

An Efficient Data Dissemination Method over Wireless Ad-Hoc Networks

He Li · Kyungsoo Bok · Kyungyong Chung · Jaesoo Yoo

Published online: 2 March 2014
© Springer Science+Business Media New York 2014

Abstract Recently, mobile data is increased sharply with the rapid development of mobile devices such as smart phone, tablet, and pocket computer. Due to mobile devices are becoming more ubiquitous and most of the mobile devices are equipped with short range wireless communication technologies, the research on data dissemination over mobile devices based mobile ad-hoc networks becomes increasingly popular. Because most of the previous data dissemination methods over mobile ad-hoc networks did not consider user preferences, they resulted in a lot of unnecessary data transmissions in the network. Although some recent studies investigated the importance of satisfying user preference during data dissemination over mobile ad-hoc networks, they either required high network management overhead or could not guarantee that the data packets can be disseminated to all of the mobile nodes that are interested in the data. In this paper, we propose an efficient method which leverages user interests and location information to support data dissemination in wireless ad hoc networks. We propose two relay nodes selection method to filter out unnecessary data transmissions in the network, such as passive relay nodes selection method and active relay nodes selection method. Compared with existing methods, the proposed data dissemination method reduces a lot of unnecessary data transmissions without requiring high network management overhead. The experimental results show that the proposed method outperforms the existing methods.

H. Li · K. Bok · J. Yoo (✉)
Department of Information and Communication Engineering, Chungbuk National University,
Cheongju 361-763, South Korea
e-mail: yjs@cbnu.ac.kr

H. Li
e-mail: lihelol@gmail.com

K. Bok
e-mail: ksbok@cbnu.ac.kr

K. Chung
School of Computer Information Engineering, Sangji University, Wonju 220-702, South Korea
e-mail: kyungyong.chung@gmail.com

Keywords Data dissemination · User interests · Location-aware service · Mobile data · Wireless network

1 Introduction

A mobile ad hoc network is a temporary self-configuring network composed of mobile devices without using any centralized infrastructure. In mobile ad hoc networks, the mobile devices can communicate with each other by using short range wireless communication protocols, such as Wi-Fi and Bluetooth. Due to the mobile devices such as smart phones, tablets, pocket PC and specific sensors have been becoming more ubiquitous, a lot of mobile applications over mobile ad hoc networks have been emerged [1–3]. And there is an increasing interest of mobile devices towards the support of mobile services. Data dissemination is important for many mobile applications of mobile ad hoc networks, such as event notification, contents sharing and advertisements. Mobile users increasingly expect to utilize Wi-Fi and Bluetooth based technologies to share contents based on their personal preferences. For instance, students can use notification service to share study and entertainment information with the other students in the campus by using their mobile devices. We consider a following real case: assume a conference about computer science is held in a hotel. The seminars about different research areas of computer science are processed in different seminar rooms. Generally, the participated scholars will select their interesting seminars to participate. The scholars can register a mobile ad hoc network service to share contents with the other scholars that have same research areas by using their mobile devices.

Because the communication range is limited, a data packet from a mobile device is not sufficient to reach all the other mobile devices. Therefore, the data dissemination can be performed across multiple hops in mobile ad hoc networks. Due to the lack of a fixed network infrastructure, a straightforward approach for data dissemination in mobile ad hoc networks is flooding [4], in which each mobile node in the network rebroadcasts data dissemination packets that it received. However, it ignored the satisfaction of user preferences. The data packets are disseminated to many mobile nodes that are not interested in the data, which cause a lot of unnecessary data transmissions. Recently, some social based data dissemination schemes [5–10] have been designed, in which each mobile node forwards data dissemination packets by considering user interests. However, high network management overheads make them not suitable for mobile ad hoc networks.

In this paper, we propose an efficient data dissemination method with low network management overhead. We target on disseminating data objects to as many mobile users who are interested in the data as possible while reducing the unnecessary data transmissions. Our key idea is to reduce the number of unnecessary relay nodes to minimize the unnecessary data transmissions in the network. Therefore, the problem of our proposed method is how to design an appropriate relay node selection strategy in a mobile ad hoc network. It is desirable from a mobile ad hoc network to select minimum relay nodes that can distribute data dissemination packets to all the mobile nodes that are interested in the data. However, [11, 12] have showed that these problems are NP-hard. In our proposed method, a decentralized solution was proposed to prune the unnecessary relay nodes. Because the absence of the global knowledge of mobile ad hoc networks, designing an efficient relay nodes selection method is a challenge. We assume that each mobile node holds its local information of interests and location. The information of 1-hop neighbor nodes of a mobile node can be obtained by exchanging ‘hello’ messages. Generally, the location information can be obtained by GPS functions [13] equipped with mobile devices. The unnecessary relay nodes can be identified

according to the interests and location information of mobile nodes. In this paper, we propose two methods to filter out the unnecessary relay nodes, passive relay nodes selection method and active relay nodes selection method. In the passive relay nodes selection method, the data transmission nodes determine the next relay nodes and prune the unnecessary relay nodes. In the active relay nodes selection method, the data recipient nodes determine whether they have to act as relay nodes. By doing this, the mobile nodes that are not interested in the disseminated data and do not affect the other mobile nodes receiving the data are pruned passively or actively. The proposed methods are decentralized, which are performed at mobile nodes in the network. Since a lot of unnecessary mobile nodes can be filtered out participating in data dissemination, the amount of unnecessary data transmissions is reduced. The experimental evaluation shows that the proposed method supports high data dissemination reachability and low network management overhead while decreasing a lot of unnecessary data transmissions.

The remainder of this paper is organized as follows. Section 2 reviews related works about data disseminations. Section 3 presents the preliminaries and motivations of this paper. In Sect. 4, we describe two methods for relay nodes selection. Section 5 discusses the performance evaluation of the proposed methods and existing methods. Finally, Sect. 6 concludes this paper.

2 Related Work

Data disseminations in mobile ad hoc networks have been received much attention in recent years due to the fast increase of ubiquitous mobile devices and the emerging wireless communication techniques. Mobile ad hoc networks integrated with 3G cellular networks to offload the traffic of data dissemination or advertisement broadcasting [14] have been also emerged. Due to the lack of a fixed communication infrastructure of mobile ad hoc networks [15–19], the way that data are transmitted in mobile ad hoc networks is quite different from the way that data are transmitted in wired networks.

Flooding as a basic method for data dissemination in mobile ad hoc networks has attached a lot of attentions. However, the major problem of flooding is the excessive number of redundant retransmissions, which may lead to consuming scarce resources such as bandwidth and power and broadcast storm problems [4]. In order to reduce the number of redundant transmissions of flooding methods, the global infrastructure based algorithms [11, 20–22] were proposed. However, the global infrastructure of a mobile ad hoc network is difficult to be obtained. The 2-hop neighbor nodes based algorithms [12, 23, 24] have been proposed to prune the dominated mobile nodes. Although some redundant transmissions can be reduced, the management overhead of the 2-hop neighborhood information of each mobile node is very large, especially when the network is dynamic or dense.

