Dealing with Mobility: Issues and Research Challenges

Evaggelia Pitoura and Bharat Bhargava
Department of Computer Sciences
Purdue University
West Lafayette, IN 47907
{pitoura, bb}@cs.purdue.edu

CSD-TR-93-070
November 1993

Abstract

Recent advances in hardware and communication technology have made mobile computing possible. It is expected, [BIV92], that in the near future, tens of millions of users will carry a portable computer with a wireless connection to a worldwide information network. This rapidly expanding technology poses new challenging problems. The mobile computing environment is an environment characterized by frequent disconnections, significant limitations of bandwidth and power, resource restrictions and fast-changing locations. The peculiarities of the new environment make old software systems inadequate and raise new challenging research questions. In this report we attempt to investigate the impact of mobility on the todays software systems, report on how research starts dealing with mobility and state some problems that remain open.

Contents

1 Introduction

Technical advances in the development of portable computers and the rapidly expanding cordless networking technology provide the base for what is called *mobile* or *nomadic* computing. Today, when a user moves, he unplugs the computer from the local area network, transports it, and plugs it back to the local area network at his destination. In a mobile environment the user should be able to remain connected to the network even while moving. This possibility raises a proliferation of new research challenges.

The purpose of this report is to identify the impact of this new computing paradigm on software systems. In the remainder of this section we introduce the architecture of the mobile distributed systems, the possible applications and finally the different modes of operation. In section 2 we discuss how different areas are affected and finally we conclude in section 3 with a list of open problems.

1.1 The Architecture

The model, see figure 1, adopted from [IB9x], consists of two distinct sets of entities: mobile hosts and fixed hosts. Some of the fixed hosts, called *base stations*, or *Mobile Support Stations (MSS)*, are augmented with a wireless interface to communicate with mobile hosts. Each mobile host communicates with one base station. The area covered by a base station is called a cell. The process during which a mobile host enters a new cell is called *hand off*. Within a cell broadcast is physically supported.

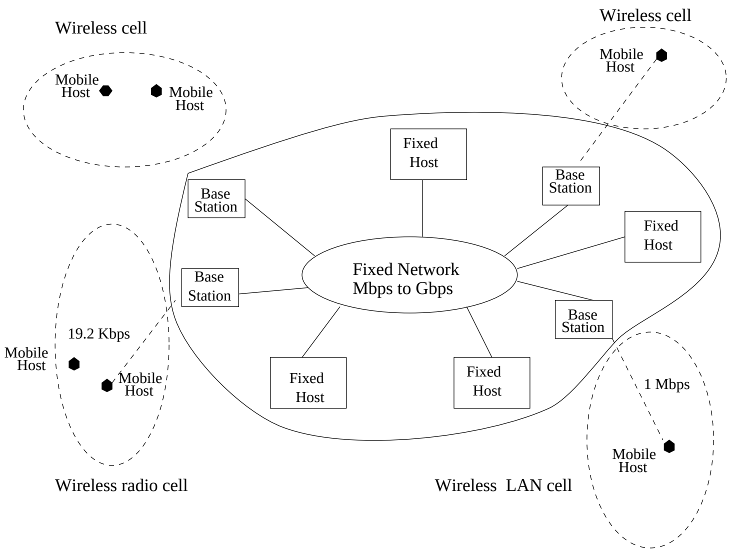

Figure 1: Mobile System Architecture

The Personal Communication Network (PCN) will include, [IB9x, AK93]: cellular architectures, radio transmission over unused portions of radio FM (e.g. Ericson GE's Mobidem), satellite services (e.g. Motorola's Iridium) and the wireless LAN (a traditional LAN extended with a wireless interface, e.g. NCR WaveLAN, Motorola's ALTAIR, Proxims Range LAN and Telesystem's ARLAN). Wireless LANs will provide connectivity within a small geographical area (at a range of a few kilometers) and their typical bandwidth ranges from 250 bps to 2Mbps. Cellular and packet radio modems will cover larger geographical areas and have a typical bandwidth of 8Kbps.

