Context-Aware Sensor Data Dissemination for Mobile Users in Remote Areas

Edith C.-H. Ngai[1], Mani B. Srivastava[2] and Jiangchuan Liu[3]
[1]Department of Information Technology, Uppsala University, Sweden
[2]Electrical Engineering Department, University of California, Los Angeles
[3]School of Computing Science, Simon Fraser University, Canada

***Abstract*— Many mobile sensing applications consider users reporting and accessing sensing data through the Internet. However, WiFi and 3G connectivities are not always available in remote areas. Existing data dissemination schemes for opportunistic networks are not sufficient for sensing applications as sensing context has not been explored. In this work, we present a novel context-aware sensing data dissemination framework for mobile users in a remote sensing field. It maximizes information utility by considering such sensing context as sensing type, locality, time-to-live, mobility and user interests. Different from existing works, the mobile users not only collect sensing data, but also upload data to sensors for information sharing. We develop a context-aware deployment algorithm and a hybrid data exchange mechanism for generic sensors and mobile users. We evaluate our solution by both analysis and simulations, and show that it can provide high information utility for mobile users at low communication overhead.**

I. Introduction

With the advancement of smart phones, mobile sensing applications have emerged recently, which leverage mobile phones as sensors to collect and report the ambient data for monitoring wildlife, pollution, social activities, and etc. [1], [2]. Most existing applications consider mobile phone users who can report and access sensing data through the Internet by 3G or WiFi connections. However, the cell phone reception is incredibly touchy, depending heavily on landscapes, carrier technologies, phone models, service providers, tower locations, and etc. Dead spots are commonly found in remote areas [3] and even in some part of major cities [4]. Nevertheless, fine-grained and real-time sensing data are essential for mobile users to make correct decisions in many activities. For example, hikers may want updated information about the weather and unusual condition of the hiking trails. Sensing data such as dangerous trails with rocks, cliff, sandy area, fallen trees from landscape changes due to climate could be captured by mobile users or pre-deployed sensors [5], [6].

In this paper, we present a novel mobile sensing framework that supports sensing data dissemination for mobile phone users in remote areas. We face several unique challenges in designing our system: (1) The mobile phones in remote areas may have poor or no cellular and WiFi connectivities. Different from existing participatory and urban sensing applications, the mobile users may not be able to report and access the sensing data through the Internet in the sensing field; (2) The sensing context such as sensing type, locality and time-to-live must be taken into account to achieve high information utility and low communication overheads for mobile users. Existing opportunistic data dissemination schemes for mobile ad hoc networks are not sufficient as they have not considered sensing context of the environment; (3) The data dissemination scheme should take care of both users sharing common mobility patterns as well as those different from the majority. The scheme needs to be robust enough to support future users with different mobility patterns from the existing ones.

To tackle the above challenges, we suggest a hybrid approach for mobile users to collect and exchange sensing data leveraging both stationary sensors and mobile phones. The mobile users can upload and download the sensing data at stationary sensors for information sharing. The uploaded sensing data will be disseminated to more mobile users within the area of interests according to their sensing context. To the best of our knowledge, we are the first to study context-aware sensor data dissemination for mobile users in remote sensing field leveraging a combination of both stationary sensors and mobile phones. We aim at maximizing the information utility for mobile users in terms of data delivery and delay considering sensing context such as sensing type, locality, time-to-live, mobility and user interests.

II. Related Work

Participatory and urban sensing have been studied recently [1], [2], which leverage mobile phones as sensors to collect sensing data from the environment. BikeNet [6] utilized an opportunistic mobile sensing platforms where data are collected and carried by cyclists with uncontrolled mobility. It operates in a delay tolerant sensing mode, where cyclists go on trips, collect sensed data, and upload their data when they return to home. The idea of high-bit-rate short-range communication has also been presented in Infostations [7] to complement the relatively high cost and low bit-rate cellular systems. With the advancement of sensing and mobile communication technologies, however, the sensing context and detailed data sharing strategies with both stationary sensors and mobile phones remain to be further explored.

