Digital Receipt

This receipt acknowledges that Turnitin received your paper. Below you will find the receipt information regarding your submission.

The first page of your submissions is displayed below.

Submission author: Bambang Soelistijanto

Assignment title: Cek karya ilmiah

Submission title: Traffic Distribution and Network Capacity Analysis in Social …

File name: d_Network_Capacity_Analysis_in_Social_Opportunistic_Netwo…

File size: 443.32K

Page count: 8

Word count: 6,911

Character count: 35,546

Submission date: 09-Jan-2023 01:29PM (UTC+0700)

Submission ID: 1990071304

Traffic Distribution and Network Capacity Analysis in Social Opportunistic Networks

by Soelistijanto Bambang

Traffic Distribution and Network Capacity Analysis in Social Opportunistic Networks

Bambang Soelistijanto and Michael Howarth
Centre for Communication Systems Research
University of Surrey, UK
{b.soelistijanto, m.howarth}@surrey.ac.uk

Abstract — Social opportunistic networks are intermittently connected mobile ad hoc networks (ICNs) that exploit human mobility to physically carry messages between disconnected parts of the network. Human mobility thus plays an essential role in the performance of forwarding protocols in the networks, and people's movements are in turn affected by their social interactions with each other. In this paper we present an analysis of the traffic distribution among the nodes of social opportunistic networks and its impact on network capacity. For our analysis, we use a human contact graph that represents a social network of individuals. We characterize the graph as a scale-free network and apply forwarding strategies based on the information required by a node to select relays for its messages, categorising this information either as isolated or complete network or local network knowledge. We use a social network property, centrality, for the forwarding strategies, additionally considering tie strength in the forwarding metric and investigate their impact on traffic distribution. We show that all the strategies result in unfair traffic distribution due to a strong non-random structure of the networks, where hub nodes process much more relay traffic than non-hub nodes. Finally, we present a mathematical model of network capacity as an upper-bound of network delivery performance where hub nodes' resources become the limiting factors, and show that including tie strength in the forwarding metric improves the network capacity.

Keywords: opportunistic networks, traffic distribution, network capacity, scale-free graph, centrality, tie strength

I. INTRODUCTION

Intermittently connected mobile ad hoc networks (ICNs) are networks which can operate in the presence of disrupted links or long transfer delays; they also in the literature sometimes referred to as delay- and/or disruption- tolerant networks (DTNs). One scenario in which ICNs can be useful is in networking for devices carried by users of mobile and portable devices, such as packet switched and mobile social networks (PSNs) [1]. PSNs use opportunistic data dissemination to facilitate data communication among users in a dynamic and frequently disconnected network environment. The nodes in PSNs make use of a store-carry-forward mechanism to carry messages to the destinations, and so a user's mobility therefore plays an essential role in the protocol's performance. People's movement behaviours are strongly affected by their social interactions with each others. Furthermore, it has been proved that knowledge of social structure can help in designing better forwarding algorithms in opportunistic networks [2].

In ICNs, there are two scenarios for node movements, i.e. deterministic and stochastic. In a deterministic scenario, future node movements and connections are completely known and hence the entire network topology is known ahead of time. In a stochastic scenario, however, node contacts are unpredictable and the network behaviour is random and unknown, and hence routing is a complex task. The simplest forwarding decision is to forward each copy of a message to any node in contact, i.e. Epidemic routing [3], while other approaches may be based on history data, mobility patterns or other information. A history contact based routing, e.g. Prophet [4], uses a probabilistic metric that indicates the likelihood of the relay to be able to deliver a message to the destination. At every contact, nodes exchange delivery probability vectors containing the delivery predictability information for destinations known by the nodes. However, given the unpredictable node mobility and frequent changes of network topology, this strategy will create much control traffic during a node contact which may be of short duration.

Recently, a novel approach of a history-based routing algorithm that uses structural information of individuals in a social network has been developed. These social-aware routing protocols use some characteristics of a social network that are less volatile than a physical network. In the networks formed by people, social relationships vary much more slowly and therefore they can be used as routing metrics. The routing algorithms may use social network properties, such as centrality and community, as routing metrics. Centrality is a measure of the relative importance of an individual within a social network and can be assessed by various metrics such as degree, betweeness, closeness centrality etc. [5]. A high centrality indicates that a person appears to be more popular and thus has more contacts than less popular people. On the other hand, people inherently form groups and this creates concept of community. People within a given community are more likely to meet each other than randomly chosen people.

Unfortunately, despite its benefits, a social-aware forwarding algorithm presents a drawback in traffic distribution among nodes in the network. Since the algorithm favours high rank (high centrality) nodes as traffic relays, a few nodes will receive much more traffic than others. So there is a need to further uncover the impact of social-aware forwarding algorithms on traffic distribution. The contribution of this paper is therefore as follows. First, we perform an analysis of the effect on traffic distribution of using centrality and tie strength as the metrics for traffic forwarding strategies. Second, we

present a model to calculate network capacity of opportunistic networks. While most of the existing models of network capacity are derived with respect to betweeness centrality, we use node degree instead. We argue that betweeness centrality is hard to calculate in ICNs since it requires complete network knowledge, while node degree is information that is locally available at a node. To the best of our knowledge, our work is the first that uses node degree to derive network capacity. Given that the focus of this paper is traffic distribution and network capacity, we do not consider here other metrics such as message delay within the network.

The remainder of the paper is organized as follows. Section II describes related work in the area of traffic distribution and balancing in social opportunistic networks. Section III presents our model of traffic distribution in social opportunistic networks for different forwarding strategies. In Section IV, we derive an analytical model of network capacity of a social opportunistic network. Finally, we discuss the results in Section V, which is followed by conclusion and future work in Section VI.

II. RELATED WORK

The uses of social network properties for designing forwarding protocols in ICNs have been broadly discussed in the literatures, e.g. in [2,6,7,8]. A social-aware forwarding scheme, LABEL [6], assumes every node possess a label indicating its community, and forwards messages to relay nodes that belong to the same community as the destination. RANK [2] uses node centrality as a forwarding metric and forwards a message to nodes that have a higher ranking (i.e. larger centrality) than the current node until the destination is reached. SimBet [7] and Bubble Rap [2] use both centrality and community as the forwarding metrics. Bubble Rap combines degree centrality and community structure to decide relay nodes. Degree centrality is the total number of links that a node has. Meanwhile, SimBet uses locally-calculated (ego) betweeness centrality and similarity as its routing metrics. The betweeness centrality of a node is expressed as a fraction and is the number of shortest paths that pass through the node divided by the number of shortest paths in the network. Similarity, on the other hand, indicates the ratio of common neighbours between individuals in social networks. When a social network displays a high degree of clustering, the probability of two nodes being connected is higher if the nodes have common neighbours. Investigations in [2,7] showed that both SimBet and Bubble Rap need lower total control traffic than Prophet but they are able to keep the delivery ratio as high as Prophet.

