A Survey of Mobile Computing Technologies and Applications

Anupam Joshi
Sanjiva Weerawarana
Ranjeewa A. Weerashighe
Tzvetan T. Drashansky
Narendran Ramakrishnan

Report Number:
95-050
A Survey of Mobile Computing Technologies and Applications

A. Joshi, S. Weerawarana, R.A. Weerasinghe, T.T. Drashansky, N. Ramakrishnan and E.N. Houstis
Department of Computer Science
Purdue University
West Lafayette, IN 47907

CSD-TR-95-050
July, 1995
A Survey of Mobile Computing Technologies and Applications *

Anupam Joshi, Sanjiva Weerawarana, Ranjeewa A. Weerasinghe, Tzvetan T. Drashansky, Narendran Ramakrishnan and Elias N. Houstis
Department of Computer Sciences
Purdue University
West Lafayette, IN 47907-1398.

July 24, 1995

1 Introduction

Mobile computing (or ubiquitous computing as it is sometimes called) is the use of computers in a non-static environment. This use may range from using notebook-type computers away from one's office or home to the use of handheld, palmtop-type PDA-like devices to perform both simple and complex computing tasks. Although there is a subtle distinction between mobile (non-static) and ubiquitous (everywhere) computing, for the purposes of this survey, we consider these to be the one and the same.

This document presents the results of a survey we did on mobile computing technologies and applications. Based on the results of this survey, we identify major academic and non-academic research institutions in the area of mobile computing, as well as the most challenging issues that need to be addressed to fully realize the dream of truly mobile computing.

This document is organized as follows: Section 2 provides information on the survey methodology used and summarizes the results. Section 3 provides brief summaries of most of the papers identified by this survey. Section 4 provides a list of uniform resource locators (URLs) that constitutes a webliography of references, a World Wide Web list of sites that contain (hypertext) information on mobile computing. A comprehensive bibliography is given in the references section.

2 The Survey

In this section we describe the survey methodology. We first consider the sources of information and then present the results of the search. Finally, we present some analyses of the survey results.

*This work was supported in part by a grant from Intel Corporation.
2.1 Information Sources

This survey was performed using several sources of information. The two primary information services were DIALOG Information Services and the World Wide Web. The DIALOG service is an “electronic library,” accessible with a personal computer (or terminal) and modem through telecommunications networks such as MCI Data Services or SprintNet. International access is also available and coverage is offered in most areas of business, legal & government, news, science, reference, social sciences & humanities, and general information.

The World Wide Web (WWW, Web) is a hyperlinked information repository on the global Internet. Information located on information servers is hyperlinked to other related pieces of information to form a giant infosphere called the Web. In addition to convenient graphical browsers to navigate the Web, there are several searching tools that are available to search the Web based on keywords. We applied these search engines in our survey as well.

We have also searched our own libraries electronic indices for relevant indices. The Purdue Libraries subscribes to the Compendex Engineering Indices and is linked to other state university libraries in Indiana as well. In addition, the authors of this survey are all doing research in the area of mobile computing, and we have used our “mental” knowledge banks to track down several web pages and papers.

2.2 Results

The DIALOG search was done with the keywords “Mobile computing”, “Ubiquitous computing”, “Wireless computing” and “Nomadic computing” for 1983 onwards. This search located 106 articles, not all of which directly related to the topic of interest even though they included the proper keyword(s). We use the same keywords to search the Web for articles and located about another 50 articles, some of which overlapped with those found via the DIALOG search.

We have thus identified over 150 pieces of information related to mobile computing. This information comes in the form of research journal articles, trade journal articles, other magazine articles, newspaper articles, books and other publications, and Web articles. In the rest of this document, we summarize this large amount of information. This is done in two forms: as several bar and pie charts illustrating the general findings and as brief reviews of (most of) the articles. The brief reviews are given in the next section.

The information in the charts is organized as follows:

- classification by area: We classify each piece of information based on what “area” it is related to. The classification of areas we used is included in this discussion. This information is used to identify the more pressing issues in mobile computing.

- ranking of areas: We identify the areas that are deemed most challenging in terms of fully realizing the dream of truly mobile computing. This is based on the quantity of published material on different areas as well as our qualitative evaluation of this material.

- classification by institution: We classify each piece of information based on the source of the information. This information is used to identify the larger players in the mobile computing arena.

- ranking of institutions: We also identify the institutions / research groups that are most active in the mobile computing arena. This is based on the quantity of published material as well as our subjective evaluation of the institutions concerned.
application areas: We identify the major application areas in which mobile computing arises.

Figure 1 depicts the academic institutions actively involved in mobile computing research. As indicated by the figure, a handful of universities (Berkeley, Carnegie-Mellon, Columbia, Lancaster, Rutgers, Waterloo and Purdue) have contributed a majority of the current literature.

The leading industrial research institutions are correspondingly portrayed in Figure 2. This shows that Xerox's Palo Alto Research Center is overwhelmingly the leader in this area.

Figure 3 depicts publication statistics in terms of the computer science area researched. We classify non-computer science areas (for example, low-power ICs for mobile hardware) in a "hardware" category. Figure 4 breaks down this information for each "major" academic institution in the mobile computing arena.

Finally, Figure 5 groups the publications in terms of the application area it relates to. As can be seen from the figure, interest in mobility is fairly widespread across all major application domains. This reflects the generally held view that truly mobile computing will revolutionize today's modus operandi in every application domain.

In addition, we have created a webliography of information. This contains links to many other sites doing work on mobile computing, both industrial as well as university. This is a living document that we keep updating as and when we learn of new sites. It can provide a useful tool for researchers, a kind of "jumpgate" into the world of mobile computing. Some of the documents found in the webliography are summarized in our narrative. This is done when the corresponding work has not (yet) appeared as a published paper, or if we think the work is important enough to merit mention multiple times. The URL for the webliography is http://www.cs.purdue.edu/research/cse/scipad/mobicomp.html and the URLs referenced are listed in Section 4.
Figure 2: Industrial (research) institutions doing research on mobile computing. The fraction of the pie is a representation of the contribution from that institution.

Figure 3: Publications categorized by area of computer science addressed. Non-computer science, hardware related publications have been categorized as "hardware."
Figure 4: Information in the previous figure per university for the “major” academic institutions.

