

RESEARCH

Open Access



# CONCOR: context-aware community-oriented routing for intermittently connected network

Rahul Johari<sup>1\*</sup>, Neelima Gupta<sup>2</sup> and Sandhya Aneja<sup>3</sup>

## Abstract

Delay-Tolerant Networks (DTNs) are characterized by time-varying and partially connected network topology. In such networks, mobility of node may be by virtue of its carrier. Social beings such as humans may act as a carrier of mobile nodes, and therefore mobility pattern of node follows a social-based movement model. Due to intermittent connectivity, end-to-end path between source and destination is rare and therefore message delivery is a challenge in such networks. In this paper, we present a new routing approach, context-aware community-oriented routing (CONCOR) that exploits the community and the context awareness of nodes, for efficient message delivery. Nodes with a common point of interest form a dynamic community. We have identified a set of node's attribute (context) and formulated a utility function to determine its capability to deliver a message. CONCOR is a multihop routing approach exploiting the nodes as message relays (store-carry-forward). Through simulation, we validated the effectiveness of node's attributes in estimation of message delivery probability. The simulations were done for our approach using the social movement-based data set on Opportunistic Network Environment (ONE) simulator. We compared our approach with dynamic social grouping (DSG)-based routing and context-aware routing (CAR) on three metrics viz. message delivery ratio, message traffic ratio and average message delay. We found that the results of our approach outperforms the DSG and CAR on all the three metrics.

**Keywords:** Intermittent connectivity; Routing algorithm; Context awareness; DTN

## 1 Introduction

A Delay-Tolerant Network is an occasionally connected network that may experience frequent, long-duration partitioning and may never have an end-to-end contemporaneous path.

Research in the area of Delay-Tolerant Networks (DTNs) has been receiving considerable attention in the last few years owing to their widespread occurrence in a variety of applications such as : a) InterPlaNetary (IPN) internet project [1], b) Wizzy digital courier service which provides asynchronous (disconnected) internet access to schools in remote villages of South Africa [2], c) transmission of information/message during mission critical operations like natural disasters or battle zones.

One of the earliest approach proposed for routing in partitioned networks is *epidemic routing* [3]. However, it is very expensive in terms of message traffic and buffer

space, which reduces the life of the network. To overcome the problem of heavy traffic and high buffer requirement, probabilistic routing [4,5] and social routing [6-10] schemes were proposed. Both the approaches provide considerable improvement in the delivery ratio under low buffer requirement.

The probabilistic routing schemes are based on the heuristic that if a node has interacted with a group of nodes in the past, it will do so again in the future. On the other hand, social routing assumes that the mobility of node is due to the mobility of its carrier (social beings, mostly humans). Therefore, movement of node follows a pattern known as social-based movement pattern. In such movement model, contacts established between nodes are not random; rather, they are based on the interest of social beings. Both probabilistic and social routing have two basic limitations. Firstly, the two approaches exploit only opportunistic contact wherein a message is forwarded by a node, say *A*, to the first node, say *B*, encountered with higher delivery probability than *A*. This is true, even if a node, say *C*, with higher message delivery probability

\*Correspondence: rahul@ipu.ac.in

<sup>1</sup>USICT, Guru Gobind Singh Indraprastha University, Sector 16C, Dwarka, Delhi 110078, India

Full list of author information is available at the end of the article

than  $B$  is available in the neighbour of  $A$ . Secondly, their performance in terms of message delivery ratio heavily depends upon the function chosen to compute the message delivery probability of nodes. In some of the routing approaches [8,11,12], delivery probability of node is computed using multiple attributes of node such as buffer space, battery level, colocation of nodes etc. Values to these multiple attributes is provided by context awareness of nodes.

Two categories of routing schemes have been proposed for DTN. First one exploits the opportunistic contact, and the other one uses the source based routing. For both the approaches, decision to select next hop node is to be made. For opportunistic contact, a decision is to be made whether to forward or not to forward a message to an encountered node. Whereas, in case of source-based routing, a decision is to be made, to which neighbouring node the message should be forwarded. This decision depends upon a node's capability (also called delivery probability) to deliver a message to destination. The routing schemes proposed in literature [4,13] suggest various node attributes for determining its delivery probability.

In this work, we present a routing approach, context-aware community-oriented routing (CONCOR), that exploits the existence of social groups [10] along with the context awareness of nodes for efficient routing of the messages in DTN. The *context* [11] has been defined as set of attributes that represent various characteristics of a system, such as available buffer space, maximum buffer space, information about node's community etc. A node may exploit the context information for efficient message delivery. The social groups known as community are formed around immobile base stations located at point of interest (POI) of social nodes. Each of the base station represents a distinct community, and initially it is the only member of its community. When a node frequently comes in contact with a base station and its contact strength [13] increases beyond certain threshold, it becomes a member of a community. We have assumed that members of community are disjoint. As communities are built around the base stations therefore number of communities in a network is equal to number of base stations. We assume that:

$n$  = number of base stations

$N$  = number of mobile nodes

$N \gg n$ .

The number of members in each community may vary. For context awareness, multiple attributes of node, such as buffer space, community, battery level etc., have been identified. Using the identified attributes, the message delivery probability of a node is computed with respect to each base station. We assume that base station serves as a sink for messages and nodes are the source and relays

for messages. For routing a message, each node maintains a delivery probability vector (DPV) table. Each record of DPV table is timestamped as detailed in Section 4.5 and represents a delivery probability vector (DPV) received from neighbouring nodes. The timestamp of a record remains valid for a limited time beyond which the DPV of a node becomes invalid and cannot be used for routing. When a node has a message to be delivered to a base station, it searches, in its DPV table, for the best possible neighbouring node whose message delivery probability for a base station is higher than itself. If such a node is found, the node forwards the message. Using simulation, we validate the effectiveness of each of the attribute in estimating the message delivery probability of node.

We have used multiple factors for computing a delivery probability vector. In [11], node's mobility, i.e. number of different connections established per unit time, is used as one of the factors for determining node's probability to deliver a message. We have shown that large number of different connections established by a node per unit time may not add to node's capability to deliver a message. Through simulation, we show that our approach outperforms both DSG (routing approach using dynamic social grouping that exploits opportunistic contacts) and CAR (Context-Aware Routing using source-based routing approach) in terms of message delivery ratio, average delay per delivered message and average message traffic per delivered message.

Although, in this paper, we have assumed the nodes to be collaborative and participating in network activities, but the impact of selfish behaviour of nodes on network performance cannot be denied specially if only a single copy of a message is maintained as it happens in first contact routing. In our approach, as multiple copies of a message are forwarded through different routes, therefore the impact of few nodes showing selfish behaviour will be small. As the study of impact of selfish behaviour of nodes on network performance is outside the scope of the paper, we would like to mention some of the studies conducted by other researchers about selfish behaviour of nodes. In [14], authors have studied the impact of distribution of selfish behaviour on performance of mobile social networks. Their work supports our view that existence of multiple paths makes social networks more robust to selfish behaviour of nodes.

The paper is organized as follows: Section 2 describes the problem. Section 3 presents the related work done in the area of routing in Delay-Tolerant Network. Section 4 presents our approach. Section 5 presents the CONCOR algorithm. Section 6 presents the experimental setup and analysis of the results generated after the implementation of the algorithm on Opportunistic Network Environment (ONE) simulator [15]. Section 7 presents the conclusion and future work.