The wireless medium has some characteristics that differentiate it from the wired technology. The bandwidth limitations place restrictions on the volume of data that can be transferred. Furthermore, the reliability of the mobile environment remains unclear, especially when users will be operating in the presence of background noise. Another important characteristic of the wireless medium is that broadcasting is physically supported within a cell and does not depend on the number of receivers.

Mobile hosts, regardless of future technology advances, will have limited computing power, memory and screen size due to their small size and weight. In addition, mobile palmtops have limited battery capacity, (two or three hours under normal use). Due to this fact, energy preservation is an important consideration. It also results in frequent disconnections.

1.2 Applications

One issue of high importance for both the research community and the industry is to foresee the typical user of those services. People who often travel, such as businessmen, field specialists etc constitute the first candidate. On the other hand, consumers may find interest on those services and use their mobile terminal for entertainment, keep their record balance or in whatever way they use their PC. The challenge is to identify the "killer" application, that will make mobile computers widely-used.

One of the most important category of applications are *mail-enabled applications*, for interpersonal communication or notification of predefined events such as a canceled flight, high traffic or super sales in one's favored department store. Simple computations (such as spreadsheets), basic record-keeping (such as notebooks and calendars), or entertainment (such as video games) will constitute standard services provided by mobile computers. Yellow page informations, (such as lists of theaters, doctors etc) will also be highly used. What will differentiate information services from similar services already provided, is their dependence on the user's location, for example answering questions asking for the *nearest* hospital which treats a specific disease. Simple form-based transactions, (such as credit card payments) will also be provided.

Mobile services for specialized applications such as for car rental companies, postal services, controlling production lines in a building, etc. are already in use.

1.3 Modes of Operation

In a non-mobile distributed system a host operates in one of two modes, either connected to the network or totally disconnected from it. In contradistinction in a mobile environment there may be various degrees of disconnection ranging from total disconnection to weak disconnection (when a terminal is connected to the rest of the network via low bandwidth). As a result of that, a mobile host may operate in more than one modes. In addition, for conserving energy, a mobile host may operate in the *doze* mode. When in that mode, the clock speed is reduced and no user computations are performed. The mobile host waits to receive any message sent to it from the rest of the network. Upon receipt of any such message the host resumes its regular mode of execution. Figure 2 summarizes the different modes of operation a mobile host can be in.

Another characteristic of the modes of operation is that they are foreseeable. While in a non-mobile distributed environment disconnections are non-predictable, in a mobile environment most

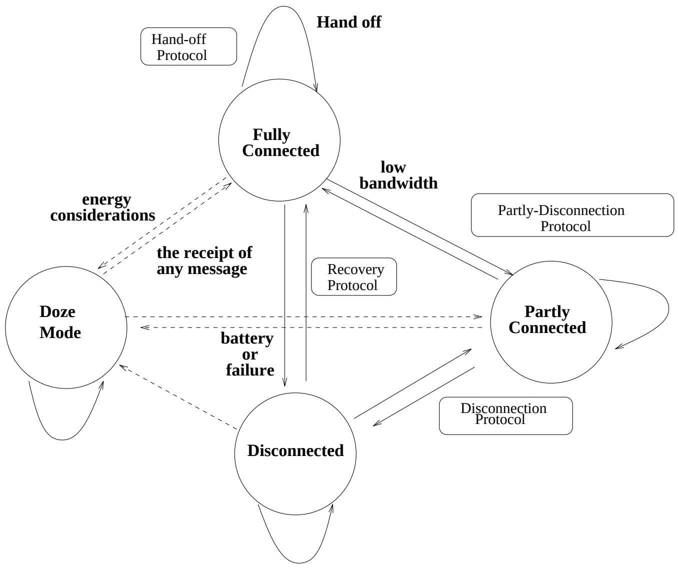

Figure 2: States of Operation of a Mobile Host

of the disconnections can be detected and thus special *protocols* may be designed to handle them. Furthermore since disconnections are frequent, ideally a mobile host should be able to operate autonomously even during total disconnections.

A mobile host can execute a disconnection protocol before it physically detaches from the network. This protocol should ensure that the mobile host has downloaded enough data and state information to continue operate and also that all interested hosts are being notified of its disconnection. A mobile host can switch to a partly-connected mode, by executing a partly-disconnection protocol. While in this mode, all communications with the network must be limited. Finally, while a mobile host is operating in any mode, it may enter a new cell. State information pertaining to the mobile computation must be transferred to the base station of the new cell, by executing a hand-off protocol.