Mobile sinks and mobile relays have been applied for improving the performance of data collection in wireless sensor networks. Shah et al. [8] presented an architecture using moving entities, called *data mules*, to collect sensing data. Gu et al. [9] proposed a partitioning-based algorithm to schedule

978-1-4673-0775-8/12/$31.00 ©2012 IEEE

the movement of mobile elements, which minimizes the required moving speed and eliminates buffer overflow. Zhao et al. [10] proposed a *message ferrying* approach to address the network partition problem in sparse ad hoc networks. Different from our work, all of the above work focus on communication between intermediate peers in ad hoc networks without any stationary sensors. Throwboxes have been proposed by Zhao et al. [11] to enhance the network capacity by deploying stationary nodes in mobile DTNs. We share a similar concept here to utilize stationary nodes to improve data dissemination. However, throwboxese aim at improving the total data rate between mobile node pairs without considering the context of the data. On the contrary, our work considers sensing context such as sensing locality, time-to-live, and user interests, which integrates sensing and opportunistic data sharing. We provide a novel sensing data dissemination scheme that maximizes the information utility for mobile users in a sensing field.

Context- and social-aware data dissemination have been explored for opportunistic networks. Boldrini et al. [12] proposed a middleware that autonomically learns context and social information of the users in order to predict their future movements. Yoneki et al. [13] proposed a socio-aware overlay over detected communities for publish/subscribe communication. Jaho et al. [14] divided users into different interest-induced social groups and locality-induced social groups to improve information dissemination in social networks. Lee et al. [15] proposed a content management mechanism for highly dynamic mobile ad hoc network environment using location-based content binding. Nevertheless, the above works mainly focus on the social and mobility patterns of mobile users, while the sensing context and the surrounding environment have not been explored.

III. Problem Descriptions

We consider a sensing field with sensors deployed to monitor events of interests in hiking trails such as high temperature, dangerous trails with rocks or fallen trees. The sensing field can be located at remote areas with poor or no cellular and WiFi coverages. Users could obtain sensing data by communicating with the stationary sensors or taking sensing measurements directly using their mobile phones. Each sensing event is associated with its location and time-to-live (TTL) according to its event type. For instance, wildlife being observed will last for shorter time, while natural fruits and water source may last for longer. We represent an event e as $< e, L_e, TTL_e >$ being detected by a sensor or a mobile phone at location L_e, where TTL_e is the time-to-live of the event. A user m_k at location k will be interested in event e if he may reach the event area in time TTL_e. The interest of the user to that event is denoted by a binary variable I_{k,L_e}.

$$I_{k,L_e} = \begin{cases} 1 & \text{if } dist(k, L_e)/v \leq TTL_e, \\ 0 & \text{otherwise,} \end{cases}$$

where $dist(k, L_e)$ is the walking distance from k to L_e and v is the walking speed of the user. We name the area with $I_{k,L_e} = 1$ as the *area of interest* denoted by A_i. It is a circular area centered at L_e with radius R_e. R_e is bounded by $TTL_e v_{max}$, where v_{max} is the maximum moving speed of hikers along a trail. We define *information utility* as the amount of updated information gained by mobile users from the sensors in the area of interests. Our system aims at maximizing the information utility for mobile users considering the sensing type, time-to-live, locality, mobility and interests of users by deploying wireless sensors in a remote sensing field. Locations of sensors with specialized functions (e.g. water quality sensors and pollution sensors) are usually suggested by the nature reserve. On the other hand, *generic sensors* (e.g. temperature sensors and humidity sensors), which have greater sensing coverage and more memory for caching and exchanging data, could be deployed at locations with greater flexibility. We focus on optimizing the deployment of generic sensors and the sensing data dissemination mechanisms in this work.