The other flooding based algorithms [4, 25–28] such as probability scheme, counter-based scheme, and distance-based scheme have also been proposed to reduce the number of redundant transmissions in the network. Instead of simply rebroadcasting the data dissemination packet to every mobile node, each mobile node rebroadcasts data dissemination packets with a certain probability. Therefore, the major problem is that some mobile nodes in the network may fail to receive the required data. Then, we can say that the data dissemination reachability is lost in these methods. In Edison et al. [29], mobile agents are used to disseminate data in mobile ad hoc networks. The geographic information is used to support the agents' decision. However, since most of the previous studies did not consider user preferences, much unnecessary data transmissions are generated during performing a data dissemination task.

Pietilainen and Diot [30] has studied the importance of social principle for efficient data dissemination based on four experimental data sets on human mobility. Li and Wu [31], Yoneki et al. [32] have been proposed to disseminate the data packet only to the mobile nodes that are interested in the data. Recently, some social based data forwarding schemes [6, 8–10] consider some social network properties including centrality, social behavior and communities have been proposed. In these methods, each mobile node forwards the data dissemination packet to the mobile nodes with high probability that belong to the same community. Boldrini et al. [7] proposed to disseminate data to the regions of the network where interested users are present. However, these approaches were based on the assumption of a global infrastructure, which is difficult to be obtained in mobile ad hoc networks.

A user-centric data dissemination method [5] which considers user interests was proposed. It disseminates data objects to the selected relay nodes. However, the effectiveness of relay nodes selection depends on the scope (the number of multiple hops) of the network information maintained at each individual mobile node. This will incur unaffordable network maintenance costs. On the other hand, some mobile nodes that are interested in the data may fail to receive the data dissemination packet. A preference-aware content dissemination method that targets on maximally satisfying users' preferences for content objects was proposed in Lin et al. [33]. However, this work focuses on selecting a suitable set of data sources and does not discuss how to select forwarding nodes for data dissemination. The goal is different compared with our proposed method.

3 Preliminaries and Motivations

In this section, we first introduce the network model and user interests description that used in this paper, and then we demonstrate the problems of data dissemination in mobile ad hoc networks and give the motivations of the proposed method.

3.1 Preliminaries

We assume that a mobile ad hoc network consists of a set of mobile nodes that communicate with each other by using short range wireless technologies. A mobile ad hoc network can be represented by a graph $G = (V, E)$, the mobile nodes are represented as a vertex set V in the graph and the links between two connected mobile nodes are presented by the edge set E . The transmission radius r of each mobile node can be considered as a circle around it. The connectivity between two mobile nodes can be determined by using locations and transmission radius. If a link exists between two mobile nodes it can be identified that the distance between them is less than r , and two mobile nodes are considered as 1-hop neighbor nodes. The neighbor nodes set $N(u)$ is used to represent all the 1-hop neighbor nodes of mobile node u . According to periodically exchanging 'hello' messages, each mobile node can obtain the information of its 1-hop neighbor nodes. The participated mobile nodes in the network may produce data objects and may also be interested in receiving data from the other mobile nodes in the network. The data dissemination among mobile nodes in mobile ad hoc networks is achieved either through 1-hop transmissions or through multi-hops transmissions by intermediate relay nodes. Without loss of generality, in this paper, we assume that each mobile node has been equipped with the same wireless capability and therefore the wireless transmission radius between mobile nodes is same.

The interest probability of a mobile node in a data object is estimated from user interest and data object descriptions. We assume that each mobile node has its local preference

information initially and each mobile node is allowed to have various interests. The interest information of each mobile node is expressed by interest keywords.

Definition 1 The **interest keywords** of each mobile node can be represented by a tuple $K(k_1, k_2, \dots, k_n)$, where k_n denotes the n th interest keyword of a mobile node. For each interest keyword $k_i \in K$, there exists a weight value w_i , which indicates the importance of keyword k_i in user’s interest keywords list. We define that $\sum_{i=1}^n W_i = 1$, where m denotes the number of the interest keywords contained by a mobile node.

A data object is described by a keywords list $K(k_1, k_2, \dots, k_m)$. The interest probability of a mobile node in a data object can be determined by keywords matching. Let $f(k_i, k_j)$ be the matching value between two keywords k_i and k_j .

$$f(k_i, k_j) = \begin{cases} 0, & \text{if } k_i \cap k_j = \emptyset \\ 1, & \text{if } k_i \cap k_j \neq \emptyset \end{cases} \tag{1}$$

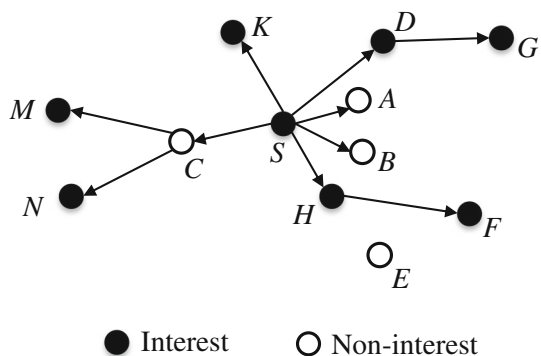
$k_i \cap k_j = \emptyset$ means that two keywords k_i and k_j are not matched and the result of $f(k_i, k_j)$ equals 0. Otherwise, the result equals 1. The interest probability of mobile node u in data object v can be evaluated by computing the summation of the matching values of the keywords. m and n denote the number of keywords contained in mobile node u and data object v . w_i is the weight value of each interest keyword k_i of mobile node u .

$$m(u, v) = \sum_{i=1}^n \sum_{j=1}^m f(k_i, k_j) \times w_i \tag{2}$$

3.2 Problem Description and Motivations

Figure 1 illustrates a data dissemination task in a mobile ad hoc network. The data dissemination packet is started from the source node S . The black nodes are interested in the disseminated data while the white nodes are not interested in the data. For flooding method, data dissemination packets are delivered to all of the mobile nodes which can receive it. For example, the source node S disseminates the data packet to all its 1-hop neighbor nodes, A, B, C, D, H , and K . On receipt of data dissemination packet, the nodes A, B, C, D, H , and K retransmit the received data packet to their neighbor nodes again. Because flooding method did not consider user interests, a large amount of unnecessary data transmissions are generated. For instance, although the mobile nodes A, B , and C are not interested in the disseminated data, they have to receive and retransmit the data. In order to resolve this problem,

Fig. 1 An example of data dissemination in a mobile ad-hoc network



some recent works [31,32] have been proposed to disseminate the data packet only to the mobile nodes that are interested in the data. That is, the source node S only disseminates the data packet to its 1-hop neighbor nodes, D , H , and K . However, this will result in another problem, e.g. some mobile nodes that are interested in the data cannot receive the data packet. For instance, because node C does not receive and forward the data dissemination packet from node S , nodes M and N cannot receive the required data. The data dissemination among mobile nodes in mobile ad hoc networks is achieved through multi-hops, then, the remote nodes cannot receive the data dissemination packet if the intermediated relay nodes do not receive and forward the data. Therefore, the data dissemination packet cannot be spread well when most of the intermediate relay nodes are not interested in the data. A global infrastructure of the mobile ad hoc network will be helpful to resolve this problem. The source node can determine the necessary relay nodes by using the whole network information before performing data dissemination. However, this will result in unaffordable network management overhead.

