Social-Aware Energy-Efficient Data Dissemination with D2D Communications

Yiming Zhao and Wei Song
Faculty of Computer Science
University of New Brunswick, Fredericton, Canada
Email: {ym.zhao, wsong}@unb.ca

***Abstract*—With the high penetration rate of smart mobile devices, it is appealing to exploit device-to-device (D2D) communications for data dissemination, *e.g.*, in disaster alerts and event notifications. The popularity of social networks also offers good opportunities to improve the efficiency of data dissemination. Though there have been some existing works on data dissemination with D2D and mobile social networks, many focus on mitigating the D2D co-channel interference to achieve high resource utilization. As mobile devices are power-limited, it is important to consider the energy efficiency and finishing time in data dissemination. In this paper, we aim at developing an effective solution for D2D data dissemination to balance between total energy consumption and transmission completion time. In particular, we propose novel algorithms for seed selection and transmission scheduling with a single seed or multiple seeds. Simulation results demonstrate that our solution outperforms two reference schemes in total energy consumption while achieving a good balance for transmission completion time.**

***Index Terms*—Data dissemination, D2D communications, social networks, energy efficiency.**

I. Introduction and Related Work

Data dissemination aims at delivering information to a group of target users in a geographical region. It has a wide range of applications, such as in disaster alert, event notification, and advertisement distribution. With the proliferation of mobile devices and evolution of wireless networks, device-to-device (D2D) communications offer promising paradigms for energy-efficient data dissemination. In D2D communications, mobile devices communicate with each other directly in a peer-to-peer (P2P) fashion bypassing the base station (BS). Since short-range communications can typically achieve high data rates with low transmit power, mobile devices can save much energy using D2D communications.

In the literature, there have been many studies on data dissemination in wireless networks. Many early works focus on data dissemination in opportunistic networks, in which the connectivity between devices is intermittent and a store-carry-and-forward method is used to exploit opportunistic contacts for delivering data between devices. In [1], a content-centric framework is proposed to facilitate data dissemination in opportunistic networks based on the characteristics of disseminated contents. Meanwhile, social relationships among end users can be further considered to improve data dissemination, *e.g.*, in mobile social networks [2]. In [3], Ioannidis *et al.* study the effect of weak ties in social networks on disseminating data in mobile social networks. In [4], Sun *et al.* propose a data dissemination scheme with D2D communications underlaying the cellular network, using a coalitional graph game. It aims to maximize the system sum rate when D2D is used to disseminate data among socially connected users.

In this work, we focus on two other important factors for data dissemination in wireless networks, *i.e.*, energy efficiency and transmission finishing time. In particular, we utilize social network information to jointly select an initial subset of D2D users as seeds for receiving data directly from the BS and schedule the further transmission among D2D users. As data dissemination consumes energy and bandwidth resources, people are generally more willing to share data with friends. To accommodate such incentive constraint, we limit the scope of D2D transmission to users with social ties as in [4]. A minimum spanning tree is used to find out the D2D links with the minimum average energy consumption so as to effectively perform seed selection. As for the transmission scheduling, the effect of parallel transmission is also explored to reduce the finishing time of data dissemination. The simulation results show that our solution achieves a good balance between energy consumption and finishing time in the single-seed and multiple-seed scenarios.

The remainder of this paper is organized as follows. In Section II, we give the system model for data dissemination with D2D communications. Section III introduces our proposed solution for seed selection and transmission scheduling. In Section IV, we present simulation results comparing our solution with reference works. Section V concludes this paper.

II. System Model

A. Data Dissemination Model

Consider a data dissemination scenario depicted in Fig. 1. The BS is requested to disseminate some data to a set of n users in an area, denoted by V. The BS first chooses a subset of those users as *seeds*, denoted by $S \subset V$, and then multicasts the data to the selected seeds at an authorized frequency channel. After that, the seeds forward the data to other users by D2D communications underlaying the cellular network, while any user that receives the data can further disseminate the data to others similarly via D2D links.