Figure 5: Publications categorized by the application area that the work is targeted to.
3 Reviews of Publications

In this section, we summarize in a few sentences the key ideas from of most of the publications used for this survey. This information may be used to select the specific articles for further reading by interested users. We note here that we were unable to locate some articles through the sources available to us. In such cases, we used the abstract of these articles to produce the summaries given here.

We first describe papers that describe the basic components and issues underlying mobile and ubiquitous computing environments.

Landay and Kaufmann [70] discuss the need for creating new environments and user interfaces for mobile computing. They argue that attempts by systems researchers to transfer workstation environments to small portable machines are not likely to meet with success. They call for a de novo analysis of what the users might want from such a system, and discuss how appropriate interfaces catering to the special need of mobile systems might be made. They also report on a mobile computing device that they have built. The paper by Newman[80] also discusses the theme of user interfaces for ubiquitous systems. It shows how the existing (and future) ubiquitous computing hardware can be combined with digital audio and video to enhance their utility. He argues for an indirect management approach to interacting with such systems, and against a traditional screen based interaction. The claim is that much of the information in such systems will flow without use intervention. The interface should therefore provide the user with a conceptual model of the systems overall behavior.

Xerox corporation has developed a series of wireless based systems, ranging from the Live-
board to the badge. Schilit et. al.[97] describe the mid sized wireless system called the PARCTAB. PARCTAB off loads most of its tasks on remote (static) hosts, and uses a reliable Infra-red (IR) communication channel. The IR network is divided into cells, and provides uninterrupted service even when the TAB moves across cells. They have also developed networking software for the system which provides a datagram service, and implemented agent based remote procedure mechanisms which are location independent.

The paper by Bellotti and Selen [20] is one of the few to address the privacy issue in ubiquitous computing. They point out that current trends in information technology are creating a scenario where ever greater personal information is being captured and stored. In the case of ubiquitous computing, certain unique characteristics can be exploited to get personal information about the users. They propose a new framework which will enhance the privacy of mobile users.

An overview of the challenges faced by mobile computing is presented in [46]. The authors, Forman and Zahorjan, emphasise that the challenges and problems involved in designing software for mobile computers can be very different from those posed by wired systems. They are especially concerned with challenges posed by wireless networking, and describe communication, mobility and portability as the fundamental issues that any mobile computing software must grapple with. They briefly describe issues such as disconnection (elective or otherwise), low bandwidth, bandwidth variability, heterogeneity of the network, security, address migration / mobile routing, location dependent information, low power design, restricted disk and memory and user interfaces. They attempt to show the kind of constraints that these factors put on the software system for mobile computers. Duchamp in [39] also provides a similar overview. He deals with the challenges that mobile computing poses to internetworking and service providing, amongst others, and lists the research issues that need to be addressed. Another review is provided by Caldwell[27]. His focus, however, is on how such systems will
reshape the way people use technology, and the manner in which ubiquity will affect managing IS needs for business and management. He also comments on the potential difficulties that will arise in mobile applications.

Distributed object management is the focus of a paper by Blair[24]. He argues that with the advent of mobility, the notion of distributed applications will have to adapt to meet the constraints imposed by mobility. He suggests for applications such as distributed multimedia systems to function, good tools for distributed object systems are necessary. The work being done at his institution (Lancaster Univ. UK) in the context of ODP standards is presented, and also related to other distributed object standards such as CORBA.

In [25], Buffone & Beck describe the use of ubiquitous computing in the area of healthcare. They claim that creating a seamless union of computing and communication resulting in a ubiquitous medical computing. They argue that this environment will include high speed networking, mobile computing platforms and work group applications and provide next generation IT services to the physician. In their opinion, standards in semantics and syntax of the higher levels of the OSI model will pose a big challenge in this area. To this end, they outline their view of the functional requirements needed from computing and communication systems to enable their use by physicians. Along similar lines, McGee describes how mobility will mean that healthcare will go where the patient does[75].

Kulkarni et. al. describe[69] the area of cost management based on operating systems. In mobile systems, where usage charges will apply to communications and other "for-fee" services, the time constants involved in cost based decisions has decreased with a concomitant increase in the complexity of the decisionmaking process. As such, the system software must know about the costs involved, and needs to involve the user in the decisionmaking process. Their paper considers several hypothetical scenarios where services, such as interfaces, are for fee. They introduce the idea of applications that are cost aware.

An important feature of mobile systems is there dynamics, which change often. Thus an application needs to be aware of such dynamics. This is the subject of a paper[98] by researchers from Xerox PARC, where they describe such a scenario. Users in there scenario interact with mobile as well as stationary systems in an office environment. There work uses dynamic environment variables to allow flexible sharing of customization information and provides efficient notification of environmental changes to applications. They also describe a sample application built at PARC which uses this system.

In a series of articles, for instance [110] and [111], Mark Weiser outlines his vision of ubiquity. He argues that in future, decreasing hardware and communication costs will make computers available everywhere. However, he argues that this by itself is not sufficient to achieve ubiquity. Computers today are tools that demand attention to operate. They distract one from the task at hand. An analogy can be made with a car. A good car is one that drives smoothly, so that the drivers sole concentration is from getting to point A to point B. After the initial stages of learning how to drive, a good car becomes "invisible". On the other hand, a car that is not in order demands constant attention, and distracts one from the task of transportation. Clearly, such a car is not desirable. Weiser argues that good tools are those that are invisible, in other words tools that don't require a (great deal of) conscious effort and though in order to operate. He brings arguments from traditional computer science fields, as well as social science endeavors such as deconstructionism, postmodernism and Polya's philosophy of the tacit dimension, to support his case. His claim is that for computers to become invisible despite being all pervasive, our interactions with them will have to take novel forms. he illustrates this by saying that the speed
of rotation of a wire hanging from the ceiling, and the temperature of a room could all carry information from the computer to the human. He specifically argues against most of the present trends in interfaces which use multimedia, virtual reality and intelligent assistants. He paints a vision of the future in which computers are literally omnipresent, and are used by humans almost subconsciously.

