Mobile Computing: Overview and Current Status

Arkady Zaslavsky [☯], Zahir Tari [◇]
[☯] School of Computer Science & Software Engineering, Monash University
900 Dandenong Road, Caulfield East 3145, Australia; Phone: +61 3 9903 2479
Fax: +61 3 9903 1071; Email: A.Zaslavsky@Monash.edu.au

[◇] Dept. of Computer Science, RMIT, Bundoora Campus, Plenty Rd., VIC 3083 Australia,
zahirt@cs.rmit.edu.au

ABSTRACT

In recent years, mobile computing has become the focus of vigorous research efforts in various areas of computer science and engineering. These areas include wireless networking, distributed systems, operating systems, distributed databases, software engineering, applications development, just to name a few. This paper introduces the conceptual overview of mobile computing, its achievements, challenges and opportunities. The current status and ongoing research projects in mobile computing worldwide are detailed. This paper also discusses the two Australian workshops on mobile computing, databases and applications held in 1996 and 1997. The selected papers from these two workshops form the basis for this special issue of Australian Computer Journal.

Key Words: mobile computing, wireless networks, data management, distributed systems

CR: C1.4 (Mobile processors), C2.1 (Wireless communication), C2.4 (Distributed systems)

1. INTRODUCTION AND BACKGROUND

Mobile computing is associated with mobility of hardware, data and software in computer applications. Mobile computing has become possible with convergence of mobile communications and computer technologies, which include mobile phones, personal digital assistants (PDA), handheld and portable computers, wireless local area networks (WLAN), wireless wide area networks and wireless ATMs. The increasing miniaturisation of virtually all system components is making mobile computing a reality (Alonso and Korth, 1993; Forman and Zahorjan, 1994). Mobile computing - the use of a portable computer capable of wireless networking - is already revolutionising the way we use computers.

Wireless networking has greatly enhanced the use of portable computers. It allows users versatile communication with other people and outright notification about important events and convenient access to up-to-date information, yet with much more flexibility than with cellular phones or pagers. It also enables continuous access to the services and resources of stationary computer networks. Wireless networking promises to do for portable computers what traditional networks have done for desktop personal computers. Networks enable stand-alone personal computers to participate in distributed systems that allow users anywhere on the network to access shared resources. With access to a wireless network, mobile users can download news or electronic documents, query a remote database, send or receive electronic mail, or even be involved in a real-time video-conference with other users.

The technical challenges that mobile computing must resolve are hardly trivial. However, some of the challenges in developing software and hardware for mobile computing systems are quite different from those involved in the design of today's stationary networked systems (Forman and Zahorjan, 1994). These challenges are discussed below in more detail. Also the implications of host mobility on distributed computations are quite significant. Mobility brings about a new style of computing. It affects both fixed and wireless networks. On the fixed network, mobile users can establish a connection from different locations. Wireless connection enables virtually unrestricted mobility and connectivity from any location within the radio coverage.

Mobile user location becomes a dynamically changing piece of data. In this case, the user updates this information, while many others may access it to find out where the mobile user is. In the mobile environment, the location of a user can be regarded as a data item whose value changes with every move. Establishing a connection requires knowledge of the location of the party we want to establish a connection with. This implies that locating a person is the same as reading the location data of that person. Such read operations may involve an extensive search across the whole network as well as a database look up. Writing the location may involve updating the location of the user in the local database as well as in other replicas of this data item (Imielinski & Badrinath, 1994).

One important characteristic about mobile computers is that they have severe power restrictions. A battery represents the largest single source of weight in a portable computer. While reducing battery weight is important, a small battery can undermine the value of portability by causing users to recharge frequently, carry spare batteries, or use their mobile computers to a minimum. Minimising power consumption can improve portability by reducing battery weight and lengthening the life of a charge. Power can be conserved not only by the design of energy-efficient software, but also by efficient operation (Douglis et al, 1994; Zhou et al, 1998). Power management software can power down individual components when they are idle, for example, spinning down the internal disk or turning off screen lighting. Applications may have to conserve power by reducing amount of computations, communication, and memory, and by performing their periodic operations infrequently to amortise the start-up overhead. The other characteristic of mobile computing is that the cost of communication is asymmetric between the mobile host and the stationary host. Since radio modem transmission normally requires about 10 times as much power as the reception operation, power can be saved by substituting the transmission operation for a reception one. For example, a mobile support station (MSS) might periodically broadcast information that otherwise would have to be requested frequently. In this way, mobile computers can obtain this information without wasting power to transmit a request.