## 2 Problem scenario

Figure 1 represents a scenario of a university system spanned over a large hilly area, with number of schools (departments) such as computer science, electronics, arts, commerce, biotechnology, administration, finance etc. The hexagons in Fig. 1 represents physical boundary of department. Each hexagon has one base station whose transmission range is shown by a dotted circle. The location of base station represents the office of head of department. The circles, triangles, rectangles and squares represent people working in different departments. People working for a department will come in frequent contact with the base station of their department. The office of the head of a department represents a point of interest (POI) for people working for that department. People working in a university may carry a communicating device such as iphone or ipad etc. The mobility of the device is guided by people movement. Over a period, the communication devices will come in contact with base station and will form a community around base station. Intuitively, nodes belonging to one department will form a community with the base station of that department.

Due to lack of wired infrastructure and hilly terrain, the existence of connected end-to-end path between the

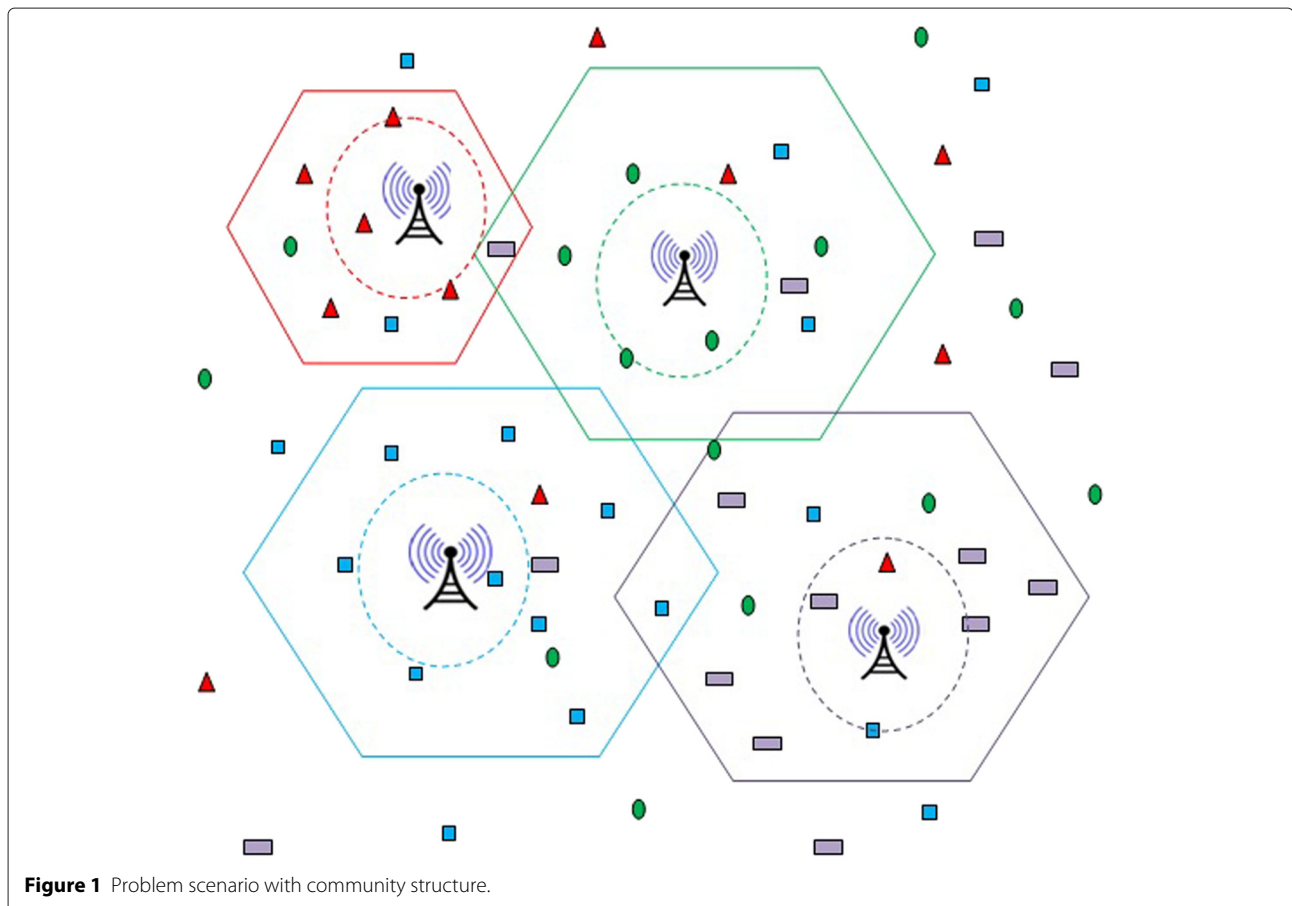
source (mobile devices) and destination (base station) of message is rare. In this scenario, routing a message from source to destination is a challenge. Using store-carry-forward approach to deliver a message is a solution, but delays are inevitable. In the absence of complete route from source to destination, choosing a next hop node is a decision problem.

## 3 Related work

One of the earliest approaches proposed for routing in partitioned networks is *epidemic routing* [3]. The goal of epidemic routing was to maximize the message delivery ratio while minimizing the message delivery time. However, it is very expensive in terms of message traffic and buffer space, which reduces the life of the network. That is, the approach does not seem to scale as the number of messages in the network grows. An intelligent buffer management scheme can improve the delivery ratio over the simple FIFO scheme.

Spray and Wait [16] routing scheme was proposed to provide controlled forwarding thus leading to reduced message traffic.

To enhance the performance of routing schemes, several researchers studied the characteristics of relay node



**Figure 1** Problem scenario with community structure.

to compute its delivery probability. It was observed by researchers [4,17-19] that the mobility is mostly not random and there is a pattern in encounters. The probabilistic routing scheme is based on the idea that nodes that have met in the past will again do so in the future and therefore nodes that have met the destination directly or through transitivity will have higher probability of delivering a message.

The routing problem in wireless networks (with intermittent connectivity) is transformed into a decision problem wherein a node having a message has to choose the next best hop (relay) node for message forwarding instead of relying on opportunistic contact. Various routing schemes for DTN evaluate the message delivery capability of a relay node for message forwarding.

Delivery probabilities are computed using the history of encounters and transitivity to reflect that if a node has been encountered in the past, it is likely to be encountered again. The approach introduced by Wang et al. in [5] is also based on probabilities of nodes of successful delivery of a message to a base station. However, the delivery probabilities are computed based on successful delivery of messages rather than regular contacts between the nodes.

As the movement of nodes is mostly not random, it was also observed that it is affected by the interest of its carrier, mostly social being. Various routing schemes exploited the social-based movement patterns for choosing the most appropriate relay node.

The SimBet routing algorithm [6] borrows ideas from social networking and contact patterns to predict paths to destinations for improving delivery ratio and time. The algorithm assumes that the groups have been formed and calculates the betweenness and centrality metrics to make routing decisions. Bubble rap [7] extends their work by allocating nodes into social groups based on direct and indirect contacts. Based on the global knowledge of the nodes' contact, they construct  $k$ -cliques/clusters to define the groups. However, grouping is static in their approach.

