Scalable Data Management in Mobile Environments: protocol design in MANETs and data management in mobile sensors.

Aris M. Ouksel
The University of Illinois at Chicago
aris@uic.edu

My presentation dealt with two areas of my current research interests: (1) Protocol design in Mobile Ad-hoc Networks, and (2) Sensor data management. Our presentation describes briefly the motivating challenges and our approach to dealing with them.

1. Protocol design in Mobile Ad-hoc Networks: Information discovery and dissemination.

Information discovery and dissemination in Mobile Ad-hoc Networks (MANETs) remains a challenging research issue, particularly in environments such as Vehicle-to-Vehicle networks where the conditions are very stringent. Regardless of the applications, new systems must: (i) handle wide density variations: traffic depends on location, day of week, time of day, etc; (ii) tolerate high mobility: vehicles typically move 10+ times faster than foot traffic (2m/s); (iii) be fault-tolerant: users will frequently connect and disconnect, communication channels are subject to enormous interference and collisions. As the external environment is impossible to completely control, it is necessary to design protocols, which can adapt to a variety of situations; (iv) avoid pure flooding as it is obviously impractical in a single network with tens of thousands of users, as will occur in a large city; (v) support efficient replication mechanisms to ensure a high level of robustness while avoiding congestion of the system; and finally, (vi) be context-aware in that it must be responsive to rapidly changing surrounding conditions, such as natural clusters, to improve performance and adapt to inevitable brokers.

Most current models show serious performance degradation at either high or low densities. For example, proactive models exhibit frequent gaps and overlaps, particularly when density and mobility are high. Others require presetting propagation parameters to outperform simpler models such as setting cluster leaders, thereby severely limiting the range of applicability. The approach assumes relatively stable networks or adequately dynamic recalculation methods, such as rapidly adjusting gossip probabilities when network density changes suddenly. Further work is still necessary to develop scalable dynamic methods. Others incur a significant overhead, exacerbated in highly mobile MANETs, which in turn impacts negatively scalability. Yet others attempt to support the delivery of specific messages to specific nodes by finding the “one best route”. These methods generally require a great amount of routing table maintenance and path pruning in a large network, particularly as the mobility increases, and are thus not efficient and/or not scalable.

Our approach has been to develop novel protocols based on market-based principles, where information discovery and dissemination in MANETs is modeled as

Dagstuhl Seminar Proceedings 06431
Scalable Data Management in Evolving Networks
http://drops.dagstuhl.de/opus/volltexte/2007/953

supply/demand of information commodities in an economic system. Each participant becomes thus a market-maker, which uses a utility to autonomously compute demand based on information gleaned from the interactions with other users. The magnitude of demand is tempered by the amount of system resources utilized to satisfy this demand. Thus, we believe economic models are likely to capture very well the behavior of information discovery and dissemination in MANETs. They can play a significant in regulating the trade-off between satisfying demand and utilizing system resources efficiently, and in developing scalable techniques.

We developed a publish/subscribe protocol, called SBSD (Self-Balanced Supply Demand), which uses a *utility function* based on subscription (demand) *age*, *frequency*, and *replication* to control propagation. Subscriptions compete for limited device memory on the basis of the *utility*. SBSD operates without central coordination overhead: no registration or routing tables, responds dynamically to network changes, and bounds the publication (supply) propagation based on the subscriptions' utilities. The main results achieved: (i) increased propagation and delivery of highly-demanded subscription *types;* (ii) robustness to changes in network size, density, high broker mobility and arrival rate, and broadcast faults, packet loss. The approach is however limited by subscriber-publisher spatial distance.

A number of issues are currently being explored, including dynamic methods for increasing replication area without sacrificing coverage, complex subscription matching, reliable storage of data with concurrent read/write (query/update) accesses, extensive push, push/pull, pull comparisons, streaming data and queries, pricing models for message forwarding, and finally, the deployment of these protocols in realistic applications such as health care, social computing, disaster recovery, tactical military control, etc.

2. **Mobile sensor networks and data management**

Applications in wireless ad hoc sensor networks require scalable localization methods and dynamic data space partitioning among sensors to support efficient query processing. These challenges are exacerbated by the inherent constraints of sensor devices in their processing speed, storage capability, battery power, and communication bandwidth. However, current stand-alone localization approaches are expensive and poorly adapted to dynamic environments. The high energy cost of these approaches led to a fundamental question: to what extent imprecise position information can be tolerated while at the same time enabling the design of efficient query processing. It turns out, as we have demonstrated in our research, that it is possible to support distributed indexing in mobile environments without requiring precise localization, thus demonstrating that substantial savings in energy and computation can be achieved while at the same time providing a means to support efficient query processing.

To meet the imperative of energy preservation in data manipulation operations without performance degradation, we propose a scalable and self-configuring scheme, which synergistically integrates localization and data space partitioning, and minimizes the otherwise prohibitive cost of separately performing localization and/or indexing. We first introduced a scheme, which enables sensors to self-configure into a consistent coordinate system and to dynamically partition the data space in a network with a small core of randomly scattered location-aware sensor within the geographical area spanning the network. Our experiments, under a variable number of location-aware sensors, show that the cost of the algorithm is significantly superior to the combination of the best previously proposed techniques for localization and indexing. As a result and because of the indexing, the overall cost of a variety of queries, such as range queries, aggregate queries, and approximation queries is significantly reduced.

We believe therefore there is need for the development of efficient distributed indexing mechanisms, which avoid the prohibitively expensive cost of precise localization. An important characteristic of this indexing must be locality preservation upon which efficient query processing is dependent.