Data-on-the-Road in Intelligent Transportation Systems

Ouri Wolfson and Bo Xu
Department of Computer Science
University of Illinois at Chicago
Chicago, IL 60532, USA
{wolfson,boxu}@cs.uic.edu

Abstract - *In this paper we examine the dissemination of information about resources in mobile ad-hoc networks, where moving objects communicate with each other via short-range wireless transmission. Each disseminated resource represents an observed spatial-temporal event, and the relevance of the resource to a moving object decays as the age of the resource and the distance from its location increase. We propose an opportunistic approach, in which an object propagates the resources it carries (namely the information that it has about these resources) to encountered objects and obtains new resources in exchange. Least relevant resources are discarded during exchanges so as to accommodate more relevant ones in the object's limited memory space. We propose two variants of the opportunistic dissemination algorithm, and a query processing scheme for finding all the resources in a geographic area.*

Keywords: data dissemination, mobile ad-hoc networks, resource discovery, intelligent transportation systems.

1 Introduction

Consider an urban area with hundreds of thousands of vehicles. Drivers and passengers in these vehicles are interested in information relevant to their trip. For example, a driver would like his/her vehicle to continuously display on a map, at any time, the available parking spaces around the current location of the vehicle. Or, the driver may be interested in the traffic conditions (e.g. average speed) one mile ahead. Such information is important for drivers to optimize their travel, to alleviate traffic congestion, or to avoid wasteful driving. The challenge is processing queries in this highly mobile environment, with an acceptable delay, overhead and accuracy. One approach to solving this problem is maintaining a distributed database stored at fixed sites that is updated and queried by the moving vehicles via the infrastructure wireless networks. Potential drawbacks of this approach are (i) the responses to queries may be outdated, (ii) the response time may not meet the real-time requirements, and (iii) access to infrastructure communication service is costly. (iv) currently there is no business model to provide a return-on-investment for setting up and operating the fixed sites. In this paper we explore a new paradigm that is based on peer-to-peer communications.

We assume that each moving object (i.e. vehicle) has the capability of communicating with its neighbors. This communication can be enabled by a local area wireless protocol such as IEEE 802.11 [6], Ultra Wide Band (UWB) [9], Bluetooth [7], or CALM [8]. These protocols provide broadband (typically tens of Mbps) but short-range (typically 50-100 meters) peer-to-peer communication. With such communication mechanisms, a moving object receives the desired information from its neighbors, or from remote objects by multi-hop transmission relayed by intermediate moving objects. Thus, resource dissemination is performed in a *mobile ad hoc network*. Compared to the traditional fixed-station based information query, this paradigm ("of the vehicles, by the vehicles, for the vehicles") has the following advantages:

1. It provides better information authenticity, accuracy, and reliability, especially for real-time information. Consider for example parking space availability. Information collected from a vehicle that is leaving a parking lot tends to be more reliable than that from the fixed site.
2. It is free of charge, assuming that vehicles are willing to relay messages for free (in exchange for their messages being relayed). Recent literature [10] discusses economic models and a currency for facilitating message communication and routing, so routers of others' messages receive credit to use towards the transmission of their own messages. A back of the envelope calculation reveals that the cost (in terms of fuel) of communicating with encountered vehicles is less than a cent per day, even if the communication is continuous throughout the day.

In this paper we propose an *opportunistic* approach to resource (parking slot, taxi-cab customer, etc.) dissemination, in which an object propagates the resources it carries to encountered objects, and obtains new resources in exchange. For example, an object finds out about available parking spaces from other objects. These spaces may either have been vacated by these encountered objects or these objects have obtained this information from other previously encountered ones. Thus the parking space information transitively spreads out across objects. Similarly, information about an accident or a taxi cab customer is propagated transitively.

The approach can also be used in dissemination of resources among pedestrians. For example, an individual wishing to sell a pair of tickets for an event (e.g. ball game, concert), may use this approach right before the event, at the event site, to propagate the resource information. Thus, we use the term *moving objects* to refer to both, vehicles and pedestrians.

