SAS: Seasonality Aware Social-Based Forwarder Selection in Delay Tolerant Networks

Amrita Bose Paul[1](✉), Akhil GV[2], Santosh Biswas[2], Sukumar Nandi[2], and Niladri Sett[2]

[1] Department of Computer Applications, Assam Engineering College, Guwahati 781013, Assam, India
amritabosepaul@gmail.com

[2] Department of Computer Science and Engineering, Indian Institute of Technology Guwahati, Guwahati 781039, Assam, India
{gv.akhil,santosh_biswas,sukumar,niladri}@iitg.ernet.in

Abstract. In social-based delay tolerant network (DTN) applications, hand-held mobile devices exchange information. The inherent social property of DTN has encouraged contemporary researchers in exploiting social metrics to devise forwarding techniques for efficient routing. This work observes evidence of seasonal behavior in contacts between node-pairs in real mobility traces, and exploits it to devise a novel seasonality aware similarity measure. We incorporate seasonality information into tie-strength, and then use it as link weight in a weighted similarity measure which we extend from Katz similarity index. We propose a Seasonality Aware Social-based (SAS) DTN forwarding technique based on the proposed similarity measure and ego-betweenness centrality. Finally we perform real trace driven simulations to show that SAS outperforms baseline social-based DTN forwarding methods significantly.

Keywords: Delay Tolerant Network · Social metrics · Ego-betweenness · Centrality · Social-based forwarding · Mobile Social Networks · Tie-strength

1 Introduction

Intermittently connected Mobile Ad hoc Networks (MANETs) lack contemporaneous end-to-end routes from source to destination. Message delivery in these networks must be delay tolerant, and so these networks are often called as Delay Tolerant Networks (DTNs). DTN was originally developed for Inter Planetary Networks (IPNs), but later its applications have been realized in terrestrial mobile networks such as Mobile Social Networks (MSNs) [1], Pocket Switched Networks (PSNs) [2], Vehicular Ad hoc Networks (VANETs) [3], which are characterized by sporadic connectivity, frequent link disturbance, existence of non-contemporaneous end-to-end route, long and unpredictable communication

S. S. Rautaray et al. (Eds.): I4CS 2020, CCIS 1139, pp. 245–265, 2020.
https://doi.org/10.1007/978-3-030-37484-6_14

latency, etc. To deal with sporadic connectivity pattern, DTNs follow a message propagation scheme referred as *store-carry-and-forward* [4], where intermediate nodes (known as carriers) store and physically carry buffered messages until they get in contact with the destination or a suitable next-hop carrier. In this scheme, each node independently makes forwarding decisions for opportunistic message exchange between them when they are in communication range of each other. In most of the terrestrial DTN applications, the mobile nodes/devices are carried and used by people and thereby making forwarding decision based on peoples' social behavioral perspectives. So, a class of DTN forwarding, namely *social-based DTN forwarding* algorithms [5] have emerged, which exploit social network properties in DTN forwarding. Our work in this paper proposes a **S***easonality* **A***ware* **S***ocial-based (SAS)* DTN forwarding mechanism, which capitalizes on seasonal behavior in human contacts.

Popular social-based DTN forwarding techniques [5] usually exploit three social network metrics: similarity between node-pairs [6], centrality of a node [7], and community of nodes [8]. Intuition behind use of these three metrics are: (i) similar nodes meet each other frequently, so a node similar to the destination node has better delivery probability of the message; (ii) central nodes act as hub, and are reachable to other nodes; and (iii) nodes inside a community meet frequently, so forwarding the message to a node that resides within the destination's community increases the chances of message delivery. SimBet [9] is a social-based DTN forwarding technique which has utilized similarity and centrality metric, whereas BubbleRap [10] has exploited centrality metric and community structure. Lack of infrastructure in DTN forces individual nodes to take forwarding decisions independently through message exchange. Unavailability of a centralized view of the network limits the social-based DTN forwarding techniques to use only locally calculable social network metrics. However, advanced social network metrics, such as random walk similarity measure [6], betweenness centrality measure [7], community detection algorithms [8] are global in nature, and can not be directly applied to DTN forwarding. So, approximated versions of the global metrics have been devised for forwarding in DTNs. The authors in SimBet [9] have used an approximated version of betweenness centrality called ego-betweenness centrality [11], that calculates the betweenness centrality of each node in their respective ego networks. BubbleRap [10]'s approximation of betweenness centrality has been a modified version of degree centrality and has used a distributed community detection algorithm for DTNs [12]. Further, it has been observed that, SimBet and BubbleRap dynamically calculate the social relationship between the nodes to choose the best relay node. SimBet models the relationship between the nodes as binary and does not consider the relative strength of its neighbors. To justify the reason, the authors of SimBet argue that ego betweenness has high correlation with sociocentric betweenness. However, by analyzing the different mobility traces [13,14] we found that the correlation of ego betweenness and social betweenness is not that high but correlation of ego betweenness and sociocentric betweenness of a node inside a community has very high correlation as shown in Table 1. Moreover, in their work, the small world created by the network of mobility traces also have very less diameter (<2)). Again, BubbleRap uses the concept of sociocentric betweenness centrality which requires the knowledge of the whole network, which in reality is not possible in DTN. Therefore, these issues of the existing state-of-the-art routing protocols of social-based DTNs motivate us to observe evidences of seasonal behavior in node contacts in real mobility traces and exploit it to devise a novel seasonality aware similarity measure.

Table 1. Characteristics of the mobility traces

Trace		Reality	Sassy	Cambridge
#Nodes		96	25	36
#Edges		3085	155	541
Average degree		64	12	30
Average clustering co-efficient		.816	.712	.892
Average shortest path length		1.324	1.503	1.141
Co-relation of sociocentric betweenness and ego betweenness	Whole network	.75	.88	.608
	Within community	.984	.987	.990

In our work, we model the contact history between node-pairs to formulate tie-strength which preserves seasonality of human contacts. Traditional approaches to model tie-strength [15–18] use variants of average separation duration between node-pairs. We observe strong seasonality, i.e., repetitive contact pattern in real mobility traces and exploit it to formulate tie-strength. Our model measures the tie-strength as weighted average of separation duration and a seasonality aware contact strength. Based on Katz [19] similarity index we define a weighted similarity index between two nodes. Our motivation of using Katz similarity metric has been its inherent property of giving more importance to the direct contacts over the indirect ones. By analyzing real mobility traces, we find that although ego-betweenness centrality is not a good substitute for sociocentric/global betweenness, but it can accurately approximate global betweenness within communities. Our proposed DTN forwarding technique SAS exploits the proposed weighted Katz based seasonality aware similarity measure and ego-betweenness centrality, where the similarity value effectively deals with intra-cluster forwarding and ego-betweenness drives the inter-cluster forwarding. We adapt the utility proposed in SimBet, which exploits similarity and centrality, and propose the Seasonality aware DTN forwarding algorithm SAS. Finally, we simulate our work on real mobility traces to demonstrate the effectiveness of SAS over state-of-the-art social-based DTN forwarding algorithms: SimBet and BubbleRap.