IV. Deployment of Generic Sensors Using Mobile Statistics

We formulate the deployment and caching problem of *generic sensors* using binary variables $z_j (j = 1, ..., n)$, where the outcome z_j signals that a generic sensor will be deployed at location j. The objective function aims at maximizing the total information utility obtained by all mobile phone users, which is denoted by $\sum_{\forall j} u_j z_j$. Given the visiting frequency of the potential locations, F_j from mobile statistics, we calculate information utility u_j as

$$u_j = \sum_{\forall i} C_{i,j} I_{i,j} F_j, \tag{1}$$

where $C_{i,j}$ is a binary variable indicating whether there is a path between i and j, $I_{i,j}$ is a binary variable equals to 1 if $dist(i, j)$ is smaller than R_i. Both $C_{i,j}$ and $I_{i,j}$ can be obtained from the map of the hiking trails. The deployment cost is bounded by budget B.

$$\begin{array}{lr} \text{maximize} & \sum_{\forall j} u_j z_j \\ \text{subject to} & \sum_{\forall j} c_j z_j \leq B, \\ \forall j \in 1, 2, ..., n & z_j \in \{1, 0\}, \\ & u_j = \sum_{\forall i} C_{i,j} I_{i,j} F_j. \end{array}$$

If the generic sensors have the same cost c, then the formulation becomes a easy case of the Knapsack problem [16]. In this case the optimal solution is to sort the items in order of increasing value and insert them into the knapsack in this order until nothing fits. We present an optimal algorithm for generic sensor deployment in Algorithm 1. Since all generic sensors' costs are identical, we maximize the total information gain by taking the locations providing the greatest information utility. Given a budget B for deployment, our system can deploy at most B/c generic sensors in the field. We sort the information utility provided by the potential locations. Then, we deploy generic sensors in this order until all the available generic sensors have been deployed. The algorithm allows us to take as much as possible the locations with the greatest information utility.

Algorithm 1 Deployment Using Mobile Statistics

Input: n potential deployment location j with corresponding information utility of u_j
Output: An optimal solution $(z_1, ..., z_n)$ for generic sensor deployment
Set all $z_j (j = 1, ..., n)$ equal to 0;
Set the remaining available generic sensors $S_b = B/c$;
while $S_b > 0$ **do**
  Choose the j' with maximum $u_{j'}$;
  $z_{j'} = 1$;
  $S_b = S_b - 1$;
end while

Property: Our algorithm exhibits the greedy choice property. A globally optimal solution can be arrived by making a locally optimal (greedy) choice.

Proof: It is sufficient to show that as much as possible of the locations providing the highest information utility must be included in the optimal solution.

Let $u_{j'}$ be the information gain of the location providing the highest information utility (location j'). We define V as the total information utility obtained by generic sensor deployment, which can be computed by $V = \sum_{q=1}^{n} u_q$. If some of location j' is left for some $j \neq j'$, then replacing j with j' will yield a higher value, i.e. $u_j \leq u_{j'}$ by definition of j'.

V. Sensing Data Dissemination for Mobile Users

Mobile users can collect sensing data from stationary sensors through short range communication when they are close to each other. They can also take sensing measurements directly using their phones like taking pictures, measuring noise level, and remarking dangerous trails. Sensing data, taken either by stationary sensors or by mobile phones, could be uploaded to the generic sensors within the area of interests. This allows other users to download sensing information about the trails ahead in their journey. We consider sensing data in the format of $< ID_s, e_s, L_s, TTL_{e_s}, t_{e_s}, D_{e_s} >$, where ID_s is the identifier of the sensor, e_s is the identifier of the detected event, L_s is the location of the sensor, TTL_{e_s} is the time-to-live of the event, t_{e_s} is the collection time of the event data, and D_{e_s} is the sensing data. A mobile phone user, who has collected new sensing data, could upload the data to all the generic sensors that he meets as long as the data is still valid, i.e. $t - t_{e_s} \leq TTL_{e_s}$, where t is the current time and t_{e_s} is the data generation time. Given the maximum moving speed of mobile users v_{max}, the generic sensors that receive the sensing data from e_s must be located within R_s of sensor s.