2 Impact of Mobility on Software Systems

This rapidly expanding technology poses new challenging problems. One important issue is to determine the role of the mobile host in a distributed system. At one extreme, mobile hosts are going to be used as dumb terminals and at the other, they are going to be used as workstations with enough computational power and memory to perform computations locally. In the former case mobile hosts will depend on some fixed host (not necessarily the same) to compute on their behalf. In the later case, mobile hosts would be active participants of the distributed system. In any case, the prevailing and most significant question is: *What are, if any, the characteristics that make mobile computation different from (non-mobile or static) distributed computation.*

Mobile hosts, will have less computing power and memory and smaller screen, due to their small size and weight. Their computational power and transmission capabilities will also be restricted because of battery conservation considerations. Furthermore bandwidth limitations impose severe restrictions on the volume of data that can be transferred. Algorithms should take advantage of the broadcast facility, since broadcast in a cell is physically supported and for a mobile host it costs less to receive than to transmit.

Battery and bandwidth limitations make *disconnections* from the network very frequent. Disconnections also have various degrees depending on the availability of the bandwidth. Because of their frequency, disconnections will have to be treated differently than failures or crashes. Compared to them, disconnections are elective in nature and can be prepared ahead of time. *Mobility* makes the location of the user a fast-changing data and the reliance on a fixed network configuration unrealistic. Furthermore the future ubiquity of mobile hosts will result in computing systems that are much larger than the distributed systems of today. *Scale* implies operating in a highly *heterogeneous* environment. It also raises questions of organizing the information efficiently.

Identifying the critical aspects of mobile computation leads to the second most crucial question, *How should software systems for mobile environments be different from software systems for non-mobile distributed environments ?* We can identify two different categories of problems. The first refers to issues of portability and extensibility; that is how existing systems should be modified or augmented to be used by mobile hosts. The second refers to building new systems taking into account the mobility of the environment and raises a proliferation of new issues.

In this paper we try to categorize changes in the way software systems are going to be designed and implemented. Researches are trying to identify the principles underlying mobile computation and the criterion for building efficient systems in terms of cost and performance. An important question is how information is going to be structured, replicated in the fixed network and efficiently transmitted. It is not clear which characteristics of the mobile environment are going to be handled by lower level software and which will be left unhidden to the higher level application software.

In the following, we try to identify the effect of mobility on the various types of software systems and then explicitly ask some questions for which extensive simulation results and experimentation are needed for evaluating proposed solutions. First we discuss distributed algorithms, then communication issues, and finally operating and database system issues. Since there is not a clear-cut between what is a database and what is an operating system issue some of the same considerations appear in both contexts.

2.1 Distributed Algorithms

Designing algorithms for distributed environments spawned a variety of research efforts. New computational models and classes have been developed, new measurements of efficiency and new techniques have and are being explored. Designing algorithms for mobile environments seems to be a stimulating task, since the parameters that must be taken into account are not yet clearly defined and seems to vary from power efficiency to dynamic reconfigurations.

Figure 3 summarizes some of the factors that must be taken into consideration when designing algorithms for mobile environments. The structure is not fixed but the network changes fast as users are relocated or disconnected, thus we can not rely on a fixed structure (such as a ring or a mesh) when designing algorithms. This fact also implies that centralized algorithms, where a site place the role of the coordinator, must be redesigned to take into consideration the new factors for determining the optimal location of the coordinating site. The nodes are not equivalent, mobile hosts and fixed hosts differ in computational power and memory. Also due to the scale, mobile and fixed hosts may be highly heterogeneous. Disconnections play an important role in a mobile environment since they are more frequent and may vary among total disconnections, weak connections through a limited bandwidth to strong connections. Disconnections are also foreseeable thus they must be treated explicitly by special protocols. Power preservation seems to be a prevailing factor, thus since transmitting consumes more energy than receiving, broadcast must be extensively used.