In [13,20], authors presented an approach that combines the advantage of social grouping and message forwarding based on message delivery probability of a node. Groups are formed dynamically as nodes come in contact with each other. The message is either forwarded using the individual delivery probability of a node, which is computed on the basis of messages forwarded by a node or by using the group probabilities of all those groups of which node is a member. In [13], the authors present an application of social grouping on network with single base station; and in [20], approach for network with  $n$  message source and sink is presented. The authors have presented an approach for identification of social groups dynamically. The approach suffers from overhead involved in creation and maintenance of groups.

To better exploit the social-based movement for message forwarding, the characteristics of nodes movement were studied. Working day movement model was proposed in [21].

In [22], authors presented the impact of movement characteristics on contact and inter-contact time. The contact time was used for determining the utility of contact being established. By capturing these characteristics of movement, the authors proposed prediction of the activities being performed by a node carrier.

Nodes capability as a relay node not only depends upon the movement pattern of its carrier but it also depends upon other attributes. In [23], routing scheme has been proposed that computes message delivery probability of a node on the basis of buffer space.

In [11], authors present an idea of CAR (Context-Aware Routing) protocol for routing of the message in DTN. The protocol is based on the idea of exploiting nodes as carriers of messages in a partitioned network to achieve message delivery. The choice of the best carrier is made using Kalman-filter-based prediction techniques and utility theory. The authors used two of the node's attribute viz-a-viz, colocation and change degree of connections, for estimating the delivery probability of a node. In the said approach, a single copy of message is forwarded. The approach does not exploit the social contact pattern of nodes.

In [24], the authors proposed a distributed scheme for sparse networks by calculating resources such as buffer, bandwidth and energy levels to take crucial routing decision. The routing decision is based on the quantum of resources available in a node's vicinity. However, authors does not exploit the existence of communities for efficient routing decisions.

In [25], authors proposed a framework for energy-efficient message forwarding using opportunistic contacts. In their work, they have proposed static and dynamic forwarding policies. Unlike the probability of forwarding a message by a node, as used in [25], in this paper, we discuss the message delivery probability of node. As per our approach, a node forwards a message only when it encounters a node that is more capable of delivering a message to its destination. This any how makes efficient use of energy by controlling the message forwarding and at the same time reduces message traffic.

In [26], authors proposed the usage of throwboxes, battery power stationary nodes, for enhancing the performance of DTN. In [27,28], authors proposed further enhancement in the architecture of throwboxes so as to make them energy efficient. The authors proposed a multitiered, multiradio hardware platform that applies heuristic for using the two tiers. By using low-power tier for listening the beacons and high-power tier for message exchange, throwboxes are able to provide energy-efficient

solutions to DTN in general and sensor networks in particular. The proposed architecture can be implemented in DTN nodes, and its effect on the performance may be studied further. However, the implementation is outside the scope of this paper.

The knowledge of nodes' physical location and a schedule of futuristic locations or contacts have also been utilized for routing in DTN. In [2], authors have proposed an approach for routing in DTN using modified Dijkstra algorithm. They assume that the DTN nodes possess the knowledge about the time and duration of contact. On the basis of this knowledge, a node may create a number of routing metrics. Their results show that the efficiency and performance increases with the amount of information used for the metric. In DTN, determining the futuristic knowledge of opportunistic contacts is extremely difficult and therefore the approach presented in [2] may only exploit the scheduled contacts for message transfer and opportunistic contact may not be used. The weakness of this approach is that each node must have access to updated schedule data.

In [29], authors have presented an approach that uses link state routing protocol. The link state packets are exchanged using epidemic approach which leads to overhead. Whenever a node has a message to transfer, it forwards message to the node which is closest to the message destination. The approach is good in the sense that it requires very small buffer space and is therefore scalable. But the approach may not be practical in the scenario wherein the nodes have sparse connectivity.

## 4 Our approach

CONCOR is a source-based routing technique. The approach exploits the existence of community structure existing in social networks and the context awareness of node to efficiently deliver the messages. When a node has a message to be delivered to the destination, it scans its neighbour. If a destination is found, a message is delivered else a next hop node is selected. A node forwards a message to that neighbouring node whose delivery probability is more than the node itself and is also highest amongst all the neighbouring nodes. Periodically, each node computes its message delivery probability with respect to each of the destination. The method to compute delivery probability is detailed in Section 4.2.

### 4.1 Community structure

The network contains social groups called community. A community is formed around a base station. The location of a base station represents a point of interest (POI). Number of communities is equal to number of base stations. Initially, each base station is the only member of its community. A node's carrier (social being) may be interested in one of POIs and may frequently visit it. A node carried

by a carrier may come in contact with the base station located at POI of the carrier. When the contact strength between a node, say  $A$  and base station, say  $B_i$ , increases above the threshold value ( $\psi$ ), node  $A$  becomes a part of the community represented by  $B_i$ . The contact strength is a function of time duration for which the two nodes were in contact and the duration for which they were not in contact. We are using the same function for computing contact strength as used in [13]. For completeness, the computation of contact strength is given in Equation 1.

$$\lambda_{A,B_i} = \lambda_{A,B_i}(1 - \phi) + \phi \frac{\text{duration}_{\text{contact}}}{\text{Total\_Duration}} \quad (1)$$

where:

$\lambda_{A,B_i}$  is the contact strength between node  $A$  and base station  $B_i$ .

$\phi$  is the rate at which contact strength changes. We are using the same value of  $\phi$  as used in [13].

$\text{duration}_{\text{contact}}$  is the time duration for which node  $A$  is in contact with  $B_i$ .

A node uses its community information in computing its DPV. A node will have a higher delivery probability for a base station of its own community. Such a node will be better capable, than the nodes of other community, for delivering a message destined to the base station of its own community. The community formation in CONCOR is different from group formation approach mentioned in [13]. The group formation and expansion may or may not involve a message destination as one of a group member. Therefore, fact that node is a member of a group does not reflect node capability to deliver a message.

This means that the purpose of group creation may be futile, and it may only result in overheads due to group formation. The group [13] may be formed between any two nodes and two groups may merge to form a larger group, but in CONCOR a node becomes a member of a group only when it comes in contact with a base station.

Further, unlike group formation in DSG, where a clusterhead of group maintains the list of member nodes, community formation is decentralized and there is no clusterhead. Each node is aware of its own community. This implementation of community formation avoids the overhead of message exchange occurring in the group formation in DSG [13]. Also, there is no concept of node resigning from a group, again reducing the overhead. In community formation, a node periodically checks its contact strength for a base station of community. If the contact strength is below the threshold value, a node ceases to be a part of community.

#### 4.2 Structure of delivery probability vector (DPV)

Delivery probability of a node for a particular base station represents its capability to deliver a message to that base station. DPV of a node is a collection of node's delivery probabilities for each of the base station. A sample DPV  $[(0.34, 0.23, 0.17, 0.95)]$  maintained by a node that has non-zero delivery probability for four base stations. In the sample DPV, 0.34 is the delivery probability (dp) for base station  $B_1$ , 0.23 is the dp for  $B_2$ , 0.17 is the dp for  $B_3$  and 0.95 is the dp for  $B_4$ . The context awareness of a node is exploited for computing its delivery probability. A node computes its delivery probability using a multi-variable utility function, detailed in Section 4.3. The following attributes are considered for computing its capability to deliver a message to a base station.