In this paper we explore this resource propagation paradigm, which we call *roaming data*. In section 2 we describe the system model. In section 3 we describe discuss spatial and temporal resources and their relevance. In section 4 we describe some possible dissemination algorithms. In section 5 we outline the research issues in roaming data. In section 6 we discuss relevant work, and in section 7 we conclude.

2 The System Model

The system consists of fixed stations and moving objects. Each station senses resources and continuously announces them by wireless broadcast. Each announcement message contains the home and the create-time of the resource. In the parking slots example, a sensor in a parking slot monitors the slot, and, while unoccupied, announces its availability to the neighboring vehicles. In the cab customer example, the customer who needs a cab may press a button in a station at the closest intersection (similar to road-crossing buttons in the USA). The station announces the location of the intersection and the time at which the button is pressed. In the car accident example, the event may be announced by the sensor that deploys the air-bag.

A station announces a resource only for the time period during which the resource is valid. So for example, the parking slot sensor announces the slot only when it is available; occupation of the slot is sensed and the broadcast ceases. The cab-call station stops announcing the request once the customer takes a cab. When the station is announcing, all the moving objects that are within the wireless coverage of the station receive the announced resource and they may choose to save the resource in their memory.

Objects use broadband (typically tens of Mbps) but short-range (typically 50-100 meters) wireless technologies for peer-to-peer communication. An object is capable of detecting the objects that are within its transmission range. We say that two objects *encounter* each other when their distance is smaller than the transmission range. If two moving objects travel within the transmission range for a period of time, after the initial exchange only newly arrived resources are exchanged.

3 Resources and Their Relevance

Resources may be spatial, temporal, or spatio-temporal. Information about the location of a gas station is a spatial resource. Information about the price of a stock on 11/12/03 at 2pm is temporal. The relevance of spatial resources decreases as distance increases, and similarly, the relevance of temporal resources decreases with age. Spatio-temporal resources generalize the spatial and the temporal resources.

A *spatio-temporal resource,* or a *resource* for short, is a piece of information about an event (e.g. the availability of a parking space, or a car accident, the speed of a vehicle at a particular time-point, the availability of a taxi-cab customer at a particular location). The event is specific to a certain location that is referred to as the *home* of the resource. For example, the home of an available parking space is the location of the space, and the home of a cab request is where the cab is requested to reach for pickup. Each resource has a time duration for which it is valid. This duration is referred to as the *valid duration*. The start point of the valid duration is referred to as the *create time* of the resource. For example, the valid duration of the resource regarding the availability of a parking space is the time period since the space becomes available until it is occupied. The valid duration of the resource regarding a cab request is the time period since the request is issued until the request is satisfied or canceled. The valid duration of an accident starts when it occurs, and lasts until it is cleaned up.

The parking problem illustrates the above issues. It is clear that both time and distance are essential decision variables. As the information about a parking slot grows stale, it becomes less and less relevant as the likelihood of its availability decreases. Its relevance should therefore be less than that of a more recent one. This comparison must however be tempered by the distance factor. A parking slot that is closer to a vehicle is certainly more relevant than one that is farther away. In general, the relevance of a resource decays as its age (i.e. the length of the period of time since its create-time) increases, and the distance from its home increases. We use the following function to compute the relevance of resource *R*:

$$F(R) = -\alpha \cdot t - \beta \cdot d \quad (\alpha, \beta \geq 0) \qquad (1)$$

t is the age of R and d is the distance from the home location of R. α and β are constants that represent the decay factors of time and distance respectively. The bigger the ratio α/β, the more the relevance is sensitive to time than to distance; conversely, the relevance is more sensitive to distance than to time.

We proposed a linear relevance function, but clearly any monotonic function of the distance and time can be used. It is interesting to investigate the cases in which various functions are appropriate.

A final comment in this section concerns spatial resources, such as gas stations, ATM machines, etc. Opportunistic dissemination of such resources is an alternative paradigm to geographic web searching (see e.g. [16]). Geographic web searching has generated a lot of interest since many search-engine queries pertain to a geographic area, e.g. find the Italian restaurants in the town of Highland Park. Thus instead of putting up a web site to be searched geographically, an Italian restaurant may decide to put a short-range transmitter and advertise via opportunistic dissemination. In mobile systems, this also solves some privacy concerns that arise when a user asks for the closest restaurant or gas station. Traditionally, the user would have had to provide her location; but she does not need to do so in our scheme.