[BAI93a, BAI93b] discuss this issue. The main premise of the approach taken in [BAI93b] is that distributed algorithms for mobile environments should be structured such as that the main bulk of the communication and computation costs is borne by the static portion of the network. They have also introduced the idea of associating with each mobile host a *proxy* on the static network, thus decoupling mobility from the design of the algorithm.

2.2 Networking and Communications

An important new aspect is organizing massive amount of information in the communication channels to provide efficient access to a large number of users. Data organization and access of digital data raises new research problems in communication and computation that are yet to be solved. There are two basic ways of communicating information, either by periodical data broadcasting or on demand. The section on broadcasting discusses some of the problems from the perspective of organizing large amount of broadcasted data.

Addressing Schemes and Routing. Most internetworking routing protocols are based on the assumption that computers rarely move. In the Internet Protocol (IP) a host's IP address depends on the network it is connected to, and packets are routed based on the network number. Such a schema is not adequate for mobile hosts, since mobile hosts are not permanently connected to the same network. Research on addressing schemes and routing concentrates: (a) on how to assign addresses to mobile hosts so that IP-based protocols may be used for routing messages, [DFGQM91, IGQM93, IDGQM91, TYT91], and (b) on how to optimize the search cost for locating a mobile host [IB9x, BIV92, AP91]. The location of a mobile host refers to the address of the base station to which it is currently connected.

In [DFGQM91] an internetworking protocol is proposed where base stations act both as gateways and routers. Mobile hosts are confined to addresses that form a single "virtual network". Each base station maintains a database describing the machines that are in its cell and those that have recently moved out of its cell. To identify the LAN in which a mobile host is currently located, a base station builds up a second database of other LANs that hold mobile hosts by using a combination of broadcast-based search and caching. Another approach, [IB9x], is to store the location of a mobile host in a specific fixed host, called Home Location Server.

To reduce the search cost for locating a mobile user, building a hierarchy of location servers on top of the base stations is proposed. Each location server is responsible for maintaining the addresses of the mobile hosts that are currently in a cell covered by a base station below it. Location servers do not have to know exactly in which cell a mobile host is currently located, but they can find it out by a process called *paging*, which is multicasting a message to a subset of the base stations below the specific location server. [BIV92] suggests defining *partitions*, by grouping together the location servers among which the user moves frequently and by separating the location servers between which the user relocates infrequently.

[BBIM93] argues that base stations should not only act as network layer routers but also as mediating agents for higher layers of the OSI model, allowing separate treatment of the wireless and wired part of a connection. According to this indirect model, data packets will traverse to some layer above the network layer at the base station before being forwarded to their destination, thus making mobility explicit at every level of the OSI model.

Compression. Compression/decompression of packets may be appropriate in mobile environment, as suggested in [DFGQM91, IB9x], since it takes advantage of a resource most in abundance in a wireless system (cycles) to utilize a scarce resource such as bandwidth. Compressed data also uses less memory but additional computation is needed to decompress.

Data Broadcasting. Broadcast in a mobile environment is different, as noted in [IVB93], for the following reasons: (a) since mobile users move among different cells, we can no longer assume that directory information is stored in the host's memory, so the broadcast should be self descriptive, for example using an index as a preample and (b) there is another performance parameter, in addition to the access time, namely tuning time, that is the time spent by the client listening to the channel. In addition achieving reliability is harder, since mobile hosts are only listening and thus there is a limited capability of acknowledgements.

In [IVB93] wireless broadcasting data is considered as a way of disseminating information to a massive number of users. Periodically broadcasted data are treated as the client's alternative "secondary storage". Directory information is also being broadcasted. Two schemas for organizing broadcasted data in a way that both tuning and access time are minimized are presented. Periodic broadcasting raises a number of questions regarding the type of communication support needed for this form of transmitting information.

Multicasting. Delivering multicast messages exactly once to a mobile host is particularly difficult for the following reasons: (a) the exact location of a mobile host may not be known to the sender or may change while the message is being transmitted and (b) copies of the same message sent by the same fixed source may reach their destination base station at different times due to the variable network latency. To overcome these problems multicast can be implemented either by unicasting copies of the same message to every destination or by sending a copy of the message to all base stations that will then forward it to the appropriate destinations, if any, in their cell.