1. Member of community represented by a base station.
2. Messages delivered per unit time to a base station.
3. Total messages forwarded per unit time.
4. Ratio of available buffer space to the maximum buffer space.
5. Ratio of available battery level to the maximum battery level.
6. Is node one hop away from destination?

If a node is a member of a community represented by a base station, it is very likely that a node will meet the base station in near future. Considering all the parameters the same, if a node say  $A$  contacts a node, say  $B$ , of the same community as the message destination, then node  $B$ 's utility to deliver that message will be higher than  $A$ .

We have considered message delivery per unit time as an attribute for computing the delivery probability of a node for a particular destination. If the value of this attribute is close to 1, a node is capable of delivering a message and its delivery probability is also high.

Total messages forwarded by a node per unit time is used for defining the goodness of a node to deliver messages. Heuristically, if a node is able to forward a large number of messages to other nodes having higher delivery probability, it means that the node is coming in frequent contact with the nodes having higher delivery probability.

We use the ratio of available buffer space to the maximum buffer space as a node attribute for computing its utility. A node having large free buffer space at any point of time is more capable to store a message for longer duration in comparison to a node with smaller free buffer space.

Another attribute which enhance the node's capability of delivering a message is the node's battery level. This attribute is more important in sensor networks. In [11], authors have highlighted the importance of the attribute in computing the utility of a node to deliver a message. A node with high battery level can remain alive for longer duration and will be able to forward or deliver the messages. The consumption of battery power by a node occurs

due to communication- and non-communication-related activities. If a node forwards or delivers more messages, then its power consumption will be more in comparison to a node not involved in forwarding and delivery of messages. Therefore, the maximization of two attributes, battery level and messages forwarded or delivered by a node seems contradictory. However, consider a network, where there are two sets of nodes. One that works on high power transmission mode and the other one that works on low power transmission mode. Suppose, number of messages forwarded by both the type of nodes is the same. The node working on high transmission mode will lose more battery power than the one working on low power transmission. As a result, a node operating on low power transmission mode have a higher delivery probability than the one operating on high power transmission mode.

Periodically, a node exchanges its delivery probability vector with the neighbouring nodes. Therefore, a node will be aware of existence of any of the base station in its neighbour. A node that is one hop away from a destination will have high utility value for forwarding a message to the destination.

In [11], authors introduce change degree of connectivity as one of the attribute for computing utility function. The change degree of connectivity specifies how frequently a node changes its contact. The high value of change in degree of connectivity means that a node is establishing connections and disconnections very frequently and probably is moving at high speed.

If a node establishes number of connections and disconnections in very short time span, it is very likely that a node comes in contact with other nodes for a small duration. In such a short duration, transfer of large-size message might not be possible. Moreover, as the change in degree of connectivity increases, less number of messages will be delivered.

A context awareness attribute used in [11] is co-location. Using community structure, we have captured the idea of co-location of node with destination. The application of prediction approach, such as Kalman filter, requires the knowledge of time at which the node was co-located with one of the destination. This knowledge is not required in community structures.

#### 4.3 Utility function

The utility value of a node to deliver a message to specific destination depends upon number of attributes already defined in Section 4.2. For a given node, there is a utility value for each of the destination. The value of utility function ranges between  $[0,1]$ . A value close to 1 means a node is highly capable of delivering a message. A node is considered to be highly capable for forwarding a message to a destination if it is one hop away from the destination.

Considering all the identified attributes, as mentioned in Section 4.2, we have designed a utility function for a node as mentioned in Equation 2.

$$U(A, b_i) = \begin{cases} .95 & \text{if } A \text{ is one hop away} \\ & \text{from the destination} \\ dp & \text{if } A \text{ is more than one hop} \\ & \text{away from the destination} \end{cases} \quad (2)$$

where

$U(A, b_i)$  is the utility of node  $A$  to deliver a message to destination  $b_i$   $dp$  is computed as:

$$dp = Cf \times (W_a \times MDT + W_b \times MFT + W_c \times BL + W_d \times BuffS + W_e \times C_B)$$

$Cf$  – compulsory factor. It is 1 if battery level and buffer space is above a threshold else it is 0.

$MDT$  – messages delivered to destination  $b_i$  per unit time.

$MFT$  – messages(with destination  $b_i$ ) forwarded per unit time.

$BL$  – ratio of current battery level to the maximum battery level.

$BuffS$  – ratio of available buffer space to the maximum buffer space.

$C_B$  – is 1 if community of node  $A$  is  $b_i$  else it is 0.

$W_a, W_b, W_c, W_d, W_e$  – weights assigned to various attributes.

#### 4.4 Initialization and update of DPV

The delivery probability vector (DPV) is initialized at the time of node startup. At this time, a node, say  $Node_A$ , is not a part of any community. It has neither delivered nor forwarded any messages.  $Node_A$  is also not aware of its one hop connectivity to a base station. In such a scenario, the utility of a node is defined on the basis of available buffer space and the battery level. At this time, the ratio of  $Node_A$ 's buffer space to the maximum buffer space holds a maximum value. Also, the ratio of current battery level to the maximum battery level is maximum, which is 1. As the weight assigned to the two attributes is 0.1 (see Table 1: Weight reference table), therefore using the utility function, the initial DPV of  $Node_A$  for each of the base stations is 0.2. If the nodes have varying buffer space, then

**Table 1 Weight reference table**

Symbol	Description	Tested value	Ideal value
$W_a$	Weight assigned to MDT	10 to 40	20
$W_b$	Weight assigned to MFT	10 to 40	10
$W_c$	Weight assigned to BL	0.1 to 0.5	0.1
$W_d$	Weight assigned to BuffS	0.1 to 0.5	0.1
$W_e$	Weight assigned to $C_B$	0.1 to 0.5	0.3

the initial DPV will be proportional to the buffer space and may range between [0.0 and 0.2] .

Periodically, a node updates its DPV. If a node is frequently involved in forwarding of the messages, its DPV will increase for the destination whose message is forwarded by the node. Similarly, if a node delivers a message to a destination, its delivery probability will increase with respect to that destination. Over a period, a node's utility to deliver a message to a base station may increase or decrease.

#### 4.5 DPVT initialization and update

Each node maintains a delivery probability vector table. A node periodically receives a DPV from its neighbouring nodes. The received DPV is stored in a table with a timestamp. If a node, say  $Node_A$ , receives a vector from another node, say  $Node_B$  for the first time, it creates a new entry for  $Node_B$  in its DPVT. If DPV of a node, say  $Node_B$ , already exists in DPVT of a node, say  $Node_A$ , then the value of DPV and the timestamp is updated. If the DPV of a node is not refreshed for long duration, the entry becomes invalid and is not used for selection of the next hop node.

The structure of DPVT is shown in Table 2. The table is prepared for  $Node_A$ . Each entry in the Table 2 is the DPV received from other nodes. The DPV received from base station has delivery probability of 1 for itself whereas the delivery probability for other base stations is 0 as the base stations are stationary.

### 5 Proposed algorithm

In this section, we first present an algorithm for detection of community structure in the social-movement-based mobility pattern of the node, followed by algorithm for computing the DPV, updating of DPVT and finally the CONCOR algorithm. In CONCOR, replicated messages are not forwarded. We have made the following assumptions:

1. A node is a source and carrier of message, and a base station is a sink for message.
2. Number of nodes in a network are much more than number of base stations.
3. A node is aware of all the base stations existing in the network.
4. A node computes its message delivery probability with respect to each of the base station.
5. Each base station represents a community.