Each of the mobile users and generic sensors will store their sensing data in an information list. When a mobile phone user and a generic sensor are within their communication range, they can exchange data with each other. Before each exchange, the user will remove outdated records from its information list, denoted by ML (Mobile List). Similarly, we denote the information list of the generic sensor as SL (Sensor List). The mobile phone and the sensor then exchange their ML and SL in the format of $r_s =< ID_s, e_s, L_s, TTL_{e_s}, t_{e_s}, D_{e_s} >$. If the data record r_s^{ML} is stored in ML but not in SL, it will be copied from ML to SL. If the data record exists in both ML and SL, it will be copied from ML to SL only when the collection time of the data t_{e_s} in ML is greater than t'_{e_s} in SL. It means that the data in ML is fresher than that in SL. The same algorithm is run on the generic sensor, which will also remove its expired data and copy new data to ML.

VI. Analysis

We analyze the data delivery rate and average data delay considering sensing data with time constraint T_c. The mobile users will upload and download sensing data with the generic sensors. The sensing data will expire after time T_c, so that only data received by users before T_c are valid.

A. Delivery Rate

We calculate $E_c[N_d]$ as the expected number of valid downloads per upload given T_c. If the next upload arrives before T_c, all the downloads will be served before the data expire. Otherwise, only the downloads arrived before T_c will be served successfully. We consider the arrivals of users follow a poisson process. The average upload and download arrival rates are λ_u and λ_d respectively. Let $f_u(t)$ be a distribution function indicating the probability that the next upload will arrive within time t. $P[X_d(t) = k]$ is the probability that k download arrivals occur during time interval t.

$$
\begin{aligned}
E_c[N_d] &= \int_{t=0}^{T_c} f_u(t)(\sum_{k=0}^{\infty} P[X_d(t) = k]k)dt \\
&+ \int_{t=T_c}^{\infty} f_u(t)(\sum_{k=0}^{\infty} P[X_d(T_c) = k]k)dt \\
&= \int_{t=0}^{T_c} f_u(t)(\sum_{k=0}^{\infty} \frac{(\lambda_d t)^k}{k!} e^{-\lambda_d t} k)dt \\
&+ \int_{t=T_c}^{\infty} f_u(t)(\sum_{k=0}^{\infty} \frac{(\lambda_d T_c)^k}{k!} e^{-\lambda_d T_c} k)dt \\
&= \int_{t=0}^{T_c} f_u(t)(\lambda_d t) e^{\lambda_d t} e^{-\lambda_d t} dt \\
&+ \lambda_d T_c \int_{t=T_c}^{\infty} f_u(t) e^{\lambda_d T_c} e^{-\lambda_d T_c} dt \\
&= \lambda_d \lambda_u \int_{t=0}^{\infty} t e^{-\lambda_u t} dt + \lambda_d T_c \lambda_u \int_{t=T_c}^{\infty} e^{-\lambda_u t} dt \\
&= \lambda_d \lambda_u [\frac{e^{-\lambda_u t}}{\lambda_u^2}(-\lambda_u t - 1)]_0^{T_c} - \lambda_d T_c [e^{-\lambda_u t}]_{T_c}^{\infty} \\
&= \frac{\lambda_d}{\lambda_u}(1 - e^{-\lambda_u T_c}).
\end{aligned}
\tag{2}
$$

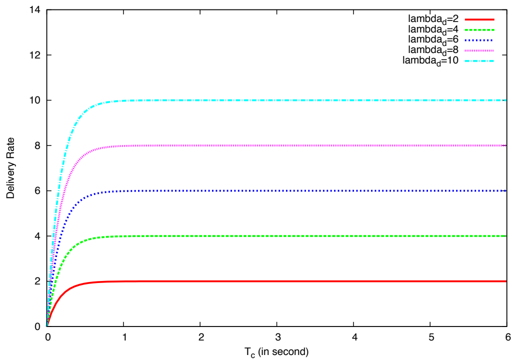

Fig. 1. Data delivery rate of successful downloads.