In [43], Eisenhart describes ubiquity from the point of view of the salespeople. With the advent of wirelessly connected PDAs, he claims, the way salespeople work will be transformed. This paper describes several PDAs, as well as the improvements in mobile computer technology such as improved storage capacity and compute power. The paper by Karcher[66] looks at the same scenario, but instead of PDAs it concentrates on laptop notebooks and wireless, portable phones. It describes how such tools can be integrated to provide managers with what amounts to a portable office. The paper talks about several existing hardware systems and their performance, and also outlines the "history" of technology in this area.

Badrinath, Acharya and Imilienski have produced significant work in the area of mobile systems, especially on how mobility affects database like systems. In this paper [27], they formulate a theoretical model of how distributed computations are affected by characteristics of mobile computers, such as elective disconnections and the differential cost of sending messages to mobile and static hosts. They argue that traditional distributed algorithms will not perform as well for mobile systems. They propose a new model for mobile hosts which uses the concept of Mobile Support Stations, static hosts which server as gateways for all mobile hosts in their "cell". It handles the dynamic nature of the network (since hosts are moving), a communications systems which is a combination of wired and wireless (broadcast), and the fact that several hosts (due to power constraints) may operate in a "doze" mode. They discuss protocols for handoff/takeover as a mobile host crosses cells, and discuss how the choice of the transfer (active/passive, leave/join) will affect the design of distributed algorithms. In a companion technical report, they demonstrate how their model can be used to modify the traditional mutual exclusion algorithm for distributed systems due to Lamport.

One of the largest efforts in the direction of mobile computing, especially at the hardware and networking levels, is being undertaken by the InfoPad group at UC Berkeley [49]. The attempt there is to develop the hardware, network infrastructure and software that will allow users to access multimedia data from wireless platforms. They are also developing technology that will allow simultaneous transmission of separate multimedia information to specific users in confined areas. Novel interfaces, based on pen and voice, are also being developed.

In [104] the authors present their views of what operating systems for PDA type devices will be like. They argue that given that PDAs will be used largely for interacting with external services and devices (rather than as computing devices themselves), the operating systems for them must also be different from today's computing devices' operating systems.

In [103] the authors present four per-session guarantee strategies for users and applications of weakly consistent replicated data. These session guarantees were developed in the context of the Bayou project at Xerox PARC in which is aimed at building a replicated data storage system to support the needs of mobile computing users who may be only intermittently connected. The intent is to present individual applications with their own view of the database that is consistent with their own actions, even if they read and write from different, potentially inconsistent servers.

In [17] Barbara and Imielinski present two different caching strategies in mobile environments. They envision a future where a large number of users armed with low-power palmtop machines will query databases over wireless communication channels. They argue that since
the wireless channels will always have low(er) bandwidth, that caching will be an important strategy to reduce contention on these channels. They present a taxonomy of cache invalidation schemes and study their effectiveness in the presence of client disconnections. They identify effective cache invalidation strategies for both sleeper and workaholic units.

In [52] Guthery argues that mobile computing requires new ways of thinking about networks. He argues against switched networks and explores possibilities for switchless networks which are named “EchoNets.”

In [92] Rekhter presents an approach to mobile internetworking that allows hosts to move around an internet in a fashion transparent to the transport layer protocols. He argues that this well-modularized architecture provides sufficient flexibility (can be gracefully introduced to the existing infrastructure) and generality (well-suited to a wide range of applications) to implement a mobile internet.

In [100] Spreitzer and Theimer present an architecture which allows them to provide location information to mobile clients in a controlled fashion. Their architecture is based on having a user agent manage the location information and a partially decentralized location query service. While making certain query operations somewhat expensive, this architecture gives the user primary control over their location information.

In [99] Schulman considers four techno-market areas that exemplify some of the newest and best mobile computing applications: multimedia, global positioning systems, pen-based computing and personal assistants/communicators.

In [108] Watson and Bershad describe their experience in building a local area mobile computing environment using stock hardware and most stock software in a graduate level operating systems class at Carnegie Mellon University. The group had design and built the infrastructure for a simple mobile computing environment for low-end palmtop machines and then used it to build a suite of mobile applications.

In [21] the authors describe their experience of building a prototype version of Sun's Solaris 2 SVR4 for mobile platforms. This requires changing the entire operating system, from the kernel level through the user command set of applications and involved enhancements in four primary areas: power management and checkpointing of system state, drivers and other support for PCMCIA devices, support for serial line networking and a new electronic mail application designed to work over slow serial connections.

In [8] Adams, Schilit, Gold and Tso describe the design and implementation of an infrared network for mobile computers. This network consists of room-sized cells each wired with a base station transceiver. Each mobile unit is represented by an agent or a proxy which resides at a well-known location. It is the responsibility of the agent to track the mobile unit as it moves from cell to cell and to deliver messages to the mobile unit.

In [106] the authors describe a mobile computing environment that provides migration transparency using packet forwarding. Each mobile host has a home address and a temporary address when they are temporarily located on some network. The home base always knows the mobile's current address and forwards packets destined to it to the temporary address. This forwarding is done either by a packet forwarding server at the home base or directly by the sender (after it has acquired the knowledge about the mobile's current location) if its software has been optionally enhanced. They found that the ability to work with existing infrastructure is a significant advantage of this method and also that the communication overhead was acceptable, especially in the case of stationary hosts with the enhancements.

In [47] Gluck describes the Open Security Architecture (OSA), which provides a basis for
the selection, design and integration of products providing security and control for a variety of hardware including mobile computers. The purposes of this architecture are to provide a framework where one can set up an acceptably secure environment without negatively affecting the data flow in the organization.

In [33] the authors study the impact of mobility on working practices and computer systems support. Their project is focusing on the electricity industry with particular emphasis on using mobile technology to enhance cooperation between field staff working on the electricity distribution network. While there is an obvious market for mobile computing technology, the authors found that there is a significant mismatch between the end user requirements of mobile systems and the technologies currently available to them.

In [112] Wood looks at the trends in portable computing equipment. He argues that the availability of add-on facilities through PCMCIA cards and the constant improvements in notebook's capabilities are making notebook computers more attractive to potential buyers.

In [44] Farber discusses directions in networking technology that can achieve fruition within the 1990s which will have a significant impact on our lives. The author believes that the following three technologies will be developed in this decade: mobile computing integrated with personal communications systems, gigabit speed regional/national and international network backbones, and the development of distributed coordination and decision making facilities.