**Table 2 Sample of DPVT**

Node	dpv	Timestamp (sec)	Valid (Y/N)
$Node_B$	[.36 .12 .45 .67]	1,560	N
$Node_C$	[.36 .12 .45 .67]	1,562	Y
$BS_1$	[1 0 0 0]	1,564	Y

6. Initially, a base station is the only member of community.
7. No node is selfish or byzantine.
8. Communities are disjoint.

---

**Algorithm 1** Community expansion

---

**Notation**

$BS_i$  – Base Station of  $i^{th}$  community

$Node_A$ : A DTN Node

$Mem_A$ : Community of which  $Node_A$  is a member where  $Mem_A = Nil, 1..n$  and  $n$ =number of Base Stations

$i$  = Community around  $BS_i$

$\lambda_{A,BS_i}$  – Contact Strength between  $Node_A$  and  $BS_i$

$\psi$  : Threshold value for joining a community

**Trigger:**  $Node_A$  contacts  $BS_i$

In  $Node_A$

**if**  $\lambda_{A,BS_i} > \psi$  **and**  $Mem_A = Nil$  **then**

$Mem_A = i$

**end if**

---



---

**Algorithm 2** Creation of delivery probability vector

---

**Notation**

$dpvTable_A$  – Delivery probability vector table of  $Node_A$

$MDT_A$  – Message delivered per unit time by  $Node_A$

$MFT_A$  – Message forwarded per unit time by  $Node_A$

$Buff_A$  – Available buffer space in  $Node_A$

$BL_A$  – Battery level of  $Node_A$

$Mem_A$  – Community of  $Node_A$

$\epsilon$  – Threshold for buffer space

$\eta$  – Threshold for battery level

$B$  – Set of Base Stations in a network

$b_i$  –  $i^{th}$  base station

$dpv_A[ ]$  – Array representing Delivery Probability Vector of  $Node_A$

**Trigger:** Periodically in  $Node_A$

**for each**  $b_i$  in  $B$

**if** isOneHopAway( $b_i$ )

$dpv_A[ i ] = .95$

**else if**  $Buff_A > \epsilon$  **and**  $BL_A > \eta$

$dpv_A[ i ] = \text{Compute } U(Node_A, b_i)$

**else**

$dpv_A[ i ] = 0$

**end if**

**end for**

---

**Function 1** is one hop away

---

boolean isOneHopAway(Base Station  $b_i$ )

**if**  $dpv_{b_i}[ ] \in dpvTable_A$  **and**  $timestamp$  of  $dpv_{b_i}$  is Valid

**return true**

**return false**

---



---

**Algorithm 3** Update of delivery probability vector table

---

**Notation**

$dpvTable_A$  – Delivery probability vector table of  $Node_A$

$Time_{curr}$  – Current Time

$\delta$  – Timer

**Trigger:**  $Node_A$  receives  $dpv_B[ ]$  from  $Node_B$  or Periodically

In  $Node_A$

**Insert**  $dpv_B[ ]$  in  $dpvTable_A$

Reset  $Timestamp = Time_{curr}$  for inserted vector

$Timestamp$  for  $dpv_B[ ]$  becomes **valid**

**for each**  $dpv_B[ ] \in dpvTable_A$

**if**  $Time_{curr} - Timestamp > \delta$  **then**

$Timestamp$  for  $dpv_B[ ]$  becomes **invalid**

**end if**

**end for**

---



---

**Algorithm 4** Context-aware community-oriented routing (CONCOR)

---

**Notation**

$Node_A$  – A DTN Node

$Node_B$  – A DTN Node

$Mem_A$  – Community of  $Node_A$

$msg_{b_i}$  – Message with  $b_i$  destination

$dpv_A[ ]$  – Delivery Probability Vector of  $Node_A$

$dpv_B[ ]$  – Delivery Probability Vector of  $Node_B$

$dpvTable_A$  – Delivery probability vector table of  $Node_A$

$BestHop$  – Best relay node for  $msg_{b_i}$

**Trigger:**  $Node_A$  creates or receives  $msg_{b_i}$

In  $Node_A$

$BestHop = Node_A$

$var = dpv_A[ i ]$

**A1: for each**  $dpv_B[ ] \in dpvTable_A$

**if**  $dpv_B[ i ] > var$  **then**

$BestHop = Node_B$

$var = dpv_B[ i ]$

**end if**

**end for**

**if**  $BestHop \neq Node_A$  **then**

**if** isNeighbour( $BestHop, Node_A$ ) **then**

{ isNeighbour( $A, B$ ) returns true if  $A$  and  $B$  are in each others transmission range else it returns false }

Send  $msg_{b_i}$  to  $BestHop$

**else**

Set  $dpv_B[ ]$  in  $dpvTable_A$  **invalid**

Goto A1

**end if**

**end if**

---



## 6 Analysis

### 6.1 Effect of timer on CONCOR

The value of timer affects the performance of CONCOR. If the value of timer is very small, then the entry in the DPVT will become invalid very frequently. As a result, a node (say Node<sub>A</sub>) will not be able to choose the relay node although a node (say Node<sub>B</sub>) capable of delivering a message is in its neighbour. On the other hand, if the value of timer is large, DPVT will continue to hold a valid DPV for a node which is not in range. As a result, the execution efficiency of routing approach will decrease. The effect of timer was studied through simulation also. The results of the same are detailed in Section 6.6.

### 6.2 Effect of DPV exchange interval

A node periodically broadcasts its DPV that is received by all the nodes in the direct transmission range of the given node. If this period of exchange is very small, then a node will frequently forward and receive the DPV message to/from its neighbouring nodes and will in turn lead to overhead on itself and network resources. On the other hand, if this period is large, a node may indicate a DPV, as valid, for other node that is not in its neighbour. This will affect the performance of CONCOR, as the algorithm will choose a relay node that is not in the range of the given node. Therefore, DPV exchange interval should be optimal, neither too short nor too long.

### 6.3 Experimental setup

The Opportunistic Network Environment (ONE) simulator implemented in Java and available as open source has been used to evaluate the algorithm. The ONE simulating environment is capable of simulating mobility patterns of nodes and message exchange between them. Many of the routing algorithm applicable to DTN environment are pre-implemented in the simulator. We implemented routing as used in our algorithm (CONCOR), the DSG-based routing algorithm [13] and context-aware routing [11]. Three metrics viz. message delivery ratio, message traffic ratio and delay per delivered message were used to compare the performance of DSG, CAR and CONCOR. The simulator generated total messages delivered, message delivery ratio, total and average message traffic per delivered message. The simulator was customized to generate total delay and average delay per delivered message. The simulation was run for non-replicated messages.

### 6.4 Data source

The simulation was run using randomly created network with eight stationary base stations and 75 mobile nodes running delay-tolerant applications. The said number of nodes were chosen after repeated experiments so that community structure is formed and its effect could be studied in the network. Of these 75 mobile nodes, 25

nodes (one third) were high transmission power nodes, whereas the remaining 50 (two thirds) were low transmission power nodes. Power consumption during communication involving high transmission power nodes is four times the power consumed during non-communication. The communication in nodes with low transmission power consumes twice the power consumed during non-communication. The base stations are assumed to be high power transmission nodes with constant power supply. The data traces from North Carolina State University (NCSU) available at [30] were used to provide the mobility pattern of nodes. The data set is the movement trace of students in a university campus. The data set contains the location of nodes at different times. The specific data set was chosen because it closely relates to the problem scenario. The data set represents students' movement pattern and therefore has been used as a social-based movement pattern. The nodes were allowed to move in an area of 5.4 km<sup>2</sup>.