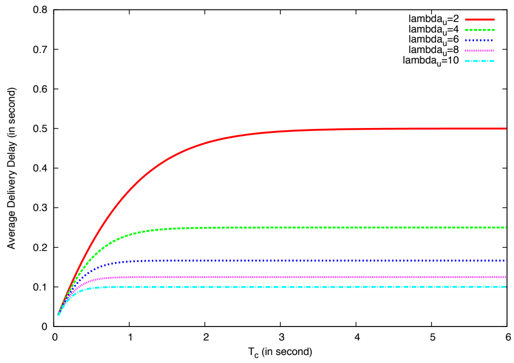

Fig. 2. Average delivery delay.

If the sensing data has a limited lifetime T_c, one piece of uploaded data can serve $E_c[N_d]$ valid downloads on average. Given the total number of download arrivals per upload be $E[N_d] = \lambda_d/\lambda_u$, the successful delivery ratio, S_c, can be computed by

$$S_c = \frac{E_c[N_d]}{E[N_d]} = 1 - e^{-\lambda_u T_c}. \tag{3}$$

Figure 1 shows the data delivery rate of successful downloads varying λ_d from 2 to 10 downloads/s with $\lambda_u = 6$ uploads/s.

B. Expected Delay

Similarly, we calculate the total data delay for all the valid downloads per upload within time constraint T_c.

$$\begin{aligned}
T_c[W] &= \int_{t=0}^{T_c} f_u(t)(\sum_{k=0}^{\infty} P[X_d(t)=k]k\frac{t}{2})dt \\
&+ \int_{t=T_c}^{\infty} f_u(t)(\sum_{k=0}^{\infty} P[X_d(T_c)=k]k\frac{T_c}{2})dt \\
&= \int_{t=0}^{T_c} f_u(t)(\lambda_d t)e^{\lambda_d t}e^{-\lambda_d t}\frac{t}{2}dt \\
&+ \int_{t=T_c}^{\infty} f_u(t)(\lambda_d T_c)e^{\lambda_d T_c}e^{-\lambda_d T_c}\frac{T_c}{2}dt \\
&= \frac{\lambda_d\lambda_u}{2}\int_{t=0}^{\infty} t^2 e^{-\lambda_u t}dt + \frac{\lambda_d T_c^2}{2}\int_{t=T_c}^{\infty} e^{-\lambda_u t}dt \\
&= \frac{\lambda_d\lambda_u}{2}[e^{-\lambda_u t}(-\frac{t^2}{\lambda_u} - \frac{2t}{\lambda_u^2} - \frac{2}{\lambda_u^3})]_0^{\infty} - \frac{\lambda_d T_c^2}{2}[e^{-\lambda_u t}]_{T_c}^{\infty} \\
&= \frac{\lambda_d}{\lambda_u^2}(1 - e^{-\lambda_u T_c}(T_c\lambda_u + 1)).
\end{aligned} \tag{4}$$

The expected delivery delay $E_c[W]$ for sensing data with limited lifetime T_c can then be computed by

$$E_c[W] = \frac{T_c[W]}{E_c[N_d]} = \frac{1 - e^{-\lambda_u T_c}(T_c\lambda_u + 1)}{\lambda_u(1 - e^{-\lambda_u T_c})}. \tag{5}$$

Figure 2 shows the expected delivery delay for successful downloads.

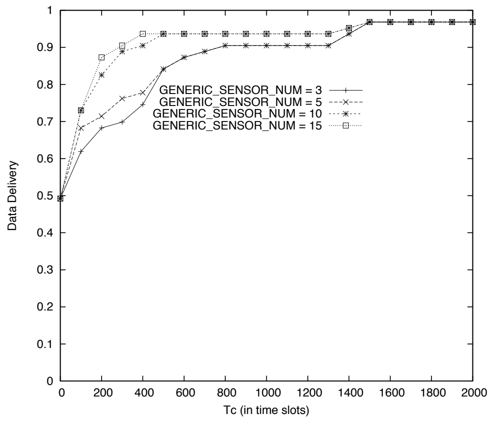

Fig. 3. Average data delivery varying number of generic sensors.