The rest of the paper has been structured as follows. Section 2 discusses related works on social-based forwarding in DTNs. The drawbacks associated with these works are listed out, which provide the motivation for the work carried out in this paper. Our proposed seasonality aware social-based forwarding in DTNs has been presented in Sect. 3. Section 4 presents the performance evaluation and analysis of the proposed social-based forwarding scheme with the benchmark SimBet [9] and BubbleRap [10] to validate its effectiveness in attaining routing objectives. Finally, in Sect. 5 we conclude our work.

2 Background and Literature Review

This section introduces the different approaches of routing techniques available in the literature of DTNs with a special focus on the social-based forwarding techniques.

The routing protocols in DTNs can be broadly classified into two categories: *flooding* and *forwarding* [20]. The protocols in the flooding family induce multiple "replicas" of each message in the network without considering the potentiality of the candidate node for being selected as a next-hop carrier [21–24]. In this routing approach, a source node tries to send all its' messages to its' neighbors if they do not have the copy of the messages. This approach does not require to store any past information about the routing or mobility of the nodes. So, flooding is the obvious choice when no information is known in advance about the movement of the nodes or about the topology of the network.

In [21], the authors have proposed "Epidemic routing" as one of the basic flooding based routing protocol in DTNs. In Epidemic, a node floods the messages to it's neighbor nodes who does not have a copy of the message. In this protocol, whenever two nodes have an encounter, they exchange their summary vector which contains the IDs of the messages they are carrying. After comparing the summary vector, each node determines the messages they are not carrying which the other nodes have and requests for those messages. Depending on this request message transfer is done between the nodes. Random pairwise exchange of messages are used to ensure eventual message delivery. This process of continuous replication flood the network with same copy of messages to guarantee maximum delivery ratio in presence of infinite storage availability for all the nodes in the network. However in reality, nodes have limited storage capacity, and a limited number of messages can be stored. Flooding the network with messages causes high overhead in term of storage and power spent on transmission and reception. This causes the degradation of network performances. In an another approach called "Two-Hop Forwarding" [25], each node is assumed to encounter every other node for some short duration of time. Within this duration, the source node replicates each message to the first encountered node and the messages are stored until they come in contact with the destination. In this protocol, routing overhead is reduced at the cost of increased message delivery latency. In addition, "Spray and Wait" [24] is a controlled flooding based routing protocol that requires no knowledge about the network. Unlike epidemic it limits the number of message copies to be forwarded in the network. The protocol works in two phases (i) *spray* and (ii) *wait*. In spray phase the source spreads $\mathcal{M}$

copies of the messages in the network. If the destination is not found in spray phase, then the relay nodes having message copies will enter into a wait phase in which they wait until the messages are delivered to the destinations directly. Relay nodes do not make any additional copies of the message, in turn reducing the resource usage.

Though, these protocols in the flooding family achieve good delivery ratio and less delivery latency, but flooding the network with duplicate messages cause high network overhead in term of storage and power spent on transmission and reception. These cause congestion leading to network performance degradation. So, another class of routing approaches called "forwarding-based" have been explored to restrict the generation of bundle replicas in the network.

The protocols in the forwarding family calculate an utility metric based on "knowledge" to qualify the candidate node as the next hop carrier on the routing path. A single copy of each message is forwarded to the qualified node. Most of these knowledge-based protocols select a suitable next-hop carrier based on contact history of potential carriers [26,27], knowledge about traffic patterns in the network [28] or on probability of encountering the destination node [29]. Furthermore, some of them have used multi-copy spraying mechanisms to improve reliability amidst intermittent connectivity [30,31].

In the basic forwarding based protocol called "First Contact" (FC) [23], the source node tries to forward the message to one of the randomly selected link among all the current contacts. The authors have tried to improve the performance of the protocol by forwarding the message in a direction closer to the intended destination node. To avoid the routing loop, a path vector has been proposed. In this scheme a single copy of each message is maintained in the network.

In an another approach, the Probabilistic Routing Protocol using History of Encounters and Transitivity (PRoPHET) [22] uses utility based replication for delivery of messages. PRoPHET uses history of encounter information to calculate the utility metric of a node. In this protocol, each source node calculates its delivery probability to every other node in the network. These probability values are updated on every contact for each known destination. The delivery probability is aged by a factor over time. It also uses, transitive relation to update the delivery predictability of a node, with whom it is not directly connected. In Rapid [32], a node calculates the utility value of each message that is present in its buffer and this utility value decides in which order it should be relayed to the next node. RAPID derives a per-packet utility function from the routing metric. At a transfer opportunity, it replicates a packet that locally results in the highest increase in utility. To calculate this utility value, it first estimates the delivery delay of the message. This estimation is based on the two or three hop's information. This limits the estimation because the destination may be present beyond two or three hops.

Recently, social-based routing is relatively a new approach and has become popular for addressing the routing problem in DTNs. It is based on the observation that in most of the terrestrial DTN applications people are carrying mobile devices (like Pocket Switched Networks, Mobile Social Networks etc.) and thereby making forwarding decision based on peoples' social behavioral perspectives. In social-based DTN applications, hand-held mobile devices exchange information. The inherent social property of DTN has encouraged contemporary researchers in exploiting social metrics to devise forwarding techniques for efficient routing. So, a class of DTN forwarding, namely *social-based DTN forwarding* algorithms [5] have emerged, which exploits social network properties in DTN forwarding. Social-based DTN forwarding has been popular in DTN specific applications like vehicular networks, mobile social networks, pocket switched networks etc. In such application domains, people carry mobile devices, whose behaviors are unpredictable from social aspects as well as from ad hoc networking aspects. Zhu et al. [5] and Wei et al. [33] have provided two recent surveys on social-based DTN forwarding techniques. "Centrality", "Similarity" and "Community" have been the most effective social network metrics used for DTN forwarding.

Authors in [34–36] explored the usefulness of community detection algorithms in DTN forwarding. The motivation of using communities has been: if the carrier encounters a node which belongs to the destination's community, the message will be delivered with high probability. The authors in [34–36] explored the possibility of community detection and interest profile based forwarding algorithms in DTNs. In these approaches, messages are forwarded to the encountered node if it belongs to the same community as the destination node or if it's interest profile matches with the destination node's interest profile. The shortcoming of these approaches is that they do not capture the dynamics of social relations among the nodes.

