Encyclopedia of GIS
Springer-Verlag 2008
10.1007/978-0-387-35973-1_795

Mobile P2P Databases

Yan Luo[1] and Ouri Wolfson[1]

(1) Department of Computer Science, University of Illinois at Chicago, Chicago, IL, USA

Without Abstract

Research supported by NASA Award NNA06AA25C, and NSF Grants OII-0611017, DGE-0549489, 0513736, and 0326284.

Definition

A mobile peer-to-peer (P2P) database is a database that is stored in the peers of a mobile P2P network. The network is composed by a finite set of mobile peers that communicate with each other via short-range wireless protocols, such as IEEE 802.11, Bluetooth, Zigbee, or ultra wide band (UWB). These protocols provide broadband (typically tens of Mbps) but short-range (typically 10–100 meters) wireless communication. On each mobile peer there is a local database that stores and manages a collection of data items or reports. A report is a set of values sensed or entered by the user at a particular time, or otherwise obtained by a mobile peer. Often a report describes a physical resource, such as an available parking slot. All the local databases maintained by the mobile peers form the mobile P2P database. The peers communicate reports and queries to neighbors directly, and the reports and queries propagate by transitive multi-hop transmissions. Figure 1 below illustrates the definition.

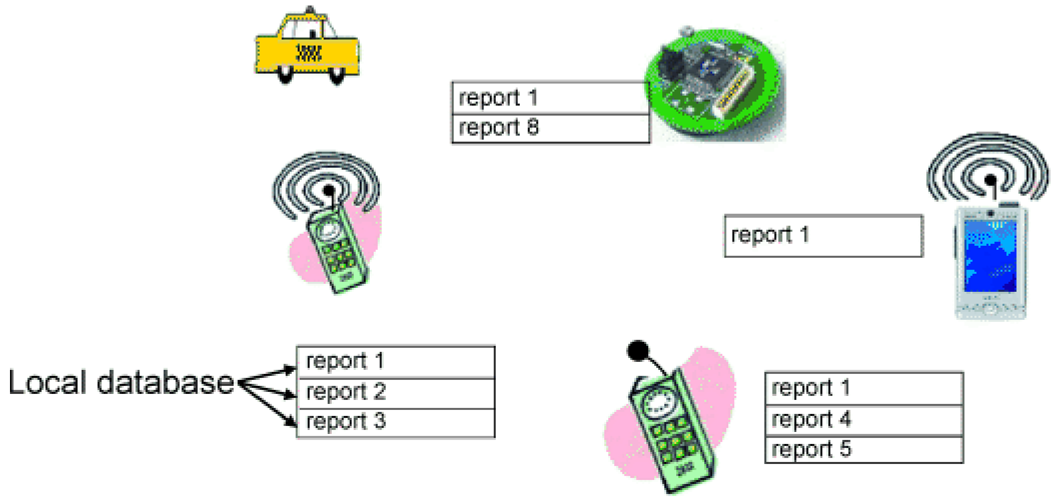

Figure 1 A mobile P2P database

In contrast to the assumptions made in the literature on mobile ad hoc networks (MANETs) and mesh networks, a peer may not know the identities of other peers in the network and the data they store. Thus, routing in the traditional MANET sense is not a common operation in mobile P2P databases.

Mobile P2P databases enable matchmaking or resource discovery services in many application domains, including social networks, transportation, mobile electronic commerce, emergency response, and homeland security.

Communication is often restricted by bandwidth and power constraints on the mobile peers. Furthermore, often reports need to be stored and later forwarded, thus memory constraints on the mobile devices constitute a problem as well. Thus, careful and efficient utilization of scarce peer resources (specifically bandwidth, power, and memory) are an important challenge for mobile P2P databases.