This research was supported by Natural Sciences and Engineering Research Council (NSERC) of Canada and NSERC Strategic Network FloodNet.

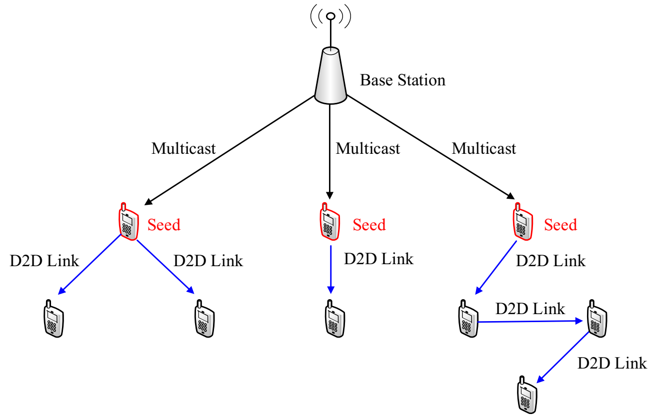


Fig. 1. System model.

To reduce co-channel interference, the D2D links share the uplink spectrum of cellular users. When the D2D underlay link from i to j shares the uplink channel of cellular user ℓ, there exists interference between D2D user j and cellular user ℓ. Then, the received signal at D2D user j can be written as

$$y_j = \sqrt{P_{t,i}} h_{ij} x_i + \sqrt{P_{t,\ell}} h_{\ell j} x_\ell + n_j \qquad (1)$$

where x_i and x_ℓ are the signals from i and ℓ transmitted with power $P_{t,i}$ and $P_{t,\ell}$, respectively, h_{ij} and $h_{\ell j}$ are the channel responses of links from i to j and from ℓ to j, respectively, and n_j is the additive white Gaussian noise (AWGN) at j. The second term in (1) is the interference from ℓ to j. Thus, the channel rate of the D2D link from i to j is given by

$$R_{ij} = B_{ij} \cdot \log_2 \left(1 + \frac{P_{t,i} \cdot |h_{ij}|^2}{P_{t,\ell} \cdot |h_{\ell j}|^2 + |n_j|^2}\right) \qquad (2)$$

where B_{ij} is the bandwidth of the resource blocks consumed by the D2D link from i to j.

Consider a Rayleigh fading channel model with log-distance path loss. Then, the received signal power at j, $P_{t,i}|h_{ij}|^2$, and the power of interference from cellular user ℓ, $P_{t,\ell}|h_{\ell j}|^2$, are exponentially distributed with mean $P_{t,i} d_{ij}^{-\alpha}$ and $P_{t,\ell} d_{\ell j}^{-\alpha}$ respectively, where d_{ij} is the distance between i and j, $d_{\ell j}$ is the distance between ℓ and j, and α is the path loss exponent. Here, we assume that a message of size F is transmitted in one time slot of length τ. Then, the data rate demand for the message is a constant, *i.e.*, $R_{ij} = F/\tau$. Accordingly, we can obtain the required transmit power from i to j, $P_{t,i}$.

B. Social Network Model

To accommodate user incentive constraint, we assume that D2D data dissemination only occurs between two socially connected users whose physical distance and D2D channel conditions can meet the data rate requirement. To characterize the social relationships among D2D users, we use *the caveman model* [5], a classic social network model which can generate synthetic social networks with realistic characteristics.

The caveman model starts with K isolated complete graphs, also known as *caves*, in which every vertex is adjacent to every other vertex. Then in a rewiring stage, every edge of a cave in the original network is randomly rewired by pointing to a node in another cave with probability p. The rewiring procedure intends to establish random inter-connections between individual nodes of different caves. Fig. 2 shows an example to generate a synthetic social network with the caveman model.