In [AB93] a protocol, called M_Cast, is presented for delivering a multicast message exactly once to a group of mobile destinations. The algorithm is based on the second approach and ensures that all necessary data structures are stored in the fixed network.

2.3 Operating Systems.

Security. Security is a very important issue, since mobile users move along different base stations. When a mobile host moves to a foreign cell, it needs to use resources of the foreign base station in order to operate. This fact causes security problems for both the mobile host and the base station who may "trust" or not the visiting mobile host.

File Systems. Disconnections are very frequent in a mobile environment and cannot be ignored or treated as erroneous behavior. On the contrary, disconnections should be made part of the computational model. In this context, file systems should be designed to support the autonomous operation of mobile hosts during disconnections. Caching can be used to exploit availability. Issues related to cache coherence, [ABGM90], prefetching and various cache management techniques are being reconsider under the new mobile environment. Furthermore, the reintegration of the disconnected mobile system with the rest of the network resembles problems handled by network partition algorithms. Coda [KS92] is an instance of a file system that supports disconnected operations.

Moreover, caching important data may be used to utilize bandwidth during weak disconnections. Degrees of consistency of cached data may be defined based on the degree of disconnection, allowing for strongly consistent data during strong connections and weakly consistent data during weak connections. Consistency of replicated files is also discussed in [DFGQM91], where the existence of two different file read operations allow the user to control consistency by stating explicitly when the most recent value is needed.

Another important issue refers to the location of files. The further clients move from their file servers the slower the communication becomes. Should files follow their users to get better response time? [DFGQM91] describes such a configuration where each client is associated with a "primary" replica that is always located near him. The client communicates asynchronously with the "primary" replica and then the "primary" replica communicates asynchronously with "secondary" replicas. File system operations should also take advantage of the broadcast nature of the medium. A simple file transfer protocol, developed according to this principle, is described also in [DFGQM91].

Figure 4 summarizes some of the issues related to mobile file systems.

Relocation Transparency. During operation a user may move to a new cell. The goal is to allow the user to continue using the files and the services that were previously available. This makes the hand off procedure transparent to the user.

2.4 Interfaces

Some of the particularities of building user interfaces for mobile hosts include: (a) the small size of the screen that asks for efficient ways of representing information (b) pen-based interfaces. In [DFGQM91] the keyboard is mapped on the screen. Issues related to database interfaces are discussed in a later section.

2.5 Data Management

Data management issues have also been discussed in [AK93, IB9x, IB93]. Figure 5 summarizes some of the effects.

Transaction Management. Transactions in a mobile environment are different: (a) Disconnections are often and predictable, and thus should be treated explicitly. Mobile computers may download prior to a disconnection enough data to operate autonomously during disconnections. They should also declare themselves down to other interested parties. (b) Disconnections vary in degree based on the available bandwidth. It may be unrealistic to ask for perfectly synchronized cached (replicated) data on a mobile host during weak disconnections. Various degrees of consistency between replicated data may be defined depending on the available bandwidth. Issues such as cache consistency, versioning, etc are of importance. (c) The memory of a mobile host can not be considered secure [AK93], so we may need to keep logs in the fixed network. (d) Due to mobility, a mobile host may move to a different cell during the execution of a transaction. The user may want to continue the execution of his transaction in the new cell or, he may want to leave his transaction to continue executing in the previous cell. In the first case data related to the host may need to migrate to the new base station.

The above discussion indicates that many of the algorithms for maintaining consistency need to be reconsidered under the new environment. Specifically, (1) ACID transactions may no longer be the appropriate model of a user's interaction with the system. A first attempt [Chr93] to define an extended transaction model for mobile transactions is [Chr93]. Reporting transactions and co-transactions are transactions that share their partial result, maintain their state and can relocate their execution from one host to another. (2) Consistency must be formally defined along with criteria for maintaining it. Integrity constraints between data stored in different mobile hosts may be too costly to be maintained. Another issue is maintaining consistency of location dependent data. (3) Recovery and disconnection algorithms, must be redesigned to take into consideration the scarcity of the bandwidth, the broadcast facility and the frequent and foreseeable disconnections. (4) Migration of transactions, when mobile hosts relocate, must explicitly be handled.