Wang et al. [9] studied the impact of social structures, in terms of centrality and community, on forwarding performance, i.e. delivery success ratio and average hop counts, in social opportunistic networks. Nevertheless, only a few papers discuss the main drawback of social-aware forwarding strategies, i.e. unfair traffic distribution among nodes in opportunistic networks. This unfair traffic distribution is not sustainable as it can quickly deplete resources in the heavy utilized (hub) nodes and eventually will degrade overall delivery ratio. The authors in [10] investigated the drawback of SimBet in term of traffic distribution in the network and then proposed Fair Route to address the issue. CAFé [11] has also

been proposed with the aim of distributing load away from hub nodes. The approach adopted by both Fair Route and CAFé is to reduce the effect of centrality by smoothing its value (in CAFé) or decreasing it in time (in Fair Route). Interaction strength, i.e. some measure of the strength of a link between two nodes, was also added in the forwarding metrics of [10,11] to improve the algorithm performance. Similarly, the authors of SimBet [7] improved their existing algorithm by adding tie strength in the routing metric. This improved algorithm, SimbetTS [12], combines the frequency of contact, contact duration and the age of contact as measures of the tie strength between two nodes in the network. Unfortunately, it has been shown in [10,11,12] that improving the contact data, i.e. centrality and tie strength, has failed to provide a proper solution for traffic balancing. Hence, we agree with the authors of [10,11] that both Fair Route and CAFé need an additional strategy, e.g. (storage) congestion control, which reduces the traffic in hub nodes. In this paper we investigate the impact of two metrics, i.e. centrality and tie strength, on forwarding algorithms in terms of traffic distribution and network capacity of social opportunistic networks. We take an analytical approach, and use the graph that resembles a human contact graph representing a social (relation) network.

III. TRAFFIC DISTRIBUTION MODEL

In communications, some nodes are usually more important than others. In a network, a node may be considered more important if many shortest paths traverse it. Typically, centrality metrics are suitable for analysis in well-connected networks, such as the Internet and MANETs. In an ICN, however, the network topology changes very rapidly and the physical connectivity of the network might be extremely unstable. Several papers simply define the topology of ICNs as unpredictable, but others argue that temporal connection models are better suited than spatial models [13]. Research topics concerned with the topology of ICNs are still wide open today. Human mobility characteristics discussed in [14] show that there exists a virtual, social (relation) graph that drives humans to move, and that this graph is less volatile than physical topology. The overlay graph represents a macroscopic property of human mobility, and the ICN steady-state (*long-term behaviour*) of protocol performance analysis can be performed over this abstract/logical layer, as in [15]. We illustrate the structural topology of a social opportunistic network in Figure 1.
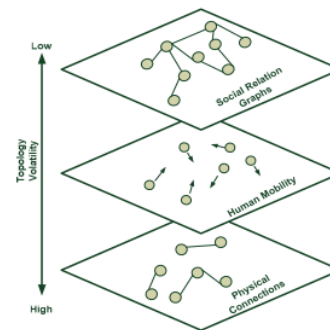

Figure 1. Structural topology of a social opportunistic network

A. Network topology generation

Hossmann et al. [16] studied several real human contact traces and argued that the derived contact graphs have a strong non-random structure (complex graph). Ferreti et al. [17] studied the feasibility of coupling between scale free graphs and social opportunistic networks. By employing real data traces, they set up contact graphs by varying the aggregated time interval and the minimum link value threshold and found that the resultant graphs possess a scale free structure. A scale free graph is one whose degree distribution, $P(k)$, follows a power law as $P(k) \sim k^{-\gamma}$, where k and γ are node degree and degree exponent, respectively. For our model, we use the Barabasi-Albert (BA) algorithm [18] to generate a binary scale free network: this binary network consists of two link states: '1' if a direct link exists between two nodes, and '0' if otherwise. The numerical analysis in [18] indicated that the BA network evolves into a scale invariant state with a range of degree exponents similar to many real social networks, i.e. $2 < \gamma_{BA} < 3$.
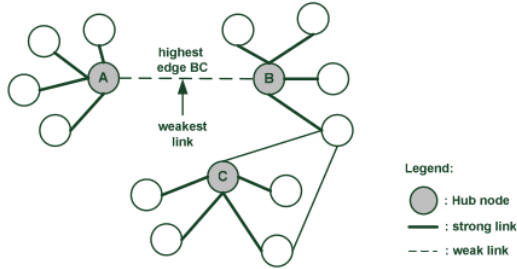

Figure 2. Illustration of the weak tie hypothesis of Onella [19]

We then improve the binary BA network to better represent a human contact graph by including tie strength in each link of the network. We call this a weighted BA network. Tie strength is a quantitative property that measures how strong a link is between two nodes. It can depend on several factors, e.g. frequency of contact, contact duration or elapsed time since last contact. We follow the 'weak-tie' hypothesis of Onella et al. [19] when assigning tie strength to the links of the binary network. This hypothesis postulates that the weakest links – *in terms of total connected time* – are those with the highest betweeness centrality (BC) in the network. This hypothesis was validated in social opportunistic networks by Williamson [20], who investigated the distribution of link connected time in real human contact traces. As illustrated in Figure 2, the link between nodes A and B has the highest edge BC and hence the link is considered as the weakest link in the network because it is the most critical link in maintaining network connectivity (since more shortest paths pass through it than through any other link). Subsequently, to assign values of tie-strength to the links of the binary network we use the work of Williamson [20] that found the distribution of link connected time in the *Reality Mining* dataset [21] fits a power law distribution as $f(x) \sim x^{-2.049}$, where x is the percentage of total link connected time. In our work, we define tie-strength as a function of the total contact duration between two nodes and thus consider $f(x)$ as the distribution of tie-strength in the network. The detailed procedure of assigning tie-strength in the links is described in Algorithm 1. This algorithm produces a weighted scale free network where the weights represent the tie strength between nodes. These weights are not the same as conventional link weights used in routing, since in our case higher weights (or tie strength) are regarded as more desirable.