In this paper, we propose a method to disseminate data packets to mobile nodes that are interested in the data with few unnecessary data transmissions and low network management overhead. We assume that each mobile node only manage its 1-hop neighbor nodes. In order to ensure the reachability of data dissemination in the network, more mobile nodes have to participate in a data dissemination task rather than only the mobile nodes that are interested in the data. The unnecessary data transmissions of data dissemination can be minimized by reducing the number of unnecessary relay nodes. As shown in Fig. 1, the non-interest nodes A , B , C and E are not interested in the disseminated data. Among them, node C can be considered as a necessary relay node as it affects the data reception of nodes M and N . Nodes A , B and E can be considered as unnecessary relay nodes as they do not affect the other mobile nodes receiving the data dissemination packet. This is because the neighbor nodes of nodes A , B and E can receive the data dissemination packet from the other mobile nodes that are interested in the data, such as nodes D and H . Therefore, the problem in this paper is how to efficiently identify these unnecessary relay nodes. Since the global knowledge of the network is hard to be obtained, identifying the unnecessary relay nodes in mobile ad hoc networks is a challenge. Given this, we propose two decentralized methods to filter out the unnecessary relay nodes in mobile ad hoc networks.

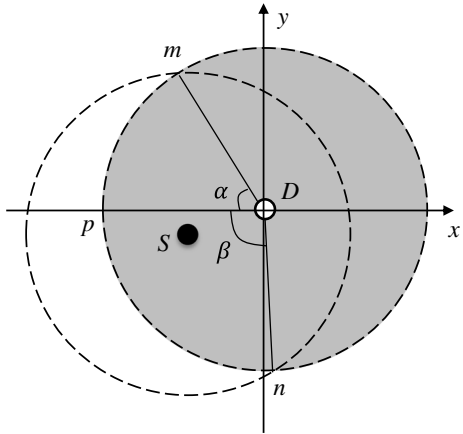
4 The Proposed Relay Nodes Selection Methods

In this section, we describe the relay nodes selection method for data dissemination in mobile ad hoc networks. In order to reduce the unnecessary data transmissions in the network, the most importance is reducing the unnecessary relay nodes. We propose two distributed relay nodes selection methods to prune the unnecessary relay nodes, the passive relay nodes selection method and the active relay nodes selection method. In passive relay nodes selection method, the data sending nodes determine the next relay nodes. In the active relay nodes selection method, the data recipient nodes determine whether they have to act as relay nodes when it receives a data dissemination packet from the other relay nodes or source node. Both of these two methods are decentralized.

4.1 Passive Relay Nodes Selection

In passive relay nodes selection method, the data sending nodes determine the next relay nodes from their 1-hop neighbor nodes. The data dissemination packet is only sent to the

Fig. 2 An example of intersection angle



selected relay nodes. The mobile nodes which receive the data dissemination packet have to act as a relay node to determine the next relay nodes from their neighbor nodes. Then, they forward the data dissemination packet to the selected relay nodes. This process is recursively executed at the selected relay nodes until the data dissemination packet is delivered to the end of the network. The multicast mechanism can be used in this procedure. We aim to deliver the data dissemination packet to as many mobile nodes that are interested in the data as possible and reduce the unnecessary data transmissions during the procedure of data dissemination. The importance is how to prune the unnecessary relay nodes from the 1-hop neighbor nodes. We propose a method that adopts user interests and location information to filter out unnecessary relay nodes.

When a source node or relay node disseminates data packet, it selects appropriate relay nodes from its neighbor nodes set. The relay nodes are selected from the mobile nodes that are interested in the data first. The interest probability of a mobile node in a data object can be determined by formula (2). If the communication range of the selected relay nodes can cover all of the other neighbor nodes within the neighbor nodes set, the remaining nodes can be identified as unnecessary relay nodes. Otherwise, the relay nodes have to be selected from the neighbor nodes set that are not interested in the data. Since the communication range of a mobile node is considered as a circle around the mobile node, the cover range of the communication range of a mobile node can be computed based on the intersection angles with its neighbor nodes. If the cover range of intersection angles surpasses 360° , we can say that the communication range $[0, 360^\circ]$ of a mobile node is completely covered. Let us consider an example of intersection angle in Fig. 2. For two mobile nodes S and D , we define the angle $\angle mDn$ (anti-clockwise) is the intersection angle of node D intersected with node S , where m and n are the intersection points of the communication circles centered at nodes S and D , respectively. According to the x coordinate axis of a horizontal coordinates, the intersection angle $\angle mDn$ (anti-clockwise) can be split into angles α and β .

Given the location coordinates of nodes S and D in a plane coordinate and the communication radius r , the position coordinates of the intersection point m and n can be computed by a quadratic equation. Then, the values of α and β can be calculated as follows.

$$\alpha = \cos^{-1} \left(\frac{|m.x|}{r} \right) \tag{3}$$

$$\beta = 180^\circ - \cos^{-1} \left(\frac{|n \cdot x|}{r} \right) \tag{4}$$

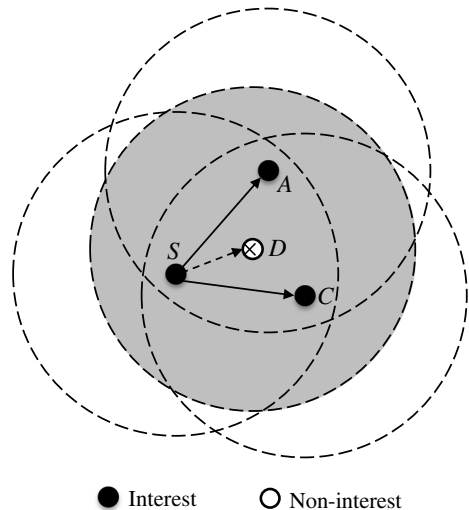
The intersection angle of node D intersected with node S can be referred as the angle interval $[180^\circ - \alpha, 180^\circ + \beta]$. Suppose that the communication range of node D is covered by a number of mobile nodes n_1, \dots, n_k with intersection angles $[\alpha_1, \beta_1], \dots, [\alpha_k, \beta_k]$, the total cover range $[\alpha, \beta]$ of node D can be calculated as follows.

$$[\alpha, \beta] = [\alpha_1, \beta_1] \cap [\alpha_2, \beta_2] \cap \dots \cap [\alpha_k, \beta_k] \tag{5}$$

Here, the overlapping part of the covered range is not involved in the calculation, e.g. the maximum cover range of intersection angles $[0, 30], [10, 90]$ and $[30, 60]$ is $[0, 90]$. If the cover range $[\alpha, \beta]$ of a mobile node is covered by $[0, 360^\circ]$, we can say that the communication range of the mobile node is completely covered. If the communication range of a mobile node is completely covered, we can guarantee that all its neighbor nodes can receive the data dissemination packet from the selected relay nodes. Therefore, if the mobile node is not interested in the data and its communication range is fully covered by the other relay nodes, the data dissemination packet does not need to send to it. The unnecessary relay nodes can be filtered out like this during data dissemination. The reduction of the relay nodes can future reduces the amount of the unnecessary data transmissions in the network.