4 Dissemination Algorithms

In this section we describe two possible opportunistic resource dissemination algorithms.

4.1 Opportunistic Resource Dissemination (ORD)

The resources in memory are ranked according to their relevance. If the number of received resources exceeds the allocated memory, less relevant resources must be shed from memory to accommodate more relevant ones. We assume that a moving object has a fixed amount of memory allocated to each application (e.g. the user allocates 10 entries for relevant parking slots. In other words, the user wants only 10 parking slots to be saved and displayed). The question is then what should be an appropriate decision criterion for this purpose.

Two types of operations may be performed at a moving object *O*. The first type is *resource acquisition*, which is performed when *O* is within the coverage area of a station while the station is announcing a resource *R*. Upon reception of *R*, if *O*'s memory space is not full, *R* is saved in memory. If the memory is full, *O* computes the relevance of *R* based on the age of R and on the distance between the current location of *O* and the home of *R*. *O* also recomputes the relevance for each stored resource. If the relevance of *R* is higher than that of any stored resource, the least relevant one is purged, and then *R* is saved. Otherwise *R* is discarded.

The second type of operations is *resource exchange*, which is performed when *O* encounters a new object *O'*. If neither *O* nor *O'* is in the middle of data exchange with a third vehicle, the resource exchange is performed between *A* and *B* as follows. *O* and *O'* first exchange their resources. Upon receiving new resources, moving object *O* computes the relevance for each received resource and re-evaluates the relevance of its own resources. If all the resources do not fit in the memory space of *O*, the least relevant ones are purged.

In either resource acquisition or resource exchange, when *O* receives a resource *R*, if *O* has a resource *R'* in its memory such that *R* and *R'* have the same home and the create-time of *R* is greater than that of *R'*, then *R'* is replaced by *R*.

4.2 Opportunistic Resource Dissemination with Invalidation (ORDI)

With ORD, a resource in an object's memory may become invalid before it is purged out. This invalid resource introduces wrong information for decision making. For example, the resource may indicate an available parking space that is actually already occupied, or it may indicate a cab request that is already satisfied. Time may be wasted if the driver uses this resource to make decisions. In order to reduce the invalid resources, we developed Opportunistic Resource Dissemination with Invalidation (ORDI). ORDI works as follows. At each station, whenever the valid duration of a resource *R* ends, the station starts to announce an invalidation message for *R* until the beginning of the announcement of the next resource. The invalidation message contains the following three data items: (i) $T_{\text{invalid}}(R)$ the time when *R* becomes invalid (which is also the time when the invalidation message is created); (ii) $T_{\text{create}}(R)$ the create-time of *R*; and (iii) $H(R)$ the home of *R*. The invalidation message is a special resource. Its home is $H(R)$ and its create-time is $T_{\text{invalid}}(R)$, and it uses the same relevance function as a regular resource. We will refer to the invalidation message as the *invalidating resource of R*, and refer to *R* as a *regular resource*. The invalidating resource is acquired and exchanged similarly to a regular resource. The only difference is as follows. When an invalidating resource ($T_{\text{invalid}}(R)$, $T_{\text{create}}(R)$, $H(R)$) is received by an object, the object uses $T_{\text{create}}(R)$ and $H(R)$ to search *R* in its memory. If *R* is found, then it is replaced by the invalidating resource.

5 Research Issues

5.1 Data Dissemination

First, it is interesting to determine how a resource spreads over time and distance. Clearly, it would be wasteful for a resource such as the availability of a parking slot at 2pm to linger around in the vehicular network for days. Clearly this information would be stale. Similarly, it would be wasteful to find in New York the information about an available parking slot in Chicago. Therefore, we are looking for dissemination algorithms and relevance functions in which the spread of a resource is local, and in which the resource disappears from the system after a reasonable duration of time. Clearly the spread of the resource depends not only on the dissemination algorithm, but on many other parameters such as traffic density and speed, transmission range, relevance of the resource to each vehicle, total number of resources in the system and the memory allocation at each vehicle for these types of resources.