VII. Simulations

We evaluate our data dissemination mechanisms with real mobile traces collected by the mobile phone users in Disney World (Orlando) [17], [18]. The human mobility traces are collected with GPS receivers carried by 41 participants at every 10 seconds. These traces are mapped into a two dimensional area and recomputed to a position at every 30 seconds by averaging three samples over that 30 second period to account for GPS errors [17]. We monitor a 4km x 4km area at the center of the theme park for more than 10 hours. The sensing area is divided into grid cells of 50m x 50m, where sensors and mobile phones in the same grid cell can communicate with each other through short range communication.

We evaluate our data dissemination mechanism considering only data exchanges between mobile users and generic sensors. Figure 3 shows the data delivery rate varying T_c with different number of generic sensors. The results illustrate that more generic sensors can provide higher uploading rate and better data delivery rate. Also, higher data delivery rate can be achieved with an increase of time constraint T_c. Similarly, Figure 4 shows that more generic sensors can achieve lower data delay. The increase of data delay with T_c follows similar shape with the curve in Figure 2 of our analysis. Figure 5 shows the average communication overheads of users. More

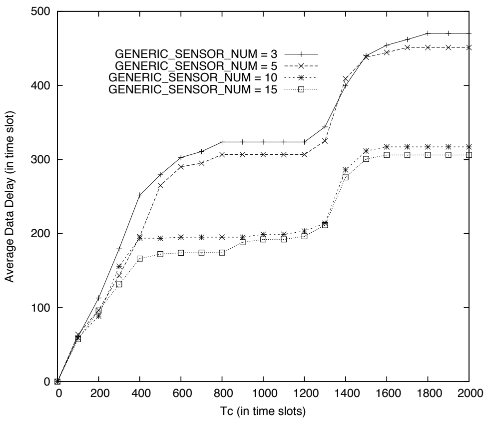

Fig. 4. Average data delay varying number of generic sensors.

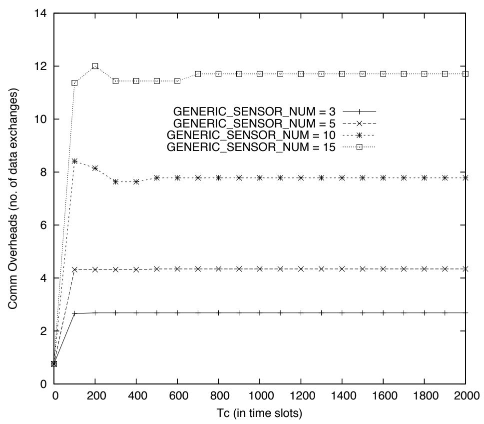

Fig. 5. Communication overheads varying number of generic sensors.

generic sensors can provide more communication opportunities, but also increase the communication overheads.

VIII. Conclusions

Sophisticated context-aware sensor data dissemination has been investigated for mobile users in sensing field with poor or no Internet connectivity. We presented a novel sensor data exchange framework for mobile phone users to obtain high information utility by uploading and downloading sensing data with stationary sensors. Generic sensors are carefully deployed according to the sensing context and mobility statistics of mobile users. We proposed a hybrid approach for opportunistic data exchange utilizing both generic sensors and mobile phones. Analysis has been conducted to evaluate the data delivery and expected data delay by varying the time-to-live, uploading and downloading rates of users. Through extensive simulations with real mobility traces, we have shown that our hybrid approach can achieve high data delivery rate and low data delay with small communication overhead.

As future works, we intend to extend the studies for hybrid networks with intermittent Internet connections in different parts of the sensing field. We will also explore the possibility to report and access sensing data through collaboration among mobile devices with different connectivities.