Figure 3 illustrates the procedure of passive relay nodes selection method. Suppose that the source node S disseminates a data packet to its 1-hop neighbor nodes A, C and D , where the black nodes A and C are interested in the disseminated data and the white node D is not interested in the data. Before disseminating the data packet, the source node S selects the appropriate relay nodes from its 1-hop neighbor nodes. Because nodes A and C are interested in the data, they are selected as relay nodes first. At this time, node S checks the communication range (gray area) of node D . Because the communication range of node D is completely covered by the communication ranges of nodes S, A and C , it is identified as an unnecessary relay node. Like this, the unnecessary relay nodes are pruned during data dissemination. Therefore, node S only disseminates the data packet to the selected relay nodes A and C . And then, the mobile nodes A and C receive the data packet and forward it

Fig. 3 An example of passive relay nodes selection



to their selected relay nodes again. The reduction of the number of relay nodes can reduce a large amount of unnecessary data transmissions in the network.

The procedure of passive relay nodes selection can be described as follows.

Algorithm 1. Passive relay nodes selection

Initialization:

source node s disseminates data packet p ;

Input:

neighbor list (NL) = the neighbor nodes of node s ;

relay list (RL) = *null*;

before sending data packet p do

for each node n of NL **do**

if $n.interest(p) == true$ **then**

n is selected as a relay node;

$RL.add(n)$;

$NL.remove(n)$;

for each node n of NL **do**

for each node m of RL **do**

 covered angle $[\alpha, \beta]$ computation;

if $[\alpha, \beta] \neq [0, 360^\circ]$ **then**

$RL.add(n)$;

$NL.remove(n)$;

end

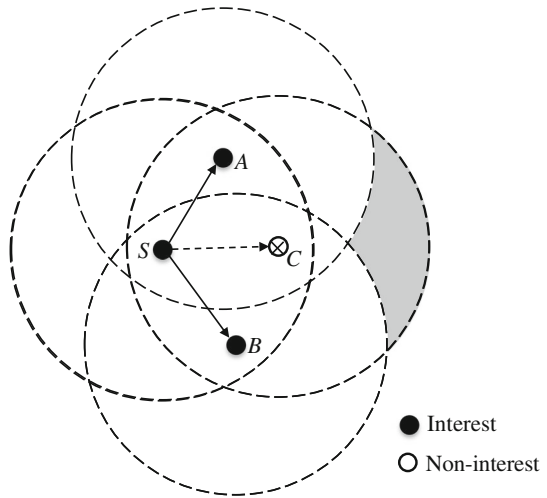
Source node s disseminates a data packet to the selected relay nodes from its 1-hop neighbor nodes. The relay nodes which receive the data dissemination packet perform the same process to select the next new relay nodes until the data dissemination packet is delivered to the end of the network. However, because the communication range of a mobile node is limited, some unnecessary relay nodes would not be identified by the data sending node. In the following, we propose another active relay nodes selection method which can prune more unnecessary relay nodes.

4.2 Active Relay Nodes Selection

In passive relay nodes selection method, the data transmission nodes identify the unnecessary relay nodes that are not interested in the disseminated data and the communication range of them are completely covered by the existing relay nodes. Actually, if the uncovered area of a mobile node does not contain any other mobile node, it can also be identified as an unnecessary relay node. For instance, as shown in Fig. 4, the source node S selecting relay nodes from its 1-hop neighbor nodes A , B and C . Although node C is not interested in the data and does not affect the other mobile nodes receiving the data packet, it is selected as a relay node by the passive relay nodes selection method. Because the communication range of the source node S is limited, it cannot obtain the knowledge of the uncovered area of node C only according to its 1-hop neighbor nodes. Then we proposed another active relay nodes selection method to resolve this problem.

In contrast to the passive relay nodes selection method in which the data sending nodes determine the next relay nodes, in active relay nodes selection method, data receipt nodes make decisions of forwarding or dropping data dissemination packets. A source node or relay

Fig. 4 The problem of passive relay nodes selection



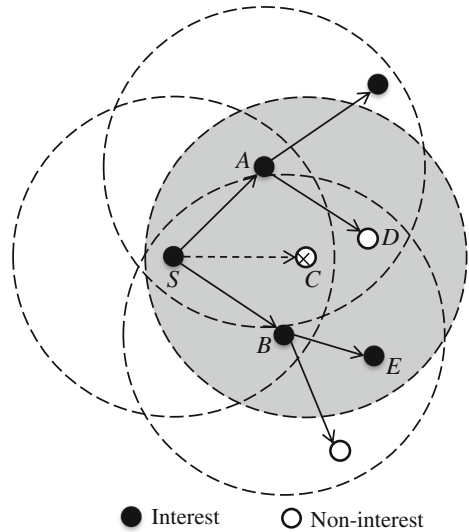
node delivers the data dissemination packet to all its 1-hop neighbor nodes. When receiving a data dissemination packet, a mobile node determines whether to re-transmit the received data packet. A mobile node is identified as an unnecessary relay node when it is not interested in the data and all of its neighbor nodes can receive the data dissemination packet from the other relay nodes. This can be implemented in a broadcast scheme. A source node or relay node first broadcasts a data dissemination packet to all its 1-hop neighbor nodes. When receiving data dissemination packets, the data recipient nodes determine whether it has to act as a relay node. The decision is made depended on whether the data recipient node is interested in the data and its 1-hop neighbor nodes can receive the data from the other relay nodes. If the recipient node is interested in the data, it receives the data packet and acts as a relay node to forward the received data. Otherwise, the data recipient node checks the location information of all its neighbor nodes to see if all of them can receive the data from the other relay nodes. If all the 1-hop neighbor nodes can receive the data dissemination packet, the mobile node is identified as an unnecessary relay node. This can be done by checking the locations and communication radiuses of its neighbor nodes. For instance, as shown in Fig. 5, the source node *S* disseminates a data packet to all its 1-hop neighbor nodes *A*, *B* and *C*. Since nodes *A* and *B* are interested in the data, they receive and forward the data packet. Since node *C* is not interested in the data, it checks its neighbor nodes (*A*, *B*, *S*, *E* and *D*) to see if all of them can receive the data by using the contained location and communication radius information. Since nodes *A* and *B* can receive data from node *S* and nodes *D* and *E* can receive data from nodes *A* and *B*, node *C* is identified as an unnecessary relay node. The absent of relay node *C* does not affect the reachability of data dissemination in the network.