The other important characteristic of mobile computing is the frequent disconnections and dozing of mobile computers. The main distinction between the disconnection and a failure is its elective nature. In traditional distributed systems, the loss of connectivity is considered to be a failure and leads to the network partitioning and other emergency procedures. Disconnections in mobile computing, on the other hand, should be treated as planned failures, which can be anticipated and prepared for. There may be various degrees of disconnection ranging from a complete disconnection to a partial or weak disconnection, eg., a terminal is weakly connected to the rest of the network via a low bandwidth radio channel. The reasons for disconnections may be due to costs involved, as it is expensive to maintain an idle wireless communication link. Also, it could happen that there are no networking capabilities at the current location. In addition, for some technologies, such as cellular modems, there is a high start-up charge for each communication session (Badrinath et al, 1993; Satyanarayanan et al, 1993). Disconnections are undesirable because they may impede computation. Moreover, the increasing scale of distributed systems will result in more frequent disconnections.

The other major issue in mobile computing is concerned with security and privacy (Forman and Zahorjan, 1994). Since mobile computers appear and disappear on various networks, prevention of impersonation of one machine by another is problematic. When a mobile computer is taken away from its local environment, the data it sends and receives are subject to possible theft and unauthorised copying. A network that allows visiting mobile computers to connect to it cannot perform the type of packet filtering now used as a security mechanism, since certain foreign packets will be legitimate packets destined for the visiting mobile host. However, the administrator of the foreign environment has security concerns as well. These concerns are much greater than the current mode of mobile computing in which a user in a foreign environment is logged into a local guest account from which the user may have a communication session (eg., telnet protocol) to his/her home environment. In the nomadic computing paradigm, a guest machine may harm its host/server - either accidentally or maliciously (Asokan, 1994). The possibility of such a harm is much greater than that likely by the typical user of a guest account on a fixed network.

Another major issue is establishing a connection when a mobile host may have no prior knowledge about the targeted network (Nzama et al, 1997). The point of entry in a network is through the physical medium or interface to the access point. The choices of physical medium include radio, infrared, wire/coaxial cable and optical means. Furthermore, a mobile host needs to communicate using one of the host network's protocols for meaningful exchange of information to occur. In addition, networks may have established security schemes. In order to join the targeted network, information on the 'code of behaviour' is normally provided to the incoming member of the community. This arrangement, characteristic of legacy computing systems, works well in a static environment. This approach does not apply to mobile hosts, which migrate within and across networks. It is important to note that complexity of connectivity depends on the variety of choices presented to the node. For example at the signal level, there are several choices to be on the medium, access method and encoding. Also, once a protocol is known, there are several ways it can be used by the upper layers. A mobile host to start communicating with a network needs to *'speak the same language'* as the targeted network. After all, the situation can be likened to visiting an unknown country where one has no prior knowledge of the language, customs, behaviour but somehow hopes to communicate and ask for directions, food or any other services. Such paradigm can be called "the ET (extraterrestrial) effect" (Nzama et al, 1997). A mobile computer that intends to establish a connection in a foreign computer network is viewed as an outsider and may have no prior knowledge of how to instigate communications. This is a situation that will arise over and over again as people demand computing anywhere without geographic barriers such as those partially achieved in GSM technology.

Wireless data networks are a natural extension and enhancement to existing wireline computer networks and services. Wireless data networks support mobile users who may require remote access to their base computer networks. Wireless data services and systems represent a rapidly growing and increasingly important segment of the telecommunications industry (McGladdery and Clifford, 1993; Motorola, 1995). For example, the three telecommunication carriers in Australia are spending billions of dollars in building wireless networks infrastructures in a bid to attract the expected migration of wireline networked information systems to the wireless applications' arena. Wireless communications for Personal Digital Assistants (PDA) are potentially enormous new markets. Figure 1 (ElWazer and Zaslavsky, 1997**)** shows the future growth in worldwide mobile users numbers. When projected to Australia, the growth is even higher given large distances and relative scarceness of populated centres.