Algorithm 1. Assigning tie strength to links for the BA network

1) Generate a binary scale free network using BA algorithm
2) Calculate an edge betweeness centrality (BC) for every link in the network and sort the links in increasing order of BC
3) Determine the range of tie strength values used in the network (e.g. we choose tie strength in the range from 1 to 10).
4) Using the tie strength distribution (e.g. $f(x) \sim x^{-2.049}$), calculate $f(x)$ (e.g. for $x = 1$ to 10) and determine the probability of each tie-strength value in the network as $f(x)/\sum_x f(x)$.
5) For each probability, multiply it by the total number of links to get the number of links that have the given tie-strength. Following the 'weak-tie' hypothesis, assign the highest fraction of the links to the largest tie strength.
6) Sort the tie strengths in decreasing order and assign them to the links in increasing order of BC, so that the link with the highest BC has the lowest tie strength.

B. Markov model of steady state traffic distribution

To study node traffic analytically, we calculate the probability of a message being found in a given node, assuming steady state traffic flow. We define this probability as the occupation ratio of a node in steady state traffic. By assuming all the steady state processes in the network, i.e. message arrival, process and departure in a node, are *i.i.d* (identical and independently distributed) and that they follow a Poisson distribution, we can use the discrete Markov process to illustrate how a message moves from one node to another with a certain transition probability. We define p_{ij} as the probability that a message from node i moves to node j. With this definition, the transition matrix P of a finite Markov chain with N nodes has the form as follows:

$$P_{N \times N} = \begin{bmatrix} 0 & p_{12} & \cdots & p_{1N} \\ p_{21} & 0 & \cdots & p_{2N} \\ \vdots & \vdots & \cdots & \vdots \\ p_{N1} & p_{N2} & \cdots & 0 \end{bmatrix} \quad (1)$$

where $\sum_j p_{ij} = 1$ and p_{ij} is determined by the network routing strategy. If $\pi(i)$ denotes the probability distribution at i-th step, then the rule of a message walk can be expressed by a simple equation:

$$\pi(t) = (P^T)^t \pi(0) \quad (2)$$

where $\pi(0)$ is the initial probability. If a chain is ergodic, then there is a unique steady state (equilibrium) distribution which solves the relation $\pi = \pi P$, where π is the steady state distribution vector. Then, the steady state probability of a message being found in node i is given by

$$\pi_i = \sum_{j=0}^{N} p_{ji} \pi_j \quad (3)$$

where p_{ji} is the transition probability a message departs from node j to i and π_j is the steady state probability of a message being found in node j.

C. Forwarding strategies for social opportunistic networks

In our study, we consider three forwarding strategies that can be divided based on the knowledge of the network structural information which is required by a node to determine its relays. We categorise these strategies as isolated, complete network and local network knowledge.

In the isolated knowledge strategy, a node selects neighbouring relays for its messages based only on its own knowledge. In the binary network, this knowledge is the number of direct neighbours and the current node will select its neighbours to be relays with equal probability, i.e. the inverse of its degree. In the weighted network, a node takes into account the tie strength of each link to the neighbours. Since the tie strength is proportional to the contact duration with the neighbours, a node requires no additional knowledge from the neighbours and hence this value of the weights can be considered as isolated knowledge. The probability of a message being relayed from node i to j, p_{ij}, for an isolated knowledge network is defined as follows

$$p_{ij} = \begin{cases} \frac{a_{ij}}{\sum_{n \in N} a_{in}}, & \text{if } a_{ij} \neq 0; \ N\text{: set of } i's \text{ neighbours} \\ 0, & \text{if } a_{ij} = 0 \end{cases} \quad (4)$$

where a_{ij} is the tie-strength between node i and j, and $\sum_j p_{ij} = 1$.

In the complete network knowledge strategy, we assume a node has knowledge of the network topology for all other nodes. Then the node can calculate the ranking of the importance of its neighbours in the network. We use node betweeness centrality (BC), which we define as the number of shortest-paths passing through the node. In the binary network, once a node has the entire network topology, it runs Dijkstra's algorithm to calculate the BC of its neighbour nodes. In the weighted network, on the other hand, where the weight is determined by tie strength, a higher weight means it is more probable that a link exists between two nodes at any time. Hence, the higher the weight the more desirable the link is and therefore it is more likely to be selected as part of the shortest path (this is the opposite of the conventional Dijkstra's algorithm). Hence, in both binary and weighted networks, the probability of a message being relayed from node i to j for a complete knowledge network can be calculated as

$$p_{ij} = \begin{cases} \frac{BC_j}{\sum_{n \in N} BC_n}, & \text{if } a_{ij} \neq 0 \ ; \ N\text{: set of } i's \text{ neighbours} \\ 0, & \text{if } a_{ij} = 0 \end{cases} \quad (5)$$

where BC_j is the betweeness centrality of node j, a_{ij} is the tie-strength between node i and j, and $\sum_j p_{ij} = 1$.

The calculation of betweeness centrality needs global knowledge of network topology, but obtaining this data is unrealistic in ICNs due to the intermittent connections and large transfer delay. More realistically, a forwarding algorithm for opportunistic networks can estimate node BC in a

decentralized manner using locally available information. Therefore, in the local network knowledge strategy, we use two different metrics, node degree and ego betweeness, to represent the BC of neighbour nodes. The correlation between node degree and node BC has been discussed in [2]. If the node degree of the neighbour nodes is used as the metric for the local network knowledge strategy (as in RANK [2]), the algorithm will favour a higher neighbour node degree as a better relay node. In this case, the probability of node j being the relay for node i, p_{ij}, is given for the local knowledge network by

$$p_{ij} = \begin{cases} \frac{k_j}{\sum_{n \in N} k_n}, & \text{if } a_{ij} \neq 0; \ N\text{: set of } i's \text{ neighbours} \\ 0, & \text{if } a_{ij} = 0 \end{cases} \quad (6)$$

where k_j is the degree of node j and $\sum_j p_{ij} = 1$. For the weighted network, we define the node degree as the sum of the tie strengths of the links attached to the node.