In [5], the authors discuss how Memorex Telex equips its field service personnel with mobile computing devices. The organization, a leading provider of networking devices and products, uses mobile computers and applications designed to run on a wireless network to increase productivity, provide better service and increase revenue.

In [7], a description of mobile computing at Price Waterhouse is given. The worldwide consulting and accounting firm is equipping thousands of its employees with portable computing devices, and considers mobile computing an important part of the company's IS architecture.

In [6], the introduction of mobile computers to Ryder Truck in described. By equipping thousands of vehicle maintenance personnel with mobile computers, the organization has seen increased productivity and better service. This case study explains how introduction of mobile computers was chosen over an automation of the previous manual processes.

In [1] the impact of mobile computers on sales and marketing personnel is discussed. The author describes the components of a virtual office, their capabilities, and the new opportunities created by them. Brown and Williamson Tobacco, for example, reports a 30% productivity as a result of providing mobile computers to its sales personnel. The author also reveals how UPS gained a competitive advantage by investing in a $150 million cellular network to collect data from its truck drivers.

In [28], Carbone describes how IBM adapted a new methodology in making mobile computers. In manufacturing the ThinkPad, IBM used outside suppliers as well as internal developments to get a product out into the market earlier. In a market expected to grow from 7.6 million in '94 to 12.1 million in '97, IBM expects its ThinkPad to be a major player. The author reveals how IBM considered mobile computing to be such a vital sector of the market, that they were willing to change their traditional design and manufacturing processes to succeed in it.

The popularity of mobile computers in different levels of the corporate hierarchy is examined in [4]. The authors discover that most portable computer users still are salespeople and field personnel. Several top-level managers' views are expressed on why executives shun laptop computer usage. Some reasons given are: the dislike of learning new technologies, the lack of human metaphors in mobile computing, and the difficulty in accessing remote data, using
current software. On the other hand, executives who have adapted to them explain how mobile computers have become a vital part of their life.

In [2], Marshall examines the important features of mobile computers. Different processor, storage, battery and modem options that are available are described, and the relative importance of these are explained from a traveller's point of view. The author what accessories are necessary, in order to improve mobile computing productivity while travelling.

An airplane's coach class cabin is used to test several laptop computers in [3]. The authors detail the features being tested, the actual experiments done, and the results obtained, thereby providing an insightful comparison of some leading laptop computers. The authors also note the key product features that were found to be productive and user-friendly during these experiments.

In [105], Vecchione claims that in many corporate environments, the management of mobile computing resources is an issue that has not been dealt with clearly. He states that though many IS managers are aware of its importance, few companies are prepared for the transition from fixed to mobile computing. He claims that a successful transition depends to a large degree on how companies handle their client-server applications.

In [95], Satyanarayanan describes the main differences between fixed and mobile computing, and how they contradict the assumptions made in current distributed systems. He details a few key features of mobile computers as: being relatively resource-poor, being more prone to loss/damage, and operating under a broader range of network conditions. The author then introduces a distributed file system, Coda, that is designed to transparently provide mobility to users. A brief explanation of Coda is also included.

In [67] the authors discuss the issue of location management in mobile computing. Search, update, and search-update are described as the operations involved in location management. They discuss simulations designed to measure the message overhead of location management using different strategies, under varying communication conditions.

In [57], the authors describe a prototyped mobile computing system using pen-based computers and cellular telephone networks. Due to the bandwidth constraints imposed by cellular networks, they propose strategies to improve communications involving the mobile components. These include distributed control, forms manipulation, keyword communication, and the usage of proxies to further compress data and reduce bandwidth usage.

In [71], Lavallee examines the state-of-the-art in mobile computing products, and how they give mobile professionals better access to information. Several mobile computing products are reviewed, and the different uses they can be put to are examined. Specific examples are given in fields such as finance where the availability of timely information is crucial.

Future applications in the area of business and enterprise with respect to new technologies in mobile computing are discussed in [89]. The emphasis is by global client-server systems, where clients may be located at any place suitable for mobile telecommunication by wire or by radio. An overview is provided for customers as well as for enterprise offering global services. Different branches are characterized in order to show the potential of commodities and services based on mobility and telecommunications.

The effects of host motion on the performance of active transport-level connections is explored in [26]. Current retransmission policies are shown to introduce unacceptably long pauses in communication (800 milliseconds and longer). A fast retransmission scheme that can reduce these pauses to levels more suitable for human interaction is proposed. This work demonstrates that reliable transport protocols must be made aware of mobility, and suggests how to adapt
these protocols to mobile computing environments.

The technical possibilities available for integrating mobile users seamlessly into fixed networks is discussed in [36]. Applications of mobile computing within the industrial environments are also described.

The architecture of the Bayou system [35] hinges on the identification of novel ways of thinking about mobile computing and using these viewpoints to derive system structures. The problem addressed is the maintenance of consistency in shared, replicated data repositories updated by mobile hosts. Clarification of the consistency guarantees and the rate of convergence are the two important issues addressed by the paper.

A case was made out by Mukherjee et. al. [78] that existing work on mobility focussed on computation and communication, to the exclusion of control. It was pointed out that a rich taxonomy of applications emerges when control is given due prominence and that the taxonomy offers valuable insights into structuring applications to function effectively under the constraints of mobility.

Another approach is addressed in [117] where the idea is to use a network as a rotating information medium by periodically retransmitting the entire contents of databases. Multiple disks spinning at different speeds on the broadcast medium are superimposed in order to support non-uniform data access. Instead of fetching data on demand, clients continuously listen to the transmissions and cache information of interest to them. This is found to be extremely valuable when the network has asymmetric bandwidth.

Techniques to cope with the performance and reliability of mobile networks are focussed in [42]. The techniques spanned three areas: deferring update propagation during periods of low bandwidth, opportunistically using high bandwidth when available, and the use of an abstraction called "dynamic sets" to reduce network latency during search. In order to use dynamic sets, applications had to be changed and they have to be able to tolerate the reordering of requests implicit in the use of dynamic sets.

Further information on the Little Work Project, an approach in contrast to the Coda System, which applies optimizations incrementally, is presented in [58]. The primary advantage is that the optimization code is a separable component; hence it is easy to apply to multiple file systems.

