ARBR: Adaptive Reinforcement-Based Routing for DTN

Ahmed Elwhishi[*], Pin-Han Ho[*], K. Naik[*], and Basem Shihada[†]

[*] Dept. of Electrical and Computer Engineering, University of Waterloo, Canada, {b7wu,pinhan}@bbcr.uwaterloo.ca
[†] Department of Computer Science, KAUST, Thuwal, Saudi Arabia, basem.shihada@kaust.edu.sa

Abstract—This paper introduces a novel routing protocol in Delay Tolerant Networks (DTNs), aiming to solve the online distributed routing problem. By manipulating a collaborative reinforcement learning technique, a group of nodes can co-operate with each other and make a forwarding decision for the stored messages based on a cost function at each contact with another node. The proposed protocol is characterized by not only considering the contact time statistics under a novel contact model, but also looks into the feedback on user behavior and network conditions, such as congestion and buffer occupancy sampled during each previous contact with any other node. Therefore, the proposed protocol can achieve high efficiency via an adaptive and intelligent routing mechanism according to network conditions. Extensive simulation is conducted to verify the proposed protocol, where a comparison is made with a number of existing encounter-based routing protocols in term of the number of transmissions of each message, message delivery delay, and delivery ratio.

Keywords-Routing, DTN, reinforcement learning.

I. INTRODUCTION

Intermittently Connected Mobile Networks (ICMNs) or Delay Tolerant Networks (DTNs) [1] is characterized by lack of end-to-end paths for a given node pair most of the time, which demonstrates a complete different design scenario from that for Mobile Ad-hoc Networks (MANETs)[2]. In this case, a routing protocol that can fully explore the network connectivity becomes the most desired, and the protocols developed for MANETs can hardly serve well. Note that a reactive protocol in MANETs which attempts to complete-path discovery when data is available for transmission may highly likely fail in DTNs, while the convergence of a proactive protocol for MANETs may result in a deluge of topology update messages.

Due to intermittent connection, a node in DTNs has to buffer a message and wait until it finds a working link to the next hop, instead of simply dropping it after a few failed transmission attempts as that done in MANETs or any other type of network. Such a store-and-forward process will continue until the message reaches its destination. Thus, the design of a DTN routing scheme demonstrates a significant difference from that for the conventional MANET routing. Specifically, a DTN routing strategy can be generally categorized as either encounter-based, store-carry-forward, or mobility-assisted routing, in which the node mobility is exploited in different degrees as a significant factor for the forwarding decision of a message.

Besides node mobility, the effect of node buffer congestion and communication capacity have been recently considered in the message forwarding decision making process. The study in [10] analyzed the effect of interference and congestion on the performance of DTN routing protocols, and the study in [11] further developed a suite of buffer management policies to improve the routing efficiency. However, all routing metrics exploited in the previously reported studies concern only on node mobility without incorporating with network states, such as congestion, interference, and available buffer space occupancy. There is obviously space to improve by providing an integrated DTN routing strategy that jointly considers node mobility and network congestion/interference states. Due to the variations in such a decentralized environment, the expected change in network topology and the links qualities can hardly be estimated, which results in unavailability of making prior assumptions about the system's environment due to network dynamics [7][17].

To overcome the lack of prior or centrally managed knowledge on the network environments, a collaborative reinforcement learning technique (CRL) has been proposed [9] for MANETs, which is characterized by an autonomous and self-organizing design for developing MANET routing protocols. The technique has been proven as a successful implementation regarding robustness and scalability in the context of MANET routing. Inspired by this, we envision that the concept of CRL can also be applied to the DTN routing protocol design in spite of the decentralized and intermittently connected characteristic in DTNs; and the use of a CRL model should be able to improve the performance and scalability of the DTN routing protocol operations.

This paper introduces a novel encounter-based routing protocol [16] by using CRL as a self-organizing technique, called Adaptive Reinforcement-Based Routing (ARBR). The main feature of the proposed protocol is the strong capability in adaptation to the fluctuation of network status and user behaviors so as reduce the number of transmissions, message delivery time, and increasing delivery ratio. The proposed ARBR jointly consider node mobility statistics, congestion, and buffer occupancy, which are taken as a

feedback in the quality-metric function. In specific, the feedback is in a form of statistical model of estimated contacts reliability based on sampling the availability of channel and buffer space during a contact between nodes. The developed quality-metric targets to facilitating decision making for each active data message, resulting in optimized network throughput. We will show via extensive simulation that the proposed ARBR protocol can achieve significant performance gain over the previously reported counterparts.

In summary, this paper is the first study which takes the CRL techniques in the design of a encounter-based DTN routing protocol. The proposed routing metric and message forwarding decision approach jointly considers not only node mobility but also network states such as wireless channels and node buffer occupancy.

The rest of the paper is organized as follows. Section II discusses some related work. Section III describes the CRL technique and its applications. Section IV introduces how DTN routing can be described as a CRL model, in which the proposed ARBR protocol is presented. Then, in Section V simulation results are presented, and a comparison is made between the proposed ARBR protocol and the previously reported counterparts using a real-world mobility model. Section VI concludes the paper.

II. RELATED WORK

With encounter-based routing protocols, the node mobility is exploited and is taken as the major factor that affects the utility function in the message forwarding decision. The node exploitation has been interpreted in different forms of quality-metrics. T. Spyropoulos et. al. in [6],[7] developed routing strategies based on different presentations of utility functions, called Most Mobile First (MMF), Last Seen First (LSF), Most Social First (MSF). These techniques have been proved to achieve some extents of performance improvement in terms of delivery delay and delivery ratio when restrictions on wireless channel condition and buffer capacity limit condition are not considered.