Acknowledgments

This work is supported in part by VINNMER program and Uppsala Excellence Center for Wireless Sensor Networks funded by VINNOVA Sweden, SICS Center for Networked Systems, and the US NSF under grant CNS-0910706. Jiangchuan Liu's research is supported by a Canadian NSERC Discovery Grant, a Discovery Accelerator Supplements Award, an NSERC Engage Grant, a MITACS Project Grant, and a China NSFC Major Program of International Cooperation Grant (61120106008).

References

[1] J. Burke, D. Estrin, M. Hansen, A. Parker, N. Ramanathan, S. Reddy, and M. B. Srivastava, "Participatory sensing," in *Workshop on World-Sensor-Web at SenSys*, Oct 2006.

[2] A. T. Campbell, S. B. Eisenman, N. D. Lane, E. Miluzzo, and R. A. Peterson, "People-centric urban sensing," in *Proc. of International Workshop on Wireless Internet (WICON)*, Aug 2006, pp. 18–31.

[3] Cellular One, "GSM voice and data coverage map," available at `http://www.icanmarketyou.com/cellularone/index.php`.

[4] CellReception, "Cell phone reception and tower search," available at `http://www.cellreception.com/`.

[5] E. Ngai, H. Huang, J. Liu, and M. Srivastava, "OppSense: Information sharing for mobile phones in sensing field with data repositories," in *Proc. of SECON*, Jun 2011.

[6] S. Eisenman, E. Miluzzo, N. Lane, R. Peterson, G. Ahn, and A. Campbell, "The bikenet mobile sensing system for cyclist experience mapping," in *Proc. of ACM SenSys*, 2007.

[7] D. Goodman, J. Borras, N. B. Mandayam, and R. D. Yates, "Infostations: a new system model for data and messaging services," in *IEEE Vehicular Technology Conference*, May 1997.

[8] R. Shah, S. Roy, S. Jain, and W. Brunette, "Data mules: Modeling a three-tier architecture for sparse sensor networks," in *Proc. of IEEE Workshop on Senosr Network Protocols and Applications (SNPA)*, May 2003, pp. 30 – 41.

[9] Y. Gu, D. Bozdag, E. Ekici, F. Ozguner, and C.-G. Lee, "Partitioning-based mobile element scheduling in wireless sensor networks," in *Proc. of SECON*, Sep 2005, pp. 386–395.

[10] W. Zhao, M. Ammar, and E. Zegura, "A message ferrying approach for data delivery in sparse mobile ad hoc networks," in *Proc. of ACM MobiHoc*, Mar 2005, pp. 187–198.

[11] N. Banerjee, M. D. Corner, and B. N. Levine, "An energy-efficient architecture for DTN throwboxes," in *Proc. of IEEE INFOCOM*, 2007.

[12] C. Boldrini, M. Conti, F. Delmastro, and A. Passarella, "Context-and social-aware middleware for opportunistic networks," *Journal of Network and Computer Applications*, 2010.

[13] E. Yoneki, P. Hui, S. Chan, and J. Crowcroft, "A socio-aware overlay for publish/subscribe communication in delay tolerant networks," in *Proc. of ACM MSWiM*, Oct 2007, pp. 225–234.

[14] E. Jaho and I. Stavrakakis, "Joint interest- and locality-aware content dissemination in social networks," in *Proc. of Conference on Wireless On Demand Network Systems and Services (WONS)*, Feb 2009, pp. 173–180.

[15] S.-B. Lee, S. H. Y. Wong, K.-W. Lee, and S. Lu, "Content management in a mobile ad hoc network: Beyond opportunistic strategy," in *Proc. of INFOCOM*, Apr 2011, pp. 266–270.

[16] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and et al., *Introduction to Algorithms*. The MIT Press, 2002.

[17] I. Rhee, M. Shin, S. Hong, K. Lee, and S. Chong, "On the levy-walk nature of human mobility," in *Proc. of INFOCOM*, Apr 2008, pp. 924–932.

[18] K. Lee, S. Hong, S. J. Kim, I. Rhee, and S. Chong, "Slaw: A mobility model for human walks," in *Proc. of INFOCOM*, Apr 2009, pp. 855–863.