### 6.5 Simulation parameters

Table 3 presents the simulation parameter as used in the experimental setup. The parameters with respect to messages, nodes, utility function and various threshold values used in algorithm were set. Table 1 presents the description and values of various weights as used in utility function. Table 4 presents the description and the ideal value for various variable as used in our approach. The

**Table 3 Simulation parameters**

Serial number	Parameter	Value
1.	Max. buffer space	1,000 MB
2.	Max. battery level	1,600 mAh
3.	Buffer space in each node	256 MB
4.	Power consumption in communication for 195 mW high power transmission node	
5.	Power consumption in communication for 112 mW low power transmission node	
6.	Power consumption in non-communication-related activity	62 mAh
7.	Message Size	4 MB
8.	Message time to live	1,440 min
9.	Rate of message (msg) generation	1 msg/min
10.	Data transfer rate	2 Mbps
11.	Time to transfer a message between nodes	16 s
12.	Speed of mobile nodes	2 to 3 km/h
13.	Transmission range at low power transmission	30 m
14.	Transmission range at high power transmission	90 m
15.	Timer	60 s

**Table 4 Variable reference table**

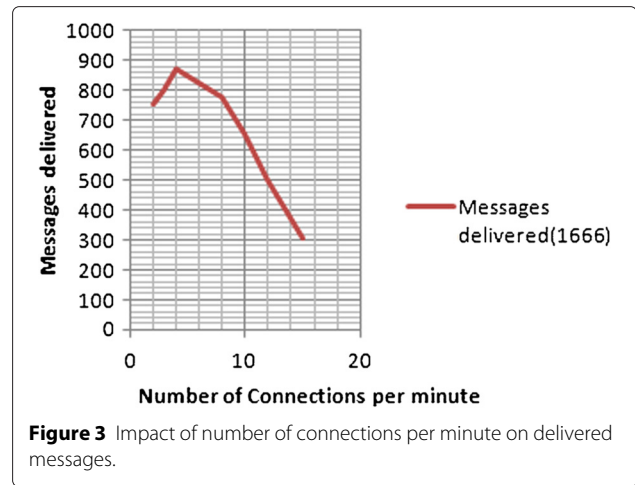
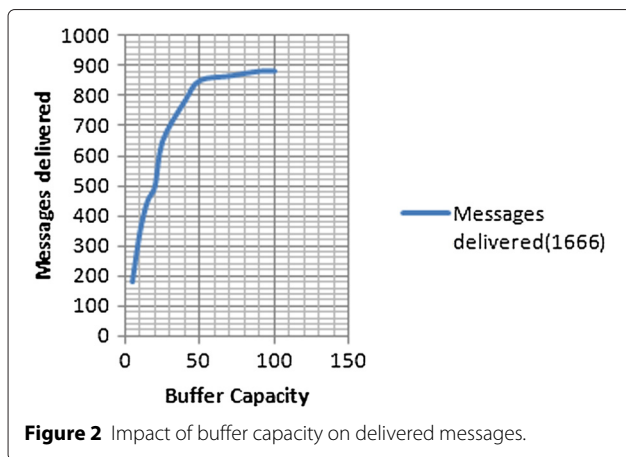
Var	Description	Tested value	Ideal value
$\epsilon$	Threshold for buffer space	2 to 10 MB	6 MB
$\eta$	Threshold for battery level	65 to 200 mAh	112 mAh
$\psi$	Threshold value for joining community	.01 to .05	0.04
$\phi$	Contact decay ratio	0.1 to 0.5	0.3

maximum buffer space and maximum battery level are constant values. It is assumed that the nodes are aware of these maximum values.

**6.6 Comparison of results**

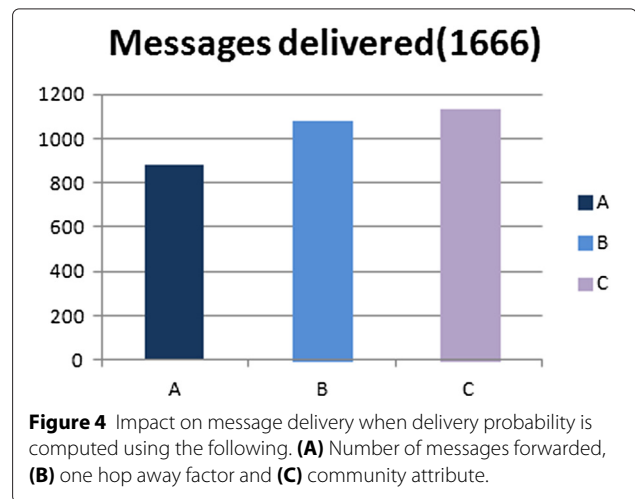
We conducted four different sets of experiments. During the first experiment, the effect of various context attributes on the routing performance was studied. In the second experiment, the effect of variation in timer value was studied. In the third experiment, the effect of number of mobile nodes on routing performance of CONCOR was compared with DSG and CAR.

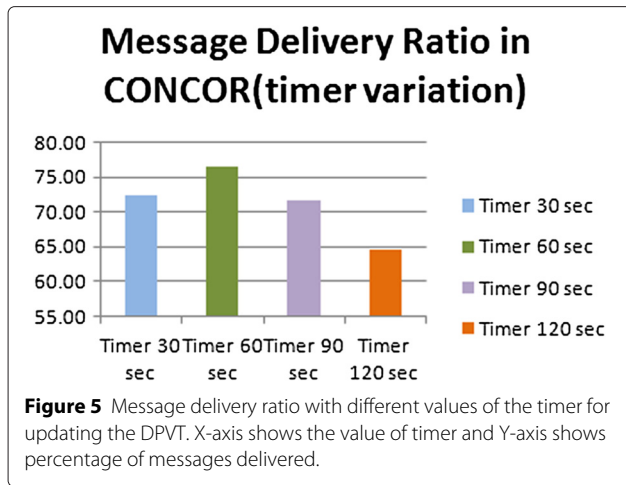
In the first set of experiment, we validated the applicability of node’s attribute in defining message delivery capability of a node. For this experiment, we varied only one attribute at a time keeping all other attributes constant. Under these experiments, firstly we validated the applicability of buffer capacity of a node. In this experiment, multiple simulations were done with varying buffer capacities, ranging from 5 to 100 MB. The results presented in Figure 2 shows that when the buffer capacity of the nodes were increased from 5 to 100 MB, the number of message delivered increased substantially, thus, justifying the applicability of buffer capacity of node in computing its delivery probability. We also validated the applicability of change degree of connectivity in estimating the delivery probability of a node. The simulation were conducted by varying the speed of nodes thereby increasing the connections on each simulation run. The results presented in Figure 3 shows that the number of messages delivered



initially increased with the increase in number of connections but with further increase in number of connections, number of messages delivered reduced considerably. The experiment justified our decision of not including change degree of connections in estimating the delivery probability of a node. The applicability of one hop away and community structure was also studied and the results are shown in Figure 4. It was found that inclusion of both the attributes helps in making better estimate of delivery probability of nodes. We did not conduct any explicit experiments for validating the applicability of number of messages delivered and forwarded per unit time because the attributes were found to be effective in estimating the delivery probability of nodes in [13]. Also, an independent experiment of applicability of battery level was not conducted because the justification of its applicability is similar to buffer capacity.