Our previous study in [15] proposed a routing protocol with its utility function obtained by manipulating the history of enter-encounter rates between nodes. Lindgren et. al. in [8] proposed a routing technique in DTNs which takes advantage of the predicted encounter probability between nodes. Evan P. C. Jones et. al. in [5] introduced a utility function for DTN routing which manipulates the minimum expected inter-encounter duration of time between nodes. Shen Ling et. al. in[4] designed feedback adaptive routing based on factors depend on the mobility of nodes, where a higher mobile node is assigned higher factor, and messages are transmitted through nodes with higher influence factors. Some forwarding techniques have considered limited buffer capacity and bandwidth [18],[3]. However, only the available bandwidth has been considered in deriving routing

metric in a form of expected delay. Although the encounter-based scheme can improve the system performance, their routing metrics have never incorporated with network states, such as congestion, and available buffer space, which are considered important when the encounter frequency is large and each node has more choices for packet forwarding.

In addition to encounter-based statistics, some studies looked into buffer management policies so as to reduce message loss[11],[12]. Although these techniques have improved the protocol performance in some situations, there is still a large room to improve by additionally considering the estimation of the channel availability and the remaining buffer space in its forwarding strategy. Even though the existing techniques have aimed to improve the data delivery, delivery ratio, and reduced number of transmissions, the channel condition and buffer occupancy state have never been considered as a factor in the derivation of routing utility functions. Motivated by above observations, this work investigates a CRL[13] based routing technique that jointly considers node mobility and the network states, including wireless channel and buffer occupancy.

III. COLLABORATIVE REINFORCEMENT LEARNING (CRL)

CRL is used for tackling the complex time-varying problems where global knowledge on system behaviors is not available. With a CRL coordination model, the agents can cooperate with each other to solve a system-wide optimization problem that could be composed of a set of discrete optimization problems (DOP). An agent can solve one or a number of discrete optimization problems via reinforcement learning by exchanging some key information with neighbor agents, which further contributes towards the solution of the system-wide optimization problem. An individual agent only possesses partial knowledge about the system-wide state and knowledge about their neighbors. As a result, each agent serves as a member of the dynamic population that joins (or leaves) the system by autonomously establishing (or tearing down) connections with their neighbors without making any use of system-wide knowledge.

With CRL, the path selection is based on the expected performance of an agent starting with initial state, s, in which the algorithm exercises an optimal state transition policy thereafter. An estimated value function $V(s)$ is employed at a CRL agent as the cost function in solving a DOP. $V(s)$ can also be presented as an optimal action-value function, $Q(s,a)$, and their relation can be expressed as $V(s) = max_a Q(s,a)$.

Two transition states are identified in CRL; local on the current agent n_i, and remote to a neighboring agent n_j. The estimation of the cost of transition from the local to the remote state takes into consideration the connection cost between the current agent and the neighboring agent.

Therefore, the estimated optimal action value function, $Q_i(s,a)$, should includes both the value function for the state $V_j(s')$ that is received from the neighboring agent, and the connection cost of the state, $D_i(s',a,s)$. The connection cost for a transition from the local state of current agent to the remote state at a neighboring agent should reflect the underlying network cost as well as the cost of transferring control from the source agent to the target agen. The transfer of control involves terminating the DOP at the originating agent and start solving a new DOP at the target agent. The cost function is given by:

$$V(s) = R(s,a) + max_a \sum_{s' \in B,P,D} P_i(s' \mid s,a)$$
$$. \left(D_i(s' \mid s,a) + Decay(V_j(s')) \right),$$
$$V(s) = max_a [Q_i(s,a)] \qquad (1)$$

where, $P_i(s' \mid s,a_j)$ represents set of transition models that describe the probability of making state transition from state s to state s' under delegation action a, $D_i(s' \mid s,a)$ is the estimated connection cost model at agent n_i of making a transition from sate s to state s' under delegation action a, $Decay(Vj(s'))$ is the decay model used at agent n_i to decay the V values of last advertised costs to given destination agents. This mechanism is used to eliminate and degrade agents that have a lower contact frequency with node n_i. Here, $R(s,a)$ is the Markov Decision Process(MDP) termination cost.

IV. PROPOSED CLR MODEL FOR DTN ROUTING

In this work, the design of DTN routing protocols is formulated as a reinforcement learning problem, in which the states, actions, transition and reinforcements of the proposed system will be explicitly identified. A learning strategy under the network constraints is then constructed and exercised in the routing protocol.

A. Background of the Model

In the terminology of CRL, a DTN is modeled as a time-varying environment in which the state at each node is determined by (1) the relative position of nodes, (2) the destination of carried messages, (3) the connectivity information on links between nodes, and (4) the buffer occupancy status. With the considered scenario of the study, the DTN routing protocol design encounters number of challenges, including the lack of global knowledge at any particular node in the DTN, and the requirement for the system-wide autonomic properties of the routing protocol to emerge from local routing decisions at routing nodes.

To trade-off between the system state observability and computation complexity, each node maintains the network status within a fixed time-window, in which the observations on terms of nodal mobility, links activities, and buffer occupancy (or referred to as *contact statistics*) are collected.
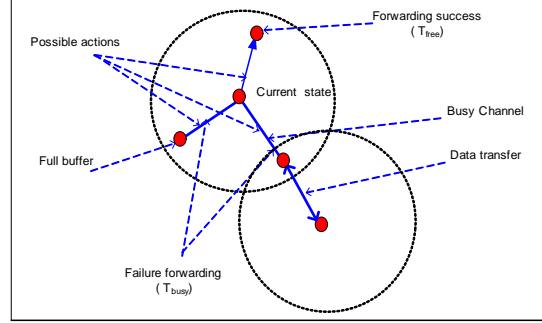

Figure 1. The reinforcement model of DTN routing.