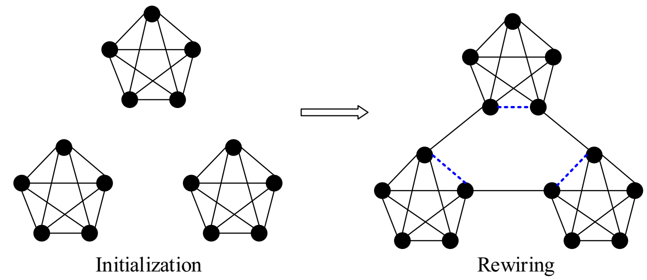


Fig. 2. Generation of a social network using the caveman model (adapted from [6]): (a) Initial network topology with 3 caves of size 5. (b) Generated social network after the rewiring stage.

As discussed in [6], the individual nodes of one cave are closely connected, while individual nodes of different caves are sparsely connected. That is, the generated social network is globally sparse but locally dense. Therefore, the social networks generated by the caveman model are featured with high clustering coefficients and low average path length. It has been proved in [5] that social networks based on this model are very close to real ones.

C. Design Objective

Given the system model in Section II-A, we can model the D2D communications among the users with social connections as a bidirectional graph $G(V, E, w)$, where V is the set of D2D users, E denotes the available D2D links between the D2D users, and each edge is labeled by weight $w(e_{ij})$, which gives the transmit power for the D2D transmission from user i to user j. Here, we aim to design a data dissemination solution, which can properly select the seeds and schedule the D2D transmission among the D2D users so as to reduce the total energy consumption of D2D transmission and the final finishing time of data dissemination, *i.e.*,

$$\min_{E' \subseteq E} \sum_{e_{ij} \in E'} w(e_{ij}) \qquad (3)$$

$$\min_{Q \in \mathcal{Q}} \max\{t_i\}, i \in V \qquad (4)$$

where t_i denotes the receiving time of user i following the transmission sequence Q. Since the two performance metrics are not independent in our scenario, it is very difficult or even impossible to minimize the total energy consumption and completion time simultaneously. Therefore, we set reducing the total energy consumption as our primary goal while achieving a balance with the transmission completion time.

III. Social-Aware Energy-Efficient Data Dissemination

As seen in the problem discussed in Section II, it is quite complex to jointly select the initial seeds and coordinate further data dissemination so as to minimize the total energy consumption and transmission completion time. Particularly, in the wireless network with limited energy budget, it is preferable to have a lightweight approximate solution which can approach the best performance. Here, we first propose a data dissemination algorithm for the single-seed scenario and then extend it to the multiple-seed scenario.

A. Single-Seed Scenario

Alg. 1 presents a single-seed tree-based dissemination (SSTBD) algorithm for selecting one seed in the bidirectional graph $G(V, E, w)$ and scheduling the data transmission. SSTBD consists of two sub-procedures, SSA and TRA, specified in Alg. 2 and Alg. 3 respectively, which correspond to two individual stages, seed selection and transmission scheduling. For easy presentation, we assume that graph G is connected. Our algorithm also works when G is not connected by regarding each connected component as an individual graph.

Algorithm 1 SSTBD($G(V, E, w)$).

Input: $G(V, E, w)$
Output: S, $(t_1, t_2, ..., t_n)$
$Tree_s \leftarrow$ SSA($G(V, E, w)$)
$(t_1, t_2, ..., t_n) \leftarrow$ TRA($Tree_s$)
return $S =$ getRootNode($Tree_s$), $(t_1, t_2, ..., t_n)$

For the SSA algorithm, it first converts the bidirectional graph $G(V, E, w)$ into an undirected graph $G'(V, E', w')$ by combining each pair of original edges $e_{ij}, e_{ji} \in E$ into an edge $e'_{ij} \in E'$ using the geometric mean $w'(e'_{ij}) = \sqrt{w(e_{ij}) \cdot w(e_{ji})}$, which can reflect the average energy consumption of the D2D link between user i and user j. Then the minimum spanning tree T of G' is generated to find out the D2D links with the minimum average energy consumption. After that, tree T is restored to a bidirectional graph $G^*(V, E^*, w)$ by replacing edge e'_{ij} in T with the original edges e_{ij} and e_{ji} in G. Then for each node i in G^*, we build a transmission tree, $Tree_i$, which is rooted at i and spans over G^*, and obtain the total energy consumption by summing up the weights of edges in $Tree_i$. Finally, the node which achieves the minimum total energy consumption is selected as the seed.