Historical Background

Traditionally search databases have been implemented by a centralized architecture. Google is preeminent example of such architecture. However, mobile P2P databases have several advantages over centralized ones. First, because short-range wireless networks utilize the unlicensed spectrum, communication to the mobile P2P database is free; there is also no cost involved in setting up and maintaining the fixed infrastructure database. Second, mobile P2P databases can be used for search in emergency, disaster, and other situations where the infrastructure is destroyed or unavailable. Third, mobile P2P databases are harder to mine for private information, and fourth, they are more reliable in the sense that failure of the central site will not render the system unavailable. Fifth, mobile P2P databases can withstand the high update rates that will be generated when representing temporary physical resources (e. g. the available parking slots), or continuous phenomena, such as the location of moving objects. The disadvantage of mobile P2P databases is that they do not provide answer guarantees. In other words, although the answer to a query exists in the database, due to mobility and lack of global coordination, the mobile P2P database may not find it.

The concept of mobile P2P database is proposed for searching local information, particularly information of a temporary nature, i. e., valid for a short duration of time [1].

Currently, there are quite a few experimental projects in mobile P2P databases. These can be roughly classified into pedestrians and vehicular projects. Vehicular projects deal with high mobility and high communication topology change-rates, whereas pedestrian projects have a strong concern with power issues. The following are several active experimental mobile P2P database projects for pedestrians and vehicles:

Pedestrians Projects

- **7DS** – Columbia University ([3])
  - http://www.cs.unc.edu/~maria/7ds/
  - Focuses on accessing web pages in environments where only some peers have access to the fixed infrastructure.
- **iClouds** – Darmstadt University ([4])
  - http://iclouds.tk.informatik.tu-darmstadt.de/
  - Focuses on the provision of incentives to brokers (intermediaries) to participate in the mobile P2P database.
- **MoGATU** – University of Maryland, Baltimore County ([5])
  - http://mogatu.umbc.edu/
  - Focuses on the processing of complex data management operations, such as joins, in a collaborative fashion.
- **PeopleNet** – National University of Singapore ([6])
  - http://www.ece.nus.edu.sg/research/projects/abstract.asp?Prj=101
  - Proposes the concept of information bazaars, each of which specializes in a particular type of information; reports and queries are propagated to the appropriate bazaar by the fixed infrastructure.
- **MoB** – University of Wisconsin and Cambridge University ([7])

  - http://www.cs.wisc.edu/~suman/projects/agora/
  - Focuses on incentives and the sharing among peers of virtual information resources such as bandwidth.
- **Mobi-Dik** – University of Illinois at Chicago ([1,2])
  - http://www.cs.uic.edu/~wolfson/html/p2p.html
  - Focuses on information representing physical resources, and proposes stateless algorithms for query processing, with particular concerns for power, bandwidth, and memory constraints.

Vehicular Projects

- **CarTALK 2000** – A European project
  - http://www.cartalk2000.net/
  - Develops a co-operative driver assistance system based upon inter-vehicle communication and mobile P2P databases via self-organizing vehicular ad hoc networks.
- **FleetNet** – Internet on the Road Project ([8])
  - http://www.ccrle.nec.de/Projects/fleetnet.htm
  - Develops a wireless multi-hop ad hoc network for intervehicle communication to improve the driver's and passengers' safety and comfort. A data dissemination method called "contention-based forwarding" (CBF) is proposed in which the next hop in the forwarding process is selected through a distributed contention mechanism based on the current positions of neighbors.
- **VII** – Vehicle Infrastructure Integration, a US DOT project
  - http://www.its.dot.gov/vii/
  - The objective of the project is to deploy advanced vehicle-to-vehicle (using the mobile P2P paradigm) and vehicle-to-infrastructure communications that could keep vehicles from leaving the road and enhance their safe movement through intersections.
- **Grassroots** – Rutgers University ([9])
  - http://paul.rutgers.edu/~gsamir/dataspace/grassroots.html

– Develops an environment in which each vehicle contributes a small piece of traffic information to the network based on the P2P paradigm, and each vehicle aggregates pieces of the information into a useful picture of the local traffic information.