The contact statistics are further fed as input to the utility function $V(s)$ (also called a cost function or quality-metric in context of DTN), which is specific to every node pairs n_i and n_j, and are maintained in a table, called *contact table*. A node n_i exchanges its contact statistics with a node n_j when they move into each other's transmission range, and the newly obtained contact statistics will be used to exchange their *contact table*. Meanwhile, the utility function $V_i(s)$ and $V_j(s)$ of a causally connected external state, s, is updated accordingly. These update actions are so-called reinforcements, which continue equipping each node with intelligence and knowledge in decision on message forwarding.

Whether a message should be forwarded from node n_i to an encountered node n_j will be determined by the value of utility function $V_j(s)$ maintained at a node n_j, the history of the quality of wireless link between n_i and n_j, and the history of buffer status of node n_j. A transmission could be failed during a contact due to bad channel condition and contention with other surrounding users, which can consume the time on-the-fly and leaves the messages to remain buffered at the node until a new contact happens. This channel busy-time can cause extra delay in message delivery or increase message loss.

Let the action of handing over a message from node n_i to n_j be denoted as A_{h_j}. If the message hand-over is successful, the transition $n_i \longrightarrow n_j$ occurs, and (or $n_i \longrightarrow n_i$) otherwise, in which the message remains in state n_i . Note that a time-to-live (TTL) is defined for each message, and a message is dropped if its TTL is expired.

B. Model Reinforcements

The cost function should reflect the cost of a given state transition, which depends on the state of the system and the action performed. In the CRL model of DTNs, the choice of forwarding action is mainly based on the routing policy and the utility function used for building information about network behavior.

The reinforcement function (i.e, the utility function) evaluates a given state transition and the corresponding

action. The action in the proposed CRL model of DTNs is whether a specific message should be forwarded to the encountered node. Fig. 1. shows the reinforcements considered for DTN routing. As a number of notations are introduced as follows. Let T_{free} denote the amount of time a channel is available during a time window interval, T^{window}, and T_{busy} denotes the amount of time that the channel is busy or the buffer of the encountered node is full during a time window interval, T^{window}. In other words, T_{free} denote the reinforcement of a successful exchange, whereas T_{busy} represents the reinforcement of a failed message exchange attempt. Let $T_{total} = (T_{busy} + T_{free})$ denote the total contact time between any node pairs during T^{window}. Each node in the network keeps track of the history of these values on every other node in the network, and are calculated at the end of each time window T^{window}.

C. Proposed Adaptive Reinforcement Based Routing (ARBR)

Based on the reinforcement model, the proposed ARBR protocol for DTNs is introduced in this section.

1) Protocol Background: In order to introduce the proposed protocol as CRL technique, a number of items; namely, states, delegation actions, transitions, connection cost, and cost function, are introduced as follows.

States definition and explanation: Each node n_i have a set of states $S_i\{B,P,D\}$, where B represents a message that is waiting in the buffer for forwarding, P represents the event that a messages has been successfully forwarded to an encountered node, and D represents the event that the message successfully has been received at node n_i. Also, n_i has internal and external states: $Ext(n_i) = B$ and $Int(n_i) = D, P$. For $P \in Int(S_i)$, there is causally connected state $B \in Ext(S_j)$, where node $n_j \in r_i$, r_i is the set of nodes in the transmission range of node n_i or nodes that are frequently encountered by node n_i in a certain time window.

Delegation actions at each node: The set of actions available at node n_i is $A = A_{h_i} \cup \{Receive\} \cup \{exchange\ of\ summary\ vectors\}$, where A_{h_i} represents the set of delegation actions, $A_{h_i} = \{H_i(r_0), ...H_i(r_{M-1})\}$, M is the set of nodes in the transmission range of node n_i. The action $H_i(n_j)$ represents an attempt to hand over a message from node n_i to node n_j. The action $\{Recieve\}$ represents an attempt to deliver a message to current custodian node n_i. The $\{exchange\ of\ summary\ vectors\}$ action is used for message vector exchange when two nodes become in transmission range of each other.

The state transition between nodes: The state transitions model for the MDP are as follows:

- The probability of successful message exchange between nodes n_i and n_j is $p_{ij}s = p_i(P \mid B, a_j)$.
- The probability of failed message exchange is $P_i(B \mid B, a_j) = 1 - p_{ij}s$.

- The probability of delivering message to destination if the current node is the destination is $P_i(D \mid B, deliver) = 1$
- The probability for all other states is $P(s' \mid s, a) = 0$.

Connection cost between nodes: The connection cost is affected by T_{free} and T_{busy}. $T_{free} \in V(s)$ represents the cost in case of successful message hand over($B \longrightarrow P$) under delegation action , and $T_{busy} \in V(s)$ is the cost in case of failed delegation action ($B \longrightarrow B$), which can be formulated as follows.