In an another approach, SimBet [9] has exploited ego-betweenness centrality and similarity to forward messages in DTN. Central nodes work as hubs and are reachable to all other nodes in the network, and nodes similar to the destination contacts with it frequently. However, the shortcoming of SimBet is that, the authors model the relationship between the nodes as binary and does not consider the relative strength of its neighbors.

Again, BubbleRap [10]'s approximation of betweenness centrality has been a modified version of degree centrality and it has used a distributed community detection algorithm for DTNs [12]. The proposed betweenness centrality of BubbleRap requires the knowledge of the whole network, which in reality is not possible in DTN.

Another set of social-based forwarding techniques have exploited the concept of tie-strength [37]. Few of these can be found in [15–18]. These techniques have modeled the change in contact patterns during time, and predicted strength of social relationships between node-pairs. The authors in [15–18] have failed to model the dynamic changes in contacts from human behavioral perspectives.

Therefore, these issues of the existing state-of-the-art routing protocols of social-based DTNs motivate us to observe evidences of seasonal behavior in node contacts in real mobility traces and exploit them to devise a novel seasonality aware similarity measure. Our work has incorporated seasonality behavior of human contacts into tie-strength towards DTN forwarding. To the best of our knowledge, our work is the first one to exploit seasonality behavior of human contacts in DTN forwarding.

3 Proposed Seasonality-Aware Forwarding Scheme

Here we present our **S**easonality **A**ware **S**ocial Based DTN Forwarding (SAS), a DTN forwarding algorithm which exploits seasonal behavior of human contacts. Our proposed measures of seasonality aware "tie-strength" is detailed in Sect. 3.1. The modified version of "similarity", and "centrality" measure with incorporation of seasonal behavior of node contacts are detailed in Sects. 3.2 and 3.3, respectively. The newly formed "utility" function to determine the node's potentiality as a next-hop forwarder in DTN routing is presented in Sects. 3.4. Finally Sect. 3.5 represents the proposed seasonality aware forwarding algorithm in social-based DTNs.

We consider a category of DTN like Pocket Switched Networks [2] or Mobile Social Networks [1], which consists of cellular devices carried by human beings. They use Bluetooth interface to exchange data among themselves. Each device can act as a source, destination, or forwarder of a message. Due to mobility of these devices, a continuous source-to-destination path may not exist. These devices communicate in *opportunistic* manner during contacts, when a sender and a receiver comes into contact at a time which is unknown beforehand. During this contact, these devices make the forwarding decision of data. In DTN the network topology changes rapidly and the nodes do not have any knowledge of future connections. The inter-node contact duration is often limited. During this duration only a limited number of messages can be transferred. Also, DTN uses multihop forwarding for messages. A large number of hops increases the probability of delivery of message, but also increases the delivery cost. So it is needed to have an efficient strategy to select the best relay nodes.

Use of social network metrics have been prominent [5] in DTNs where the network is formed with hand held mobile devices carried by humans. The reason for this is that mobility in such networks is driven by the social network properties, which are less volatile than the traditional metrics. In this work, we exploit the seasonality/repetitive pattern in human contacts and have incorporated it with the other state-of-the-art social network metrics towards proposing a **S**easonality **A**ware **S**ocial-based forwarding called SAS. Similar to SimBet [9], we select the best relay node based on a utility metric which exploits two social network metrics: centrality and similarity. We model the seasonality in human contact to propose a novel formulation for calculating tie-strength, and incorporate it as link weight into the proposed weighted similarity metric based on Katz similarity index [19]. By analyzing real mobility traces, we find that ego-betweenness [11] can be a good approximation for sociocentric betweenness [7] inside communities. We combine the proposed seasonality aware similarity and ego-betweenness in a utility function and propose the forwarding mechanism SAS.

3.1 Strength of Tie

Strength of tie [37] measures the strength of social relationship between two individuals. A simple way to measure the tie-strength may be the total number of contact or the total duration of contact. Motivation of using tie-strength in DTN forwarding has been: if a node carrying the message gets into contact with a node which is strongly connected to the destination (i.e., has met with the destination many times or for long time in the past), they may meet again and may deliver the message with high probability. Tie-strength can be regarded as a similarity measure for directly connected node-pairs. However, the destination may not be directly connected to every node the carrier meet, so multi-hop similarity measure is required. We discuss the multi-hop similarity in the next subsection.

Traditional approaches to model node-pair's tie-strength use variants of average separation duration [15–18]. In general, the average separation duration between two nodes x and y during the time interval $[0, T]$ is given as:

$$S_{avg}^{[0,T]}(x, y) = \frac{\int_{t=0}^{T} f(t)dt}{T} \tag{1}$$

where $f(t)$ represents the estimated time remaining for the next encounter between the nodes x and y at time t. Strength of tie is usually formulated as a function which is inversely proportional to the average separation duration. This approach assumes that the node-pairs which have come into contact for longer duration in past are tied with stronger relationship, and are likely to come into contact in future.

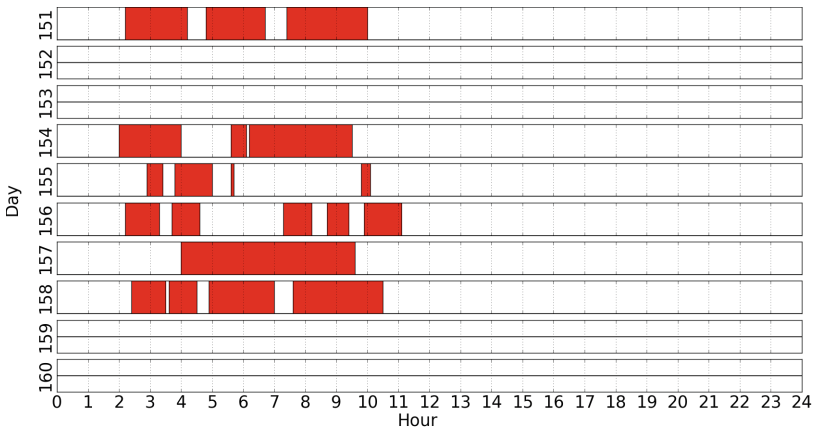

Fig. 1. Seasonality pattern in `Reality` trace

Our proposed measure of tie-strength has been encouraged by observed seasonality pattern in node-pairs' contact history. Figure 1 shows the contact pattern of two nodes in the `Reality` trace during 10 days. Each rectangle in the vertical dimension represents a day, and each day is divided into 24 parts which represent hours. The duration of a day filled with red is the contact duration between the two nodes. The figure shows that the six days which have some contacts, follow similar contact pattern. The bursts of contacts happen during the same 9 h period of these days. It might be explained as, the two persons workplace may be same and this nine hour duration might be their working hours. The days which observe no contact may be holidays, which repeat in every seven days. It is also observable that the contact pattern of the day at the top of the figure is very similar to the eighth day from top. It indicates that the contact pattern repeats every week. We exploit this seasonality pattern of human contact to measure of link strength, which is a weighted average of the traditional average separation duration and seasonality aware tie-strength. We describe below how each node calculates their tie-strength with its' neighbors. For simplicity and as per the requirement for the mobility traces in hand, we explain this method with two granularities of seasonality, daily and weekly. However, this method is trivially extendable for more levels of seasonality, such as monthly, quarterly, yearly, etc.

