Proceedings of the Conference for the e-Democracy, pp. 203-211, Athens, Sept. 2003

Data Management Issues and Technologies for Location-Aware Computing

Dimitris Katsaros *Yannis Manolopoulos* *Apostolos N. Papadopoulos*

Department of Informatics
Aristotle University of Thessaloniki
54006 Thessaloniki, GREECE
{dkatsaro, yannis, apostol} @ delab.csd.auth.gr

Abstract

Modern applications pose high quality requirements with respect to the type of processing needed and the expected system performance. A modern research direction is *location-aware computing*, which aims at providing services to users taking into consideration the location of the user in space. In order for location-aware computing to be feasible, several data management issues must be addressed. Geographical data must be available in order to determine the position of user terminal devices. Efficient techniques must be available to compute efficiently the location of a terminal device. Moreover, algorithmic techniques that take into consideration the mobility of users must exist in order to predict a future location. Finally, effective and efficient data mining algorithms are needed in order to extract useful knowledge with respect to user demands and therefore increase the quality of services. In this paper we address the data management issues involved towards providing location-aware services.

Keywords

Data Management, Spatial Databases, Data Mining, Location-Aware Computing

1. Introduction

Much of the information that underpins the challenges facing society today — terrorist activities, global environmental change and natural disasters — is geospatial in nature. An everyday example of key geospatial data is the location of a cell phone user who has dialed 100 in an emergency. In terrorist situations, the origins and destinations of phone calls, travel patterns of individuals, dispersal patterns of airborne chemicals, assessment of places at risk, and the allocation of resources are all geospatial data. As the volume of geospatial data increases by several orders of magnitude over the coming years, so will the potential for corresponding advances in knowledge of our world and in our ability to react to change.

Thus, an interdisciplinary research and development initiative focusing on challenges in location-aware computing, databases, data mining and human-computer-interaction technologies has emerged. Its contribution lies in integrating diverse perspectives, especially on future cross-disciplinary research directions. As Figure 1 shows, the convergence of advances in these three areas, combined with a sharp increase in the quality and quantity of geospatial information, promises to transform our world. Geospatial data have become critically important in areas ranging from crisis management and public health to national security and international commerce (Patterson 2003).

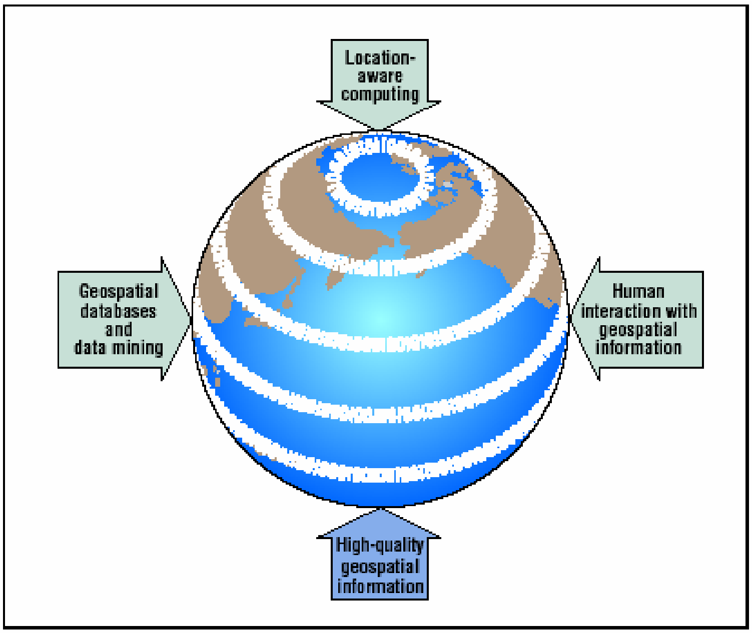


Figure 1. The convergence of four independent technological advances can transform the world.

Earlier we used the term location-aware computing, but what is exactly the meaning of this term? Location-aware computing is a special case of *context-aware* computing, which describes the special capability of an information infrastructure to recognize and react to a real-world context. Context, in this sense, comprises any number of factors, including user identity; current physical location; weather conditions; time of day, date, or season; and whether the user is asleep or awake, driving or walking. The most critical aspects of context are location and identity. Location-aware computing systems respond to a user's location, either spontaneously (for instance, warning of a nearby hazard) or when activated by a user request (for example, is it going to rain in the next hour?).