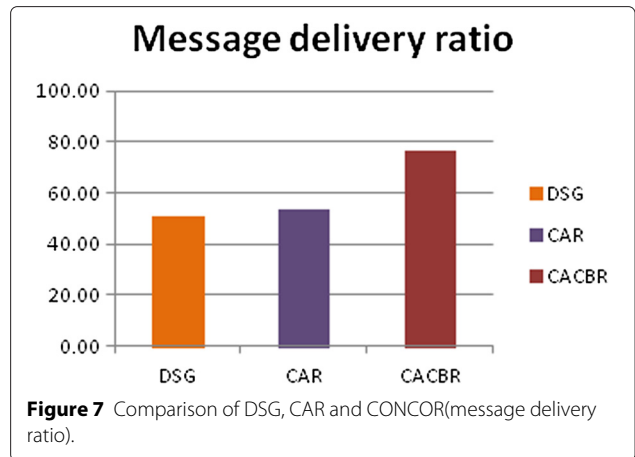
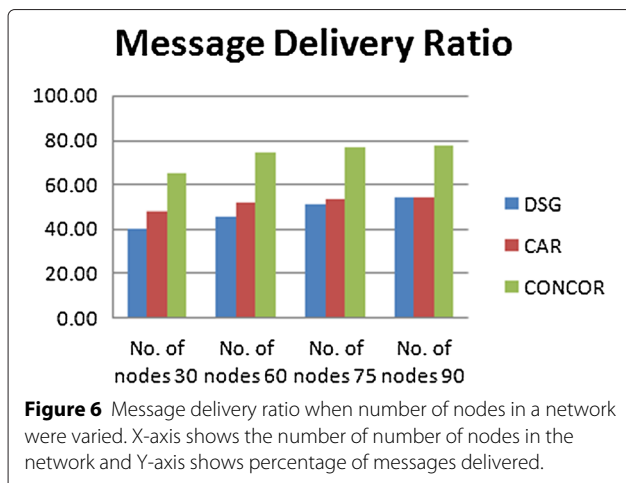
In the second experiment, we studied the impact of timer value. Figure 5 shows the message delivery ratio in CONCOR with different timer values. The value of timer,





controlling the update of values in DPVT, was varied from 30 to 120 s. It was observed that when the value of timer was kept at 30 s, the DPV value in DPVT was invalidated in 30 s only and therefore, a node does not forward a message to a highly capable node present in its neighbour as its DPV is invalid. This reduces the message delivery ratio. At higher values of timer, i.e. 90 and 120 s, the message delivery ratio is again reduced because the values in DPVT are old and not presenting the current network state. The CONCOR, for the given set of data and network parameters, performs best when timer value is 60 s.

In the third experiment, we studied the impact of number of nodes on network performance. We run DSG, CAR and CONCOR with the same network parameters but varying the number of nodes from 30 to 90. Figure 6 shows the message delivery ratio with different number of nodes. When the number of nodes is reduced, it is observed that the performance of CONCOR and DSG is more affected than CAR. This is because message delivery/forwarding in both CONCOR and DSG exploits the social contact of

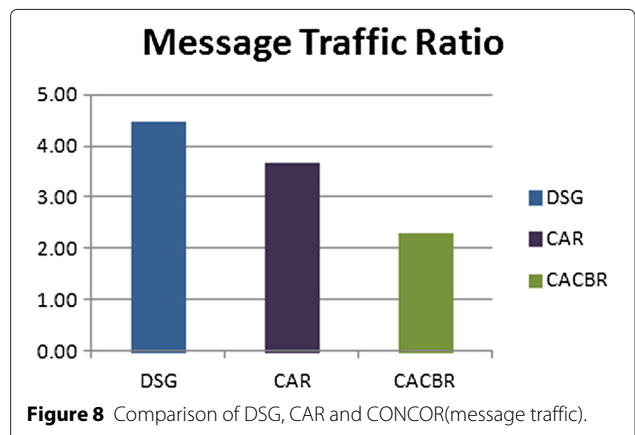


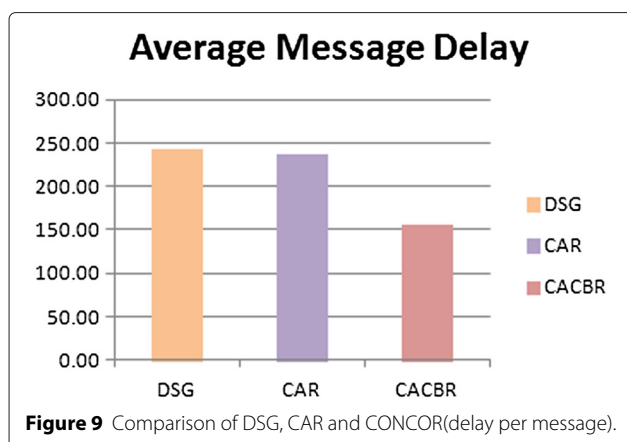
the node; therefore when number of nodes is reduced, the social contacts also reduced whereas when the number of nodes was increased the message delivery ratio increased. Whereas, in case of CAR, although there is no use of community- and social-based contact but with less number of nodes in a network, a node is getting fewer opportunities to forward the messages and therefore message delivery ratio for CAR also reduces.

In the fourth experiment, we run the simulation on DSG, CAR and CONCOR independently. A total of 5,000 messages were generated. The CONCOR algorithm outperformed both DSG and CAR algorithm. There was an increase of 51% and 43% (see Figure 7) in the number of messages delivered for CONCOR over DSG and CAR, respectively. There was an improvement of 29% and 11% (see Figure 8) in message traffic for CONCOR over DSG and CAR, respectively. Also an improvement of 57% and 53% (see Figure 9) in message delay for CONCOR over DSG and CAR, respectively.

### 7 Conclusions

The proposed algorithm exploits the existence of community in social networks. It also makes use of context





awareness of nodes to improve the estimate of message delivery probability of a node. The simulation results shows an improvement of message delivery ratio, message traffic ratio, and reduces delay significantly over both DSG and CAR. The work also highlights the effectiveness of node's attribute at estimating its message delivery probability. Further, as future work, an application of CONCOR may be widened to a network with  $N$  message source and sink. Also, the performance of the CONCOR will be studied through its deployment in the real-time scenario where a node's message delivery capabilities are likely to get affected by geographical terrain and environmental interferences.

#### Competing interests

The authors declare that they have no competing interests.

#### Author details

<sup>1</sup>USICT, Guru Gobind Singh Indraprastha University, Sector 16C, Dwarka, Delhi 110078, India. <sup>2</sup>Department of Computer Science, University of Delhi, Delhi 110007, India. <sup>3</sup>Institute of Informatics and Communication, University of Delhi, Delhi 110021, India.