We divide the duration of a day into equal size time-window Δ_0, say an hour. $\Delta_1 = a \times \Delta_0 =$ duration of a day, $\Delta_2 = b \times \Delta_1 = a \times b \times \Delta_0 =$ duration of a week. Note that, in our case $a = 24$ and $b = 7$, but for maintaining generality we use the variables a and b. $f(t)$ represents the time remaining to the next encounter between two nodes (x, y) at time t. Each node x maintains a seasonality matrix $m^{(x,y)}$ for each of its contacts y, $m^{(x,y)}[i, j]$ be the elements of the seasonality matrix $m^{(x,y)}$. The dimension of $m^{(x,y)}$ is $b \times a$. When two nodes x and y come to contact for the first time, both of the nodes initialize $m^{(x,y)}$, and all of its elements are initialized as 0. The nodes keep a variable p which stores the total number of time-windows elapsed, and is initialized as 0. They also keep the variables q and s, initialized as 0, which keep track of the offset of the current time window in the seasonality matrix for row and column, respectively. The variable $\mathbb{T}$ representing non-seasonal strength of the link (x, y) is initialized as 0. After each Δ_0 amount of time, both of the nodes trigger the following steps, which update an matrix element, the variables, and calculate the tie-strength of (x, y) for the next time window.

$$m^{(x,y)}[q, s] = \frac{m^{(x,y)}[q, s] + \Delta_0 / \int_{t=p\times\Delta_0}^{p\times\Delta_0+\Delta_0} f(t)dt}{p \times \Delta_0 + \Delta_0} \tag{2}$$

$$\mathbb{T} = \frac{p \times \Delta_0 + \Delta_0}{\mathbb{T} + \int_{t=p\times\Delta_0}^{p\times\Delta_0+\Delta_0} f(t)dt} \tag{3}$$

where p is incremented as $p = p+1$ and the offsets of the next time window in the seasonality matrix for row and column are updated as $q = (p - p \mod (a \times b)) \mod b$ and $s = p \mod b$, respectively.

Finally, the seasonality aware tie-strength of the link (x, y) for the next time-window is calculated as the weighted average of average separation duration and seasonality components:

$$w^p(x,y) = \alpha \times m^{(x,y)}[q,s] + (1-\alpha) \times \frac{1}{\mathbb{T}} \tag{4}$$

where the parameter $0 \le \alpha \le 1$ regulates the weight of the seasonality aware component in the tie-strength formulation.

3.2 Similarity

The motivation of using similarity measure in DTN forwarding is that similar nodes meet frequently, and a node similar to the destination node is highly likely to deliver the message to the destination node. In SimBet, similarity between two nodes is calculated as the number of common neighbors between them. It treats direct and indirect contacts in a similar manner. We argue that the nodes which have met the destination at past, are more similar to the destination than those which are two hop away. We adopt Katz index [19] to define the similarity metric. Katz similarity index between the nodes x and y is given as:

$$Katz(x,y) := \sum_{l=1}^{\infty} \beta^l \times |paths_{x,y}^{<l>}|, \tag{5}$$

where $paths_{x,y}^{<l>}$ represents the set of all paths of length l between nodes x and y. $\beta > 0$ is a constant that regulates the amount of importance given to higher length paths. As $\beta \to 0$, Katz index starts behaving like common neighbor.

We modify Katz index to accommodate tie-strength. We consider upto 2 length paths to make it locally calculable. The Similarity measure between nodes x and y is given as:

$$Sim(x,y) = \beta \times w(x,y) + \beta^2 \times \sum_{k \in N(x) \cap N(y)} w(x,k) + w(k,y), \tag{6}$$

where $N(x)$ is the set of neighbors of a node x, and $w(x,y)$ is the weight of a link (x,y) for the current time window, given by Eq. (4).

3.3 Centrality

Centrality measures the importance/accessibility of a node in the network. Central nodes are considered as highly reachable to the other nodes in the network. Betweenness [7] is one of the widely used centrality measure used in social-based DTN forwarding techniques [5]. Nodes with high betweenness centrality fall into large number of shortest paths linking to other node-pairs in the network. Thus, these nodes act as bridges to reach to all other nodes in the networks. Betweenness Centrality is calculated as:

$$Bet_C(x) = \sum_{y \neq z \neq x, (y,z) \in \text{Nodes}} \frac{g_{y,z}(x)}{g_{y,z}} \tag{7}$$

where $Bet_C(x)$ is the global/socio centric betweenness centrality of node x, $g_{y,z}$ is the total number of geodesics (shortest paths) between nodes y and z, and $g_{y,z}(x)$ is the number of shortest paths between node y and z passing through x.

Socio centric betweenness is a global measure, and is difficult to measure in DTN forwarding because the nodes in DTN have access to the local information only. Flooding may be one solution, but it will increase the message cost exponentially. Moreover, due to sparse and dynamic nature of DTN, message may take long to reach the destination. Consequently, in DTN it is impossible to achieve consistent values of the global measures like socio centric betweenness throughout the network. SimBet [9] has capitalized the concept of Ego networks [11] in DTN forwarding, which approximates socio centric betweenness by calculating betweenness centrality locally, within the node's ego network. Ego network of a node is defined as a network which consists of the node, its neighbors, the links of the node with its neighbor, and the connections among its neighbors. The ego-betweenness of a node x is calculated as:

$$Bet_E(x) = \sum_{y \neq z \neq x, (y,z) \in N(x)} \frac{g_{y,z}(x)}{g_{y,z}} \tag{8}$$

where $Bet_E(x)$ is the ego-betweenness centrality of x, $g_{y,z}$ is the total number of geodesics (shortest paths) between nodes y and z, and $g_{y,z}(x)$ is the number of shortest paths between node y and z passing through x. $N(x)$ is the set of neighbors of x.

Marsden [38] has observed that ego-betweenness and socio centric betweenness are highly correlated in social networks. We investigate the relationship between socio centric and ego-betweenness in the real mobility traces discussed in Sect. 1. Table 1 shows that correlation between ego-betweenness and socio centric betweenness in the whole network is insignificant. However, when the network is partitioned into communities, ego-betweenness and socio centric betweenness correlate highly. So, we argue that a node with high ego-betweenness acts as a good hub inside its community, and can be useful in forwarding the message when the destination is inside its community.