The other local metric that can be used to determine the BC of neighbour nodes is ego betweeness, which can be calculated using the concept of the '*ego network*' [22]. Here, the node BC is calculated by computing the number of nodes that are indirectly connected through the node. Then, the ego betweeness of a node is calculated as the sum of the reciprocal of the entries of $A^2[1 - a_{ij}]$, where A is an adjacency matrix and a_{ij} is the link weight between node i and j. This strategy favours the neighbour having a higher ego betweeness as a better relay. Hence, in this case the probability of node j being the relay node for node i, p_{ij}, can be calculated for the local knowledge network as follows

$$p_{ij} = \begin{cases} \frac{EB_j}{\sum_{n \in N} EB_n}, & \text{if } a_{ij} \neq 0; \ N\text{: set of } i's \text{ neighbours} \\ 0, & \text{if } a_{ij} = 0 \end{cases} \quad (7)$$

where EB_j is the ego betweeness of node j and $\sum_j p_{ij} = 1$.

IV. Network Capacity Model

In a non-random structure network, there exist a few hub nodes that have a very large node degree and which process much more traffic than other nodes in the network. As a result, the main resources of these nodes, i.e. storage and power, will quickly be depleted and this will eventually degrade the network performance. Hence, a network capacity model is required to understand the maximum possible message transfer rate in a network. Suppose λ is the average rate of messages generated in a network node. We are then interested in a critical value λ_c (measured by the number of messages created at a node per unit time) where a 'phase transition' takes place in the network from free flow to congested flow. The value λ_c thus reflects the network's maximum traffic capacity.

Our network capacity model is derived with respect to node degree. Although most of the published network capacity models are based on betweeness centrality, we argue that node degree is locally available data and is easily obtained in the ICN context. We assume the network consists of N nodes and for each node we define the following variables: λ_i is the mean number of generated messages at node i in each time step and is assumed to follow a Poisson distribution and be independent

for each node; c_i is the maximum number of messages that can be forwarded by node i in each time step; and k_i is the degree of node i. We also assume the message length to be constant. Traffic at a node's network layer consists of two parts, i.e. arriving and departing traffic. Moreover, we can divide a node's arriving traffic into messages created by the node (local traffic) and messages that are transferred from other nodes (relay traffic). Hence, for each node j, we define the mean number of arrivals as

$$\alpha_j = \lambda_j + \sum_{i=1}^{N} p_{ij}\beta_i \tag{8}$$

where λ_j is the local traffic originating at node j, p_{ij} is the probability that messages will be forwarded from node i to j and β_i is the total relay traffic departing from node i. By assuming steady state flow in the network, i.e. $\alpha_i = \beta_i$, and by assuming the information generation rate to be uniform in all nodes as λ, then (8) becomes

$$\alpha_j = \lambda + \sum_{i=1}^{N} p_{ij}\alpha_i \tag{9}$$

We see that the summation $\sum_{i=1}^{N} p_{ij}\alpha_i$ in (9) reflects the total relay traffic that arrives at a node. Meanwhile, [23] shows that typically the load distribution in a scale free network follows a power law as

$$Prob_l(\ell) \sim \ell^{-\delta} \tag{10}$$

where δ is the load exponent. The authors in [23] also argue that the load is highly correlated with node degree k, scaling as

$$\ell_k \sim k^{\eta} \,, \quad \text{for } \eta = (\gamma - 1)/(\delta - 1) \tag{11}$$

Barthelemy [24] stated that δ tightly depends on the network's degree exponent γ and the routing strategy. For shortest path forwarding, he found δ has the range 1.8 to 2.3. Applying (11) to (9), we are able to approximate the mean arrival rate of traffic in node j as

$$\alpha_j \approx \left(\varepsilon k_j^{\eta}\right)\lambda \tag{12}$$

where λ, k, η and ε are the average node message generation rate, node degree, degree-load exponent and load scaling constant, respectively.

To calculate the network capacity we model the traffic process in the network of N nodes as N queuing systems of $M/M/1$. As described in Section III.B, we assume that all the steady state processes in the network are i.i.d (identical and independently distributed). To be specific, we assume that the contact duration between two nodes is uniformly distributed and sufficiently long to exchange all messages and control data (e.g. network neighbourhood metrics). Furthermore, we assume that the inter-contact times between a node and any potential relay are i.i.d. and hence all the processes in the network, i.e. message arrival, service and departure, can be considered as exponential/Poisson processes. In queuing theory, typically the behaviour of the queue depends on the arrival process α_i and the service process c_i at node i. Our aim is to determine the critical message generating rate λ_c, where if $\lambda < \lambda_c$ the network can sustain a free flow state. Initially, we define the

mean delay in the network as the summation of the delay of N queue systems as follows

$$T = \frac{1}{S}\sum_{j=1}^{N} \frac{c_j}{c_j - \alpha_j} \tag{13}$$

where $S = \sum_{j=1}^{N} c_j$. We solve the optimization problem, i.e. minimize the average delay T and maximize message generation rate λ subject to the constraint of network resources S. Then, the problem can be formulated as

$$\min_{c_j} \; \sigma T + (1 - \sigma)\,{1}/{\lambda} \; \text{ subject to } S \tag{14}$$

where $0 \le \sigma \le 1$ is a weighted parameter. Subsequently, we devise the optimization problem in the form of the Lagrangian method with Lagrangian multiplier φ as

$$L = \frac{\sigma}{S}\sum_{j=1}^{N} \frac{c_j}{c_j - (\varepsilon k_j^{\eta})\lambda} + (1 - \sigma)\frac{1}{\lambda} + \varphi\left(\sum_{j=1}^{N} c_j - S\right) \tag{15}$$

The set of c_j and λ maximizes L under the conditions $\frac{\partial L}{\partial c_j} = 0$, $\frac{\partial L}{\partial \lambda} = 0$ and $\frac{\partial L}{\partial \varphi} = 0$. Finally, we get $\varphi = \frac{1-\sigma}{\lambda S}$ and arrive at the equation

$$c_j = \left(\left(\varepsilon k_j^{\eta}\right) + \sqrt{\frac{\sigma}{1-\sigma}\left(\varepsilon k_j^{\eta}\right)}\right)\lambda \tag{16}$$

If we are only interested in maximizing λ, we can set $\sigma = 0$, in which case (16) becomes

$$c_j = \left(\varepsilon k_j^{\eta}\right)\lambda \tag{17}$$

We assume that the node local delivery capacity is constant, i.e. $c_j = c$ and without loss of generality we assume $c = 1$. Hence we obtain the node's critical message generation rate as

$$\lambda_{cr} = \frac{1}{\varepsilon k_{cr}^{\eta}} \tag{18}$$

where k_{cr} is the maximum degree k in the network and hence corresponds to the node that has the largest node degree. The authors in [25] describe how the maximum node degree in a scale free network varies with the total number of nodes N as $k_{max} \sim N^{1/(\gamma-1)}$ and thus we can rewrite (18) by substituting k_{max} into k_{cr} as

$$\lambda_{cr} = \frac{1}{\varepsilon N^{1/(\delta-1)}} \tag{19}$$

where N, δ and ε are the total number of nodes, load exponent and a scaling constant, respectively.