As seen in Alg. 2, the seed selection is performed on the bidirectional graph $G^*(V, E^*, w)$ generated from tree T which is connected and acyclic. Once the seed is selected, the data transmission paths in G^* are determined. The transmission paths are actually a directed tree rooted at the seed, in which a directed edge e_{ij} represents that node i transmits data to node j. The finishing time of data dissemination depends on the time order that a node transmits the data to its children in the tree. Fig. 3 shows an example that the finishing time is decreased when each node first transmits data to its child node in a subtree of more descendants and higher depths.

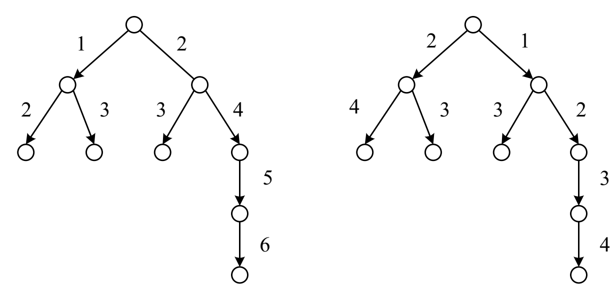
Here, we consider two important factors in determining the transmission order, *i.e.*, the number of descendants of node i in tree $Tree_s$, $descendant(i)$, and the depth of the subtree rooted at that node, $depth(i)$. Therefore, in Alg. 3, we determine the transmission order using an influence score for each node i in G^*, which is defined as a weighted sum [7] of the two attributes $descendant(i)$ and $depth(i)$, *i.e.*,

$$infScore_i = \beta \cdot descentant(i)_{norm} + (1 - \beta) \cdot depth(i)_{norm}.$$

Here, $descentant(i)_{norm}$ and $depth(i)_{norm}$ are the normalized values of the two attributes, while β $(0 < \beta < 1)$ is a weighting factor. Thus, each node transmits the data to its children following the ranking of their influence scores.

Algorithm 2 SSA($G(V, E, w)$).

Input: $G(V, E, w)$
Output: $Tree_s$
$E' = \emptyset$
for $e_{ij}, e_{ji} \in E$ **do**
  // *Define a new weight for each undirected edge*
  $w'(e'_{ij}) = \sqrt{w(e_{ij}) \cdot w(e_{ji})}$
  $E' = E' \cup e'_{ij}$
end for
Define graph $G'(V, E', w')$
// *Compute minimum spanning tree of* G'
$T =$ GetMinSpanTree($G'(V, E', w')$)
$E^* = \emptyset$
for $e'_{ij} \in T$ **do**
  // *Replace undirected* e'_{ij} *by directional* e_{ij} *and* e_{ji}
  $E^* = E^* \cup e_{ij} \cup e_{ji}$
end for
Define graph $G^*(V, E^*, w)$
for $i \in V$ **do**
  Generate tree $Tree_i$ rooted at node i based on G^*
  $totalEnergy_i = \sum_{i,j \in T_i} w(e_{i,j})$
end for
$s = \text{argmin}_{i \in V} totalEnergy_i$
return $Tree_s$



(a) Finishing time = 6 (b) Finishing time = 4

Fig. 3. An illustrative example on the impact of transmission order. The label on each edge is the time slot in which the transmission occurs.