Scientific Fundamentals

There are two main paradigms for answering queries in mobile P2P databases, one is report pulling and the other one is report pushing.

Report pulling means that a mobile peer makes an explicit request for the report it is interested in receiving, and the whole network is flooded with queries, the interested report will be pulled from the mobile peers that have them. Report pulling is widely used in resource discovery, such as route discovery in mobile ad hoc networks and file discovery by query flooding in wired P2P networks like Gnutella. Flooding in a wireless network is in fact relatively efficient as compared to wired networks because of wireless multicast advantage.

Another possible approach for data dissemination is report pushing. Report pushing is the dual problem of report pulling; reports are flooded, and consumed by peers whose query is answered by received reports. So far there exist mechanisms to broadcast information in the complete network, or in a specific geographic area (geocast), apart from to any one specific mobile node (unicast/mobile ad hoc routing) or any one arbitrary node (anycast). Report pushing paradigm can be further divided into stateful methods and stateless methods. Most stateful methods are topology-based, i. e., they impose a structure of links in the network, and maintain states of data dissemination. PStree, which organizes the peers as a tree, is an example of topology-based methods.

Another group of stateful methods is the cluster- or hierarchy-based method, such as [*13*], in which moving peers are grouped into some clusters or hierarchies and the cluster heads are randomly selected. Reports are disseminated through the network in a cluster or hierarchy manner, which means that reports are first disseminated to every cluster head and each cluster head then broadcasts the reports to the member peers in its group. Although cluster- or hierarchy-based methods can minimize the energy dissipation in moving peers, these methods will fail or cost more energy in highly mobile environments since they have to maintain a hierarchy structure and frequently reselect cluster heads.

Another stateful paradigm consists of location-based methods (see [*14*]). In location-based methods, each moving peer knows the location of itself and its neighbors through some localization techniques, such as GPS or atomic multilateration (see [*14*]).

The simplest location-based data dissemination is greedy forwarding, in which each moving peer transmits a report to a neighbor that is closer to the destination than it is. However, greedy forwarding can fail in some cases, such as when a report is stuck in local minima, which means that the report stays in a mobile peer whose neighbors are all further from the destination. Therefore, some recovery strategies are proposed, such as greedy perimeter stateless routing

(GPSR) [15]. Other location-based methods, such as geographic adaptive fidelity (GAF) [16] and geographical and energy aware routing (GEAR) [17]), take advantage of knowledge about both location and energy to disseminate information and resources more efficiently.

In stateless methods, the most basic and simplest one is the flooding-based method, such as [10]. In flooding-based methods, mobile peers simply propagate received reports to all neighboring mobile peers until the destination or maximum a hop is reached. Each report is propagated as soon as is received. Flooding-based methods have many advantages, such as no state maintenance, no route discovery, and easy deployment. However, they inherently cannot overcome several problems, such as implosion, overlap, and resource blindness. Therefore, other stateless methods are proposed, such as gossiping-based methods and negotiation-based methods.

Gossiping-based methods, such as [11], improve flooding-based methods by transmitting received reports to a randomly selected neighbor or to the neighbors that are interested in the particular content. The advantages of gossiping-based methods include reducing the implosion and lowering the system overhead. However, the cost of determining the particular interests of each moving peer can be huge and transmitting reports to a randomly selected neighbor can still cause the implosion problem and waste peers' memory, bandwidth and energy. Furthermore, dissemination and, thus, performance are reduced compared to pure flooding.

Negotiation-based methods solve the implosion and overlap problem by transmitting first the IDs of reports; the reports themselves are transmitted only when requested (see [12]). Thus, some extra data transmission is involved, which costs more memory, bandwidth, and energy. In addition, in negotiation-based methods, moving peers have to generate meta-data or a signature for every report so that negotiation can be carried out, which will increase the system overhead and decrease the efficiency.