V. RESULTS AND EVALUATION

A. *Traffic distribution analysis*

In Section III.B, we defined the occupation ratio as the probability of a message being found in a network node in steady state traffic flow. We can also consider the occupation ratio as the fraction of the total messages (traffic) that arrive at a node in steady state flow. In other words, the occupation ratio is the ratio of the number of times a node acts as a relay node divided by the total relay traffic in the network. We now describe the metrics used in our performance comparison of forwarding strategies, i.e.:

- PAR (peak-to-average ratio): the ratio of maximum to mean occupation ratio (traffic) of nodes in the network. A lower PAR is desirable since this means that traffic is more evenly distributed across the network.
- The percentage of total traffic carried by the busiest nodes.

For our study, we generate a binary scale free network in MATLAB using the BA model for $N = 100$ nodes, $m_o = 5$ seed nodes, $m = 3$ edges and subsequently apply all three forwarding strategies (described in Section III.C) to it. We also create a weighted network by calculating tie strength for the links of the binary network using Algorithm 1. Again, we apply all three forwarding strategies to the weighted network. Hence, we can compare the performance of the forwarding strategies when tie strength is excluded (on the binary networks) and included (on the weighted networks) in our analysis. Our results presented here are the average of results obtained for ten different BA network topologies.
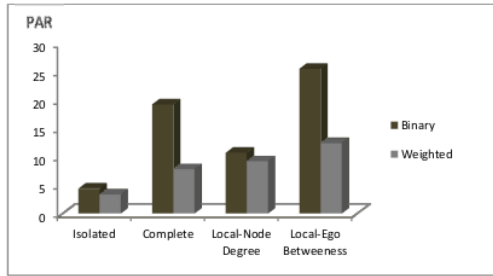

Figure 3. The average PAR in the binary and weighted network

In Figure 3, we show as a histogram the PAR for all three forwarding strategies in both the binary and weighted networks. In the case of the local network knowledge strategy we present separate results for node degree and ego betweeness metrics. We note that the higher the PAR, the worse the traffic distribution will be in the network. We notice that the isolated knowledge strategy shows the best performance among the strategies in distributing the traffic in the network. In this strategy, a node selects a relay node based on its own knowledge (regardless of the social properties of its neighbours) and Newman [26] argued that this type of forwarding can be considered as a 'random walk' in both binary and weighted networks. We also see that including tie-strength in the routing metric significantly improves the performance of the complete knowledge and ego-betweeness local knowledge strategies, and makes a modest improvement in the other two cases. As we have mentioned, these first two strategies (complete, and local ego-betweeness) favour nodes with higher betweeness centrality and ego betweeness as relay nodes for most of the traffic in the network. By adding tie strength, we significantly reduce the traffic relayed through hub nodes and redirect it to the well connected neighbour nodes.

We also measure the performance of the forwarding strategies based on the total traffic carried by the busiest nodes in the networks. In Figure 4 and 5, we show histograms that represent the number of the busiest nodes that carry a given fraction of the total traffic (either 25%, 50% or 75%) in the

binary and weighted networks (for $N = 100$ nodes), respectively. In the binary network (Fig. 4), we note the trend in traffic distribution from fairest to least fair, i.e. isolated knowledge (fairest), local network knowledge (node degree), complete network knowledge and local network knowledge (ego-betweeness) (least fair). By comparison, all the forwarding strategies show performance improvements in the weighted networks (Fig. 5). Here, the performance of the complete network knowledge and local knowledge (ego betweeness) forwarding strategies increase significantly when tie strength is added in the routing metric. On the other hand, only modest improvements in traffic distribution are shown in the isolated and local knowledge (node degree) forwarding strategies.
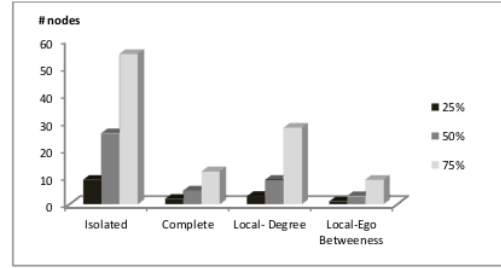

Figure 4. The percentage of total traffic carried by the busiest nodes in the binary network
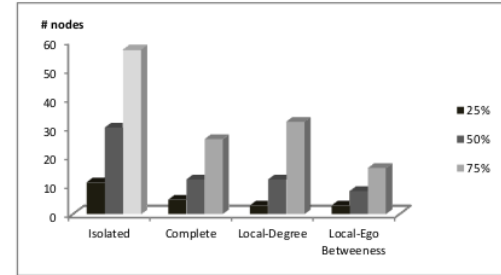

Figure 5. The percentage of total traffic carried by the busiest nodes in the weighted network

We also investigate the effect of increasing the number of nodes in the networks on the traffic distribution. We generate weighted networks for $N = 200$ to 1000 nodes and apply all the forwarding strategies to them. We present the results in Figure 6, which shows the percentage of the busiest nodes that between them are on average at any time carrying 50% of the traffic. As we can see, isolated knowledge forwarding exhibits the fairest traffic distribution, and this does not vary significantly as the number of nodes increases. As we commented above, according to Newman [26], the isolated knowledge forwarding (taking into account tie strength) can be considered as a 'random walk' and exhibits fairer traffic distribution compared to the other strategies. Meanwhile, in the local knowledge forwarding strategies (both node degree and ego betweeness) the percentage of nodes that carry 50% of the traffic decreases noticeably with the increase in the number of

828

nodes in the network. The performance of the local knowledge strategies is significantly worse than the complete knowledge strategy in large networks. This is to be expected, since it seems reasonable that in larger networks local network information is not a good predictor of the overall path from the source to the destination.
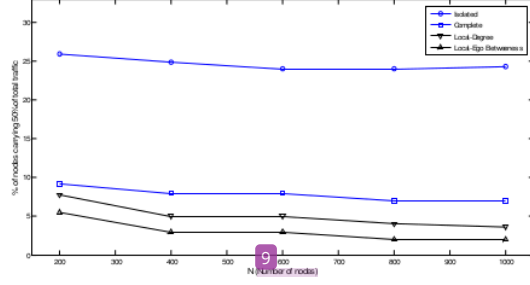