Query Processing. Work on query processing focusses on (a) the handling of new types of queries and (b) the introduction of new parameters regarding query optimization. Queries may include location dependent data and furthermore data that depend on the direction of movement. Location dependent data may change during query evaluation. Furthermore they may be imprecise because of the overhead of continuously maintaining the location information up-to-date. Another new issue is querying broadcasted data.

To control the volume of the location updates, [IB92] introduces the idea of *bounded ignorance*, where even if we do not know the real position of a user, the location we know and the real position are guaranteed to be in the same partition. Partitions are defined based on the user mobility patterns (user profile) and group together cells between which the user relocates very often. In that environment, in order to obtain an exact answer, a combination of data search and additional data acquisition during the run time of the query is required.

Some of the new considerations for query optimization include (a) minimizing the financial expenses of transactions, (b) minimizing the search cost for locating data, (c) preserving bandwidth, (d) limiting the power consumption. An optimization algorithm for selecting plans based on their energy consumption while keeping the overall throughput in an acceptable range, is described in [AG93]. (e) producing the first set of output values fast and (f) controlling the precision of data to limit computational and communication expenses.

Interfaces Some of the desirable characteristics of user interfaces for mobile databases, as identified by [AHK92, AHK93] include: (a) graphical representation of queries and of the schema (b) high-level semantic data models that would facilitate the representation of information from heterogeneous resources and also hide from the user low-level structuring unrelated to his interests. New data output methods for the result of the queries are also necessary. A prototype is described in [AHK92, AHK93].

Data Replication Schemes. As in the case of file systems, dynamic replication schemes, where data move closer to the active readers, are needed. [BI92] discusses replication schemes for data produced by mobile servers for mobile clients. The various schemas are evaluated based on the read and write patterns for the specific data and on the search cost for locating them.

3 Directions of Future Research

3.1 Modeling the new Environment

Defining an abstract model for the mobile environment is a difficult task because of the uncertain of the communication environment, the dynamic nature of the configuration network and the various resource restrictions. The model must take into account the mobility of the users, the resource limitations of the mobile hosts and the particularities of the wireless medium.

- **Modeling Mobility**
  The frequent relocations of the mobile hosts result in a fast changing network configuration. Instead of associating each host with a fixed location, we associate each host with the probability of being in a fixed cell. Some of the parameters that determine this probability include:
  - *the user's history*, that is the typical mobility behavior of the user, for example how often and between which locations he moves. This is similar to the traditional notion of a *user's profile*.

- *the mobility ratio*, [IB92], this is defined for each host as the ratio of its relocations to the searches for its location. The smallest the mobility ratio the higher the probability of correctly locating the host.

- **Communication Environment**
  The communication environment is different than that of a wired distributed network because of:
  - the limited bandwidth and
  - the frequency of disconnections

  Experimental data are necessary for determining the exact frequency of the disconnections as well as the degrees of disconnection in terms of the available bandwidth.

- **Modeling Mobile Hosts**
  Due to the limitations of weight, size and power, the mobile hosts have limited resources, including memory, screen size, CPU clock, CPU time, and transmission capabilities. This fact has the implication that mobile hosts should not be considered as equivalent participants in the distributed environment but they should be modeled as second-class objects.

3.2 What has to be Answered

3.2.1 Model of Operation

A general question, whose possible answers must be evaluated through performance measurements, refers to how mobile hosts are going to be used. Their likely role in a distributed system varies from that of a dumb terminal with no active participation in the computation to that of a relatively independent component with enough memory and computing power to perform part of the distributed computation locally.

The factors that will determine their future use include performance and cost issues. Performing operations locally costs in terms of power consumption and complicates the data replication and the integrity preservation schemes. Furthermore, since, as already mentioned, the memory of a mobile host cannot be considered stable, logging operations may be necessary. Also, if a mobile host is an active participant in a distributed algorithm, then the search cost for locating the host and the data it produces or consumes, may be unbearable. On the other hand, this type of use, may result in a better performance, it will allow autonomous operations during disconnections and it may limit bandwidth use. Performing operations in the fixed network consumes bandwidth and power for the transmission and probably results in worst performance. As shown from the above discussion various the relative importance of various tradeoffs must be estimated.