3.4 Utility

A carrier having a message must choose another node to forward it, so that the message reaches the destination with high probability. When a carrier comes into contact with a node, it calculates an utility function of the node with respect to the destination. The carrier forwards the message to the node based on this utility function. Like SimBet [9], we define the utility as a combination of two utilities: similarity and centrality.

Utility of a node y (which comes into contact with the carrier x) for delivering a message to node d is calculated as:

$$Utility(y, d) = \gamma \times SimUtility(y, d) + (1 - \gamma) \times BCUtility(y) \tag{9}$$

where,

- $SimUtility(y,d) = \frac{Sim(y,d)}{Sim(x,d)+Sim(y,d)}$ is the similarity utility of the node y with the destination d with respect to the career x,
- $BCUtility(y) = \frac{Bet_E(y)}{Bet_E(x)+Bet_E(y)}$ is the betweenness utility of the node y with the destination d with respect to the career x,
- $\gamma \in [0,1]$ is a balancing parameter, which allows for setting the relative importance of Betweenness utility and Similarity utility,
- $Sim(-,-)$ and $Bet_E(-)$ are calculated using Eqs. (6) and (8) respectively.

3.5 Forwarding Algorithm

Here we present our proposed forwarding algorithm based on ego-betweenness centrality and seasonality aware similarity index, which extends the forwarding algorithm of SimBet [9]. It evaluates a nodes' utility for being chosen as a potential forwarder. This algorithm makes no pre-assumption of global knowledge of the network, and makes the forwarding decisions on the fly based on locally exchanged information. For this to happen, on encountering a node y, node x verifies whether it is carrying any messages destined to y. If this is found to be true, then all messages destined for y are delivered. Subsequently, the encounter vectors are received from node y. The encounter vector contains information (list of contacts and tie-strength of the links with their contacts) about the nodes that each of them have encountered. This encounter information is then used to update the ego-betweenness value on node x and similarity value as described in Eqs. (8) and (6) respectively. Further, the two nodes x and y exchange a summery vector that contains a list of destination nodes for whom they are carrying messages, and their betweenness and similarity values. Thereafter, node x calculates the Utility value of its own and of node y for each destination in the received summery vector following Eq. (9). If the node y's utility is higher than x's, x forwards the message to y in greedy fashion. We summarize the algorithm as follows.

1. On encountering y, if node x has messages destined for y, it delivers them to y.
2. x receives the encounter vector of node y, which contains y's contacts and $w^p(y,k)$'s where $k \in N(y)$.
3. Node x and y exchange the summery vector information containing list of messages carried by them for each destination node.
4. For each message in the Message list calculate $Utility(x,d)$ and $Utility(y,d)$ for each destination d.
5. If $Utility(y,d) > Utility(x,d)$, node y becomes the forwarder and receives messages from x.

4 Performance Evaluation of SAS

This section detail the performance evaluation and analysis of the proposed social-based forwarding scheme (SAS) with the benchmark SimBet [9] and

BubbleRap [10] to validate its' effectiveness in attaining routing objectives. The different evaluation metrics under consideration are described in Sect. 4.1. Section 4.2 provides a brief description of the data traces used in the experiments and summarizes characteristics of the social network induced by the contacts in the mobility traces. The experimental setup used for generation of mobility traces through trace-driven test with dataset from the `Reality` [39] and `Cambridge` [40] datasets are represented in Sect. 4.3. Finally, the experimental results and their analysis are summarized in Sect. 4.4.

4.1 Routing Objective and Evaluation Metrics

Routing Objective of DTN routing protocol depends on application. Generally the objective is to increase the delivery ratio while not increasing the cost of delivery much. Generally, DTN routing protocols are evaluated based on the following metrics, which we follow in this work:

- **Delivery Ratio:** It is the ratio between the number of messages delivered and the total number of messages generated.
- **Delivery Cost:** It is the ratio between the number of message transmission required for delivery to the total number of messages delivered.
- **Average Latency:** It is the time duration between the message generation and its delivery, averaged over all messages.

4.2 Data Sets

We perform our experiments on three real mobility traces, namely `Cambridge`, `Reality` and `Sassy`. A brief description of the three traces are given next. The characteristics of the social network induced by the contacts in the mobility traces are already summarized and discussed in Table 1 of Sect. 1.

- `Cambridge`: This dataset [13] includes the traces of Bluetooth sightings by groups of users carrying iMotes for 11 number of days. The iMotes devices were distributed among the doctoral students and faculty comprising a research group at the University of Cambridge Computer Laboratory.
- `Reality`: The MIT's Reality Mining experiment [14] conducted in 2004 was aimed at studying community dynamics. The study consist of one hundred Nokia 6600 smart phones having Bluetooth network connectivity and were distributed among the students and staff at MIT. The study generated data, collected by these 100 human carried devices over the course of nine months, include call logs, Bluetooth devices in proximity (i.e. contact logs), cell tower IDs, application usage, and phone status. The study resulted in the first mobile data set with rich personal behavior and interpersonal interactions.
- `Sassy`: This dataset [41] is an outcome of the experiments carried out by a group of participants (22 undergraduate students, 3 postgraduate students, and 2 members of staff) forming a mobile sensor network at University of St Andrews. The experimental set up was made of 27 T-mote invent devices

(mobile IEEE 802.15.4 sensors) carried by human users and Linux-based base stations for bridging the 802.15.4 sensors to the wired network. The participants were asked to carry the devices whenever possible over a period of 79 days. The data set contains information about the participants' encounter records as well as their social network data generated from Facebook data.

4.3 Experiment Setup

We have used Opportunistic Networking Environment (ONE) simulator [42] for simulation purpose. It is specifically designed for evaluation of DTN routing and application protocols. We have evaluated our simulation through trace-driven test with dataset from the `Reality` [39] and `Cambridge` [40] datasets, described in Sect. 4.2. `Reality` dataset spans for about six months. During the simulations for reality datasets 1000 messages were generated during 5–6 month period by randomly choosing the source and destination nodes. `Cambridge` dataset spans for about 11 days. During the simulations with `Cambridge` dataset, 1000 messages were generated after 9–11 day period by randomly choosing the source and destination nodes. Each simulation is repeated 10 times with different random seeds, and the average evaluation results are reported. The parameters for the simulations for the datasets are summarized in Table 2.