Figure 6. The percentage of total nodes carrying 50% of total traffic (weighted networks)

B. Network capacity analysis

As described in (10), typically the load distribution in a scale free network follows a power law. However, for the isolated knowledge forwarding we found that the occupation ratio increases linearly with the node degree in both the binary and weighted networks. These results are in line with the simulation results in [27] in that the mean arriving traffic of nodes increases linearly with the increasing node degree for random forwarding. As previously mentioned, to some extent isolated knowledge forwarding can be considered as random forwarding that will disregard the centrality of neighbour nodes when selecting relay nodes. We show in Figure 7 the relation between node degree and occupation ratio for the binary network case only, due to limited space.
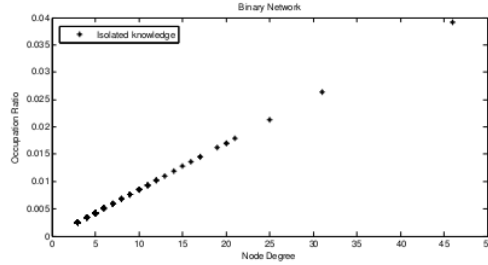

Figure 7. Occupation ratio vs. degree for isolated knowledge forwarding.

However, for the other forwarding strategies we found that the occupation (load) distributions follow a power law. We generate a network with $N = 500$ nodes, $m_o = 5$ nodes, $m = 3$ edges, apply the forwarding strategies and calculate the occupation (load) distribution for each run for both the binary and weighted networks. Our results are again the average of ten different network topologies. Subsequently, we get the load exponent δ for each forwarding strategy as listed in Table I. Furthermore, using (19) and the average values of δ of each forwarding strategy, we plot the critical message generation

rate λ_{cr} as a function of network size, for $N=100$ to 1000 and $\varepsilon = 0.01$, in Figure 8 to 10.

TABLE I. Load exponent in the networks (ten different networks)

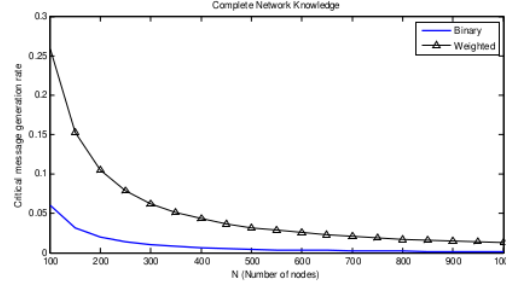| Forwarding strategy	Load exponent (δ)	
	Binary network	Weighted network
Complete network knowledge	$1.51 < \delta < 1.732$ (Avg. 1.621)	$1.659 < \delta < 1.887$ (Avg. 1.779)
Local knowledge (node degree)	$1.621 < \delta < 1.853$ (Avg. 1.713)	$1.635 < \delta < 1.864$ (Avg. 1.742)
Local knowledge (ego betweeness)	$1.319 < \delta < 1.616$ (Avg. 1.416)	$1.527 < \delta < 1.866$ (Avg. 1.639)

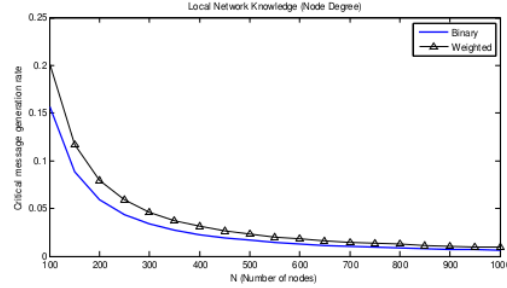**Figure 8.** λ_{cr} vs. N, for complete network knowledge forwarding

Figure 9. λ_{cr} vs. N, for local network knowledge (node degree) forwarding

As we note in Figure 8, adding tie strength (i.e. the weighted network) significantly increases the critical message generation rate λ_{cr} for complete network knowledge forwarding. In the other words, adding tie strength in the routing metric in this strategy will significantly reduce the occupation ratio in the most important hub nodes in the network. This result is in line with the result in Figure 3, where the PAR in the complete knowledge network also significantly decreases. Meanwhile, in the local network knowledge (node degree) forwarding (Fig. 9), we can see that adding tie strength slightly increases the critical message generation rate, and therefore slightly reduces the occupation ratio in the most popular hub nodes (in Fig. 3, PAR also reduces slightly in this forwarding strategy). Finally, in Figure 10 we see that the performance of the local network knowledge (ego betweeness) forwarding strategy improves dramatically when tie strength is

added in the routing metric. This result is in line with the PAR for this strategy where the occupation ratio of the most important hub node is reduced significantly by adding tie strength. However, this strategy shows the worst performance among other strategies, in terms of both traffic distribution (highest PAR) and network capacity (lowest λ_{cr}). Hence, we can argue that the future application of forwarding strategy based on ego betweeness especially in large networks is questionable. Researchers should seek other local metrics that may have better impact on traffic distribution in social opportunistic networks.
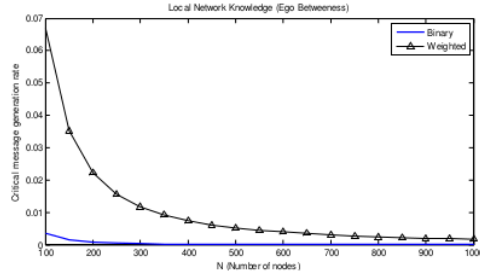

Figure 10. λ_{cr} vs. N, for local network knowledge (ego betweeness) forwarding

VI. CONCLUSIONS AND FUTURE WORK

In this paper we have investigated the impact on traffic distribution and network capacity of including centrality and tie strength in forwarding strategies for social opportunistic networks. We showed that adding tie strength to the routing metric improves the performance of all forwarding strategies (isolated, complete and local knowledge) in term of traffic distribution. However, we have shown that all forwarding strategies still result in unfair traffic distribution, where some nodes process much more traffic than others. Furthermore, we found that the capacity of opportunistic networks to deliver messages strongly depends on the largest degree hub nodes. We also showed that including tie strength (i.e. the weighted networks) improves the network capacity of social opportunistic networks.

In the future, we will include community structure in the routing metric of the forwarding strategies and investigate its impact on global (in whole network) and local (within community) traffic distribution in social opportunistic networks.