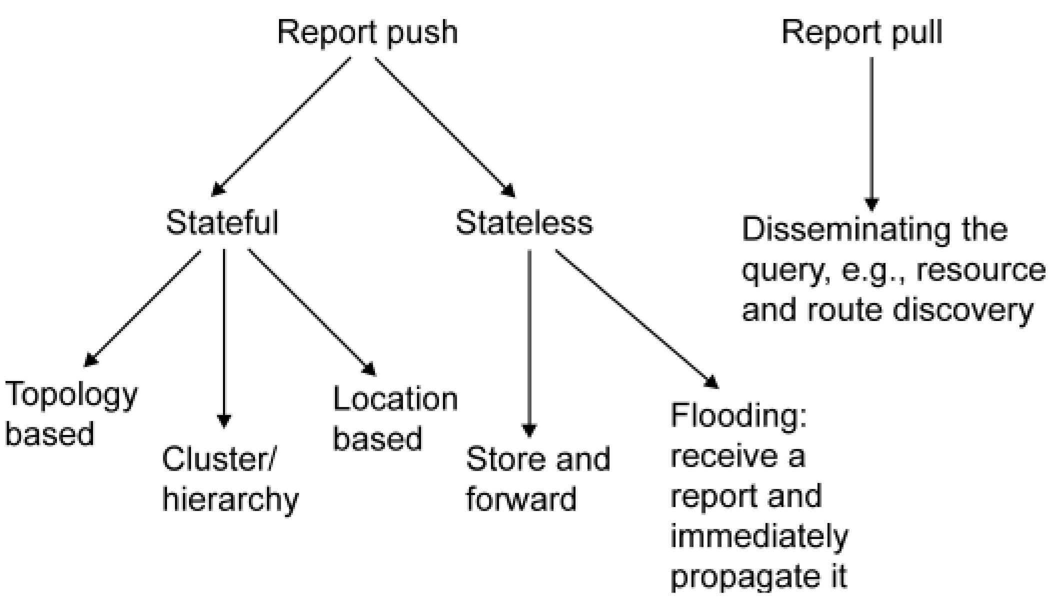

Figure 2 Query answering methods in mobile P2P databases

Another important stateless paradigm for data dissemination in mobile P2P networks is store-and-forward, such as [2], which to ranks all the reports in a peer's database in terms of their relevance or expected utility, and then the reports are communicated and saved in the order of their relevance. Alternatively, the reports requested and communicated are the ones with the relevance above a certain threshold. The notion of relevance quantifies the importance or the expected utility of a report to a peer at a particular time and at a particular location. Other store-and-forward methods include PeopleNet [6] and 7DS [3].

In summary, the paradigms for data dissemination in mobile P2P databases are summarized in Fig. 2 below.

Key Applications

Mobile P2P databases provide mobile users a search engine for transient and highly dynamic information in a local geospatial environment. Mobile P2P databases employ a unified model for both the cellular infrastructure and the mobile ad hoc environments. When the infrastructure is available, it can be augmented by the mobile P2P database approach.

Consider a mobile P2P database platform, i. e., a set of software services for data management in a mobile P2P environment; it is similar to a regular database management system, but geared to mobile P2P interactions. Such a platform will enable quick building of matchmaking or resource discovery services in many application domains, including social networks, emergency response, homeland security, military, airport applications, mobile e-commerce, and transportation.

Social Networks

In a large professional, political, or social gathering, mobile P2P databases are useful to automatically facilitate a face-to-face meeting based on matching profiles. For example, in a professional gathering, mobile P2P databases enable attendees to specify queries (interest profiles) and resource descriptions (expertise) to facilitate conversations, when mutual interest is detected. This opportunistic matchmaking can greatly enhance the value of networking events allowing users to connect with targeted, interested parties without a priori knowledge of their name, title, phone number, or other personal information. A face-to-face meeting can be setup by including in the resource description the identification information of the resource (person), such as cell-phone number, name, screen name, picture, physical description, etc. This information may be used together with the (possibly imprecise) location to help set up the face-to-face meeting. Thus, the individual's profile that is stored in mobile P2P databases will serve as a “wearable web-site”. Similarly, mobile P2P databases can facilitate face-to-face meetings in singles matchmaking.