3.2.2 Communication Questions

One of the most active research areas is the development of efficient search algorithms and routing protocols. There is a need to evaluate the relative performance of the various algorithms.

Another important issue is to clearly define how reliable wireless communication is. Research is needed to specify the rate of disconnections, the degree and the frequency of weak disconnections.

An issue of general interest concerns the study of broadcasting as a facility for disseminating large amount of data. Some of the research challenges in this area include: (a) the development and the performance analysis of communication facilities for supporting broadcasting, including an efficient switch to an "on demand" mode, an efficient allocation schema of the communication channels to index and broadcast information, (b) the specification of the structure of the broadcast information, (c) the specification of the periodicity of broadcasting and how often it should be refreshed to include up-to-date information (d) the specification of the structure and the precision of the index information and (e) a specification language for the user to specify the type of information he wants.

3.2.3 Data Management Questions

Mobile computing environments are environments characterized by frequent disconnections, significant resource limitations, bandwidth restrictions and fast-changing locations. A key requirement of a data management system for such a computing environment will be the efficient access of critical data regardless of location. How are data going to be distributed and replicated on the fixed network and at the mobile users to maximize availability and maintain efficiency remains an open problem.

Furthermore the need to access shared data implies interdependencies between the elements of a mobile computing system. Caching of data at mobile hosts is necessary for saving bandwidth and for allowing the autonomous operation of mobile units during disconnections. On the other hand, caching also implies the need for maintaining consistency. New algorithms for maintaining consistency, that take into account the particularities of the environment, need to be developed.

Query languages must be extended to take into account the mobility of the users and the fact that data may be highly distributed. Besides, new interfaces, appropriate for small screens, need to be developed.

References

[AB93] Arup Acharya and B. R. Badrinath. Delivering Multicast Messages in Networks with Mobile Hosts. In *Proceedings of 13th Intl. Conf. on Distributed Computing Systems*, pages 292–299, May 1993.

[ABGM90] Rafael Alonso, Daniel Barbara, and Hector Garcia-Molina. Data Caching Issues in an Information Retrieval System. *ACM Transactions on Database Systems*, 15(3):359–384, September 1990.

[AG93] Rafael Alonso and Sumit Ganguly. Query Optimization for Energy Efficiency in Mobile Environments. In *Proceedings of the 1993 1993 Workshop on Optimization in Databases*, Aigen, Austria, September 1993.

[AHK92] Rafael Alonso, Eben M. Haber, and Henry F. Korth. A Database Interface for Mobile Computers. In *Proceedings of the IEEE Globecomm 1992 Workshop on Networking of Personal Communications Applications*, Orlando, Florida, December 1992.

[AHK93] Rafael Alonso, Eben M. Haber, and Henry F. Korth. A Mobile Computer Interface for Heterogeneous Databases. In *Proceedings of the RIDE-IMS'93 Workshop*, Vienna, Austria, April 1993.

[AK93] Rafael Alonso and Henry F. Korth. Database System Issues in Nomadic Computing. In *Proceedings of the 1993 SIGMOD Conference*, Washington, D.C., May 1993.

[AP91] Baruch Awerbuch and David Peleg. Concurrent Online Tracking of Mobile Users. In *Proceedings of the ACM SIGCOMM Symposium on Communications, Architectures and Protocols*, pages 221–233, September 1991.

[BAI93a] B. R. Badrinath, Arup Acharya, and Tomasz Imielinski. Impact of Mobility on Distributed Computations. *Operating Systems Review*, 27(2):15–20, April 1993.

[BAI93b] B. R. Badrinath, Arup Acharya, and Tomasz Imielinski. Structuring Distributed Algorithms for Mobile Hosts. Technical Report DCS -TR- 298/WINLAB -TR-55, Rudgers xx, 1993. Submitted for Publication.

[BBIM93] B. R. Badrinath, A. Bakre, T. Imielinski, and R. Marantz. Handling Mobile Clients: A Case for Indirect Interaction. In *Fourth Workshop on Workstation Operating Systems*, Aigen, Austria, October 1993.

[BI92] B. R. Badrinath and T. Imielinski. Replication and Mobility. In *Proceedings of the 2nd IEEE Workshop on Management of Replicated Data*, pages 9–12, November 1992.