Predictive caching techniques are elaborated in [68]. The goal of this work is to reduce the burden on users of specifying files to be hoarded in anticipation of disconnection. The system uses a list of observed file references and a set of clustering algorithms to construct a plausible mapping of those references into different tasks. Hoarding is then performed on tasks rather than on individual files.

A paper on teleporting, [22], reported on experience with using a system that allows the display of an application to follow the user around as he moves, leaving program execution at the original site. This ability is especially convenient when combined with an active badge system that tracks user location.

Work at Xerox PARC on making ParcTab applications sensitive to the current physical location of the user are described in [96]. Each ParcTab was associated with a user, and that the loss of a Tab was as serious as losing a key, though some additional security could be provided via a PIN code. The privileges of a Tab could be easily revoked by killing the proxy server associated with it. The typical bandwidth was 19.2Kbps, although bandwidths up to 1 Mb/s were possible.

Experiences with designing applications for wireless computing are described in [107]. The theme is that developers should exploit application-specific knowledge to address mobile resource constraints. In certain cases, it was shown to be desirable to offer alternative actions to
the user, allowing them to make performance versus cost decisions.

The use of group communication primitives for mobile computing is detailed in [30]. It was emphasized that the performance overhead of this approach is indeed acceptable. A video transport protocol for wireless networks is also presented in [76]. The novel feature of this protocol is its ability to dynamically adapt the bandwidth required to the current content of the video. Specifically, video segments with a large amount of motion can be rendered in a lossy manner without noticeable degradation of picture quality. In fact, this is the opposite of MPEG, where segments with lots of motion tend to result in higher bandwidth requirements.

The problem of end-to-end TCP adaptation in mobile environments is addressed by Yavatkar et al. [114]. They observed that such communication often involves a wireless segment and a much longer LAN or WAN segment. A solution is described in which intermediary code allows TCP to independently adapt to the characteristics of the two segments. This solution provides substantially better performance, while preserving complete upward compatibility with existing clients and servers.

The problem of wireless communication between mobile hosts in locations where there are no base stations or other mobile infrastructure is presented in [60]. A protocol is described in which the hosts themselves serve as forwarding agents and thus constitute an impromptu mobile infrastructure.

The implementation of a Web browser on an Apple Newton, communicating via a low-bandwidth wireless link is described in [19]. The strategy adopted was that of partitioning applications so that the CPU-intensive processing occurred on powerful servers. This was also shown to be critical to good performance.

A very different approach to mobile access of the Web is described in [65]. Web documents are now programs in TCP/IP that are executed at the client by an interpreter that enforces safety. At present, the system works on IBM ThinkPad clients and Sparcstation servers over a 2 Mb/s WaveLAN wireless link.

Three mobile host protocols from Sony, IBM and Columbia University are compared in [79]. These are compatible with the TCP/IP suite. A set of basic requirements for a mobile host protocol is also suggested and it is observed that none of the three proposals entirely satisfy these requirements. Each proposal is shown to have faults in their implementation of the functionality of the mobile network layer.

Database system issues in nomadic computing are discussed in [11]. This new generation of mobile computers can be an integrated part of a distributed computing environment, one in which users change physical location frequently. Several issues are dealt with in detail, including the reliance on short-lived batteries, power consumption, resource utilization and the likelihood of temporary disconnection.

An overview of wireless systems and technologies is given in [29]. It covers the key technologies and innovations for emerging digital cellular systems; describes advanced signal-processing technologies, which hold promise of further increasing the quality and capability of digital wireless systems. It also reviews the technical requirements and solutions of a high-speed, wireless LAN.

The state-of-the-art in Wireless LANs is discussed in [72],[53] and [45]. New Agent based WANs have been shown to presage the future of connected computing in [91]. Instead of relying on real-time, connection-oriented sessions, smart networks make extensive use of store and forward messaging transports. They are designed to host software agents, or proxies, that move around the network, routing or filtering messages sent to a user and seeking out information or
services on the user's behalf.

In [83] Ohsawa points out that a large gap between the mobile technology and mobile applications exists and argues for the use of a systematic design and development approach. A multiple access technology to realize mobile computing and communication is proposed and discussed.

In [81] Noble and Satyanarayanan present and discuss some empirical study results of the high availability aspects of the Coda file system. They demonstrate that the support provided by Coda can be used for mobile computing supplying evidence from the users. They also explore the effectiveness and resource costs of key aspects of server replication and disconnected operation (the two high availability mechanisms in Coda).

In [73] Maass and van der Crujs point out that the introduction of mobility into the communication networks requires the provision of network-wide information services in order to: enable the mobility service itself; support applications on mobile computing and communications devices; and support users of mobile equipment. An important aspect of these new services is their availability, since in many communications networks the availability requirements are extremely high. The authors argue that a common platform consisting of a distributed directory service like X500 can be used to implement such services network-wide.

In [34] the authors discuss the issues associated with developing distributed system services to operate in an environment where the end systems can experience different degrees of connectivity during typical operational cycles. They argue that the current file systems, including those developed for use in a mobile environment, contain assumptions about their underlying communications infrastructures that are unlikely to hold in a mobile environment. To underline this argument, they examine in-depth a specific file-system issue – the support of shared libraries. Then they propose a new service to support shared libraries in mobile environments and discuss the integration of this service into a wider architecture of reactive services being developed to support distributed mobile computing.

In [102, 101] Spreitzer and Theimer discuss the necessity of providing and using of user location information in a mobile and wireless computing environment and the corresponding security and privacy threats of revealing such information to unauthorized agents. The location information about "public" objects – services, buildings, information depots – is considered desirable while the location information about individual users must be restricted. Some individual user location information is necessary – to route messages directed to a user, to answer queries about "nearby" services or "how to get there", etc. Therefore, the designers of mobile information systems have to concentrate on the problem how to obtain, store, manage, and distribute such location information. The authors assume that there is a "user agent" – the software responsible for release of the user location information who may or may not trust the location and information service for not revealing such information to unauthorized agents. The paper suggests several designs each appropriate for different security goals and environments. The authors describe them starting from the most "trusted" environment and moving toward the most "hostile" one identifying the requirements and some solutions to them in the system design.