- $D_i(B \mid P, a) = T_{free}$ (time in which channel is free), where $a \in A_{h_i}$.
- $D_i(B \mid B, a) = T_{busy}$, where $a \in A_{h_i}$. It includes the cases when the buffer of the encountered node is full or the channel is busy.
- $D_i(s' \mid s, a) = 0$, where $a \notin A_{h_i}$

Cost function calculation: Given the estimated models $P(s' \mid s,a)$, $D(s' \mid s,a)$, and $(V_j(s'))$ the optimal value function can be calculated by solving the set of modified Bellman equations [14] for distributed model-based reinforcement learning as follows:

$$V_i(s) = max_a[Q_i(s,a)], \qquad (2)$$

$$Q_i(s,a) = \sum_{s' \in B,P,D} P_i(s' \mid s, a_j) \left(D_i(s' \mid s, a_j) + (V_j(s')) \right) \qquad (3)$$

Note that the calculation of Q-value is quite simple since each action has only two possible outcomes. For the event of handing over the message to encountered node n_j, given the message at buffer of custodian node n_i (state B), the Q-value at the node n_i is:

$$Q_i(B, a_j) = P_{ij}S[T_{free} + V_j(s')] - P_{ij}F[V_i(s') + T_{busy}]$$
$$= P_{ij}S[T_{free} + V_j(P)] - P_{ij}F[V_i(B) + T_{busy}] \qquad (4)$$

where the minus sign represents the network backward feedback from link to the node, T_{free} represents the duration of time in which the channel is free at node n_i during time-window period T^{window}, T_{total} is the total contact time during time-window period T^{window}. $P_{ij}S = \frac{T_{free}}{T_{total}}$ is the probability that the link between node n_i and node n_j is available during time-window T^{window}. $P_{ij}F = 1 - P_{ij}S$ is the probability that the link is busy during time-window T^{window}. The buffer effect takes place when the buffer of the encountered node is full. In this case, the time in which the buffer is full during a contact is added to the busy time of the channel, i.e., the busy time composed of the channel busy time and the time in which the buffer of the encountered node is full. The busy-time must be represented as "backward feedback" in the cost function Q. Note that the Q value is calculated at each custodian node regarding to the destination node, then the custodian node will make the decision of forwarding according to the routing forwarding policy. This routing strategy is mainly based on $max_a Q(B, a_j)$ values which can be simplified as:

$$Q_i(B,a_j) = \frac{Ps(T_{free} + Vj(P) - P_{ij}F.T_{busy}}{1 - PijF}$$

$$= V_j(P) + T_{free} - \frac{P_{ij}F}{PijS}T_{busy} \qquad (5)$$

where $V_j(p)$ is quality-metric, (cost function) represents the average cost of contact time between node n_j and the destination node n_d. Note that $V_i(B)$ is set to zero since we need to consider only the advertised $V_j(P)$ from node n_j regarding the destination node.

2) Information Exchange and Update During Contacts: The proposed technique performs two updates, update during encounters which includes summary vector and transitivity update, and update at end of the time-window.

The update during encounters: When two nodes are within their transmission range, they exchange V vectors regarding the message destination. With the update, the custodian node decides whether or not the message should be forwarded to the encountered node. This exchange of summary vectors is followed by another update, called transitivity update. The proposed algorithm has transitivity property[8], based on the observation that if node n_i frequently encounters node n_j and n_j frequently encounters node n_d, then n_i has good ability to forward messages to n_d through n_j. Thus, the congestion history of buffer and link availability of node n_j plays a key role in using transitivity property. In order to maximize the average contact time ,$V(s)$, between node n_i and n_d and make any message destined to node n_d goes through node n_j, a proper update using transitivity property should be made. To deal with this maximization problem (maximizing $V(s)$) using transitivity update, we formulate the information presentation first as inter-encounter time[15], then the corresponding average contact time is obtained. Using inter-contact time in the transitivity update is simpler and can adaptively update values only when $V_{(n_i,n_d)} < V_{(n_j,n_d)}$ in order to ensure that node n_i reaches n_d through n_j. Otherwise, if $V_{(n_i,n_d)} > V_{(n_j,n_d)}$, the transitivity property is not useful since node n_i is a better candidate for forwarding messages directly to node n_d rather than forwarding them through n_j. The inter-encounter time between nodes n_i and n_d, $R_{(n_i,n_d)}$, is calculated by $\frac{T^{winodw}}{V^{window}_{(n_i,n_d)}}$, where T^{window} represents duration of sliding window, $V^{window}_{(n_i,n_d)}$ is the average contact time duration ($V_i(s)$) between n_i and n_d during T^{window}. $R_{(n_i,n_j)}$, and $R_{(n_j,n_d)}$, are obtained using the similar way. The calculation of the $V(s)$ is given by equation (5). The new updated inter-encounter time is calculated as follows:

$$R^{new}_{(n_i,n_d)} = wR_{(n_i,n_d)} + (1-w)(R_{(n_i,n_j)} + R_{(n_j,n_d)}) \qquad (6)$$

where w is a weighting factor calculated from

$$w = \frac{R_{(n_i,n_j)} + R_{(n_j,n_d)}}{R_{(n_i,n_d)}} \qquad (7)$$

Note that w must be less than 1; that is $R_{(n_i,n_d)} > R_{(n_i,n_j)} + R_{(n_j,n_d)}$.

The new contact time is obtained by applying the following relation:

$$V^{new}_{(n_i,n_d)} = \frac{T^{window}}{R^{new}_{(n_i,n_d)}} \qquad (8)$$

V^{new} represents the new values of $V(s)$ that is obtained form the transitivity update. The introduced transitivity-update rule has great impact on protocol performance.

Time-window update: As discussed earlier, the connectivity between any two nodes is measured as the amount of meeting time intervals during a time-window T^{window}, which is mainly based on the time in which the wireless channel is busy or the buffer is full, and the time in which the channel is free and the buffer is available. These contact period components are time varying. They can change largely from time window to time window which have huge impact on the protocol forwarding policy. Hence, a representative smooth transfer of $V(s)$ values between consecutive time-windows is needed. We propose adaptive parameter that reflects the rate of change of connectivity between nodes and can be used to make smooth transfer from a previous time window to a current time window. We use T_{total1} and T_{total2} to refer to the total contact time (T_{total}) durations obtained during time windows T_1 and T_2, respectively. Three cases are identified: 1) the contact time duration in the previous time-window is less than the current time-window; 2) the contact time duration in the previous time-window is greater than current time-window; 3) no contact happened during one of the time-windows. The accumulated total contact time at the end of current time window is based on both the total contact time durations in current time-window and the last accumulated total contact time obtained in previous time-window. To combine these changes of connectivity values when transferring from previous time window to current time window, the three cases are formulated as follows:

- if $T_{total1} > T_{total2}$

$$V(s)^{new} = \beta.V(s)^{old} + (1-\beta)V(s)^{current} \qquad (9)$$

- if $T_{total1} < T_{total2}$

$$V(s)^{new} = (1-\beta)V(s)^{old} + \beta.V(s)^{current} \qquad (10)$$

The parameter β is given by

$$\beta = 1 - \frac{|T_{total1} - T_{total2}|}{max(T_{total1}, T_{total2})}, T_{total1}, T_{total2} > 0 \qquad (11)$$

- if $T_{total1} = 0$, or $T_{total2} = 0$ then $V(s)^{new} = \frac{V(s)^{old} + V(s)^{current}}{4}$. This case represents worst case scenario, i.e. unstable node behavior, or low quality of node mobility. Hence, the $V(s)^{new}$ value should be low.

3) The Weighted Copy Rule: The forwarding of message copies is based on the goodness of the encountered node regarding the destination[15]. The source of a message initially starts with L copies; any node n_i that has $N_i > 1$ message copy tokens (source or relay) and that encounters another node n_j with no copies, hands over to node n_j a number of message copy tokens according to its goodness for the destination node n_d. Node n_i hands over some of the message copy tokens to node n_j and keeps the rest for itself according to the following formula:

$$N_j = \frac{V(n_i, n_d)}{V(n_i, n_d) + V(n_j, n_d)} \qquad (12)$$

where N_i is the number of message tokens that node n_i has, $V_{(n_j, n_d)}$ and $V_{(n_i, n_d)}$ are the average of contact time between node n_j and node n_d, and node n_i and n_d, respectively. Node n_i hands over message tokens to node n_j only if the value of $V(s)$ at node n_j regarding to the destination is better than that of node n_i. This formula guarantees that the largest number of message copies is spread to relay nodes that have better information about destination node. After L messages have been copied to custodian nodes, each of the L nodes carrying a copy of the message performs according to the forwarding rule as descried in the following section.

4) The Forwarding Rule in ARBR :

- If the destination node is one hop away from an encountered node, the custodian node hands over the message to the encountered node.
- If more than one node exist in the transmission range of custodian node n_i, a node with the highest value of Q will be chosen among all other nodes, according to the relation $V_i(s) = max_a[Q_i(s,a)]$.
- If the value of $V(s)$ of the encountered node regarding to the destination node is greater than that of the custodian node by threshold value, $Th1$, a custodian node hands over the message to the encountered node.

V. PERFORMANCE EVALUATION

To evaluate the performance of the ARBR, extensive simulations were performed considering different scenarios. A DTN simulator based on that used by [6] was developed to suite ARBR mechanism. ARBR was compared with previously reported approaches under different variations of buffer capacities, traffic loads, and connectivity levels, using a realistic mobility scenario. The comparisons are in terms of the average delivery delay, the total transmissions, and delivery ratio.

A. Mobility Model

This study uses Community-Based Mobility Model[6] which is known to well resemble real node movements. The network consists of a number of nodes moving independently on a 2-dimensional torus in a geographical region. Each node transmits up to distance $K \geq 0$ meters away, and each message transmission takes one time unit. Euclidean distance is used to measure the proximity between two nodes (or their positions) n_i and n_j. Additionally, we assume that the network is intermittently connected most of the time. A slotted collision avoidance MAC protocol with Clear-to-Send (CTS) and Request-to-Send (RTS), has been implemented in order to arbitrate between nodes contending for the shared channel. The network model used for the proposed protocol varies from being extremely sparse to almost connected networks for a suite of different application scenarios. The performance of the ARBR were compared to those of the other protocols with heterogeneous node mobility and mobility showing a strong location preference. The proposed technique is compared to the following protocols.

- Epidemic routing [7],
- Spray and Focus (S&F) [7],
- Self-Adaptive Routing Protocol (SARP) [15],
- The proposed technique (ARBR),
- Most Mobile First strategy [6].

For all the protocols, an attempt has been made to tune the parameters in each scenario separately, in order to achieve good transmission-delay. The utility's threshold parameter for Spray and Focus, and SARP is set to 150 and 30, respectively.

B. Evaluation Scenarios

In this section, the performance of the protocols is evaluated first with respect to the low transmission range and varying buffer capacity under high traffic load (i.e., Scenario A). Second, with respect to the high-level of connectivity and varying traffic load (i.e., Scenario B). Finally, the performance of the protocols is examined in terms of the level of connectivity changes (Scenario C). In all scenarios, 100 nodes move according to the community-based mobility model [6] in a 500×500 2-dimensional torus in a geographical region. The message inter-arrival time is uniformly distributed during simulation time in such a way that the traffic can be varied from low traffic (10 messages per node) to high traffic (70 generated messages per node).