Since the location and communication radius *r* of each mobile node are available, the distance between two mobile nodes *m* and *n* can be measured by using the following distance function *dist(m, n)*, where *x* and *y* indicates the 2-axis location coordinate of the mobile node.

$$dist(m, n) = \sqrt{(x_m - x_n)^2 + (y_m - y_n)^2} \tag{6}$$

If $dist(m, n) < r$, the mobile node *m* is determined can receive data packet from node *n*. The procedure of the active relay nodes selection method for a given node *n* is described as follows:

Fig. 5 An example of active relay nodes selection



Algorithm 2. Active relay nodes selection

Initialization:

source node s broadcasts data dissemination packet p ;
 neighbor list (NL) = the neighbor nodes of node n ;
 relay list (RL) = null;

```

upon reception of data packet  $p$  do
if  $n.interest(p) == true$  then
    node  $n$  receives  $p$  and acts as a relay node
if  $n.interest(p) == false$  then
    for each neighbor node  $m$  of mobile node  $n$  do
        if  $dist(s, m) < r$  then
             $NL.remove(m)$ ;
            if  $m.interest(p) == true$  then
                 $RL.add(m)$ ;
    if  $NL \neq null$  then
        for each node  $m$  in  $NL$  do
            for each node  $u$  in  $RL$  do
                if  $dist(u, m) < r$  then
                     $NL.remove(m)$ ;
    if  $NL \neq null$  then
        node  $n$  receives  $p$  and acts as a relay node
end
    
```

Upon receiving a data dissemination packet, a mobile node n in the network may activate the above procedure. If the mobile node n is interested in the disseminated data it receives the data and acts as a relay node. Otherwise, it checks whether all its neighbor nodes can receive the data by using the distance function (5). Let r represent the communication radius and $dist(u, v)$ denote the distance between nodes u and v , then $dist(u, v) < r$ denotes that

node u can get data dissemination packet from node v , and vice versa. If all the neighbor nodes can receive the data from the existing relay nodes, mobile node n does not receive and forward the data. This process repeats iteratively until the data dissemination packet is delivered to the end of the network.

5 Performance Evaluation

In this section, we evaluate the performance of the proposed method compared to the pure flooding method, the 2-hop based flooding method, the user interest based method and the user-centric method. The performance of data dissemination in mobile ad hoc networks depends on how to efficiently disseminate the data packet to as many mobile nodes that are interested in the data as possible with low network management overhead and high reachability. To evaluate the performance, we perform a simulation in mobile ad hoc network environments.

5.1 Experimental Environment

We conducted our experiments on a desktop PC. All of the experiments were coded in Java. We evaluate the performance of the proposed method compared with the existing methods in terms of the following four metrics:

Network management overhead is the total number of messages that have been generated during the maintenance of the network.

Total amount of data transmissions is the total number of data packets that have been transmitted in the network during a data dissemination task.

Reachability of data dissemination packet is measured by the ratio of the number of mobile nodes that received the data dissemination packet to the total number of mobile nodes in the network.

Ratio of user preference satisfaction is measured by the number of interest nodes that receive the data packet to the total number of mobile nodes that receive the data packet in the network.

The performance evaluation is conducted on a dataset which is generated by modifying the networked-based generator [34]. The development areas are set to $100\text{ m} \times 100\text{ m}$ and $300\text{ m} \times 300\text{ m}$. The communication radius of each mobile node is set to 30 m. The simulation was lasted in a period of 400 s. The moving speeds of the mobile nodes are set from 0 to 3 m/s. The network is dynamic that each mobile node arrives and departs freely and the total number of the mobile nodes is set to 500.

We generate a keyword set based on a keyword space with 100 categories of keywords in total. We assume that the order of the keyword in the keyword set represent the popularity of the keyword in the network. According to Definition 1, the user interest in each keyword is drawn from the keyword set. And each mobile user is assigned 5 interest keywords. The weight of each interest keyword is assigned according to the descending order of the keyword in the keyword list. Then, the weight values of the interest keywords are $\{5/15, 4/15, \dots, 1/15\}$. Each data item is also described by 5 keywords from the keyword set, but they are not assigned weight values. The data dissemination task is started from the arbitrary mobile node within the network.

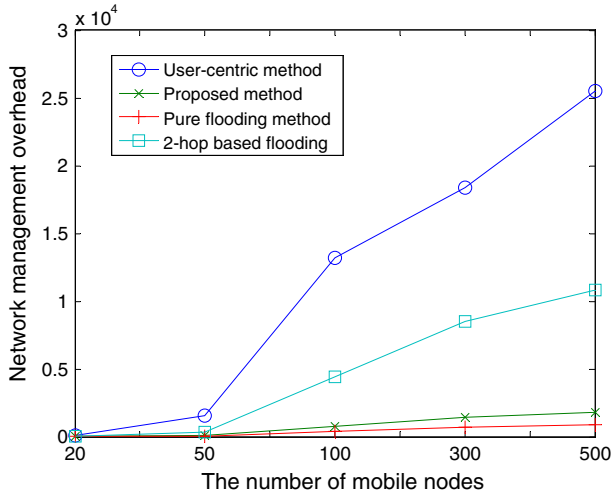


Fig. 6 Network management overhead according to the number of mobile nodes

5.2 Experimental Results

In the first experiment, we compare the performance the proposed method with the existing methods in network management overhead according to the number of mobile nodes. Figure 6 shows the result of network management overhead with the increase of mobile nodes. The network management overhead is low when the number of mobile nodes is small. This is because the network is not connected well when the distribution of mobile nodes is sparse in the network. Because each mobile node only manages its local information, the pure flooding method has lowest network management overhead. Since each mobile node has to manage the neighbor nodes with k ($k = 3$) hops, the user-centric method has the highest network management overhead. In the 2-hop based flooding method, each mobile node manages its 2-hop neighbor nodes. Therefore, the network management overhead of the 2-hop based flooding method is lower than that of the user-centric method. Since each mobile node only needs to manage its 1-hop neighbor nodes, the network management overhead of the proposed method is less than those of the user-centric method and the 2-hop based method and a little higher than that of the pure flooding method. Because the network management overhead of the user interest based method is the same as that of the proposed method, it is omitted in the graph.