It would also be nice if the same algorithm disseminates purely spatial and purely temporal resources, by modifying the relevance function. For example, we would like the information about the location of a gas station not to decay, but to remain within a limited geographic area around the gas station.

Spread of a resource is also important for query processing. For example, consider queries that find all the resources, or the number of resources, within a particular geographic region R. For example, find all the cab requests within five blocks of the Sears Tower, or find the number of available parking spaces within the UIC campus. Determining how a resource spreads will help defining the area to which such a query should be propagated.

Spread of a resource can be analyzed theoretically or experimentally by simulations. On the theoretical side, a promising approach is considering opportunistic resource spreading as a form of epidemics [1, 3, 4, 5], where objects with resources are viewed as infectious individuals and vehicles accepting resources are similar to susceptible individuals. Starting from the classical epidemic Kermack-Mckendrick model [17], we investigated ways of adapting and expanding it to our more complex application.

For several reasons, resource spreading in mobile ad-hoc networks cannot be modeled using the simple SIS (Susceptible, Infectious, Susceptible) or the SIR (Susceptible, Infectious, Removed) models [17]. The propagation in these models concerns a single disease in a population where individuals are initially all susceptible to the disease, then possibly infected for some time, and later immune. In our case, we need to investigate the propagation of multiple diseases, and an individual is infected depending on the characteristic of each disease and those infecting him at a given point in time. This characteristic in our case is expressed by the relevance function. An individual may not be susceptible to an infection at one time and become infected at a later time. Because the relevance of a resource for a moving object is dynamic, an object may refuse a resource at one time and then later find it relevant. Further, an individual may become immune to a disease at one point and later become susceptible. An object may shed away a resource at one time just to determine it to be relevant at another time.

We have developed differential equations that model the dissemination of a temporal resource, and are working on extending them to the spatial and spatio-temporal cases.

5.2 Querying

Consider queries that find all the resources within a particular geographic region R. For example, find all the available parking spaces within the UIC eastern campus, or find all the cab requests within five blocks of the Sears Tower. When a moving object receives such a query from the user, it sends the query to all the objects that may have information about resources located inside the queried area. It then computes the answer to the query from the answers that it receives. The problems that need to be addressed include, (i) how to determine the set of objects to which the query is sent, (ii) how to disseminate the query to them; (iii) how are the answers collected and delivered to the query originator?

Consider problem (i). Observe that information about the resources in the region R may travel beyond the region. Thus, the query destination area P, i.e. the region to which the query is propagated, may be larger than the queried region R. Consider the following approach to determine the query destination area. Suppose that from the analysis of data dissemination, we know that the maximum distance to which a resource is propagated is bounded by some constant b. Assume that R is a polygonal area, and D_b denotes a disk with radius b. Then P is $\{R \cup$ interior of $R \cup$ the points which are in the "sweep" of D_b when its center moves along the edges of $P\}$. In computational geometry the region is called the *Minkowski sum* of R and D_b.

Now consider problem (ii), i.e. how the query is disseminated to all the objects in the destination area. One distinguishes among several situations, depending on what each moving object knows about the future motion plans of other objects in the system. One situation is that each object knows the trajectories of each other object. The trajectories can be known by the objects exchanging their trajectories, and trajectories of neighboring vehicles they are aware of, as resources. Another situation is that objects do not know the future expected trajectories of other objects; and there can be intermediate situations where some trajectories are known but some are not. In each one of these situations the propagation mechanism is different.

Finally, to propagate the answer back to the query originator there can be several strategies. First, each moving object can send to the query originator, o, the resources of R it is aware of; in turn, o consolidates the results (e.g. eliminates duplicates). Second possibility is that a leader is elected in the region P; the leader collects and consolidates the answers of the responding vehicles, before delivering them to o.

5.3 Networking

Traditional MANET routing protocols (see [2] for a survey) are not suited to the high mobility environment of vehicular networks. For deployment in vehicular networks, topology-based routing protocols would require a large number of routing states and incur large routing overheads for updating topology changes. Furthermore, routing is often data-centric with group communications referenced by attribute-based names, rather than being node-centric with communications referenced by global node identifiers. For example, a vehicle would query which locations ahead have average speed below a specific threshold, rather than the average speed of a particular vehicle.