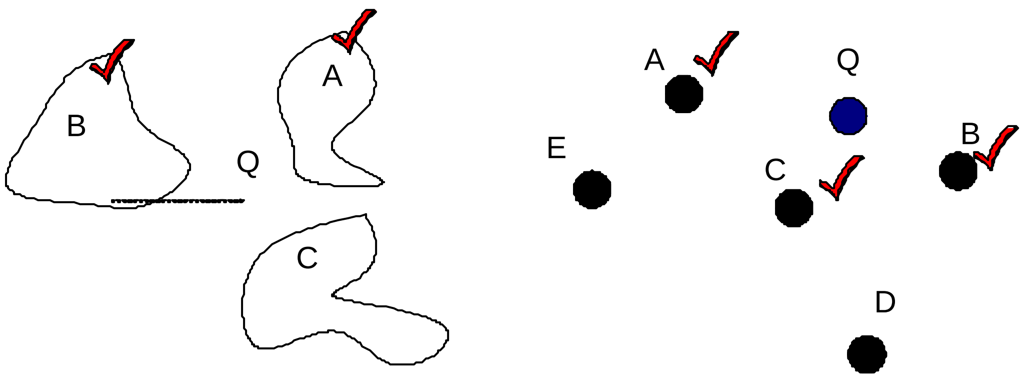
The paper is organized as follows. In Section 2 we cover the important issues that are related to data management. Several new trends are presented and explained along with their impact on location-aware computing. In Section 3 a detailed presentation is performed for the technologies that support location-aware computing. Finally, Section 3 concludes the paper.

2. Data Management Issues

Modern applications require the storage and management of diverse data types that are not so simple as number and characters (Silberchatz 1997, Manolopoulos 2000). For example, in order to be able to apply full text retrieval in a digital library application the database management system (DBMS) must provide advanced tools (Baeza-Yates 1999). As another example, consider an image database system storing medical images (e.g. MRI scans). Again the DBMS must provide techniques for efficient retrieval of similar images, given an image as input (image retrieval by content). In this section we focus on two data management research directions that are related to location-aware computing: *geographical information systems* and *data mining.*

2.1 Geographical Data Management

Geographical Information Systems (G.I.S.) are tools that enable the storage and retrieval of information about the spatial properties of objects (Rigaux 2001). Consider for example an application that requires the retrieval of objects based on the exact or approximate location of a reference object. A user may request for all objects that are contained in a specific geographic area (range query) or request for the 10 objects that are closer to a reference position (nearest-neighbor query). Usually, the data volume in such application is huge, and this poses another challenge for the DBMS. Figure 1 provides examples of range and nearest-neighbor queries.



(a) range query:
find the areas that intersect rectangle Q

(a) nearest-neighbor query:
find the three cities closer to traveler Q

Figure 1. (a) range query example, (b) nearest-neighbor query example.

In order to answer such queries efficiently we require:

- effective representation of the geospatial objects in the database
- effective and efficient algorithms and access methods in order to provide the required answers fast.

With respect to the first issue, advanced data types have been proposed to capture the semantics of the spatial properties (position, size, geometry). These data types must be supported by the

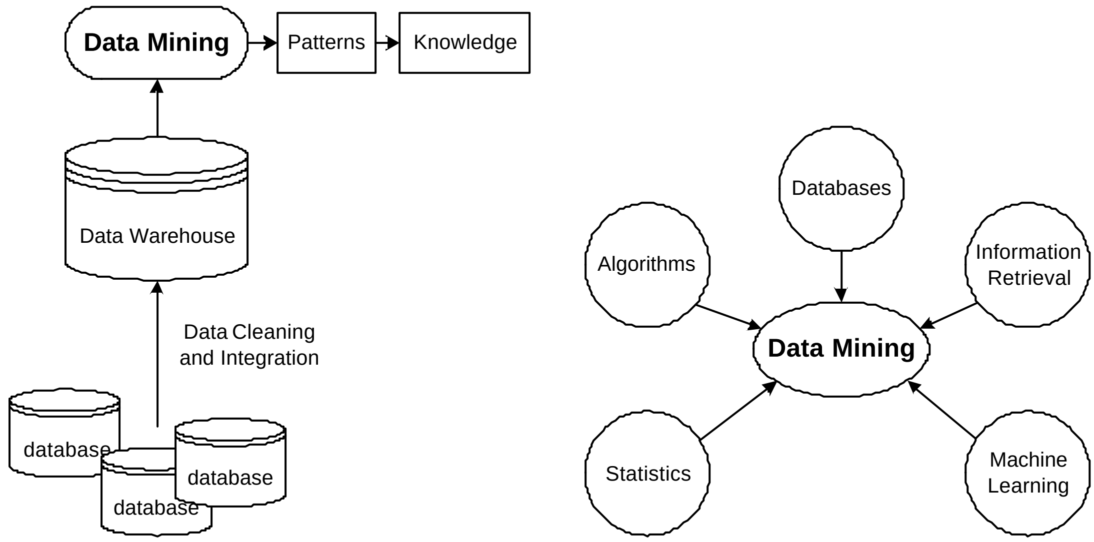
DBMS as ordinary data types (numbers and characters). With respect to the second issue, advanced data structures have been investigated towards increased query processing efficiency.