Scenario A: Effect of Buffer Size: In this scenario the performance of ARBR is examined under the situation the network is sparse (under low transmission range i.e., K = 20), with high traffic load (60 messages generated per node), and varying buffer space capacity. In this scenario we examine the effect of the backward feedback that is resulted when the buffer of a node reaches its capacity limit. If a node encounters another node and has limited remaining buffer space, a portion of the messages that should be forwarded can be delivered even though the encountered node metric is better than the custodian node. This situation results in extra queuing delay, especially in the case of flooding schemes. The performance of the ARBR was examined with respect to different buffer space values. Fig. 2 compares, with different buffer space capacities, the delivery delay, number
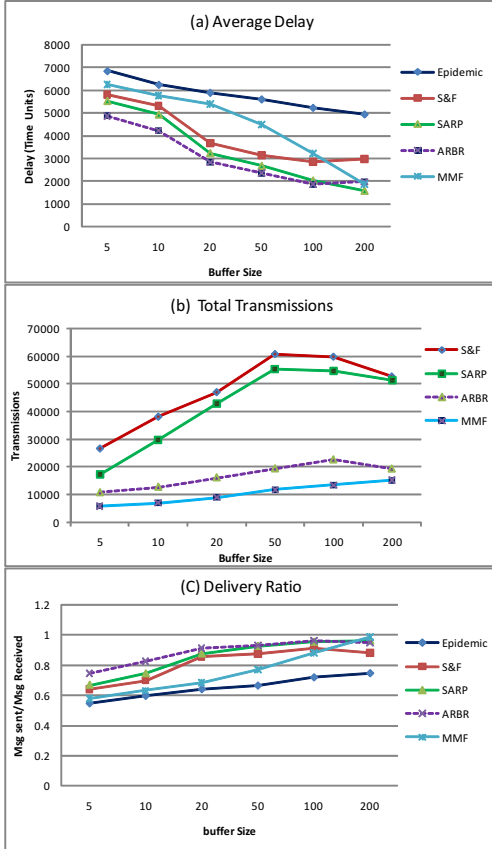
Figure 2. (a)(b)(c): The effect of buffer size on message delivery delay, number of transmissions and delivery ratio.

of transmissions performed, and the delivery ratio produced by the considered protocols. The buffer space was varied from 5 (limited capacity) to 200 (relatively high capacity) messages to reflect the performance of the protocols under the considered traffic load. As shown, when the buffer size is small (less than 100 message in our case) the performance of the protocols is highly affected by the capacity of the buffer, especially the Epidemic routing.

Epidemic routing produces the worst delivery delay in all scenarios, since it is affected by both the limited buffer size and the mobility model; ARBR scheme takes nodes that their buffers are full in its consideration by degrading their corresponding metric. As a result it produces the lowest delivery delay and highest delivery ratio. ARBR can produce delay shorter than that of SARP by 1.4 times, and delivery ratio higher than SARP by 1.8 times. Although it performs more transmissions than MMF, it produces delivery delay lower than that of MMF by 1.7 times. As the buffer size is expanded, the performance of all protocols improves. The SARP and MMF schemes produce the highest performance, since the effect of the constraint that

affects their performances is relaxed. However, ARBR still has very competitive performance with lower transmissions. It produces transmissions less than that of SARP by 3.5 times.

Scenario B: Effect of Traffic Load: This scenario is similar to scenario A but with variation in the traffic load. Each node attempts to randomly select a destination node, begins generating messages and continues to increase the rate which results in average traffic loads, i.e., the total number of messages generated throughout the simulation. The main goal of this scenario is to increase contentions on wireless channel and observe how the feedback mechanism that is resulted from this contention affects the performance of the ARBR. The protocols have been examined under low bandwidth value (i.e., one messages per unit of time), which makes it insufficient to enable some contacts to forward all intended messages. The protocols have been examined for two-buffer sizes: 1) unlimited capacity; and 2) low capacity (5 messages). Fig.3 shows the performance of all the routing algorithms in terms of the average delivery delay, total number of transmissions, and delivery ratio. In this scenario the buffer size effect is relaxed, and transmission range is set to 80 meters (highly connected network).

Epidemic routing produces the longest delivery delay and requires higher number of transmissions compared to all other schemes. It produces delivery delay at-least three times longer than the S&F does and require number of transmissions at least an order of magnitude higher than the ARBR scheme. Thus its not included in the figures. From the results shown in Fig. 3(a), 3(b), and 3(c), we observe that as the traffic load increases, the available bandwidth decreases, as a result the performance of the protocols decreases. Epidemic routing has the lowest performance. When the traffic load is moderate (less that 50 messages), it is clear that the MMF outperforms all existing multiple-copy routing protocols. This is because in MMF, the effect of buffer size is relaxed making all the nodes that are marked as roaming nodes buffer unlimited number of messages while roaming between communities. The ARBR scheme shows the second best delivery delay. ARBR can produce delay up to 0.97 of that of MMF, faster than that of SARP and S&F by 1.5 and 1.9 times, respectively. ARBR and MMF produce the best delivery ratio compared to all existing schemes. ARBR and MMF can achieve delivery ratio above 96% and 97% respectively, while the epidemic routing degrades below 40% for high traffic loads. SARP and S&F protocols can achieve delivery ratio above 92%, 79% respectively.

As the traffic load exceeds 50 messages per node, the contention on wireless channel become higher. ARBR outperforms all schemes since the ARBR mechanism takes in its consideration busy links by reducing their corresponding routing-metric, resulting in rerouting messages through low
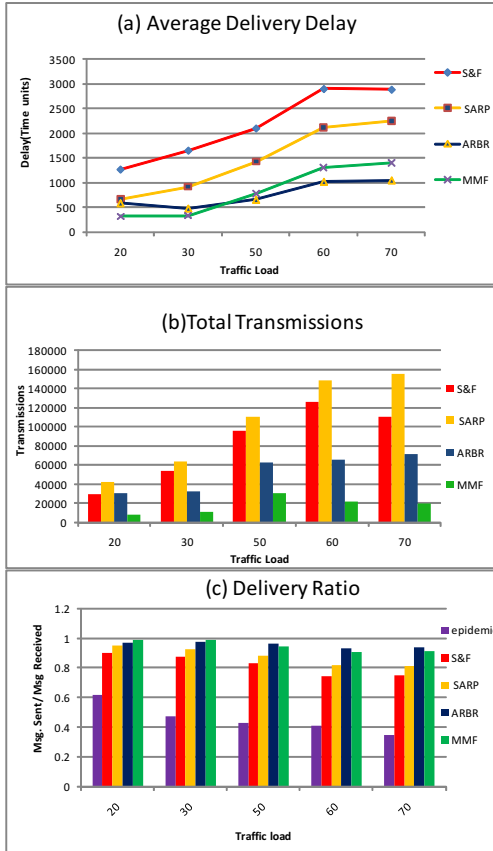
Figure 3. (a)(b)(c). The effect of traffic load on message delivery delay, number of transmissions, and delivery ratio.