Table 2. Parameters for simulation setup

Dataset	`Reality`	`Cambridge`
Number nodes	97	36
Transmission range	10 m	10 m
Transmission speed	250 kBps	250 kBps
Message size	10–100 kb	10–100 kb
Time To Live (TTL)	1–12 days	2 min–24 h

4.4 Results and Discussion

We compare the performance of the proposed forwarding algorithm SAS with the state-of-the-art social-based DTN forwarding algorithms: SimBet [9] and BubbleRap [10]. We vary the parameter α to tune the effect of seasonality in SAS. We set β, the parameter of the Katz similarity measure to a typical value .05 [6].

Figures 2, 3, 4, 5, 6 and 7 summarize the comparative performance of SAS with BubbleRap and SimBet for the three evaluation metrics viz., "delivery ratio", "delivery cost" and "average latency". Further, to evaluate the effects of the seasonality component on the performance of SAS, we set three different values for the parameter α (i.e., $\alpha = 0$, $\alpha = 1$, $\alpha = 0.3$) and have obtained

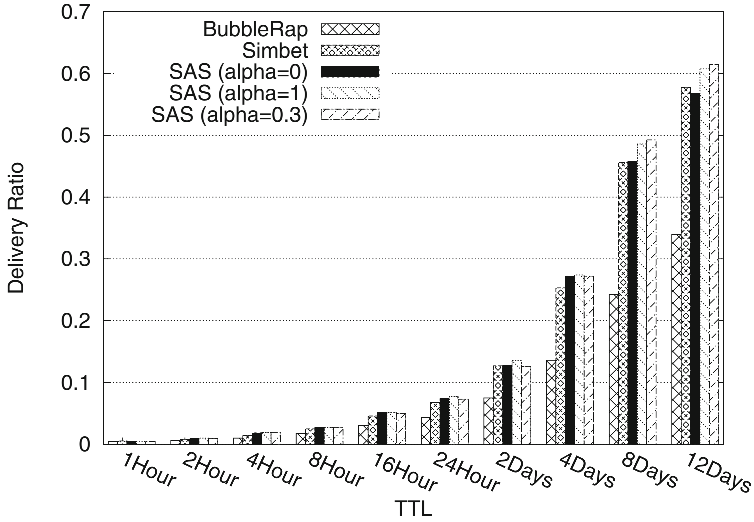

Fig. 2. Message delivery ratio Vs TTL in Reality data set

the simulation results for three different performing versions of SAS. Results for SAS (alpha = 0) measure the performance of SAS when the tie-strength does not contain the seasonality component, SAS (alpha = 1) represents SAS when the tie-strength is calculated with a high weighted value set for the seasonality component, and SAS (alpha = 0.3) represents SAS where the tie-strength is calculated with a low weighted seasonality component. Here we detail the results of these three performing versions of SAS with varying TTL values (as represented in Table 2). We also varied the utility parameter γ for SAS and SimBet, but found that $\gamma = 0.5$ gives best performance in general.

Figures 2 and 3 show that all the three different versions of SAS (i.e., SAS (alpha = 0), SAS (alpha = 1), SAS (alpha = 0.3)) outperform SimBet and BubbleRap significantly with respect to delivery ratio over the two traces (i.e., "reality" and "cambridge") and the TTL values. SAS(alpha = 0.3) outperforms SimBet by 6.50% and BubbleRap by 81.10% for TTL=12 Days in `Reality` trace. SAS(alpha = 0.3) outperforms SimBet by 5.41% and BubbleRap by 21.70% for TTL=12 Days in `Cambridge` trace. SAS (alpha = 1) always outperforms SAS (alpha = 0), which indicates the usefulness of the seasonality component of tie-strength. Again, it is also notable from Figs. 4 and 5 that all the performing versions of SAS (i.e., SAS (alpha = 0), SAS (alpha = 1), SAS (alpha = 0.3)) do not incur much delivery cost as compared to SimBet to achieve the gain in delivery ratio in `Reality`, and achieves better delivery cost than SimBet in `Cambridge`. From Figs. 6 and 7, it has been observed that for all of the forwarding techniques under consideration in the simulation study (except BubbleRap), average latency values are almost same for all the protocols.