In some cases there is a need to predict the position of a set of objects in the near future. For example, consider the query "*give the locations of the two nearest gas stations five minutes from now*". To answer such queries the concept of time must be captured in the DBMS. Combining spatial and temporal information in a DBMS is an active research area, aiming at providing efficient representation and processing for spatio-temporal queries (queries that need spatial and temporal information to be answered).

From the above discussion it is evident that efficient representation and retrieval of objects based on their location is of vital importance for high-quality location-aware services.

2.2 Data Mining

Data Mining is another active research direction and aims at discovering knowledge from the underlying data (Han 2001, Dunham 2003). A major problem that arises in databases with huge data volumes, is that, data become unmanageable for humans. There is hidden information in these data that only specialized knowledge discovery algorithms can reveal. Figure 2 shows the role of data mining in the knowledge discovery process and the disciplines that has been affected by.



(a) Data Mining as a part of the Knowledge Discovery Process

(b) Data Mining has been affected by many disciplines

With efficient data mining techniques the user can discover hidden patterns in the data and therefore draw conclusions about the data and their meaning. Some of the tools that data mining techniques offer are:

- **data clustering**, grouping data into meaningful clusters,
- **state prediction**, predicting future data states according to past and present states,
- **association rule mining**, discovering relationships among the data.

3. Technologies Supporting Location-Aware Computing

Having, briefly, described the major issues of geographical information systems and data mining, we proceed with a more detailed discussion on location-aware computing. Location-aware computing is a new class of computing, which emerged due to the evolution of location sensing, wireless networking, and mobile computing. A location-aware computing system must be cognizant of its user's state and surroundings, and must modify its behavior based on this information. A user's context can be quite rich, consisting of attributes such as physical location, physiological state (e.g., body temperature and heart rate), emotional state (e.g., angry, distraught, or calm), personal history, daily behavioral patterns, and so on. If a human assistant were given such context, s/he would make decisions in a proactive fashion, anticipating user needs. In making these decisions, the assistant would typically not disturb the user at inopportune moments except in an emergency.

3.1 Location sensing

3.1.1 The Global Positioning System

The Global Positioning System (GPS) is the most widely known location-sensing system today (GPSWorld). Using time-of-flight information derived from radio signals broadcast by a constellation of satellites (see Figure 2) in earth orbit, GPS enables a relatively cheap receiver (on the order of 100€ today) to deduce its latitude, longitude, and altitude to an accuracy of a few meters. Although certainly important, GPS is not a universally applicable location sensing mechanism, for several reasons. It does not work indoors, particularly in steel-framed buildings, and its resolution of a few meters is not adequate for many applications. GPS uses an absolute coordinate system, whereas some applications (for example, guidance systems for robotic equipment) require coordinates relative to specific objects. In addition, the specialized components needed for GPS impose weight, cost, and energy consumption requirements that are problematic for mobile hardware. Consequently, several other location sensing mechanisms have been developed, which vary significantly in their capabilities and infrastructure requirements. System costs vary as well, reflecting different trade-offs among device portability, device expense, and infrastructure needs.

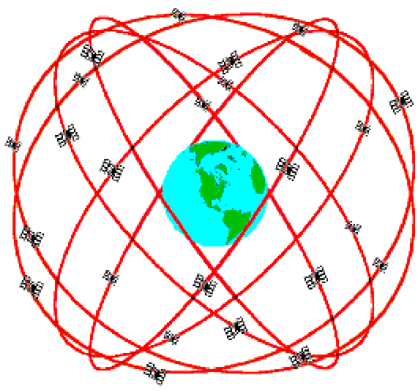


Figure 2. The GPS operational constellation.