[BIV92] B. R. Badrinath, T. Imielinski, and A. Virmani. Locating Strategies for Personal Communication Networks. In *Proceedings of the IEEE Globecomm 1992 Workshop on Networking of Personal Communications Applications*, Orlando, Florida, December 1992.

[Chr93] Panos K. Chrysanthis. Transaction Processing in Mobile Computing Environment. In *Proceedings of the IEEE Workshop on Advances in Parallel and Distributed Systems*, pages 77–83, Princeton, New Jersey, October 1993.

[DFGQM91] Daniel Duchamp, Steven K. Feiner, and Jr Gerald Q. Maguire. Software Technology for Wireless Mobile Computing. *IEEE Network Magazine*, pages 12–18, November 1991.

[IB92] T. Imielinski and B. R. Badrinath. Querying in Highly Mobile Distributed Environments. In *Proceedings of 18th VLDB*, pages 41–52, August 1992.

[IB93] Tomasz Imielinksi and B. R. Badrinath. Data Management for Mobile Computing. *SIGMOD Record*, 22(1):34–39, March 1993.

[IB9x] Tomasz Imielinksi and B. R. Badrinath. Wireless Mobile Computing: Solutions and Challenges in Data Management. *Communications of the ACM*, x(x), 199x.

[IDGQM91] John Ioannidis, Dan Duchamp, and Jr Gerald Q. Maguire. IP-based Protocols for Mobile Internetworking. In *Proceedings of the ACM SIGCOMM Symposium on Communications, Architectures and Protocols*, pages 235–245, September 1991.

[IGQM93] John Ioannidis and Jr Gerald Q. Maguire. The Design and Implementation of a Mobile Internetworking Architecture. In *Proceedings of the 1993 Winter Usenix*, pages 491–502, San Diego, CA, January 1993.

[IVB93] T. Imielinski, S. Viswanathan, and B. R. Badrinath. Data on the Air: Organization and Access. Technical Report xx, Rudgers xx, 1993. Submitted for Publication.

[KS92] James J. Kistler and M. Satyanarayanan. Disconnected Operation in the Coda File System. *ACM Transactions on Computer Systems*, 10(1):3–25, February 1992.

[TYT91] Fumio Teraoka, Yasuhiko Yokote, and Mario Tokoro. A Network Archtecture Providing Host Migration Transparency. In *Proceedings of the ACM SIGCOMM Symposium on Communications, Architectures and Protocols*, pages 209–220, September 1991.

Mobility	Hand-off	Not a fixed underlying structure
	Cost for a fixed source, destination pair includes search cost to locate participants	
Scale	distances between nodes may be large nodes may be heterogeneous	
Disconnections	Doze Mode	A message to a node in doze mode may cause it to return to normal operation
	Degrees of Disconnection	the connection between nodes may vary from strong to weak
	Elective	disconnectionsmust be handled explicitly as a "graceful exit" of the participant
	Frequent	disconnections are more frequent than in non-mobile distributed environments
Medium	bandwidth limitations less reliable Thus: minimize the number of the wireless messages exchanged	
Palmtops	Power	algroithms that preserve energy Ex. since the cost of transmitting is higher than the cost of receiving, exploit the broadcast facilities
	asymmetry among the nodes of the network, mobile nodes have less resources, including memory and computational power	

Figure 3: Impact of Mobility on Designing Algorithms

Mobility	Where are the files located? **Some Approaches:** Files that follow their users
Scale	Files stored under heterogeneous file sytems Highly distributed file systems
Disconnections	cache data at the mobile host **Some Issues and Approaches:** cache consistency: varying with the degree of disconnection user-defined by special file operations reintegration : favor optimistic algorithms that allow more availability
Medium	File operations that take advantage of the broadcast facility

Figure 4: Impact of Mobility on Operating Systems

TM	Issues : Transaction Models : Handle Transaction Migration Long-lived Transactions Integrity - Consistency: degrees of consistency handle frequent, predictable disconnections
Query Proc	Issues : Types of Queries query location dependent data combine data acquisition techniques Query Optimization: optimize search cost for location optimize energy consumption
Interface	small-size screen pen-based, graphical

Figure 5: Impact of Mobility on Data Management