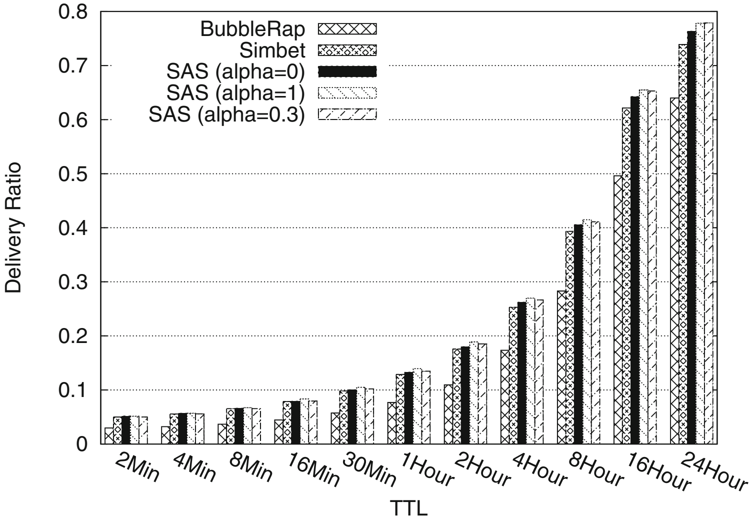

Fig. 3. Message delivery ratio Vs TTL in Cambridge data set

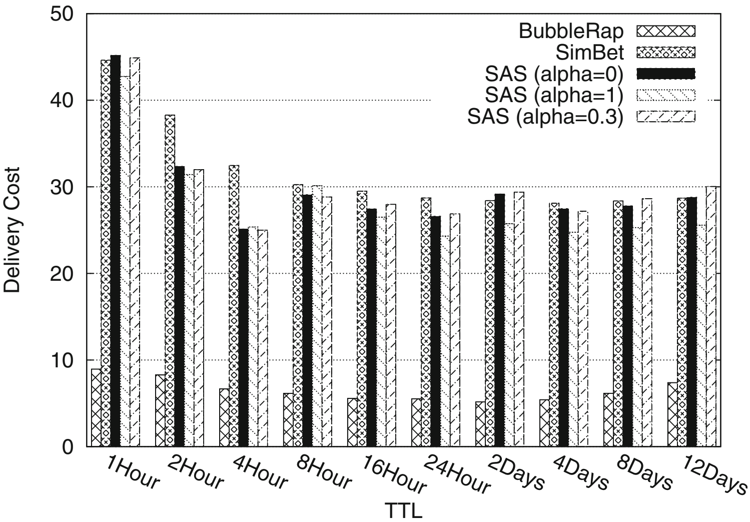

Fig. 4. Message overhead ratio Vs TTL in Reality data set

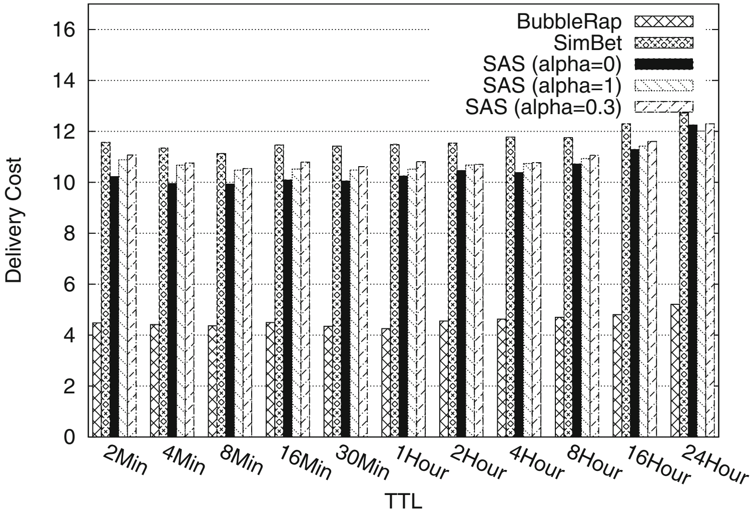

Fig. 5. Message overhead ratio Vs TTL in Cambridge data set

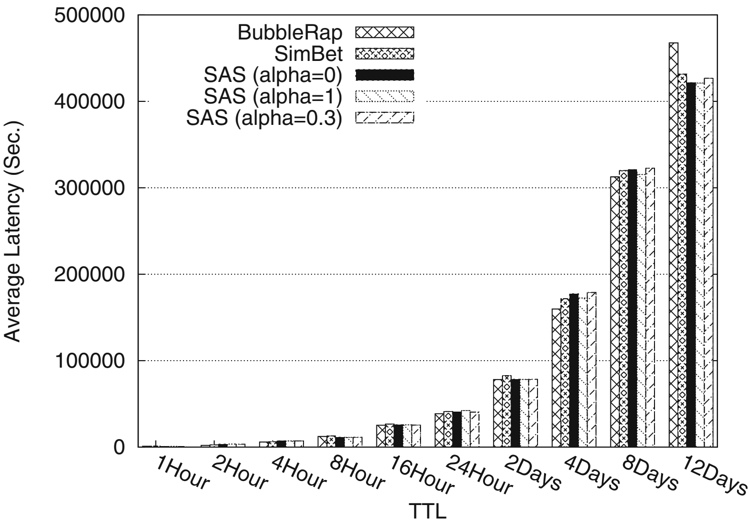

Fig. 6. Message average latency Vs TTL in Reality data set

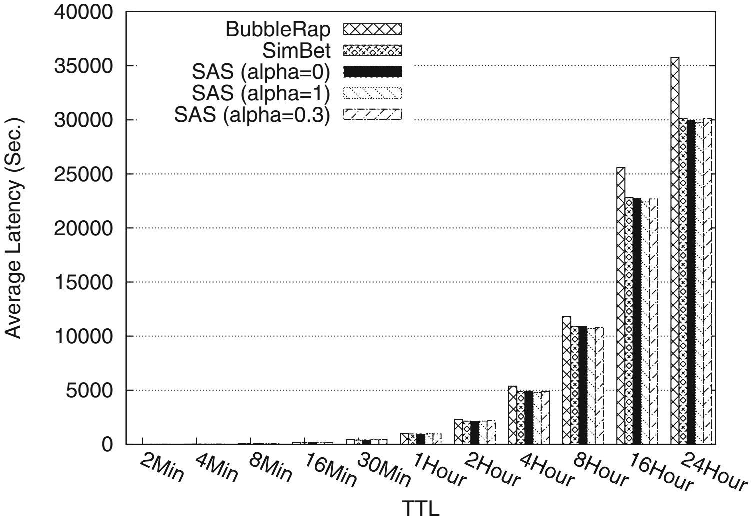

Fig. 7. Message average latency Vs TTL in Cambridge data set

5 Conclusion

This work has proposed SAS, a novel seasonality aware adaptive forwarding technique in social DTNs. The work is based on the observation of existence of seasonal behavioral pattern in node contacts in real mobility traces. SAS invoked a weighted Katz based similarity measure and ego-betweenness centrality to evaluate a utility value of an encountered node. Based on this utility, it decides the competency of a candidate node for being selected as a next hop message forwarder in DTN routing. The proposed method has been evaluated against different routing metrics through extensive set of simulation study with real mobility trace data sets. The performances of SAS has been found to get enhanced compared to the existing baseline social-based forwarding schemes available for DTNs.

Acknowledgements. The work has been carried out as a part of the Collaborative Research Scheme awarded for the project titled "On Ensuring Reliable Communication over Mobile Social Networks (MSNs)", of Assam Science and Technology University (ASTU) under TEQIP III program of MHRD.

References

1. Vastardis, N., Yang, K.: Mobile social networks: architectures, social properties, and key research challenges. IEEE Commun. Surv. Tutorials **15**(3), 1355–1371 (2013)
2. Hui, P., Chaintreau, A., Scott, J., Gass, R., Crowcroft, J., Diot, C.: Pocket switched networks and human mobility in conference environments. In: Proceedings of the