Figure 4. (a)(b)(c). The effect of traffic load on message delivery delay, number of transmissions, and delivery ratio.

contention areas. However, this mechanism makes messages take long routes which result in more transmissions than that of MMF. The delivery delay obtained by the ARPR is faster 1.7 times than that of MMF.

As the buffer capacity reduced to low capacity (e.g. 5 messages), and the traffic load increases; the available bandwidth decreases and the buffer occupancy increases. When the traffic increases, it is clear that the proposed approach outperforms all existing multiple-copy routing protocols. The ARBR scheme obtains the fastest delivery delay and the best delivery ratio compared to all existing schemes. It is faster than MMF by 1.6 time, SARP by 1.3, and S&F by 1.5. Although the ARBR scheme requires more transmissions compared to the MMF, the number is still smaller than that produced by SARP. ARBR can achieve delivery ratio above 88% for all traffic loads, while the epidemic routing degrades below 50%, SARP and S&F below 80%, and MMF below 70% when the traffic load is high. When the traffic load is low (below 50 messages), SARP outperforms all other schemes in terms of delivery delay. It is faster than ARBR by 1.2. That's because the contention on wireless channel is low, making SARP employing its
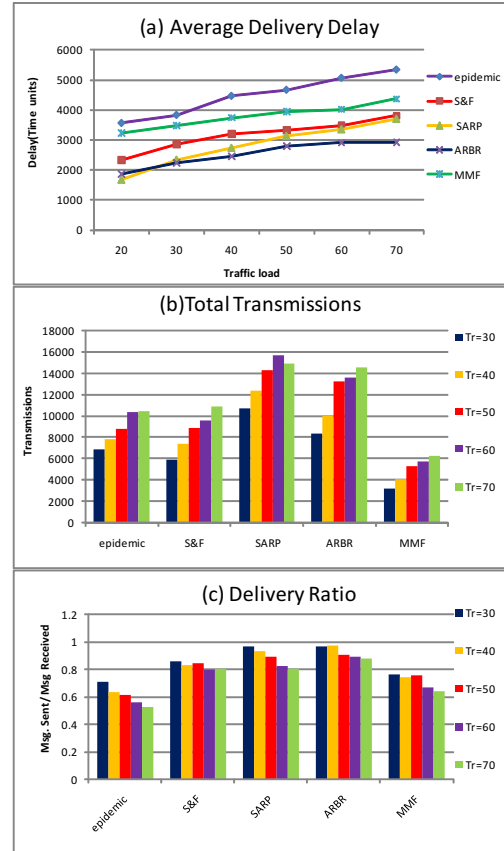
routing policy more efficiently. Figure 4(a), 3(b), and 3(c) shows the performance of all techniques under this scenario.

Scenario C: The Effect of Connectivity: This scenario studies the performance of the protocols at different connectivity ranges. The level of connectivity ranges from very sparse to highly connected networks by varying the value of K. This scenario examines the effect of feedback mechanism of the ARBR that is resulted from the cases when the network is highly congested and the demand on the wireless channel is very high. The buffer capacity is kept low (10 messages), and the traffic load is considerably high (60 messages). Figure 5 shows the average delay and the number of transmissions as a function of transmission range.

From the results shown in the figures, a number of interesting observations can be made about these figures. First, although Epidemic routing performs too many transmissions, it is still far from achieving competitive delays because of the contention caused by increasing the demand on the wireless channel. Second, the ARBR scheme outperforms all protocols in terms of delivery delay with fewer transmissions than the Spray and Focus and SARP schemes,
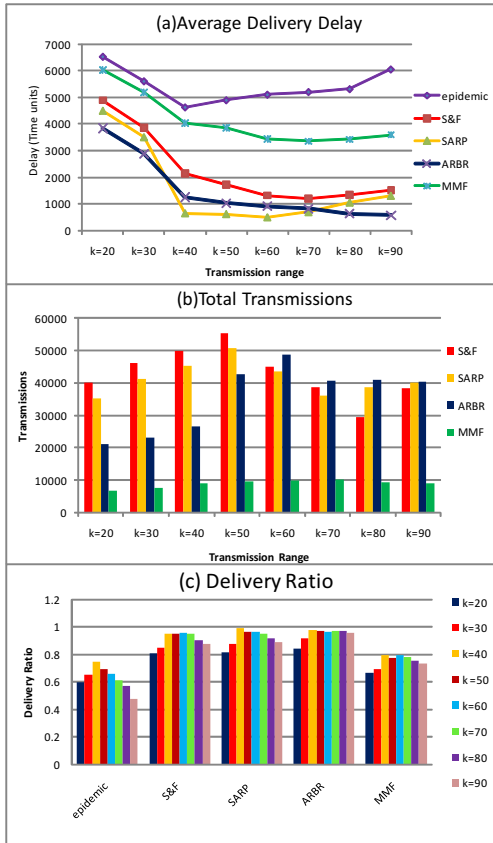
Figure 5. 5(a)(b)(c): The effect of connectivity on delivery delay, the number of transmissions, and delivery ratio.

for either low levels or high levels of connectivity. When the network is sparsely connected, the ARBR scheme can achieve shorter delivery delay than all other schemes, that is because the performance of other schemes is affected by the uncertainty of buffer occupancy status. On the other hand, when the network is moderate-connected, the SARP scheme can achieve a competitive-level of delivery delay compared to the proposed ARBR scheme with more transmissions. As the network becomes almost connected and the traffic load is high, the uncertainty of both buffer occupancy status and the availability of bandwidth affect the performance of the other techniques. As a result, the ARBR outperforms all other schemes in terms of delivery delay and delivery ratio.

