Query Processing in Mobile Environments

Jignesh M. Patel
University of Michigan

Introduction

We are on the verge of a new revolution in computing in which a large number of devices, ranging from small personal digital assistants (PDAs) to "invisible" embedded devices will be pervasive in our physical world. The hand-held computing devices and the mobile phones that are available today are precursors to the types of devices that we are likely to see in the future. In addition to these computing devices that are currently available, we are also likely to encounter computing devices embedded in every day articles like appliances, clothes, watches, environmental sensor devices, etc. Besides having computing and storage resources, these devices will also have some connectivity to a data network. In many cases as the device moves, it may change the form of network connectivity; for example within a room a hand-held PDA might use a 802.11b or Bluetooth wireless connection, and may switch to using a cellular wireless connection when outside in an open space. These devices may want to query data that could be spread across the system. Many of these devices will have substantial computing and storage resources and will keep data cached locally. In addition, many of these devices will be capable of gathering data (as is done by sensor devices) and will produce data sets that may be of interest to other devices in the system.

From the perspective or managing and querying data in these systems, the following characteristics of this system need to be taken into account:

- Large number of devices: The number of devices in is system is likely to be large, and if we view the network of these devices as a distributed system, this system will be much larger than (in terms of number of nodes) than the distributed systems that exist today.

- Network variability: The network characteristics such as connectivity, bandwidth, network congestion, etc. are likely to vary often. *Ad hoc* networks may be formed by a collection of devices. In addition, devices may frequently change the form of network connectivity that they use.

- Heterogeneity in device characteristics: The system will have a wide variety of devices including traditional servers, sensor nodes that have limited computing and storage capabilities, personal devices (such as hand-held devices and cell phones) with moderate storage and computing capabilities. In many cases these devices will draw their power from a battery.

- Location-aware querying: Many of the personal devices in the system are likely to be mobile and will be *location-aware*. The devices will often subscribe to data and pose queries that are based on the location of the device.

Data distribution and query processing techniques that are designed for such environments must be scalable with respect to the large number of nodes, and data set sizes in the system. In addition these techniques must quickly adapt to changes such as changing network characteristics, or location of a mobile user, and must conserve the key resources, which for many battery-powered devices is the energy consumption.

Query Processing and Data Management Issues

One way to view this system is to think of it as having a *backend* and a *frontend* component. The backend consists primarily of traditional servers connected by traditional high-bandwidth and reliable communication networks. The frontend consists of *edge* devices that connect to the backend system, potentially at different points at different times. Scalable and efficient solutions for delivering data to the edge devices are required. In many cases the data that needs to be delivered to the edge devices will depend on attributes such as the current location of the device. For example a driver of an electric car, when driving on the road, may want to be *notified* when she is within two miles of a gas station that can recharge her electric car. The data delivery mechanisms must not only be scalable in handling a large number of data delivery requests, but must also be able to quickly adapt to factors such as changes in the network characteristics, coping with server delays, congestion, etc.

In many cases, data sets can be cached in the edge devices. Data management software running in the edge devices must be optimized based on the characteristics of the device. For example, if a device is running on battery, then optimizing for the power consumption is critical.

With location-aware devices, as the device moves around in the physical space, it is possible to collect the *trajectory* of the device (and the user of the device). Large repositories of trajectory data are likely to emerge in the future, and these will find use in a variety of applications. Techniques for efficiently storing and analyzing the trajectory data will be essential to extract useful information from these data sets. For example, consider a patient that is diagnosed with a disease that is caused by a rare form of bacteria. To identify the source of the infection, and to control the spread of the disease, epidemiologists need to reconstruct the past history of the patient. If the patient was a cell phone user and subscribed to a service that was E-911 enabled (which requires that the location of the cell phone user be identified within a distance of approx 100 meters), then the epidemiologists can access the *trajectory* describing the recent movements of the patient, and can also identify others who might have recently come in contact with the patient.