ACM SIGCOMM Workshop on Delay-tolerant Networking, pp. 244–251. ACM (2005)
3. Pereira, P.R., Casaca, A., Rodrigues, J.J., Soares, V.N., Triay, J., Cervelló-Pastor, C.: From delay-tolerant networks to vehicular delay-tolerant networks. IEEE Commun. Surv. Tutorials **14**(4), 1166–1182 (2012)
4. Conti, M., Giordano, S.: Mobile ad hoc networking: milestones, challenges, and new research directions. IEEE Commun. Mag. **52**(1), 85–96 (2014)
5. Zhu, Y., Xu, B., Shi, X., Wang, Y.: A survey of social-based routing in delay tolerant networks: positive and negative social effects. IEEE Commun. Surv. Tutorials **15**(1), 387–401 (2013)
6. Liben-Nowell, D., Kleinberg, J.: The link-prediction problem for social networks. In: Proceedings of the Conference on Information and Knowledge Management (CIKM 2003), pp. 556–559 (2003)
7. Freeman, L.C.: Centrality in social networks conceptual clarification. Soc. Netw. **1**(3), 215–239 (1978)
8. Fortunato, S.: Community detection in graphs. Phys. Rep. **486**(3), 75–174 (2010)
9. Daly, E.M., Haahr, M.: Social network analysis for routing in disconnected delay-tolerant MANETs. In: Proceedings of the 8th ACM International Symposium on Mobile Ad hoc Networking and Computing, pp. 32–40. ACM (2007)
10. Hui, P., Crowcroft, J., Yoneki, E.: Bubble rap: social-based forwarding in delay-tolerant networks. IEEE Trans. Mob. Comput. **10**(11), 1576–1589 (2011)
11. Freeman, L.C.: Centered graphs and the structure of ego networks. Math. Soc. Sci. **3**(3), 291–304 (1982)
12. Hui, P., Yoneki, E., Chan, S.Y., Crowcroft, J.: Distributed community detection in delay tolerant networks. In: Proceedings of 2nd ACM/IEEE International Workshop on Mobility in the Evolving Internet Architecture. ACM (2007)
13. Scott, J., Gass, R., Crowcroft, J., Hui, P., Diot, C., Chaintreau, A.: CRAWDAD dataset cambridge/haggle (v. 2006–09-15), September 2006. http://crawdad.org/cambridge/haggle/20060915
14. Eagle, N., Pentland, A.S.: Reality mining: sensing complex social systems. Pers. Ubiquit. Comput. **10**(4), 255–268 (2006)
15. Li, F., Wu, J.: LocalCom: a community-based epidemic forwarding scheme in disruption-tolerant networks. In: International Conference on Sensor, Mesh and Ad hoc Communications and Networks, pp. 1–9. IEEE (2009)
16. Zhou, T., Choudhury, R.R., Chakrabarty, K.: Diverse routing: exploiting social behavior for routing in delay-tolerant networks. In: International Conference on Computational Science and Engineering, vol. 4, pp. 1115–1122. IEEE (2009)
17. Wei, K., Guo, S., Zeng, D., Xu, K., Li, K.: Exploiting small world properties for message forwarding in delay tolerant networks. IEEE Trans. Comput. **64**(10), 2809–2818 (2015)
18. Bulut, E., Szymanski, B.K.: Exploiting friendship relations for efficient routing in mobile social networks. IEEE Trans. Parallel Distrib. Syst. **23**(12), 2254–2265 (2012)
19. Katz, L.: A new status index derived from sociometric analysis. Psychometrika **18**(1), 39–43 (1953)
20. Jones, E.P., Ward, P.A.: Routing strategies for delay-tolerant networks. ACM Comput. Commun. Rev. (CCR) (2006)
21. Vahdat, A., Becker, D., et al.: Epidemic routing for partially connected ad hoc networks. Technical report, Technical report CS-200006, Duke University (2000)

22. Lindgren, A., Doria, A., Schelén, O.: Probabilistic routing in intermittently connected networks. ACM SIGMOBILE Mob. Comput. Commun. Rev. **7**(3), 19–20 (2003)
23. Jain, S., Fall, K., Patra, R.: Routing in a delay tolerant network, vol. 34. ACM (2004)
24. Spyropoulos, T., Psounis, K., Raghavendra, C.S.: Spray and wait: an efficient routing scheme for intermittently connected mobile networks. In: Proceedings of the ACM SIGCOMM Workshop on Delay-tolerant Networking, pp. 252–259. ACM (2005)
25. Grossglauser, M., Tse, D.: Mobility increases the capacity of ad-hoc wireless networks. In: Proceedings of the INFOCOM, vol. 3, pp. 1360–1369. IEEE (2001)
26. Ciobanu, R.I., Reina, D., Dobre, C., Toral, S., Johnson, P.: JDER: a history-based forwarding scheme for delay tolerant networks using Jaccard distance and encountered ration. J. Netw. Comput. Appl. **40**, 279–291 (2014)
27. Ayub, Q., Rashid, S., Zahid, M.S.M., Abdullah, A.H.: Contact quality based forwarding strategy for delay tolerant network. J. Netw. Comput. Appl. **39**, 302–309 (2014)
28. Shin, K., Kim, K., Kim, S.: Traffic management strategy for delay-tolerant networks. J. Netw. Comput. Appl. **35**(6), 1762–1770 (2012)
29. Yuan, Q., Cardei, I., Wu, J.: An efficient prediction-based routing in disruption-tolerant networks. IEEE Trans. Parallel Distrib. Syst. **23**(1), 19–31 (2012)
30. Bulut, E., Wang, Z., Szymanski, B.K.: Cost-effective multiperiod spraying for routing in delay-tolerant networks. IEEE/ACM Trans. Netw. (TON) **18**(5), 1530–1543 (2010)
31. Niu, J., Wang, D., Atiquzzaman, M.: Copy limited flooding over opportunistic networks. J. Netw. Comput. Appl. **58**, 94–107 (2015)
32. Balasubramanian, A., Levine, B., Venkataramani, A.: DTN routing as a resource allocation problem. ACM SIGCOMM Comput. Commun. Rev. **37**(4), 373–384 (2007)
33. Wei, K., Liang, X., Xu, K.: A survey of social-aware routing protocols in delay tolerant networks: applications, taxonomy and design-related issues. IEEE Commun. Surv. Tutorials **16**(1), 556–578 (2014)
34. Hui, P., Crowcroft, J.: How small labels create big improvements. In: Proceedings of the International Conference on Pervasive Computing and Communications Workshops, pp. 65–70. IEEE (2007)
35. Wu, J., Wang, Y.: Social feature-based multi-path routing in delay tolerant networks. In: Proceedings of the INFOCOM, pp. 1368–1376. IEEE (2012)
36. Mei, A., Morabito, G., Santi, P., Stefa, J.: Social-aware stateless forwarding in pocket switched networks. In: Proceedings of the INFOCOM, pp. 251–255. IEEE (2011)
37. Granovetter, M.S.: The strength of weak ties. Am. J. Sociol. **78**(6), 1360–1380 (1973)
38. Marsden, P.V.: Egocentric and sociocentric measures of network centrality. Soc. Netw. **24**(4), 407–422 (2002)
39. Eagle, N., Pentland, A.S., Lazer, D.: Inferring friendship network structure by using mobile phone data. Natl. Acad. Sci. **106**(36), 15274–15278 (2009)
40. Scott, J., Gass, R., Crowcroft, J., Hui, P., Diot, C., Chaintreau, A.: Crawdad dataset Cambridge/haggle (v. 2006–09-15). CRAWDAD: Wireless Network data archive (2006)

41. Bigwood, G., Henderson, T., Rehunathan, D., Bateman, M., Bhatti, S.: CRAWDAD dataset st_andrews/sassy (v. 2011–06-03), June 2011. http://crawdad.org/st_andrews/sassy/20110603/mobile
42. Keränen, A., Ott, J., Kärkkäinen, T.: The ONE simulator for DTN protocol evaluation. In: Proceedings of the 2nd International Conference on Simulation Tools and Techniques, (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering) (2009)