Figure 7 shows the performance of network management overhead according to the moving speeds of the mobile nodes. In this experiment, the average moving speeds of mobile nodes are varied from 1 to 3m/s and the number of mobile nodes is fixed in 500. From the results, we can see that the network management overhead increases sharply with the increase of moving speeds. This is because the faster the movement of the mobile nodes is, the higher the frequency of the network topology re-organizing is. Because each mobile node only manages its local information, the increase of the moving speeds of mobile nodes does not affect the network management overhead of the pure flooding method. The results show that the user-centric method has the highest network management overhead. This is because the maintenance of k -hop ($k = 3$) neighbor nodes results in high network management overhead when the mobile nodes move fast. Since each mobile node manages its

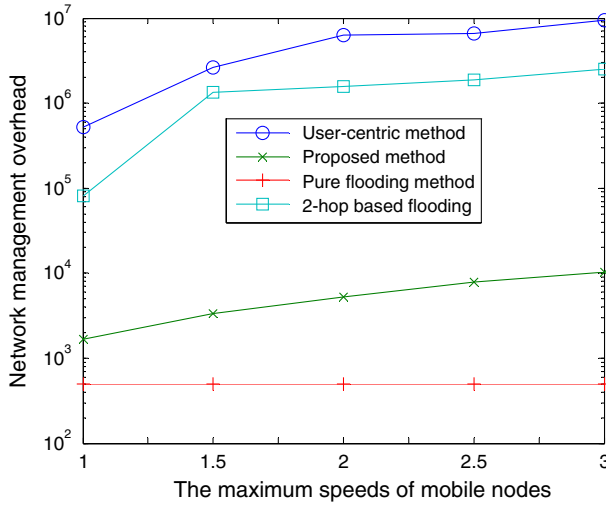


Fig. 7 Network management overhead according to moving speeds of mobile nodes

1-hop neighbor nodes, the network management overhead of the proposed method is less than those of the 2-hop based flooding method and the user-centric method. Notice that, in order to clearly distinguish the differences among the methods, the y-axis in the chart is represented in logarithmic scale.

Figure 8a shows the total amount of transmissions according to the number of mobile nodes within an area of 100 m × 100 m. The total amount of transmissions is measured by the number of data packets that are transmitted in the network during data dissemination. The data dissemination is initiated from a source node, which is selected from the nodes in the network randomly. Clearly, the total amount of data transmissions is increased with the increase of mobile nodes. The total amount of data transmissions of the pure flooding method is much higher than that of the other methods. This is because the data dissemination is performed by disseminating the data packet to all the mobile nodes within the network. The mobile nodes which receive the data packet retransmit the receive data. This results in a large amount of data transmissions in the network. Since some dominated relay nodes have been pruned, the total amount of data transmissions of the 2-hop based flooding method is less than that of the pure flooding method. In the user interest based method, the data dissemination packet is only sent to the neighbor nodes that are interested in the data. The mobile nodes which receive the data retransmit the data packet again. Then, it has the lowest total amount of data transmissions. The total amount of data transmissions of the user-centric method, the proposed passive method and the proposed active method are similar. Because the user-centric method adopts the neighbor nodes of k hops ($k = 3$) to select the appropriate relay nodes, more unnecessary relay nodes can be filtered out than the proposed methods. The proposed active method can filter out more unnecessary relay nodes than that of the proposed passive method.

Figure 8b shows the total amount of data transmission according to the number of mobile nodes within an area of 300 m × 300 m. The total amount of transmissions is increased with the increase of the number of mobile nodes. When the number of the mobile nodes is small, nearly all of the mobile nodes are not connected with the other mobile nodes, which results in low total amount of data transmissions. The overall result of Fig. 8b is lower than that of

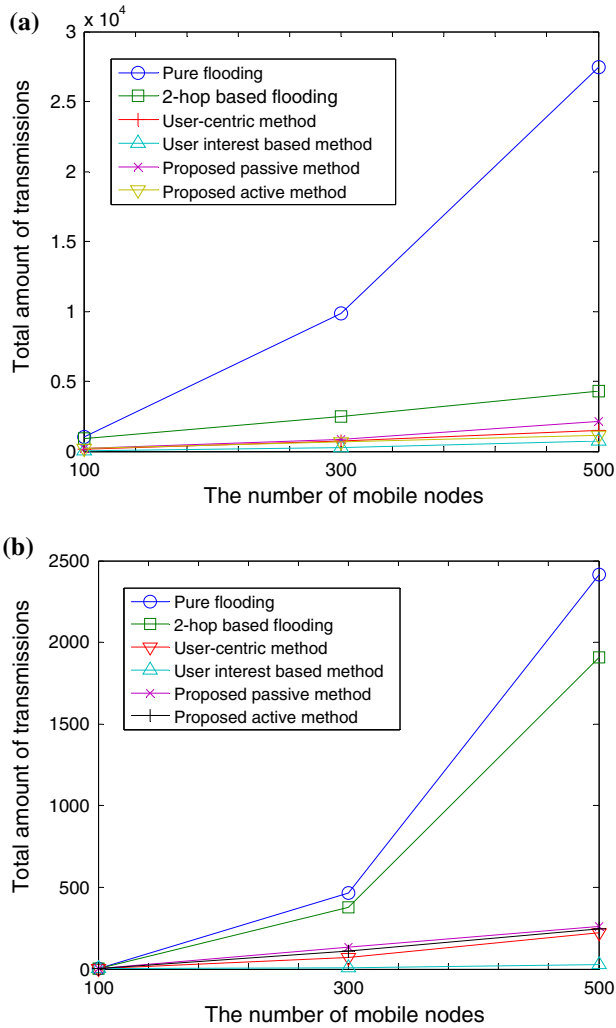


Fig. 8 Total amount of transmissions according to the number of mobile nodes (a) 100 m \times 100 m, (b) 300 m \times 300 m

Fig. 8a. This is because the number of mobile nodes is the same but the development area is different. The pure flooding method has the highest total amount of data transmissions, where the user interest based method has the lowest total amount of data transmissions. The user-centric method has a little lower total amount of transmissions than the proposed methods.

In order to study the different mixes of the number of interest nodes and the number of non-interest nodes that affect the data dissemination, we adjust the ratio (the number of interest nodes)/(the number of non-interest nodes) within the network. Figure 9 shows the total amount of data transmissions of the methods according to the different ratios. The ratio is varied from 1 to 4. We fix the total number of mobile nodes to 500 and vary the ratios. When the number of interest nodes is the same as the number of non-interest nodes the ratio is set to

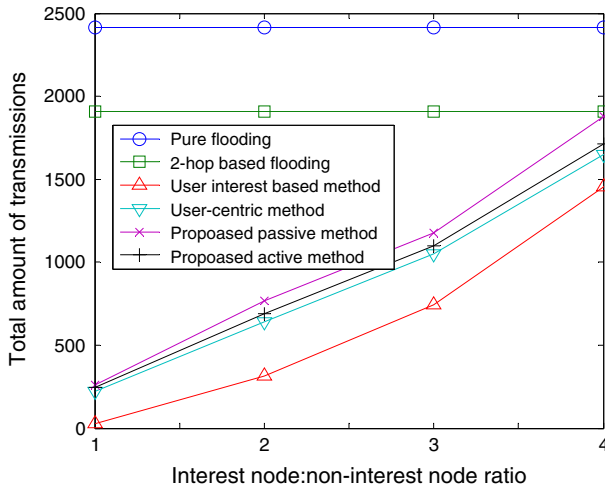


Fig. 9 Total amount of transmissions according to the different ratios (interest nodes)/(non-interest nodes)

1. In this experiment, the development area is 300 m × 300 m. From the result, we can see that the user interest based method, the user-centric method and the proposed methods show better performance compared with the pure flooding method and the 2-hop based flooding method. Since the total number of mobile nodes is constant, the results of pure flooding method and the 2-hop based flooding method do not change. The total amount of data transmissions of the user interest based method, the user-centric method and the proposed methods increased fast with the increase of the ratio. With high (the number of interest nodes)/(the number of non-interest nodes) ratio, the results of the user interest based method, the user-centric method and the proposed methods are almost equivalent to the 2-hop based flooding method.