For vehicular networks we need to explore communication paradigms that take advantage of emerging standards such as Ultra Wide Band (UWB) and the CALM network architecture. Medium Access Control problems are expected to be severe, and require innovative solutions. The paradigms need to support the mobile point-to-multipoint and multipoint-to-point communication types.

5.4 Economic Models

Unlike other mobile ad hoc networks in which all the nodes belong to a single authority (e.g., a single military unit or rescue team) and have a common goal, a vehicular network typically consists of vehicles owned by different authorities (different companies, different private owners) and each vehicle has its own goal. In such an environment, the driver of a vehicle may decide not to cooperate in resource exchange. However, our preliminary work suggests that the intensity of moving object interactions has fatal effects on the performance of the system. Specifically, the ratio between the valid resources and the invalid resources [1] in an object's memory decreases dramatically as the intensity of interactions decreases. Therefore, some incentive mechanism is needed to encourage resource exchange.

One possibility is to build the incentive model upon virtual currency. Each mobile node carries virtual currency in the form of a counter that is protected from illegitimate manipulation by a trusted and tamper resistant hardware module [18]. There can be many possibilities for the incentive scheme, depending on who pays, who charges, and how much is paid or charged. For example, the producer of a resource can pay an "advertisement" fee by attaching a certain amount of virtual currency in the announcement message. The amount of the advertisement fee would depend on how far away and for how long the producer wishes the resource data to roam. Each mobile node that transmits the resource withdraws a "commission" from the advertisement fee. A query that retrieves the average speed a mile ahead is a good example of a producer paid resource. In other words, in this case the producer of the query will pay for the propagation of the query, collection of the answers, and the delivery of the results. A gas station that advertises to neighboring vehicles is another example of producer-paid communication.

Another possibility is that the consumer of a resource, rather than the producer, pays the fee. Parking slot advertisements are examples of such resources. Consumer- and producer- paid resources can be combined, i.e. the communication cost can be covered partly by the consumer, and partly by the producer. Taxi-cab requests and advertisements are examples of this possible hybrid scheme.

This incentive model is inspired by [10] which proposed a similar incentive model for Mobile Ad-hoc Networks. However, that incentive model is geared towards MANET's, characterized by point to point communication, with known source and destination for each message.

5.5 Privacy and security

Important privacy concerns arise when a vehicle has to provide its location or future trajectory. In our case, this situation occurs when a vehicle generates a query and needs to specify where the answer needs to be returned. Anonymization techniques can be used to address this problem.

Security issues arise in the economic model we proposed. How do we prevent cheating? For example, how do we prevent a vehicle from generating fictitious resources, and be paid for them? How do we prevent a vehicle from advertising resources for which it has received invalidation messages? How do we make sure that a consumer always pays for the resources it consumes? It seems that public key crypto-systems can resolve these issues, but the details need to be worked out. Furthermore, frameworks in which such protocols can be proven secure need to be developed.

6 Relevant Work

Different resource discovery architectures have been developed for ubiquitous computing environments over the last few years. Typically these architectures consist of a dedicated directory agent that stores information about different services or data, a set of protocols that allow

[1] An invalid resource is a resource whose valid duration has ended.

resource providers to find a directory agent and to register with it, and a naming convention for resources. Examples are the Service Location Protocol (SLP), Jini, Salutation, and UPnP. In highly mobile environments, due to high variability of the network topology we cannot rely on any one component to be always available. Therefore, it is important to develop approaches that rely more on opportunistic exchanges of resources than on a dedicated resource directory.

A lot of work has been done on sensor networks from database community (see for example [14, 15]). However, the existing work considers only stationary sensors, whereas in our model there are both stationary sensors (stations) and moving sensors (moving objects).

Our work is also relevant to geographic routing (e.g. [11]), where the destination of a message is a geographic location or area. However, in most of the existing works in this area, message delivery is possible only if the source and destination are connected, namely there exists a path from the source to the destination. Li [12], Chen [13], and Vahdat [5] are among the first few works on routing in disconnected ad hoc networks. They either do not fit our context (e.g. [13] requires a moving object to actively move to reach the next object), or too aggressive on bandwidth consumption (e.g. [5] uses flooding). Furthermore, the existing work does not discuss how the destination area is decided.

