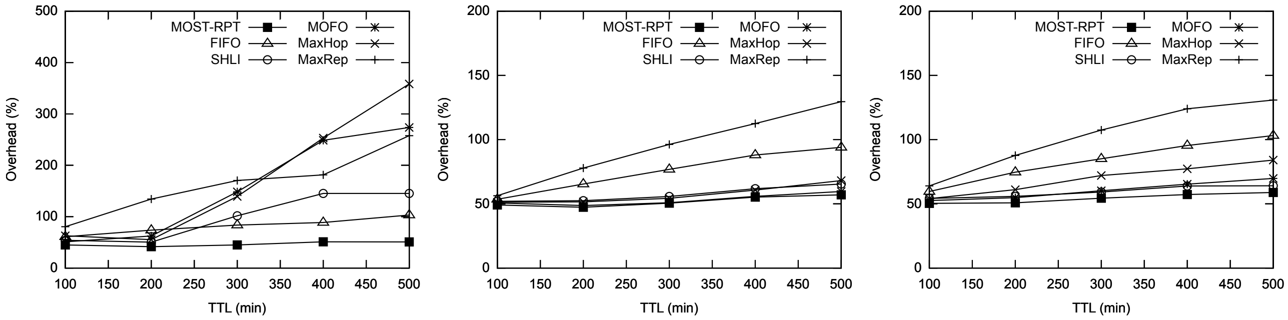
This paper studies the replication in Epidemic Routing in DTN environments by formulating the problem of replication control as a Markov chain model. The paper here considers the replication in Epidemic Routing as heuristic problem to obtain the optimal policy as stopping rule. We solve the problem based on an ordinary differential equation approximation for finding the optimally controlled Markov chain, under different network sizes. The proposed *MOST-RPT* forwarding and dropping policy is constructed based on an analytic study which considers the trade-off between delivery probability and resource consumption for Epidemic Routing as optimal stopping rule of replication in DTNs. The numerical results of this paper show that, based on an our scenarios, it gives help to design new forwarding scheme for utility based routing protocols to achieve better performance, namely in obtaining the best delivery ratio in our evaluation while having the lowest message overhead and delay.

For the future, we aim to extend our model beyond the idea of ordinary differential equations, which is based on the fact that all nodes in the opportunistic network are homogeneous. This assumption is made to simplify the function formulation of some issues in opportunistic networks such as message dissemination, congestion control and utility based forwarding decisions. This assumption built on the fact that the inter-contact time of pair wise node interactions is approximated and exponentially distributed. But in some DTN scenarios the mobility pattern is different inside the same group which yields to different contact and mobility modeling instead of i.i.d. modeling. Therefore, as future work will study and look at the modeling of heterogeneous modeling based on Markov Chain to optimize the solution of problems in opportunistic network issues and challenges.



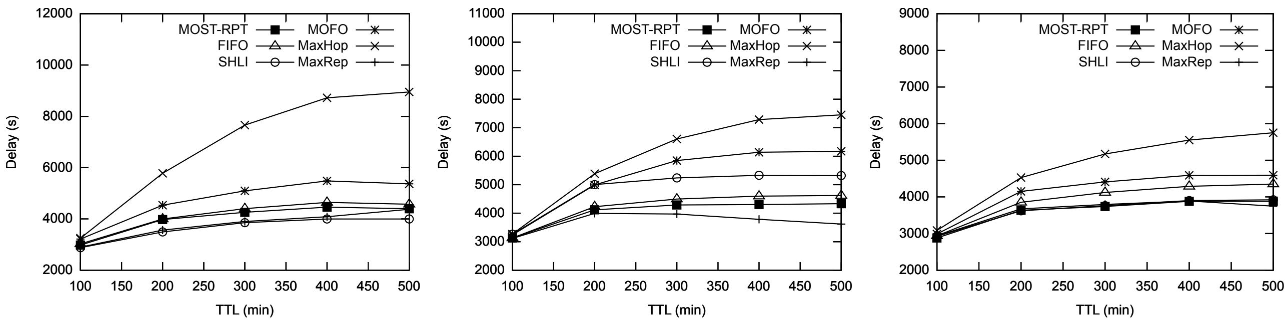
(a) Scenario 1 (b) Scenario 2 (c) Scenario 3

Fig. 2. Delivery Ratio



(a) Scenario 1 (b) Scenario 2 (c) Scenario 3

Fig. 3. Overhead Ratio



(a) Scenario 1 (b) Scenario 2 (c) Scenario 3

Fig. 4. Delay

REFERENCES

[1] A. Lindgren, A. Doria, E. Davies, and S. Grasic, "RFC 6693: Probabilistic Routing Protocol for Intermittently Connected Networks," *IETF*, 2012.

[2] A. Vahdat and D. Becker, " Epidemic Routing for Partially-Connected Ad Hoc Networks," *Technique Report, Department of Computer Science, Duke University, USA*, vol. vol.20, no. 6, 2000.

[3] T. Spyropoulos and K. Psounis, " Spray and Wait: An Efficient Routing Scheme for Intermittently Connected Mobile Networks ," *Proceedings of the ACM SIGCOMM Workshop on Delay-Tolerant Networking,* , 2005.

[4] T. Spyropoulos, K. Psounis, and C. S. Raghavendra, "Spray and Focus: Efficient Mobility-Assisted Routing for Heterogeneous and Correlated Mobility," in *In Proceedings of the IEEE International Conference on Pervasive Computing and Communications - Workshops (PerCom Workshops)*, 2007, pp. 79–85.

[5] A. Spyropoulos, K. Psounis, and C. S. Raghavendra, "Single-Copy Routing in Intermittently Connected Mobile Networks," in *Proceedings of the Annual IEEE Communications Society Conference on Sensor and Ad Hoc Communications and Networks (SECON)*, 2004, pp. 235–244.

[6] R. Groenevelt, P. Nain, and G. Koole, "The Message Delay in Mobile Ad Hoc Networks," *Performance Evaluation*, vol. 62, no. 1-4, 2005.

[7] X. Zhang, G. Neglia, J. F. Kurose, and D. F. Towsley, "Performance Modeling of Epidemic Routing," *Computer Networks*, vol. 51, no. 10, pp. 2867–2891, 2007.

[8] G. Neglia and X. Zhang, "Optimal Delay-power Tradeoff in Sparse Delay Tolerant Networks: A Preliminary Study," in *Proceedings of the SIGCOMM Workshop on Challenged Networks, (CHANTS@SIGCOMM)*, 2006, pp. 237–244.

[9] J. Xu, X. Feng, W. Yang, R. Wang, and B. Q. Han, "Optimal Joint Expected Delay Forwarding in Delay Tolerant Networks," *International Journal of Distributed Sensor Networks* , 2013.

[10] C. Liu and J.Wu, " An Optimal Probabilistic Forwarding Protocol in Delay Tolerant Networks," *International Journal of Distributed Sensor Networks* , pp. 105–114, 2009.

[11] D. P. Bertsekas, "Dynamic Programming and Suboptimal Control: A Survey from ADP to MPC," *European Journal on Control*, vol. 11, no. 4-5, pp. 310–334, 2005.

[12] A. Keränen, J. Ott, and T. Kärkkäinen, "The ONE Simulator for DTN Protocol Evaluation," in *Proceedings of the International Conference on Simulation Tools and Techniques for Communications, Networks and Systems (SimuTools)*. ICST/ACM, 2009.

[13] A. Keränen, T. Kärkkäinen, and J. Ott, "Simulating Mobility and DTNs with the ONE (Invited Paper)," *Journal of Communications*, vol. 5, no. 2, pp. 92–105, 2010.