Next, we investigate the performance of reachability of data dissemination packets in the network. Figure 10 shows the reachability of data dissemination packets according to the number of mobile nodes within the development area of 300 m × 300 m. The reachability of data dissemination is increased with the increase of the number of mobile nodes. Since the network is not connected well when the number of mobile nodes is small, the poorly connected mobile nodes may fail to receive the data dissemination packets. Because the flooding methods try to flood the data dissemination packets to all the mobile nodes in the whole network, the pure flooding method and the 2-hop based flooding method have higher reachability of data dissemination than the other methods. The reachability of the user-centric method depends on the scope (the number of multiple hops) of the network information maintained at each individual mobile node. Since the data dissemination packets are only sent to the 1-hop neighbor nodes that are interested in the data, the user interest based method has the lowest reachability of data dissemination.

Finally we discuss the ratio of user preference satisfaction in data dissemination. In this experiment, the development area is 300 m × 300 m. The ratio of user preference satisfaction is measured by the number of interest nodes that receive the data packet to the total number of mobile nodes that receive the data packet in the network. Figure 11 shows that the user interest based method has the highest performance of user preference satisfaction in data dissemination. This is because the data packet is only disseminated to the mobile nodes that are interested in the data. Since the proposed methods and user-centric method consider

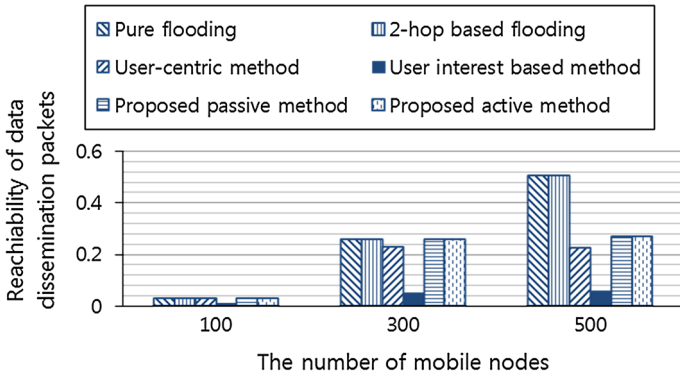


Fig. 10 Reachability of data dissemination according to the number of mobile nodes

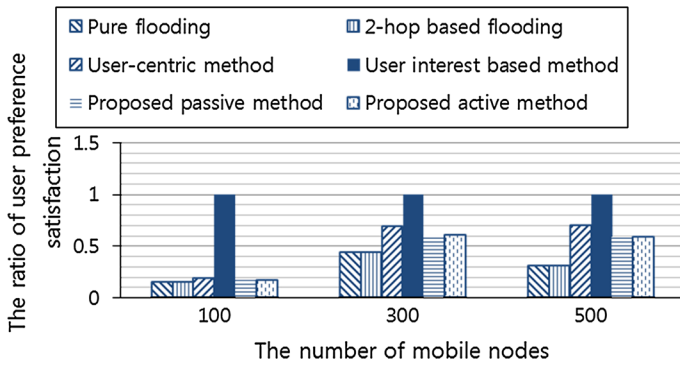


Fig. 11 The ratio of user preference satisfaction according to the number of mobile nodes

satisfying user preference in data dissemination, they are more efficient than the flooding methods.

6 Conclusions

In this paper, we studied data dissemination methods over mobile ad-hoc networks. Because most of the existing data dissemination methods over mobile ad-hoc networks failed to consider user preferences, they resulted in a lot of unnecessary data transmissions in the network. Although some recent methods have considered satisfying user interests when performing data dissemination, they either required high network management overhead or cannot guarantee that the mobile nodes that are interested in the data can receive the data dissemination packets. We proposed two relay nodes selection methods to prune the unnecessary relay nodes. The passive relay nodes selection method allows each data transmission node to select a partial set of its neighbor nodes as relay nodes to participate in data dissemination. In the active relay nodes selection method, the data receipt nodes determine whether it has to act as a relay node to participate data dissemination. The reduction of the unnecessary relay nodes can decrease a lot of unnecessary data transmissions in the network. The experimental results have shown that the proposed method achieves low network management overhead

and small data transmissions for data dissemination, while the existing methods achieve either low network management overhead or small data transmissions. In the future work, we additionally plan to investigate selecting appropriate data objects to achieve efficient data dissemination.

Acknowledgments This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government (MSIP) (No. 2013R1A2A2A01015710) and the MSIP (Ministry of Science, ICT and Future Planning), Korea, under the ITRC (Information Technology Research Center) support program (NIPA-2013-H0301-13-4009) supervised by the NIPA (National IT Industry Promotion Agency).