B. Multiple-Seed Scenario

In this section, we extend Alg. 1 to the multiple-seed scenario with k seeds $(k \geq 2)$. Alg. 4 gives a multiple-seed tree-based dissemination (MSTBD) algorithm. The main rationale of MSTBD is to break $(k - 1)$ bidirectional edges in G to convert the original graph into k connected subgraphs and then perform single-seed selection and transmission scheduling

Algorithm 3 TRA($Trees_s$)

Input: $Trees_s$
Output: (t_1, t_2, ..., t_n)

for $i \in Trees_s$ **do**
  $N_i \leftarrow$ getChildNode(i)
  for $n_i \in N_i$ **do**
    $descendant(n_i) \leftarrow$ number of descendants of n_i
    $depth(n_i) \leftarrow$ depth of subtree rooted at n_i
  end for
  $maxDesc(i) = \max_{n_i \in N_i} descendant(n_i)$
  $maxDepth(i) = \max_{n_i \in N_i} depth(n_i)$
  for $n_i \in N_i$ **do**
    $infScore_{n_i} \leftarrow \beta \cdot descendant(n_i)/maxDesc(i) + (1-\beta) \cdot depth(n_i)/maxDepth(i)$
  end for
end for
$t_{seed} \leftarrow 0$
for $i \in Trees_s$ **do**
  Sort N_i into L_i in descending order of $infScore_{n_i}$
  for $n_i \in N_i$ **do**
    t_{n_i} = t_i + n_i's position number in L_i
  end for
end for

Algorithm 4 MSTBD($G(V, E, w), k$)

Input: $G(V, E, w), k$
Output: $S = \{s_1, ..., s_k\}$, (t_1, t_2, ..., t_n)

$l = 0$ *// l is index of choices for breaking edges*
for $E_{k-1} \subseteq E$ **do** *// E_{k-1} is any $(k-1)$ edges in E*
  $l = l + 1$
  Generate k subgraphs $G_g^l(V_g^l, E_g^l, w_g^l)$, $1 \leq g \leq k$ by breaking edges in E_{k-1}
  $|\overline{V}^l| = \frac{1}{k} \sum_{g=1}^{k} |V_g^l|$
  $Var^l = \frac{1}{k} \sum_{g=1}^{k} (|V_g^l| - |\overline{V}^l|)^2$
end for
$z = \text{argmin}_{1 \leq m \leq l} Var^m$
for $g = 1 : k$ **do**
  $(s_g, (t_{g_1}, ..., t_{g_{|V_g^z|}})) \leftarrow$ SSTBD($G_g^z(V_g^z, E_g^z, w_g^z)$)
end for
return $S = \{s_1, ..., s_k\}$, (t_1, t_2, ..., t_n)

in each subgraph using SSTBD. Intuitively, the total energy consumption can be potentially reduced by breaking edges of large weights. However, it may be detrimental to the transmission finishing time. For example, the subgraphs can be quite unbalanced such that the parallelization with multiple seeds cannot be effectively utilized to reduce the completion time. Since we intend to balance between the total energy consumption and the transmission finishing time, we use an alternative approach which breaks the edges such that the variance of the size of so generated k subgraphs is minimized.

IV. Simulation Results

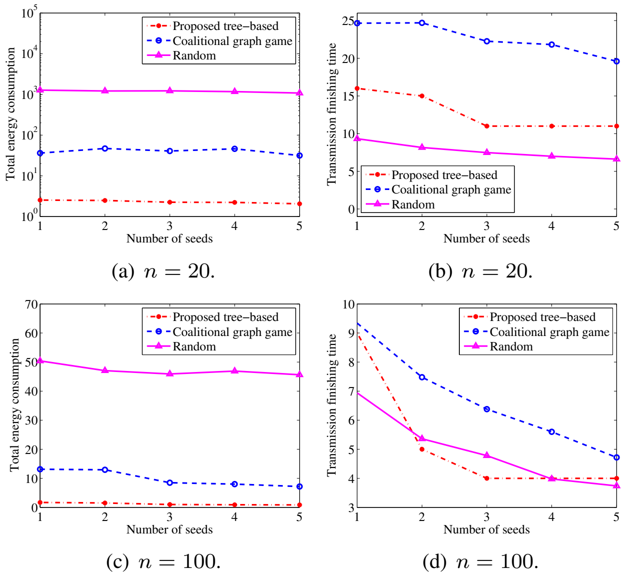
To evaluate the performance of our proposed algorithms, we conduct computer simulations with a 200m×200m square region in which cellular users and D2D users are uniformly deployed. The number of cellular users is fixed at 150, while the number of D2D users varies in the simulations. Each D2D user is assigned an arbitrary cellular uplink channel for its D2D communications. The D2D transmission range is set to 100m. We use the caveman model with different rewiring probabilities to generate the social relationships among D2D users.