In [55] the authors classify and examine various promising approaches for research and development in visual languages; among them – the ubiquitous computing approach. Descriptions of representative languages and research work are provided as well.

In [61] Johnson and Williams consider mobile computing as an important application of high definition display systems. In many mobile applications the computer workstations have to pro-
vide flat panel, projection, and virtual image high definition displays with graphics and imagery for the user at the point of action. They use as a case study the manned space exploration and science program. The authors discuss the development of advanced mobile computing systems which utilize high definition displays and support the user in the remote field environments anticipated by the space exploration program. Such systems can greatly improve worker efficiency through the concept of telepresence.

In [18] Barbara-Milla and Garcia-Molina point out that in a mobile computing environment, it is important to have dynamic replicated data management algorithms that allow for instance copies to migrate from one site to another or for new copies to be generated. They demonstrate that such dynamic algorithms can be obtained simply by letting a transaction update the directory that specifies the sites holding copies and argue that no fundamentally new algorithms are need to cope with this aspect of the mobility. The authors also discuss how the existing algorithms have to be “tuned” for a mobile environment.

In [116] Zadok and Duchamp describe a mechanism for replacing files, including open files, of a read-only file system while the file system remains mounted; the act of replacement is transparent to the user. Such a replacement mechanism can improve fault-tolerance, performance, or both. Their mechanism monitors, from the client side, the latency of operations directed at each file system. When latency degrades, the client automatically seeks a replacement file system that is equivalent to but hopefully faster than the current file system. The files in the replacement file system then take the place of those in the current file system. This work has particular relevance to mobile computers, which in some cases might move over a wide area. Wide area movement can be expected to lead to highly variable response time, and give rise to three sorts of problems: increased latency, increased failures, and decreased stability. If a mobile client moves through regions having partial replicas of common file systems, then the mobile client can depend on their mechanism to provide increased fault tolerance and more uniform performance.

In [84] Parmar reviews how the power management has been done traditionally in the desktop personal computers. Then he argues that the traditional methods do not work properly for the mobile computers and presents some new power management techniques to use in the portable computer designs.

In [56] the authors identify and examine four components as the critical building blocks of a mobile computer environment - the computer, the operating system, the file system, and the communication. They conclude that in the area of the personal mobile computers the predominantly used MS-DOS operating system is not suitable for mobile computer applications. Development of a prototype mobile machine is described.

In [32] Dalton expresses the opinion that the mobile computing puts an unnecessary high burden on the people using it and on the people supporting the mobile environment. The users will be always unhappy with the power, interface, and the accessible applications of the mobile devices, no matter how sophisticated they become, simply because the mobile computers will be at least a step behind the mainstream computers. The supporting personnel will have to work harder to maintain a reasonably good mobile environments where the people and their computers may wander wherever they wish.

In [40] Duchamp and Tait argue that, assuming low-sharing workloads at the mobile computers, it is possible to build more efficient implementations that provide the same (or stronger) semantics by making fullest use of the assumption. Accordingly, they take an extreme point on the spreading/searching tradeoff: the searcher is made responsible for all the work. Thus, spreading can be made asynchronous and hence very fast from the caller’s viewpoint. They
employ an important optimization to eliminate most of the searching work when there is little sharing. Their technique carries the additional advantage of simplicity and flexibility of the recovering algorithms.

In [15] Badrinath, Acharya, and Imielinski discuss the communication and synchronization issues in distributed systems when the underlying network contains mobile, as well as static, hosts. The mobility of the hosts and the corresponding properties of the network such as dynamic architecture, low communication bandwidth, broadcast capabilities, power limitations, and disconnections of the mobile hosts, require a model for distributed systems with mobile hosts that explicitly incorporates the mobility. The authors present a model in which a distributed system consists of a static/fixed network and a wireless network associated with each fixed host (called mobile support station or MSS) for communication with appropriately located mobile hosts. The impact of mobility on the distributed computations is discussed. In order to send messages to mobile hosts, one needs first to search for the host’s location which affects the cost of communication. On the other hand, the medium’s broadcasting capabilities allow sending a message to all “nearby” located mobile hosts at low cost. The change of the network architecture when a host moves causes a need to reassign the logical map between the network and the distributed algorithm. The frequent disconnection of the mobile hosts requires graceful exit from the current computations.

In [59] Imielinski and Badrinath identify and discuss the research issues in the data management with respect to the mobile, wireless computing paradigm as they can be addressed by the database community. There are four basic and orthogonal problems in the mobile computing data management – mobility, disconnection, new data access modes, and scale. They arise from the unique physical characteristics of the wireless computing – small size of the screen, low bandwidth, power restrictions. The mobility describes the fact that the user’s location and access to computing power and the data location can change quite rapidly. The research issues include locating users, handling requests for location-dependent data, replicating data, addressing, etc. The disconnection is considered a “planned failure” and the issues here are the consistency of the local and global data and the recovery of the systems. There is a need for new, more “concise” and energy efficient data access methods prompted by the low communication bandwidth and the battery power limitations. Also, the properties of the medium makes the passive broadcasting of information a desirable approach for disseminating public data. The large number of potential users brings up the issues of scaling the services available to mobile users and their appropriate organization.

In [115] Yokote presents the object-oriented operating system Apertos and its underlying architecture which allows it to accommodate mobile environments. Apertos is built upon a set of objects which are characterized by a set of properties, methods, storage, etc., called metaobjects. A group of metaobjects form a metaspace. The Apertos objects migrate between different metaspaces when their environment changes (e.g., due to the mobility of the host). The author presents a practical implementation of this idea by using a reflector for programming the metaobjects and a MetaCore – a small operating system microkernel which can be run on mobile hosts.

In [41] the authors discuss from the systems software point of view some of the possibilities and requirements for mobile computing on wireless local area networks (LANs). The challenges to the software support for mobile operations (mostly evident in the internetwork routing protocols), the user interface design for the small mobile computers, and the hardware tradeoffs, are identified. The design of the Student Electronic Notebook is sketched to provide a partial cata-

16
log of problems in building a real system for wireless mobile computing. This project has been initiated to investigate the potential of wireless mobile computing to reshape education. Some of the key directions for research in software technology for wireless technology are examined. They include different routing for mobile networks, non-traditional devices and software for user I/O (notebook/pen metaphors and virtual workspaces), file systems suited for the mobile users (with “mobile” files that follow the users), and dealing with the natural characteristics and limitations of the mobile computers. The authors relate some of their experience with LANs.