VI. CONCLUSION

In this work we introduced a novel multi-copy routing scheme based on Collaborative Reinforcement Learning technique, called ARBR, for intermittently connected mobile networks. ARBR can successfully recognize and take advantages of potential opportunities for forwarding messages to their destinations.

The appropriate design of the utility function enables ARPR to capture the node mobility and network contention in an efficient way such that the forwarding rule of this scheme is highly accurate. The efficiency of the ARBR is characterized by rerouting messages around nodes experience either high buffer occupancy, wireless interference, or congestion. The proposed scheme exhibits great stability and performs few transmissions across all scenarios, while achieving a delivery delay that faster than all existing spraying and flooding based schemes when the the network experience contention on wireless links or buffer space.

REFERENCES

[1] "Disruption Tolerant Networking. http://www.dtnrg.org"

[2] J. Broch, D. A. Maltzsh, D. B. Johnson, Y.-C. Hu, and J. Jetcheva. A performance comparison of multi-hop wireless ad hoc network routing protocols. Proc. ACM MobiCom. Oct. 1998.

[3] V. Erramilli, M. Crovella. Forwarding in opportunistic Networks with Resource constraints. CHANTS'08, September 15, 2008, San Francisco, California, USA.

[4] Shen Ling, Wu Wei, "Feedback Adaptive Routing Algorithm for DTN," cmc, vol. 2, pp.267-271, 2009 WRI International Conference on Communications and Mobile Computing, 2009.

[5] Evan P. C. Jones,Lily Li, and Paul A. S. Ward.Practical Routing in Delay-Tolerant Networks.In SIGCOMM'05 Workshops, August 22–26, 2005, Philadelphia, PA, USA.

[6] Thrasyvoulos Spyropoulos, Thierry Turletti, Katia Obraczka, "Routing in Delay-Tolerant Networks Comprising Heterogeneous Node Populations," IEEE Transactions on Mobile Computing, vol. 8, no. 8, pp. 1132-1147, Aug. 2009, doi:10.1109/TMC.2008.172.

[7] T. Spyropoulos, K. Psounis, and C.S. Raghavendra, "Efficient Routing in Intermittently Connected Mobile Networks: The Multiple-Copy Case," IEEE/ACM Trans. Networking, vol. 16, no. 1, pp. 77-90, Feb. 2008.

[8] A. Lindgren, A. Doria, and O. Schel´en. Probabilistic routing in intermittently connected networks. SIGMOBILE Mobile Computing Communications Review, 7:19–20, July 2003.

[9] Dowling, J.; Curran, E.; Cunningham, R.; Cahill, V. (2005). Using Feedback in Collaborative Reinforcement Learning to Adaptively Optimise MANET Routing, IEEE Transactions on Systems, Man and Cybernetics (Part A), Special Issue on Engineering Self-Orangized Distributed Systems, pp. 360-372, Vol. 35, No. 3, ISSN 0018-9472.

[10] Jindal, A. and Psounis, K. (2007) 'Contention-aware analysis of routing schemes for mobile opportunistic networks', Paper presented in the Proceedings of the

1st International ACM MobiSys Workshop on Mobile Opportunistic Networking (MobiOpp 2007), pp.1–8.

[11] A. Krifa, C. Barakat, and T. Spyropoulos, "Optimal buffer management policies for delay tolerant networks," in Proc. of IEEE SECON, 2008.

[12] A. Lindgren and K. Phanse, Evaluation of Queueing Policies and Forwarding Strategies for Routing in Intermittently Connected Networks, First International Conference on Communication System Software and Middleware, 2006.

[13] J. Dowling, R. Cunningham, E. Curran, and V. Cahill. Building autonomic systems using collaborative reinforcement learning. Knowl. Eng. Rev., 21(3):231–238, 2006.

[14] Jim Dowling and Seif Haridi, "Decentralized Reinforcement Learning for the Online Optimization of Distributed Systems", Reinforcement Learning: Theory and Applications, Advanced Robotic Systems Journal, Editors Cornelius Weber, Mark Elshaw and Norbert Michael Mayer, I-Tech Education and Publishing, ISBN 978-3-902613-14-1, 2008: 142-167.

[15] Ahmed Elwhishi and Pin-Han Ho, "SARP - A Novel Multi-Copy Routing Protocol for Intermittently Connected Mobile Networks". GLOBECOM 2009: 1-7.

[16] Nelson, S. C., Bakht, M., Kravets, R., and Harris, A. F. 2009. Encounter based routing in DTNs. SIGMOBILE Mob. Comput. Commun. Rev. 13, 1 (Jun. 2009), 56-59. DOI= http://doi.acm.org/10.1145/1558590.1558602.

[17] V. Erramilli, M. Crovella, A. Chaintreau , C. Diot, Delegation forwarding, Proceedings of the 9th ACM international symposium on Mobile ad hoc networking and computing, May 26-30, 2008, Hong Kong, Hong Kong, China.

[18] A. Balasubramanian , B. Levine , A. Venkataramani, DTN routing as a resource allocation problem, Proceedings of the 2007 conference on Applications, technologies, architectures, and protocols for computer communications, August 27-31, 2007, Kyoto, Japan.