7 Conclusion

In this paper we devised a model for dissemination of spatio-temporal resources in an infrastructure-less environment, in which the database is distributed among the moving objects. The moving objects also serve as routers of queries and answers. We discussed two possible resource dissemination algorithms which differ in their treatment of invalidation messages. We discussed the research issues for future investigation and proposed possible approaches for solving some of them.

8 References

[1] A. Beaufour, M. Leopold, P. Bonnet, Smart-Tag Based Data Dissemination, in *Proceedings of the First ACM International Workshop on Wireless Sensor Networks and Applications (WSNA)*, Atlanta, Georgia, 2002.

[2] E. Royer and C. Toh, A Review of Current Routing Protocols for Ad Hoc Mobile Wireless Networks, *IEEE Personal Communications*, pages 46--55, Apr. 1999.

[3] A. Khelil, C. Becker, J. Tian, K. Rothermel. An Epidemic Model for Information Diffusion in MANETs, MSWiM'02, pages 54-60, Atlanta, Georgia, September, 2002.

[4] A. Demers, D. Greene, C. Hauser, et al. Epidemic Algorithms for Replicated Database Maintenance, in *Proceedings of the 6th Annual ACM Symp. On Principles of Distributed Computing*, pages 1-12, Vancouver, BC, August 1987.

[5] Vahdat and D. Becker. Epidemic Routing for Partially Connected Ad Hoc Networks, *Technical Report CS-200006*, Duke University, April 2000. http://citeseer.nj.nec.com/vahdat00epidemic.html

[6] IEEE Computer Society. Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications.1997.

[7] J. Haartsen, M. Naghshineh, J. Inouye, O. J. Joeresson, and W. Allen. Bluetooth: Vision, Goals, and Architecture. *ACM Mobile Computing and Communications Review*, 2(4):38-45, October 1998.

[8] http://www.etsi.org/frameset/home.htm?/technicalactiv/ITS/ITS.htm

[9] http://www.ubisense.net/technology/uwb.html

[10] L. Buttyan and J. Hubaux. Nuglets: a Virtual Currency to Stimulate Cooperation in Self-Organized Mobile Ad Hoc Networks. *Technical Report DSC/2001/001*, Swiss Federal Institute of Technology – Lausanne, Department of Communication Systems, 2001.

[11] Y. Ko, N. Vaidya. GeoTORA: A Protocol for Geocasting in Mobile Ad Hoc Networks. *2000 International Conference on Network Protocols*, Osaka, Japan., Nov. 2000.

[12] Q. Li and D. Rus. Sending Messages to Mobile Users in Disconnected Ad-hoc Wireless Networks. In *Proc. 6th Annual ACM/IEEE International Conference on Mobile Computing*, MOBICOM, 2000.

[13] X. Chen, A. L. Murphy, Enabling Disconnected Transitive Communication in Mobile Ad Hoc Networks, *ACM Principals of Mobile Computing*, POMC'01, 2001.

[14] P. Bonnet, J. Gehrke and P. Seshadri. Towards Sensor Database Systems. *Proceedings of the Second International Conference on Mobile Data Management*, Hong Kong, January 2000.

[15] C. Intanagonwiwat, R. Govindan and D. Estrin. Directed Diffusion: A Scalable and Robust Communication Paradigm for Sensor Networks. *Proceedings of MobiCOM*, Boston, MA, 2000.

[16] Markowetz, A., Brinkhoff, T., Seeger, B., Exploiting the Internet As a Geospatial Database, *International Workshop on Next Generation Geospatial Information*, October 19-21, 2003, Boston, Mass.

[17] N. Bailey. *The Mathematical Theory of Infectious Diseases and its Applications* (second edition), Hafner Press, 1975.

[18] A. Pfitzmann, B. Pfitzmann, and M. Waidner. Trusting Mobile User Devices and Security Modules. *IEEE Computer*, February 1997.