REFERENCES

[1] G. Hui, A. Chaintreau, J. Scott, R. Gass, J. Crowcroft, C.Diot, "Pocket Switched Networks and Human Mobility in Conference Environments", *Proc. 2005 ACM SIGCOMM*, NY, 2005.

[2] P. Hui, J. Crowcroft, E. Yoneki, "BUBBLE Rap: Social-based Forwarding in Delay Tolerant Networks", *Proc. 9th ACM International Symposium on Mobile Ad Hoc Networking and Computing*, Hong Kong, China, 2008.

[3] A. Vahdat, D. Becker, "Epidemic Routing for Partially Connected Ad Hoc Networks", *Tech Report CS-200006*, Duke University, 2000.

[4] A. Lindgren, A. Doria, O. Schelen, "Probabilistic Routing in Intermittently Connected Networks", *ACM SIGMOBILE Mobile Computing and Communication Review*, vol.7, no. 3, pp. 19-20, 2003.

[5] L.C. Freeman, "Centrality in Social Networks: Conceptual Clarification", *Social Networks*, vol. 1, no. 3, pp. 215-239, 1979.

[6] P. Hui, J. Crocroft, "How Small Labels Create Big Improvements", *Proc. CoNEXT'06 ACM Conference*, NY, USA, 2006.

[7] E. Daly, M. Haahr, "Social Network Analysis for Routing in Disconnected Delay Tolerant MANETs", *Proc. 8th ACM International Symposium on Mobile Ad Hoc Networking and Computing*, Montreal, Canada, 2007.

[8] P.Hui, E. Yoneki, S.Y. Chan, J. Crowcroft, "Distributed Community Detection in DTNs", *Proc. 2nd ACM/IEEE Int'l Workshop on Mobility in the Evolving Internet*, NY, USA, 2007.

[9] N. Wang, E. Yoneki, "Impact of Social Structure on Forwarding Algorithms in Opportunistic Networks", *Proc. iCOST Mobile and Wireless Networking Conference*, Shanghai, China, 2011.

[10] J.M. Pujol, A.L. Toledo, P. Rodriguez, "Fair Routing In Delay Tolerant Networks", *Proc. INFOCOM 2009*, Rio de Janeiro, 2009.

[11] A. Grundy, M. Radenkovic, "Promoting Congestion Control In Opportunistic Networks", *Proc. International Conference WiMob2010*, 2010.

[12] E. Daly, M. Haahr, "Social Network Analysis for Information Flow in Disconnected Delay-Tolerant MANETs", *IEEE Transactions on Mobile Computing*, vol. 8, no. 5, pp. 606-621, 2009.

[13] C. Boldrini, M. Conti, M. A. Passarella, "Modelling Data Dissemination in Opportunistic Networks", *Proc. 3rd ACM Workshop on Challenged Networks CHANTS 08*, 2008.

[14] V. Borrel, F. Legendre, M. Amorim, S. Fdida, "SIMPS: Using Sociology for Personal Mobility", *IEEE/ACM Transactions on Networks*, vol. 17, no. 3, pp. 831-842, 2009.

[15] T. Hossmann, T. Spyropoulos, F. Legendre, "Social Network Analysis of Human Mobility and Implications For DTN Performance Analysis and Mobility Modeling", *TIK-Report No. 323 Computer Engineering and Networks Laboratory, ETH Zurich*, Switzerland, 2010.

[16] T. Hossmann, T. Spyropoulos, F. Legendre, "A Complex Network Analysis of Human Mobility", *Proc. IEEE Conference on Computer Communications Workshops 2011*, Shanghai, 2011.

[17] S. Ferreti, V. Ghini, "Scale-Free Opportunistic Network: is it Possible?", *Technical Report. arXiv:1107.1937*, July 2011.

[18] R. Albert, A.L. Barabasi, "Statistical Mechanics of Complex Networks", *Reviews of Modern Physics*, vol. 74, no. 47, pp. 47-97, January 2002.

[19] J.P. Onnela, J. Saramaki, J. Hyvonen, G. Szabo, D. Lazer, K. Kaski, J. Kertesz, A.L. Barabasi, "Structure and Tie Strength in Mobile Communication Networks", *Proc. the National Academy of Sciences*, vol. 104, no.18, 2007.

[20] G. Williamson, "Routing in Human Contact Networks", *PhD thesis of National University of Ireland*, Dublin, 2010.

[21] N. Eagle, A. Pentland, D. Lazer, "Inferring Social Network Structure using Mobile Phone Data", *Proc. the National Academy of Sciences*, vol. 106, no. 36, 2009.

[22] M. Everett, S.P. Borgatti, "Ego Network Betweeness", *Social Networks*, vol. 27, no. 1, pp. 31-38, 2005.

[23] K.I. Goh, B. Kahng, D. Kim, "Universal Behavior of Load Distribution in Scale Free Networks", *Physical Review Letters*, vol. 87, no. 27, 278701, 2001.

[24] M. Barthelemy, "Comment on Universal Behaviour of Load Distribution in Scale Free Networks", *Physical Review Letters*, vol. 91, no. 18, 189803, 2003.

[25] P.L. Krapivsky, S. Redner, F. Leyvraz, "Connectivity of Growing Random Networks", *Physical Review Letters*, vol. 85, no. 21, 4629, 2000.

[26] M.E.J. Newman, "Analysis of Weighted Networks", *Physical Review E*, vol. 70, no. 5, 056131, 2004.

[27] R. Germano, A.P.S. de Moura, "Traffic of Particles in Complex Networks", *Physical Review E*, vol. 74, no. 3, 036117, 2006.