3.1.2. Wireless Indoor Geolocation systems

The main elements of such a system (see Figure 3) are a number of location sensing devices that measure metrics related to the relative position of a mobile terminal (MT) with respect to a known reference point (RP), a positioning algorithm that processes metrics reported by location sensing elements to estimate the location coordinates of MT, and a display system that illustrates the location of the MT to users. The location metrics may indicate the approximate arrival direction of the signal or the approximate distance between the MT and RP. The angle of arrival (AOA) is the common metric used in direction-based systems. The received signal strength (RSS), carrier signal phase of arrival (POA), and time of arrival (TOA) of the received signal are the metrics used for estimation of distance. As the measurements of metrics become less reliable, the complexity of the position algorithm increases.

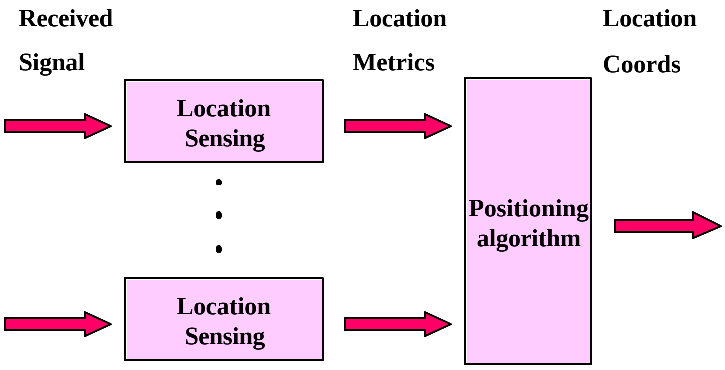


Figure 3. Functional diagram of an indoor geolocation system.

3.2 Wireless communications

The past decade has seen explosive growth in the deployment of wireless communication technologies. Although voice communication (cell phones) has been the primary driver, data communication technologies have also grown substantially. The IEEE 802.11 family of wireless LAN technologies is now widely embraced, with many vendors offering supporting hardware. Bluetooth is another standard, backed by a growing number of hardware and software vendors; although it offers no bandwidth advantage over 802.11, it was designed to be cheap to produce and frugal in its power demands. Infrared wireless communication is the lowest-cost wireless technology available today, primarily because it is the mass-market technology used in TV remote controls. Most laptops, many handheld computers, and some peripheral devices such as printers are manufactured today with built-in support for IrDA (the Infrared Data Association standard for wireless data transfer via infrared radiation). Infrared wireless communication must be by line of sight, with range limited to a few feet. High levels of ambient light also affect it adversely, such as prevail outdoors during daylight hours.

3.3 Mobile computing

Mobile computing is commonly associated with small-form-factor devices such as PDAs and tetherless (wireless) connectivity. Such devices provide access to information processing and communication capabilities, but do not necessarily have any awareness of the context in which they operate. The premise of mobile computing is to provide connectivity in an “anyplace, anytime” fashion.

Mobile computing emerged from the integration of cellular technology with the Web and it is not a special case of distributed computing, because it is characterized by four constraints:

- Mobile elements are resource-poor relative to static elements. For a given cost and level of technology, considerations of weight, power, size and ergonomics will exact a penalty in computational resources such as processor speed, memory size, and disk capacity. While mobile elements will improve in absolute ability, they will always be resource-poor relative to static elements.

- Mobility is inherently hazardous. In addition to security concerns, portable computers are more vulnerable to loss or damage.
- Mobile connectivity is highly variable in performance and reliability. Some buildings may offer reliable, high-bandwidth wireless connectivity while others may only offer low-bandwidth connectivity. Outdoors, a mobile client may have to rely on a low-bandwidth wireless network with gaps in coverage.
- Mobile elements rely on a finite energy source. While battery technology will undoubtedly improve over time, the need to be sensitive to power consumption will not diminish. Concern for power consumption must span many levels of hardware and software to be fully effective.

3.4 Data delivery alternatives in location-aware environments

As the world and its inhabitants are increasingly "wired," individuals traveling through and between places have real-time access to an increasing variety of information, much of it geospatial in nature. Freeing users from desktop computers and physical network connections will bring geospatial information into a full range of real-world contexts, revolutionizing how people interact with the world around them. Imagine, for example, the ability to call up place-specific information about nearby medical services, to plan emergency evacuation routes during a crisis, or to coordinate the field collection of data on vector-borne diseases, using just a handheld device.

This new setting changed the way the information is delivered to the clients. Support for different styles of data delivery allows a distributed information system to he optimized for various server, client, network, data, and application properties. We can identify three main characteristics that can be used to compare data delivery mechanisms:

- push versus pull.
- periodic versus aperiodic.
- unicast versus 1-to-N.