Two reference data dissemination schemes are compared with our proposed algorithms. First, we consider a random approach, which randomly selects the seed(s), and then have each user who has received the message forward the data to an arbitrary friend who is within the D2D range but has not received the data. This procedure continues until all users have received the data. Second, we consider a coalitional graph game approach based on [4]. Here, the first k users with the largest number of friends within the D2D range are selected as seed(s). To coordinate the D2D data dissemination, a coalitional graph game is formulated to build a transmission graph $\hat{G}(V, \hat{E})$ from the initial bidirectional graph $G(V, E)$. Different from [4], we define the utility function as $U(\hat{G}) = \sum_{e_{ij} \in \hat{E}} (m_i - m_j)(w_{max} - w(e_{ij}))$, where m_i (or m_j) is a binary indicator denoting whether i (or j) has possessed the data, and w_{max} is the maximum weight of edges in E.

First, we run simulations with different numbers of D2D users and seeds. The social relationships are modeled by the caveman model with two caves and a rewiring probability $p = 0.3$. The size of each cave is adapted with the number of D2D users. As the random approach and coalitional graph game based approach involve random decisions, we run the corresponding simulations for 50 times to remove the randomness effect. Fig. 4 shows the results of total energy consumption and finishing time for different scenarios. As seen in Fig. 4(a) and Fig. 4(c), the proposed tree-based algorithms (SSTBD and MSTBD) can effectively reduce the total energy consumption when the number of D2D users is scaled up.

Moreover, when the network size is small ($n = 20$), Fig. 4(b) shows that the finishing time of our solution is relatively lower than that of the coalitional graph game based approach. Though the random approach achieves the lowest finishing time, it suffers a significantly higher energy cost. In the random approach, all users who have received the message randomly transmit the data to their friends that they have not transmitted to. As a result, the number of receiving users will grow at a nearly exponential rate when the network is connected. However, the random approach has the highest energy consumption to attain the lowest finishing time. Nonetheless, when there is a large D2D user population and more than one seed, Fig. 4(d) shows that our solution achieves almost the lowest finishing time. This is mainly because our solution can effectively balance the energy cost and transmission time in MSTBD when building the subgraphs for the seeds.

Further, we test the performance of the data dissemination algorithms with different social connection structures. Here, we fix the number of D2D users to 50, the number of caves



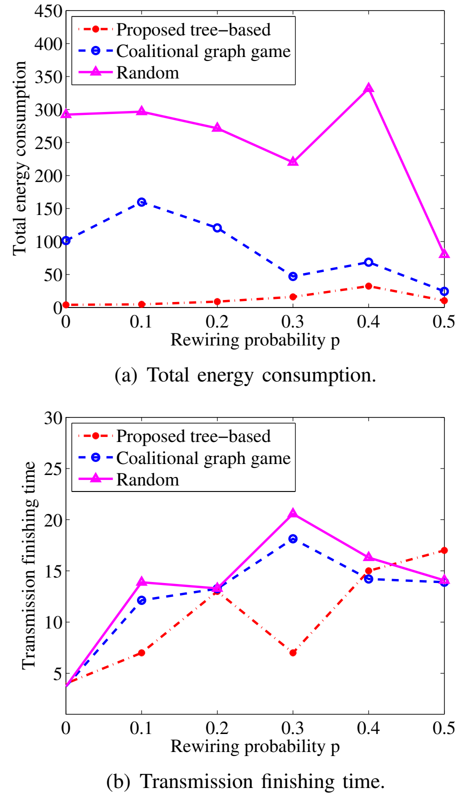
(a) $n = 20$. (b) $n = 20$.