In [13], the author discusses the issue of anonymity in mobile computing environments. He argues that maintaining anonymity is an important issue and surveys the nature of anonymity provided in proposed or existing mobile computing environments. Example solutions for providing limited, but practical, anonymity using standard cryptographic techniques are also given.

In [82], an architecture to simplify communicating applications is presented. The author views mobile, handheld computers as “mobile companions” of the desktop computer running Microsoft Windows. The architecture is based on each desktop having exactly one “sync partner” handheld system. The partners synchronize their object stores each time they are together and conflicting changes are flagged as such for user intervention. This architecture supports a uniform model for application activation and (event) notification. The socket-based approach makes the media/network type being used for the communication (for example, dial-up line, local docking station, infrared or wireless LAN) transparent to the application.

In [10], the authors present three algorithms for maintaining causal ordering on message delivery in mobile systems. The first algorithm handles resource constraints introduced by mobile hosts, but does not scale well and is not graceful to host disconnections and connections. The second algorithm eliminates these problems while inhibiting some messages. The final algorithm is a trade-off between the first two algorithms.

In [16], the author discusses how changing communication requirements are handled in the MosquitoNet project. Their goal is to allow portable computers to move seamlessly from one communication medium to another, for example from an Ethernet connection to a wireless modem, without rebooting or restarting applications.

In [74], the authors discuss a client–side–only approach to providing access to a heterogeneous set of network file services. This is different from the approach in systems such as Coda [58] which rely on changes to both the client and the server. Rather than committing to a specific file access protocol, they support access to multiple file services.

In [77], the authors present a design for supporting file access in a mobile environment. The main design goals they address are to reduce communication cost and to provide data consistency. The amount of communication is reduced by a combination of caching, profiling, proxied access, delayed writes and loose reads. Data consistency is provided by using the proxy services, a centralized data manager with callbacks and strict reads.

In [9], the authors present techniques for detecting mutual consistency between shared objects. Since objects must be cached at the client to ensure disconnected operation, it is necessary to ensure that the objects are mutually consistent. How to use the approach to support causal consistency is also described.

In [51], the authors describe their approach to dealing with the challenges of providing disconnected operation in the Thor object-oriented database system. The difficulties they deal with are: the small size of objects, the richness and complexity of their interconnections, the large number of them and the fact that they are accessed within atomic transactions. They propose three approaches to address these challenges: using the database query language for hoarding,
using dependent commits to tentatively commit transactions at the disconnected client and using the high-level semantic of objects to avoid transaction aborts.

In [93], the authors discuss a hybrid client-server and peer-to-peer model for file systems based on multi-level replication. They use the relationship between mobile computing personae and file systems to support the file system needs of mobile users.

In [14], the authors describe an indoor wireless system for personalized shopping assistance. This assistance is provided in the form of two components: a hand-held wireless communications device, the Personal Shopping Assistant (PSA), and a centralized server that maintains the various associated databases. The server maintains all the customer and store information and provides audio/visual responses to inquiries from tens to hundreds of customers in real-time over a small area wireless network.

In [94], the authors describe a method for providing identity privacy to mobile users during authentication. This work is aimed at avoiding unauthorized tracking of users' migration. The basic solution is to use aliases, which ensures non-traceability by hiding the user's real identity and also his relationship with domain authorities. The paper presents a classification of the different degrees of non-traceability and presents a new, efficient method for alias computation. Using this technique during authentication of mobile users avoids the drawbacks of existing solutions such as GSM and CDPD.

In [12], the authors present the implementation of a pen-based database interface for a pen computer (using the PenPoint operating system) with a built in cellular phone. Using the cellular phone capabilities, the interface connects to remote databases and displays the schema information of the database chosen by the user. In response to gestures made by the user, the interface automatically generates queries and retrieves results. The universal relation concept is used to aid in automatically generating queries based on the attributes chosen by the user.

In [23], the authors discuss alternative approaches to transparent resource discovery for mobile computers. They consider the mobile-IP approach to transparency (having the mobile machine have some "virtual" location) not acceptable as the associated latencies are location dependent.

In [54], the authors discuss a wide range of issues related to anonymity in mobile environments. They review current state-of-the-art approaches and propose several potential solutions to the problem of keeping user's movements private.

3.1 Research at Purdue

This section describes the research in mobile environments at Purdue.

Joshi et al [64], enumerate the difficulties paramount in building a ubiquitous access system on mobile high performance platforms. The obstacles considered include the user interface for these walkstations, the ability of sniffing information across heterogeneous geographically distributed information systems, the ability of processing sensing data for monitoring and control, and the dynamic reconfigurability of computations between the mobile unit and the stationary servers. Intelligent models that address the aforementioned problems are also presented.

The architectural design of "ubiquitous" Problem Solving Environments (UPSEs) on wireless notebook platforms supported by stationary high performance computing (HPC) servers is addressed in [38]. PSEs are being mapped onto mobile and stationary computational engines. A partial differential equation based PDE (PDELab) is extended and a proxy-based methodology is used to partition the application components across the dual (wireless and stationary) network.
of computational units. This helps in building a distributed PSE architecture. This architecture is currently being validated by building a ubiquitous version of PDELab (UPDELab) on a two mega-bit wireless ethernet network of notebook platforms and heterogeneous parallel machine.

A set of artificial / computational intelligence techniques used by the cooperating agents that constitutes SciAgents, an agent oriented model of computing, is presented in [37]. These agents using local knowledge perform local computations and communicate only with "neighboring" agents. They cooperate in solving a global, complex problem, and none of them exercises centralized control over the computations. This is designed in the context of scientific computing models based on partial differential equations to permit the non-expert user to cost-effectively and easily develop software for solving complex mathematical problems. This approach can be combined with the proxy approach described in [38] to allow wirelessly connected hosts with limited compute power to access the power of networked High Performance Computing platforms. This approach also requires that the cooperation mechanism of the agents be dynamic and adaptable. The issues of learning and adaptation in interactive agents under "distributed problem solving environments" is discussed in [62]. A combination of neuro-fuzzy learning, static adaptation, and an epistemic utility based formulation is shown to be useful in converting a standalone, single agent system into a collaborative multiagent one. It is argued that such systems are the natural choice for the creation of systems that would model the complexities of the physical world. The design and architecture of a multiagent system, SciAgents, for cooperative and distributed scientific computing is presented. A language (SKI-F), useful for interagent coordination is also outlined.