Traffic Distribution and Network Capacity Analysis in Social Opportunistic Networks

Internet Source

1%

6   ndl.ethernet.edu.et
Internet Source

1%

7   repository.exst.jaxa.jp
Internet Source

1%

8   Shanshan Lu, Yanliang Liu, Yonghe Liu, Mohan Kumar. "LOOP: A location based routing scheme for opportunistic networks", 2012 IEEE 9th International Conference on Mobile Ad-Hoc and Sensor Systems (MASS 2012), 2012
Publication

<1%

9   downloads.hindawi.com
Internet Source

<1%

10   tel.archives-ouvertes.fr
Internet Source

<1%

11   eprints.nottingham.ac.uk
Internet Source

<1%

12   J. M. Pujol. "Fair Routing in Delay Tolerant Networks", IEEE INFOCOM 2009 - The 28th Conference on Computer Communications, 04/2009
Publication

<1%

13   en.wikipedia.org
Internet Source

<1%

14 mafiadoc.com
Internet Source
<1 %

15 Christophe Diot. "Pocket switched networks and human mobility in conference environments", Proceeding of the 2005 ACM SIGCOMM workshop on Delay-tolerant networking - WDTN 05 WDTN 05, 2005
Publication
<1 %

16 www.skoltech.ru
Internet Source
<1 %

17 www.social-nets.eu
Internet Source
<1 %

18 dokumen.pub
Internet Source
<1 %

19 web-ext.u-aizu.ac.jp
Internet Source
<1 %

20 Bambang Soelistijanto. "Improving Node Popularity Calculation Using Kalman Filter in Opportunistic Mobile Social Networks", 2018 6th International Conference on Information and Communication Technology (ICoICT), 2018
Publication
<1 %

21 Kun Yang, Xueqi Cheng, Liang Hu, Jianming Zhang. "Mobile social networks: state-of-the-
<1 %

art and a new vision", International Journal of Communication Systems, 2012
Publication

22   Sourav Das, Dongjin Lee, Dae Hyun Kim, Partha Pratim Pande. "Small-World Network Enabled Energy Efficient and Robust 3D NoC Architectures", Proceedings of the 25th edition on Great Lakes Symposium on VLSI - GLSVLSI '15, 2015
Publication   <1%

23   www.tct.hut.fi
Internet Source   <1%

24   Liu Qilie, Li Xin, Li Yun, Wang Lei, Pan Yingjun. "Probabilistic Sequence Routing in intermittently connected wireless networks", 2010 2nd International Conference on Computer Engineering and Technology, 2010
Publication   <1%

25   Niu, Jianwei, Mingzhu Liu, Yazhi Liu, Lei Shu, and Dapeng Wu. "A venues-aware message routing scheme for delay-tolerant networks : Venues-aware message routing scheme", Wireless Communications and Mobile Computing, 2013.
Publication   <1%

26   Submitted to Higher Education Commission Pakistan
Student Paper   <1%

27 Stefano Ferretti. "Shaping opportunistic networks", Computer Communications, 2013
Publication

<1 %

28 hai.ics.uci.edu
Internet Source

<1 %

29 www.csroc.org.tw
Internet Source

<1 %

30 Hu, Tong, Feng Hong, Xiqing Zhang, and Zhongwen Guo. "BiBUBBLE: Social-based Forwarding in Pocket Switched Networks", 2010 7th International Conference on Ubiquitous Intelligence & Computing and 7th International Conference on Autonomic & Trusted Computing, 2010.
Publication

<1 %

31 Online Information Review, Volume 38, Issue 3 (2014-09-16)
Publication

<1 %

32 Shi Chen, Jun Zhang, Qiang Gao. "An efficient hybrid routing based on contact history in delay tolerant networks", 2010 Seventh International Conference on Wireless and Optical Communications Networks - (WOCN), 2010
Publication

<1 %

33 citeseerx.ist.psu.edu
Internet Source

<1 %

34 reality.media.mit.edu
Internet Source
<1 %

35 www0.cs.ucl.ac.uk
Internet Source
<1 %

36 B. Jedari, F. Xia, A. Mohammed Ahmed, P. Pirozmand, Y. Najaflou. "Using social network analysis (SNA) to design socially aware network solutions in delay-tolerant networks (DTNs)", Elsevier BV, 2021
Publication
<1 %

37 etheses.whiterose.ac.uk
Internet Source
<1 %

38 Flavio Fabbri, Roberto Verdone. "A Sociability-Based Routing Scheme for Delay-Tolerant Networks", EURASIP Journal on Wireless Communications and Networking, 2010
Publication
<1 %

39 Kuang Xu, Guang-Hua Yang, Victor O.K. Li, Shu-Yan Chan. "Detecting dynamic communities in opportunistic networks", 2009 First International Conference on Ubiquitous and Future Networks, 2009
Publication
<1 %

40 Xiaoguang Fan, Victor O. K. Li, Kuang Xu. "Fairness Analysis of Routing in Opportunistic Mobile Networks", IEEE Transactions on Vehicular Technology, 2014
<1 %

| 41 | ijcai.org
Internet Source | <1% |
| --- | --- | --- |
| 42 | lion.sjtu.edu.cn
Internet Source | <1% |
| 43 | nozdr.ru
Internet Source | <1% |
| 44 | www.mattstabeler.co.uk
Internet Source | <1% |
| 45 | Eira, Antonio, Joao Pedro, and Joao Pires. "Cost-optimized dimensioning of translucent WDM networks with Mixed-Line-Rate spectrum-flexible channels", 2012 IEEE 13th International Conference on High Performance Switching and Routing, 2012.
Publication | <1% |
| 46 | Suvadip Batabyal, Parama Bhaumik. "Mobility Models, Traces and Impact of Mobility on Opportunistic Routing Algorithms: A Survey", IEEE Communications Surveys & Tutorials, 2015
Publication | <1% |
| 47 | ipfs.io
Internet Source | <1% |
| 48 | ns2.thinkmind.org
Internet Source | <1% |

49 siti2.ulusofona.pt
Internet Source <1%

50 www.eurecom.fr
Internet Source <1%

51 www.hindawi.com
Internet Source <1%

52 www.research-collection.ethz.ch
Internet Source <1%

53 Jingwei Miao, Omar Hasan, Sonia Ben Mokhtar, Lionel Brunie, Kangbin Yim. "An Analysis of Strategies for Preventing Selfish Behavior in Mobile Delay Tolerant Networks", 2012 Sixth International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing, 2012
Publication <1%

54 Milind R. Penurkar, Umesh A. Deshpande. "CONHIS: Contact history-based routing algorithm for a vehicular delay tolerant network", 2014 Annual IEEE India Conference (INDICON), 2014
Publication <1%

55 Socievole, Annalisa, and Salvatore Marano. "Evaluating the impact of energy consumption on routing performance in delay tolerant networks", 2012 8th International Wireless
<1%

Communications and Mobile Computing Conference (IWCMC), 2012.
Publication

56 Zhensheng Zhang. "Routing Protocols in Intermittently Connected Mobile Ad Hoc Networks and Delay-Tolerant Networks", Algorithms and Protocols for Wireless and Mobile Ad Hoc Networks Boukerche/Mobile, 03/26/2008
Publication

<1 %

Exclude quotes   On   Exclude matches   < 5 words
Exclude bibliography   On