While there are numerous other dimensions that should he considered, such as fault-tolerance, error properties, network topology, etc. we have found that these three characteristics provide a good basis for discussing many popular approaches. Figure 4 shows these characteristics and how several common mechanisms relate to them.

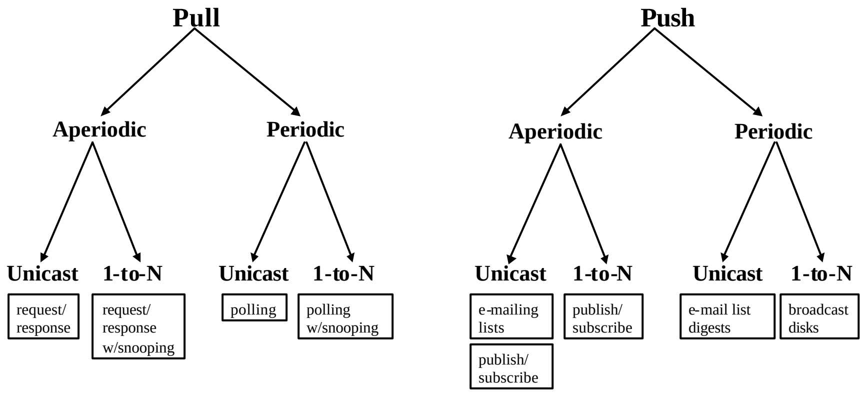


Figure 4. Data delivery options.

3.4.1 Client pull versus server push

Current database servers and object repositories manage data for clients that explicitly request data when they require it. When a request is received at a server, the server locates the information of interest and returns it to the client. This request-response style of operation is pull-based. In contrast, push-based data delivery involves sending information to a client population in advance of any specific request. With push-based delivery, the server initiates the transfer.

3.4.2 Aperiodic versus periodic

Both push and pull can be performed in either an aperiodic or periodic fashion. Aperiodic delivery is event-driven - a data request (for pull) or transmission (for push) is triggered by an event such as a user action (for pull) or data update (for push). In contrast, periodic delivery is performed according to some pre-arranged schedule. This schedule may he fixed, or may he generated with some degree of randomness. An application that sends out stock prices on a regular basis is an example of periodic push, whereas one that sends out stock prices only when they change is an example of aperiodic push.

3.4.3 Unicast versus 1-to-N

The third characteristic of data delivery mechanisms is whether they are based on unicast or 1-to-N communication. With unicast communication, data items are sent from a data source (e.g., a single server) to one other machine, while 1-to-N communication allows multiple machines to receive the data sent by a data source. Two types of 1 -to-N data delivery can he distinguished: multicast and broadcast. With multicast, data is sent to a specific subset of clients who have indicated their interest in receiving the data. Since the recipients are known, given a two-way communications medium it is possible to make multicast reliable; that is, network protocols can be developed that guarantee the eventual delivery of the message to all clients that should receive it. In contrast, broadcasting sends information over a medium on which an unidentified and possibly unbounded set of clients can listen.

Concluding Remarks

Location-aware computing is a special case of *context-aware* computing, which describes the special capability of an information infrastructure to recognize and react to a real-world context. Context, in this sense, comprises any number of factors, including user identity; current physical location; weather conditions; time of day, date, or season; and whether the user is asleep or awake, driving or walking.

In this paper we briefly discussed how geospatial information processing and data mining tools can be combined with location-aware computing in order to provide high quality location-aware services. In our days, we are witnessing a tremendous research effort towards combining several disciplines together in order to provide more useful and robust systems with worldwide impact.

References

R. Baeza-Yates and B. Ribeiro-Neto, (1999) "Modern Information Retrieval", Addison-Wesley.

M.H. Dunham (2003), "*Data Mining: Introductory and Advanced Topics*", Prentice Hall.

GPSWorld: http://www.gpsworld.com

J. Han and M. Kamber, (2001) "*Data Mining: Concepts and Techniques*", Morgan Kaufmann.

Y. Manolopoulos, Y. Theodoridis and V.J. Tsotras, (2000) "Advanced Database Indexing", Kluwer Academic Publishers.

C.A. Patterson, R.R. Muntz, C.M. Pancake, (2003) "Challenges in Location-Aware Computing", *IEEE Pervasive Computing*, Vol.2, No.2, pp.80-89.

P. Rigaux, M. Scholl and A. Voisard, (2002) "*Spatial Databases with Applications to GIS*", Morgan Kaufmann.

A. Silberschatz, H.F. Korth and S. Sudarshan, (1997) '*Database System Concepts*'', McGraw-Hill.