For creating "intelligent" interfaces to work with mobile hosts, AI techniques sensitive to the domain of application are needed. PYTHIA, an intelligent agent to guide PDE (Partial Differential Equation) solving is introduced in [109]. PYTHIA attempts to solve the problem of determining an optimal strategy (i.e., a solution method and its parameters) for solving a given PDE problem within user specified resource and accuracy requirements. It uses the performance behavior of solution methods on previously solved problems as a basis for predicting a solution strategy for solving a new problem.

It has been recognized that many of the decision making processes involved in the above methodology are not 'crisp'. Rather, they are 'fuzzy'. Several recent computationally intelligent approaches have been proposed that address this situation. Ramakrishnan et al. [90], propose a neuro-fuzzy system that automate the "class-selection problem", the first step in PYTHIA's method, that determines the class(es) to which the given PDE problem belongs. A wide comparison of computationally intelligent approaches, under the broad umbrella of 'soft computing', is presented in [63], where it is shown that such methods outweigh traditional and naive procedures of classification.

The CrossPoint project [48, 31] at the Dept. of Computer Sciences, Purdue University aims at providing a seamless internet environment for a large university campus that allows mobile, wireless communication. The primary problem addressed is the dynamic management of routes when mobile computers move from the range of one base station to another. The approach combines the wireless network technology with high-speed Asynchronous Transfer Mode (ATM) switching technology to solve the problem. ATM switches not only provide sufficient aggregate bandwidth to handle both data transfers and routing updates, but also facilitate sending routing update packets with little overhead using a single cell. This is shown to be feasible to construct a wireless internet which scales to allow each individual on a campus to carry a mobile computer.

Data management in mobile distributed environments is another area of research at Purdue.
General issues and concerns about dealing with mobile environments are enumerated in [88]. The focus of this research is to identify the impact of wireless communications on distributed computing systems. Maintaining consistency of data in mobile distributed environments [86] is achieved in the following way. A database is partitioned into a set of clusters in a dynamic manner. While all data inside a cluster are mutually consistent, degrees of inconsistency are allowed among data at different clusters. When clusters are merged, strict consistency is restored.

Means for providing transaction support appropriate for mobile environments are discussed in [87]. The general architecture of an information system for mobile environments is described in [85]. System support for maintaining consistency of replicated data, and for providing transaction schemas that account for the frequent but predictable disconnections, the mobility, and the vulnerability of the environment is also detailed. Reliable stream transmission in mobile computing environments is also discussed [113].

Research in Mobile environments of telemedicine at Purdue is detailed in [50]. This work is being done by the Telematics group at Purdue, which consists of several faculty members of the Computer Science Department, and collaborators from Electrical Engineering, Agricultural Engineering, and County Extension office. The project consists of developing and analyzing infrastructure for telemedicine, and prototyping it in the Indiana Rural Medical Network (IRMN). A prototype telemedicine environment MediPad will be implemented on a platform consisting of mobile hosts and a high-speed network facility connecting seven identified counties, the Indiana University Medical Center and Purdue. Performance evaluation will be done in the context of the medical needs of these highly rural counties. This network will also be used for distance learning, especially in the medical context.

4 Webligraphy

In addition to this static report, we have created a webligraphy of information on mobile computing. This webligraphy, which can be found at the URL

http://www.cs.purdue.edu/research/cse/scipad/mobicomp.html

is an active document that we expect to continue to maintain and update for an indefinite period.

In this following, we provide a list of the major URLs that are contained in this webligraphy.

University Research Groups

- Center for Information Technology and Integration, University of Michigan
  http://www.citi.umich.edu/mobile

- Dataman Research Group, Rutgers University
  http://paul.rutgers.edu/acharya/dataman.html

- Mobile Computing Laboratory, Department of Computer Science, Columbia University
  http://www.mcl.cs.columbia.edu/

- InfoPad Group, Department of Electrical Engineering and Computer Science, University of California, Berkeley
  http://infopad.eecs.berkeley.edu/

- Mobile Computing at the University of Washington
  http://www.cs.washington.edu/research/mobicomp/mobile.html
• Mobile Computing at Macquarie University
  http://www.mpce.mq.edu.au/ andrewm/mobile.html

• Wearable Computers Group, Carnegie Mellon University
  http://www.cs.cmu.edu/afs/cs.cmu.edu/project/vuman/www/home.html

• Walkstation II Project, Royal Institute of Technology, Sweden
  http://www.it.kth.se/TSlab/WS/ws.html

• Mobile Computing Home Page, Lancaster University, UK
  http://www.comp.lancs.ac.uk/computing/users/nigel/most.html

• Mobile Computing Project, Shoshin Lab, University of Waterloo
  http://ccnga.uwaterloo.ca/mobile/

• Wireless LAN Group, University of Massachusetts, Amherst
  http://www.ecs.umass.edu/ece/wireless/

• Index of Mobile and Wireless Computing, Department of Computer Science, University of Washington
  http://snapple.cs.washington.edu:600/mobile/

Industry/Military Research

• The Active Badge System, Olivetti Research Laboratory, UK
  (http://www.cam-orl.co.uk/ab.html

• Monet Home Page
  http://fury.nosc.mil/

• Mobile Computing at BBN
  http://malachite.bbn.com/Departments/DistributedSystems/Mobile.html

• Ubiquitous Computing, Xerox PARC

• The Apple Virtual Campus: Mobility

• WAMIS Program, Electronic System Technology Office
  http://esto.sysplan.com/ESTO/WAMIS/

5 Acknowledgements

We would like to thank Intel Corporation and The Intel Foundation for their gracious support of this survey. Specially, Dr. Pat Mitchell of the Mobile Architecture Labs has been very supportive of our efforts and we greatly appreciate his useful comments and insights as we progressed on this work. We also acknowledge the support provided to us by Purdue University in the form of computing and other facilities to realize the survey.
References