Emergency Response, Homeland Security, and the Military

Mobile P2P databases offer the capability to extend decision-making and coordination capability. This finds applications in emergency environments, an area of particular concern to the government trying to find technologies that can be exploited to support the more than eight million first responders in US homeland security. Consider workers in disaster areas, soldiers and military personnel operating in environments where the wireless fixed infrastructure is significantly degraded or non-existent. They would welcome a capability that lets them automatically propagate messages, pictures, or resource information to other workers, based on matching profiles, security, and attribute values rather than node-id. As mobile users involved in an emergency response naturally cluster around the location of interest, a self-forming, high-bandwidth network that allows secure point-to-point or point-to-multipoint communication without the need of potentially compromised infrastructure could be of great benefit. For instance, a picture of a wanted person could be propagated to all those involved in a targeted search at the scene.

Consider a related emergency response application. Scientists are developing cockroach-sized robots or sensors that are carried by real cockroaches, which are able to search victims in exploded or earthquake-damaged buildings. These robots or sensors are equipped with radio transmitters. When a robot discovers a victim by sensing carbon dioxide, it may not have the transmission power to reach the outside rescuers; it can use local data dissemination to propagate the information to human rescuers outside the rubble. Sensors can also be installed on wild animals for endangered species assistance. A sensor monitors its carrier's health condition, and it disseminates a report when an emergency symptom is detected.

Airport Applications

Airports provide several different opportunities for the use of mobile P2P databases. From the point of view of commerce, airports have stores and kiosks where merchandise is sold similarly to a mall. Imagine arriving at a large airport and realizing you do not have the computer power cord you need for your presentation. Mobile P2P databases will enable a user to search for the needed product - just like in a mall. Merchants can similarly provide their location information and offer promotional incentives to passengers.

Mobile P2P databases can also be used by airport personnel to coordinate their activities. This is especially important when there is a communication failure due an emergency that degrades the infrastructure. Like the case of early responders, airport personnel can continue to coordinate their activities through the use of the mobile P2P network that is available even though the infrastructure is not functioning. Another potential opportunity that will benefit both the consumer and the airport operations is the dissemination of real-time information regarding flight changes, delays, queue length, parking information, special security alerts and procedures, and baggage information. This can augment the present audio announcements that often cannot be heard in nearby restaurants, stores, or restrooms, and the limited, expensive displays.

Mobile E-commerce

Consider short-range wireless broadcast and mobile P2P dissemination of a merchant's sale and inventory information. It will enable a customer (whose cell phone is mobile P2P databases enabled) that enters a mall to locate a desired product at the best price. When a significant percentage of people have mobile devices that can query retail data, merchants will be motivated to provide inventory/sale/coupons information electronically to nearby potential customers. The information will be provided and disseminated in a P2P fashion (in, say, a mall or airport) by the mobile P2P databases software.

Transportation Safety and Efficiency

Mobile P2P databases software can improve safety and mobility by enabling travelers to cooperate intelligently and automatically. A vehicle will be able to automatically and transitively communicate to trailing vehicles its “slow speed” message when it encounters an accident, congestion, or dangerous road surface conditions. This will allow other drivers to make decisions, such as finding alternative roads. Also, early warning messages may allow a following vehicle to anticipate sudden braking or a malfunctioning brake light, thus preventing pile-ups in some situations. Similarly, other resource information, such as ridesharing opportunities, transfer protection (transfer bus requested to wait for passengers), will be propagated transitively, improving the efficiency of the transportation system.