(c) $n = 100$. (d) $n = 100$.

Fig. 4. Comparison of data dissemination algorithms with different network sizes and numbers of seeds.

to 10, and the size of each cave to 5. Then, we change the rewiring probability p ranging from 0 to 0.5 to generate social networks of different characteristics. In particular, the number of social communities varies with the rewiring probability p. With a larger value of p, the D2D users become more densely connected in the social network and there are less isolated social communities. Based on the so generated social networks, we conduct simulations to compare the three data dissemination approaches again.

Fig. 5 shows the total energy consumption and finishing time with different social structures. Here, we set the number of seeds selected for each D2D cluster to 1. As seen in Fig. 5(a), our solution always achieves the lowest total energy consumption for different values of p. In addition, with the increase of p, the total energy consumption of our algorithm only varies slightly, while the other two schemes fluctuate dramatically. This implies that our solution is more tolerant of variations of social structures. In Fig. 5(b), we can see that our algorithm also effectively reduces the finishing time of data dissemination. Specifically, when the rewiring probability $p < 0.4$, the finishing time of our algorithm is the lowest among the three schemes. It is worth mentioning that the sharp drop of finishing time at $p = 0.3$ is due to the particular network topology for this case. Because the network topology is determined by the physical distance between D2D users and the rewiring probability p, this case happens to result in very balanced transmission trees of similar size for different D2D clusters. As a result, a high level of parallelization leads to quite low finishing time.



(a) Total energy consumption.

(b) Transmission finishing time.

Fig. 5. Comparison of data dissemination algorithms with different social structures.

V. Conclusion and Future Work

In this paper, we have studied how to take advantage of D2D communications and social connections to achieve energy-efficient data dissemination. The proposed solution can effectively balance between energy consumption and transmission finishing time via seed selection and transmission scheduling. In the simulations, we validate the effectiveness of our algorithms and compare them with two reference data dissemination schemes. The simulation results show that our solution can improve the two aspects in various scenarios with different network sizes and social structures. In the future, we are interested in extending this work to more dynamic scenarios with user mobility. It is promising to further enhance the data dissemination efficiency by jointly exploiting social-awareness and mobility patterns.

References

[1] L.-J. Chen, C.-H. Yu, C.-L. Tseng, H.-H. Chu, and C.-F. Chou, “A content-centric framework for effective data dissemination in opportunistic networks,” *IEEE J. Select. Areas Commun.*, vol. 26, no. 5, pp. 761–772, 2008.

[2] N. Kayastha, D. Niyato, P. Wang, and E. Hossain, “Applications, architectures, and protocol design issues for mobile social networks: A survey,” *Proceedings of the IEEE*, vol. 99, no. 12, pp. 2130–2158, 2011.

[3] S. Ioannidis and A. Chaintreau, “On the strength of weak ties in mobile social networks,” in *Proceedings of ACM EuroSys Workshop on Social Network Systems*, 2009, pp. 19–25.

[4] Y. Sun, T. Wang, L. Song, and Z. Han, “Efficient resource allocation for mobile social networks in D2D communication underlaying cellular networks,” in *Proc. IEEE ICC*, 2014, pp. 2466–2471.

[5] D. J. Watts, *Small Worlds: The Dynamics of Networks Between Order and Randomness*. Princeton University Press, 1999.

[6] M. Musolesi and C. Mascolo, “Designing mobility models based on social network theory,” *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 11, no. 3, pp. 59–70, 2007.

[7] R. V. Rao, “Introduction to multiple attribute decision-making (MADM) methods,” *Decision Making in the Manufacturing Environment: Using Graph Theory and Fuzzy Multiple Attribute Decision Making Methods*, pp. 27–41, 2007.