Received: 30 March 2014 Accepted: 10 April 2015

Published online: 26 May 2015

#### References

1. S Burleigh, A Hooke, L Torgerson, K Fall, V Cerf, B Durst, K Scott, H Weiss, Delay-tolerant networking: an approach to interplanetary internet. *IEEE Commun. Mag.* **41**(6), 128–136 (2003)
2. S Jain, K Fall, R Patra, Routing in a delay tolerant network. *SIGCOMM Comput. Commun. Rev.*
3. A Vahdat, D Becker, et al., Epidemic routing for partially connected ad hoc networks. Technical report, Technical Report CS-200006, Duke University (2000)
4. A Lindgren, A Doria, O Schelén, Probabilistic routing in intermittently connected networks. *SIGMOBILE Mob. Comput. Commun. Rev.*
5. Y Wang, H Wu, Delay/fault-tolerant mobile sensor network (DFT-MSN): a new paradigm for pervasive information gathering. *IEEE Trans. Mobile Comput.* **6**(9), 1021–1034 (2007). doi:10.1109/TMC.2007.1006
6. EM Daly, M Haahr, in *Proceedings of the 8th ACM International Symposium on Mobile Ad Hoc Networking and Computing*. MobiHoc '07. Social network analysis for routing in disconnected delay-tolerant manets (ACM New York, 2007), pp. 32–40. doi:10.1145/1288107.1288113. http://doi.acm.org/10.1145/1288107.1288113
7. P Hui, J Crowcroft, E Yoneki, in *Proceedings of the 9th ACM International Symposium on Mobile Ad Hoc Networking and Computing*. MobiHoc '08. Bubble rap: social-based forwarding in delay tolerant networks (ACM New York, 2008), pp. 241–250. doi:10.1145/1374618.1374652. http://doi.acm.org/10.1145/1374618.1374652
8. P Costa, C Mascolo, M Musolesi, GP Picco, Socially-aware routing for publish-subscribe in delay-tolerant mobile ad hoc networks. *IEEE J. Selected Areas Commun.* **26**(5), 748–760 (2008)
9. CE Palazzi, A Bujari, Social-aware delay tolerant networking for mobile-to-mobile file sharing. *Int. J. Commun. Syst.* **25**(10), 1281–1299 (2012)
10. K Yang, X Cheng, L Hu, J Zhang, Mobile social networks: state-of-the-art and a new vision. *Int. J. Commun. Syst.* **25**(10), 1245–1259 (2012)
11. M Musolesi, C Mascolo, Car: context-aware adaptive routing for delay-tolerant mobile networks. *IEEE Trans. Mobile Comput.* **8**(2), 246–260 (2009)
12. U Shevade, HH Song, L Qiu, Y Zhang, in *Network Protocols, 2008. ICNP 2008. IEEE International Conference On*. Incentive-aware routing in DTNs (IEEE, 2008), pp. 238–247
13. R Cabaniss, S Madria, G Rush, A Trotta, SS Vullin, in *Proceedings of the 2010 Eleventh International Conference on Mobile Data Management*. MDM '10. Dynamic social grouping based routing in a mobile ad-hoc network (IEEE Computer Society Washington, DC, 2010), pp. 295–296. doi:10.1109/MDM.2010.94. http://dx.doi.org/10.1109/MDM.2010.94
14. P Hui, K Xu, VO Li, J Crowcroft, V Latora, P Lio, in *INFOCOM Workshops 2009, IEEE*. Selfishness, altruism and message spreading in mobile social networks (IEEE, 2009), pp. 1–6
15. A Keränen, J Ott, T Kärkkäinen, in *SIMUTools '09: Proceedings of the 2nd International Conference on Simulation Tools and Techniques*. The ONE simulator for DTN protocol evaluation (ICST New York, 2009)
16. T Spyropoulos, K Psounis, CS Raghavendra, in *Proceedings of the 2005 ACM SIGCOMM Workshop on Delay-tolerant Networking*. Spray and wait: an efficient routing scheme for intermittently connected mobile networks (ACM, 2005), pp. 252–259
17. L McNamara, C Mascolo, L Capra, in *Proceedings of the 14th ACM International Conference on Mobile Computing and Networking*. Media sharing based on colocation prediction in urban transport (MobiCom '08, ACM New York, 2008), pp. 58–69. doi:10.1145/1409944.1409953. http://doi.acm.org/10.1145/1409944.1409953
18. SC Nelson, M Bakht, R Kravets, AFIII Harris, Encounter: based routing in dtns. *SIGMOBILE Mob. Comput. Commun. Rev.*
19. Q Yuan, I Cardei, J Wu, in *Proceedings of the Tenth ACM International Symposium on Mobile Ad Hoc Networking and Computing*. MobiHoc '09. Predict and relay: an efficient routing in disruption-tolerant networks (ACM New York, 2009), pp. 95–104. doi:10.1145/1530748.1530762. http://doi.acm.org/10.1145/1530748.1530762
20. R Cabaniss, JM Bridges, A Wilson, S Madria, in *Wireless Communications and Networking Conference (WCNC), 2011 IEEE*. Dsg-m<sup>2</sup>: A group-based social routing algorithm (IEEE, 2011), pp. 504–509
21. F Ekman, A Keränen, J Karvo, J Ott, in *Proceedings of the 1st ACM SIGMOBILE Workshop on Mobility Models*. Working day movement model (ACM, 2008), pp. 33–40
22. KA Hummel, A Hess, Movement activity estimation and forwarding effects for opportunistic networking based on urban mobility traces. *Wireless Commun. Mobile Comput.* (2012)
23. R Johari, N Gupta, S Aneja, in *Quality, Reliability, Security and Robustness in Heterogeneous Networks*. Dsg-pc: Dynamic social grouping based routing for non-uniform buffer capacities in dtn supported with periodic carriers (Springer, 2013), pp. 1–15
24. G Sandulescu, P Schaffer, S Nadjm-Tehrani, Exploiting resource heterogeneity in delay-tolerant networks. *Wireless Commun. Mobile Comput.* (2012)
25. Y Li, Y Jiang, D Jin, L Su, L Zeng, DO Wu, Energy-efficient optimal opportunistic forwarding for delay-tolerant networks. *IEEE Trans. Vehicular Technol.* **59**(9), 4500–4512 (2010)
26. W Zhao, Y Chen, M Ammar, M Corner, B Levine, E Zegura, in *Mobile Adhoc and Sensor Systems (MASS), 2006 IEEE International Conference On*. Capacity enhancement using throwboxes in DTNs (IEEE, 2006), pp. 31–40
27. N Banerjee, MD Corner, BN Levine, in *INFOCOM 2007. 26th IEEE International Conference on Computer Communications*. An energy-efficient architecture for DTN throwboxes (IEEE, 2007), pp. 776–784

28. N Banerjee, MD Corner, BN Levine, Design and field experimentation of an energy-efficient architecture for dtn throwboxes. *IEEE/ACM Trans. Netw. (TON)*. **18**(2), 554–567 (2010)
29. EP Jones, L Li, JK Schmidtke, PA Ward, Practical routing in delay-tolerant networks. *IEEE Trans. Mobile Comput.* **6**(8), 943–959 (2007)
30. I Rhee, M Shin, S Hong, K Lee, S Kim, S Chong, CRAWDAD data set ncsu/mobilitymodels (v. 2009-07-23). Downloaded from <http://crawdad.cs.dartmouth.edu/ncsu/mobilitymodels> (2009)

**Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:**

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Immediate publication on acceptance
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

---

Submit your next manuscript at ▶ [springeropen.com](http://springeropen.com)

---