Inefficiencies in the transportation system result in excessive environmental pollution, fuel consumption, risk to public safety, and congestion. Ridesharing (i. e., vehicles carrying more than one person, either publicly provided such as transit, a taxi, or a vanpool, or prearranged rides in a privately owned vehicle) and car sharing (i. e., a program that allows registered users to borrow a car on an hourly basis from fixed locations) have the potential to alleviate these problems. Currently the matchmaking required in ridesharing is performed offline. However, the success of ridesharing will depend largely on the efficient identification and matching of riders/drivers to vehicles in real time in a local environment, which is where the benefit of our technology lies, providing information that is simultaneously relevant in time, location, and interest. Mobile P2P databases incorporated in navigational devices and PDA's can be used to disseminate to other devices and PDA's information about relevant resources, such as ridesharing partners, free parking slots, and available taxicabs or taxicab customers.

Future Directions

There are many challenges and directions for the future research in mobile P2P databases in mobile P2P networks:

1. **Prolong network lifetime** How to maximize the network life is a common but difficult problem in mobile P2P databases. Currently, some approaches as discussed above, e. g., ranking and cluster-based-methods, are proposed to address this problem and prolong the lifetime of sensor networks, mobile ad hoc networks, and mobile P2P databases. The future research question is how to employ the redundancy of networks and the density of peers in order to maximally extend the network lifetime.
2. **Sparse networks** Currently, the performance of many algorithms and systems heavily depends on the density of peers in mobile P2P networks. They do not perform very well if the network is sparse. Therefore, understanding how to design and develop mobile P2P databases for sparse networks is an important and difficult challenge. Recent work that

heads in this direction includes delay tolerant networks, store and forward flooding, and mobile peers whose sole function is to provide connectivity.

3. **Rapid topology changes** Another challenge for designing and developing mobile P2P databases is high mobility of peers. This poses problems to mobile P2P databases, e. g., how to efficiently disseminate queries and answers, and how to reconfigure rapidly when the topology of networks changes frequently. Stateless approaches seem most suitable to address these problems.
4. **Emergent global behavior from local knowledge** Mobile P2P databases can be treated as a special type of distributed system. Each peer maintains a local database and all the local databases form the virtual mobile P2P database. Therefore, peers can only use the local knowledge to predict or affect the global behavior of the whole mobile P2P database. The future research direction will be how to employ the local knowledge and propose the adaptive local algorithms to direct or affect the global behavior of mobile P2P databases.
5. **(Self-) localization techniques** Location-based approaches are more and more popular and necessary, and location information of peers is useful for efficiently storing and managing information. However, self-localization techniques are still not efficient and effective enough due to the limitation of peers or localization techniques. For example, GPS is not available indoors and the accuracy of GPS is not enough for some mobile P2P databases. Therefore, creating efficient and effective self-localization technique for mobile P2P databases is an important research direction.
6. **Integration of mobile P2P databases and infrastructure** As discussed above, mobile P2P databases do not guarantee answer completeness. In this sense, the integration with an available infrastructure, such as the Internet or a cellular network may improve performance significantly. This integration has two aspects. First, using the communication infrastructure in order to process queries more efficiently in the mobile P2P database; and second, using data on the fixed network in order to provide better and more answers to a query. The seamless integration of mobile P2P databases and infrastructure databases introduces important research challenges. Recent work on data integration in the database community can provide a starting point for is research.
7. **Specialized queries** Existing mobile P2P query processing methods deal with simple queries, e. g. selections; each query is satisfied by one or more reports. However, in many application classes one may be interested in more sophisticated queries. For example, in mobile electronic commerce a user may be interested in the minimum gas price within the next 30 miles on the highway. Processing of such queries may present interesting optimization opportunities.
8. **Mathematical modeling of data dissemination** Many query processing and data dissemination algorithms may benefit from a mathematical model of data propagation. For example, a formula giving the number n of mobile peers having a report that was generated at time t at location l would be very useful in ranking of such a report. The number n is a function of the density of mobile peers, motion speed, bandwidth and memory availability at the peers, memory management, etc. Related work done in epidemiology about the spread of infectious diseases would be a good starting point for this research. Results in random graphs may also be applicable.