References

1. Ye, L. (2010). MP2P based on social model to serve for LBS. In *Proceedings of the international conference on E-Business and E-Government* (pp. 1679–1682).
2. Magalhaes, J., & Holanda, M. (2011). EIKO: A social mobile network for MANET. In *Proceedings of the information systems and technologies* (pp. 15–18).
3. Emre, S., Oriana, R., Patrick, S., & Gustavo, A. (2009). Enabling social networking in ad hoc networks of mobile phones. *VLDB Endowment*, 2(2), 1634–1639.
4. Ni, S. Y., Tseng, Y. C., Chen, Y. S., & Sheu, J. P. (1999). The broadcast storm problem in a mobile ad hoc network. In *Proceedings of the 5th ACM/IEEE international conference on mobile computing and networking* (pp. 151–162).
5. Gao, W., & Cao, G. (2011). User-centric data dissemination in disruption tolerant networks. In *Proceedings of the IEEE INFOCOM* (pp. 3119–3127).
6. Hui, P., Crowcroft, J., & Yoneki, E. (2008). Bubble rap: Social-based forwarding in delay tolerant networks. In *Proceedings of the 9th ACM international symposium on mobile ad hoc networking and computing* (pp. 241–250).
7. Boldrini, C., Conti, M., & Passarella, A. (2008). ContentPlace: Social-aware data dissemination in opportunistic networks. In *Proceedings of the 11th international symposium on modeling, analysis and simulation of wireless and mobile systems (MSWiM '08)* (pp. 203–210).
8. Gao, W., Li, Q., Zhao, B., & Cao, G. (2009). Multicasting in delay tolerant networks: A social network perspective. In *Proceedings of the 10th ACM international symposium on mobile ad-hoc networking and computing* (pp. 299–308).
9. Boldrini, C., Conti, M., & Passarella, A. (2008). Exploiting users' social relations to forward data in opportunistic networks: The HiBOP solution. *Pervasive and Mobile Computing*, 1(4), 633–657.
10. Gao, W., Li, Q., Zhao, B., & Cao, G. (2012). Social-aware multicast in disruption-tolerant networks. *IEEE/ACM Transactions on Networking*, 20(5), 1553–1566.
11. Gandhi, R., Parthasarathy, S., & Mishra, A. (2003). Minimizing broadcast latency and redundancy in ad hoc networks. In *Proceedings of the 4th ACM international symposium on mobile ad hoc networking and computing* (pp. 222–232).
12. Lim, H., & Kim, C. (2001). Flooding in wireless ad hoc networks. *Computer Communication*, 24(3–4), 353–363.
13. Dommety, G., & Jain, R. (1996). Potential networking applications of global positioning systems (GPS). Technical report TR-24. The Ohio State University.
14. Han, B., Hui, P., Kumar, V. A., Marathe, M. V., Pei, G., & Srinivasan, A. (2010). Cellular traffic offloading through opportunistic communications: A case study. In *Proceedings of the 5th ACM workshop on challenged, networks* (pp. 31–38).
15. Krishan, K., & Singh, V. P. (2014). Power consumption based simulation model for mobile ad-hoc network. *Wireless Personal Communications*. doi:10.1007/s11277-013-1589-7.
16. Babar, S., & Kim, K. (2013). Towards enhanced searching architecture for unstructured peer-to-peer over mobile ad hoc networks. *Wireless Personal Communications*. doi:10.1007/s11277-013-1560-7.
17. Kok, G. X., Chow, C., Xu, Y., & Hiroshi, I. (2013). EAOMDV-MIMC: A multipath routing protocol for multi-interface multi-channel mobile ad-hoc networks. *Wireless Personal Communications*, 73(3), 477–504.
18. Kim, B. K., & Eom, D. S. (2013). A proposal of B-tree based routing algorithm for monitoring systems with regular movements in MANETs. *Wireless Personal Communications*, 73(3), 517–534.
19. Zhao, H., Zhu, H., Li, D., & Liu, N. (2013). Quality-driven capacity-aware resources optimization design for ubiquitous wireless networks. *Wireless Personal Communications*. doi:10.1007/s11277-013-1508-y.

20. Stojmenovic, I., Seddigh, S., & Zunic, J. (2002). Dominating sets and neighbor elimination based broadcasting algorithms in wireless networks. *IEEE Transactions on Parallel and Distributed Systems*, 13(1), 14–25.
21. Peng W., & Lu, X. C. (2000). On the reduction of broadcast redundancy in mobile ad hoc networks. In *Proceedings of the 1st ACM international symposium on mobile ad hoc networking and computing* (pp. 129–130).
22. Agbaria, A., Hugerat, M., & Friedman, R. (2011). Efficient and reliable dissemination in mobile Ad Hoc networks by location extrapolation. *Journal of Computer Networks and Communications*, 2011, 1–11.
23. Lou, W., & Wu, J. (2002). On reducing broadcast redundancy in Ad Hoc wireless networks. *IEEE Transactions on Mobile Computing*, 1(2), 11–123.
24. Qayyum, A., Viennot, L., & Laouiti, A. (2002). Multipoint relaying: An efficient technique for flooding in mobile wireless networks. In *Proceedings of the 35th annual Hawaii international conference on system sciences* (pp. 3866–3875).
25. Pleisch, S., Balakrishnan, M., Birman, K., & Renesse, R. (2006). MISTRAL: Efficient flooding in mobile ad-hoc networks. In *Proceedings of the 7th ACM international symposium on mobile ad hoc networking and computing* (pp. 1–12).
26. Meyer, H., & Hummel, K. A. (2009). A geo-location based opportunistic data dissemination approach for MANETs. In *Proceedings of the 4th ACM workshop on challenged, networks* (pp. 1–8).
27. Kokuti, A., & Simon, V. (2012). Location based data dissemination in mobile ad hoc networks. In *Proceedings of the 35th international conference on telecommunications and signal processing* (pp. 57–61).
28. Waluyo, A. B., Taniar, D., Rahayu, W., Aikebaier, A., Takizawa, M., & Srinivasan, B. (2012). Mobile Peer-to-Peer data dissemination in wireless ad-hoc networks. *Information Science*, 230, 3–20.
29. Edison, P., Tales, H., Flavio, R., Carlos, E., & Tony, L. (2013). Exploring geographic context awareness for data dissemination on mobile ad hoc networks. *Ad Hoc Networks*, 11(6), 1746–1764.
30. Pietilainen, A. K., & Diot, C. (2012). Dissemination in opportunistic social networks: The role of temporal communities. In *Proceedings of the 13th ACM international symposium on mobile ad hoc networking and computing* (pp. 165–174).
31. Li, F., & Wu, J. (2009). Mops: Providing content-based service in disruption tolerant networks. In *Proceedings of the international conference on distributed, computing systems* (pp. 526–533).
32. Yoneki, E., Hui, P., Chan, S., & Crowcroft, J. (2007). A socio-aware overlay for publish/subscribe communication in delay tolerant networks. In *Proceedings of the 10th international symposium on modeling, analysis and simulation of wireless and mobile systems* (pp. 225–234).
33. Lin, K. C., Chen, C. W., & Chou, C. F. (2012). Preference-aware content dissemination in opportunistic mobile social networks. In *Proceedings of the IEEE INFOCOM* (pp. 1960–1968).
34. Thomas, B. (2002). A framework for generating network-based moving objects. *GeoInformatica*, 6(2), 153–180.



He Li received the B.S. degree in the department of computer science from Yunnan University, China in 2006. He received his M.S. degree in the department of Information and Communication Engineering at Chungbuk National University, Korea in 2010. He currently is a Ph.D. student in the department of Information and Communication Engineering at Chungbuk National University, Korea. His main research interests include query processing, distributed computing, data dissemination, mobile ad hoc network.



Kyungsoo Bok received the B.S. in Mathematics from Chungbuk National University, Korea in 1998 and also received M.S. and Ph.D. in Computer and Communication Engineering from Chungbuk National University, Korea in 2000 and 2005. He is now a research professor in Information and Communication Engineering, Chungbuk National University, Korea. His research interests are database system, storage management system, social network service, mobile ad-hoc network, and big data processing.



Kyungyong Chung has received B.S., M.S., and Ph.D. degrees in 2000, 2002, and 2005, respectively, all from the Department of Computer Information Engineering, Inha University, Korea. He has worked for Software Technology Leading Department, Korea IT Industry Promotion Agency (KIPA). He is currently a professor in the School of Computer Information Engineering, Sangji University, Korea. His research interests include Medical Data Mining, Healthcare, Knowledge System, HCI, and Recommendation.



Jaesoo Yoo received his B.S. in Computer Engineering from Chonbuk National University, Korea in 1989, and he also received his M.S. and Ph.D. in Computer Science from the Korean Advanced Institute of Science and Technology, Korea in 1991 and 1995. He is now a professor in Information and Communication Engineering, Chungbuk National University, Korea. His main research interests include database systems, sensor data management, location based services, distributed computing and storage management system.