Other important research directions include incentives for broker participation in query processing, and transactions/atomicity/recovery issues in databases distributed over mobile peers.

Recommended Reading

1. Xu, B., Wolfson, O.: Data Management in Mobile P2P Networks, Springer Verlag Lecture Notes in Computer Science, Proc. of the 2nd International Workshop on Databases, Information Systems, and P2P Computing (DBISP2P'04), Toronto, Canada, Aug 2004

2. Wolfson, O., Xu, B., Yin, H.B., Cao, H.: Search-and-discover in mobile P2P network databases. In Proc. of the 26th IEEE International Conference on Distributed Computing Systems (ICDCS'06), Lisbon, Portugal (2006)

3. Papadopouli, M., Schulzrinne, H.: Design and implementation of a P2P data dissemination and prefetching tool for mobile users, First NY Metro Area Networking Workshop, IBM TJ Watson Research Center, Hawthorne, NY (2001)

4. Heinemann, A., Kangasharju, J., Lyardet, F., Mühlhäuser, M.: iClouds – P2P information sharing in mobile environments, International Conference on Parallel and Distributed Computing (Euro-Par 2003), Klagenfurt, Austria (2003)

5. Perich, F.: On P2P Data Management in Pervasive Computing Environments, Ph. D. thesis, UMBC, May 2004

6. Motani, M., Srinivasan, V., Nuggehalli, P.: PeopleNet: engineering a wireless virtual social network, International Conference Mobile Computing and Networking (MobiCom'05), Aug 2005

7. Chakravorty, R., Agarwal, S., Banerjee, S., Pratt, I.: MoB: A mobile bazaar for wide-area wireless services, International Conference Mobile Computing and Networking (MobiCom'05), Aug 2005

8. Franz, W., Eberhardt, R., Luckenbach, T.: FleetNet – Internet on the road, The 8th World Congress on Intelligent Transportation Systems, Sydney, Australia, Oct 2001

9. Goel, S., Imielinski, T., Ozbay, K., Nath, B.: Grassroots: A scalable and robust information architecture, Technical Report DCS-TR-523, Department of Computer Science, Rutgers University, June 2003

10. Oliveira, R., Bernardo, L., Pinto, P.: Flooding techniques for resource discovery on high mobility MANETs, Workshop on Wireless ad hoc Networks (2005)

11. Datta, A., Quarteroni, S., Aberer, K.: Autonomous Gossiping: A self-Organizing epidemic algorithm for selective information dissemination in wireless mobile ad hoc networks. In: The International Conference on Semantics of a Networked World (2004)

12. Kulik, J., Heinzelman, W., Balakrishnan, H.: Negotiation-based protocols for disseminating information in wireless sensor networks, Wireless Networks, vol. 8, pp. 169–185 (2002)

13. Visvanathan, A.,Youn, J.H., Deogun, J.: Hierarchical data dissemination scheme for large scale sensor networks. In: IEEE International Conference on Communications (ICC'05), pp. 3030–3036, May 2005

14. Mauve, M., Widmer, A., Hartenstein, H.: A survey on position-based routing in mobile ad hoc networks, IEEE Network, vol. 15, no. 6, pp. 30–39 (2001)

15. Karp, B., Kung, H.T.: GPSR: Greedy perimeter stateless routing for wireless sensor networks. In: The 6th Annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom'00), pp. 243–254, Aug 2000

16. Xu, Y., Heidemann, J., Estrin, D.: Geography-informed energy conservation for ad hoc routing. In: The ACM International Conference on Mobile Computing and Networking, pp. 70–84, Rome, Italy, July 2001

17. Yu, Y., Govindan, R., Estrin, D.: Geographical and energy aware routing: A recursive data dissemination protocol for wireless sensor networks, Technical Report UCLA/CSD-TR-01-0023,UCLA, May 2001