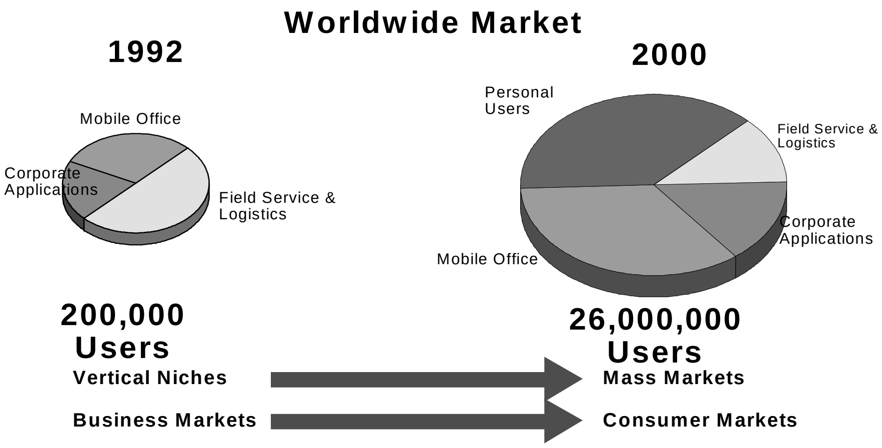

Fig.1. Future growth in worldwide mobile users numbers

It is easy to notice that current computer applications follow the rapid advancements in the telecommunications industry. Eventually, information systems will be influenced by the rapid evolution of the wireless segment of this industry. Since mobility affects many assumptions upon which today's distributed systems are based, such systems will have to move to where tomorrow's technology can support them. Wireless data technology is foreseen to be a main infrastructure platform for future applications, which are naturally distributed, dynamic and require much flexibility and mobility.

In mobile computing systems, the underlying network infrastructure is somewhat different from traditional distributed systems. Designers of mobile information systems have much less control over wireless networks since not only the communication media is provided by telecommunications providers, but also base stations and servers are part of a proprietary wireless network. For example, location of base stations is considered commercial information and is unavailable to application developers. The layered conceptual architecture of mobile computing systems is illustrated in figure 2. The following subsections of will focus on these three layers.

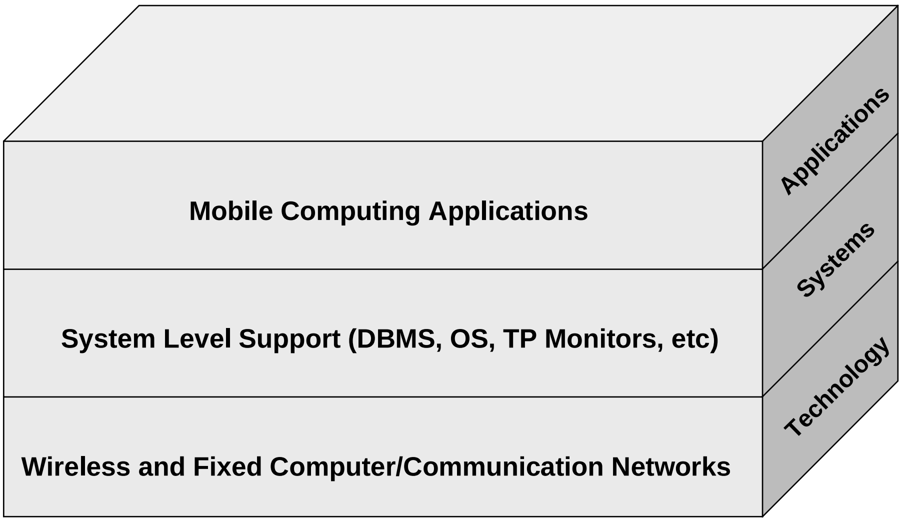

Fig.2. Layered View on Mobile Computing Systems

1.1 Wireless Networks and Hardware

The components of mobile computing infrastructure are illustrated in figure 3. The architectural model consists of two distinct sets of entities: mobile hosts and fixed hosts. Some of the fixed hosts, called mobile support stations (MSS) (Imielinski and Korth, 1996) or home base nodes (HBN) (Yeo and Zaslavsky, 1994) have a wireless interface to communicate with mobile hosts. Mobile hosts can connect to any other fixed host where it can register as a visitor. This node is called the visitor base node (VBN). The VBN routes all transactions, messages and communication calls to and from the mobile host to its appropriate HBN. The segment of a larger computer network or a geographical area controlled by a corresponding HBN is called its zone of influence. Fixed hosts and communication links between them constitute the static or the fixed network, and can be considered to be the trusted part of the infrastructure. Thus, the general architecture for the network with mobile hosts is a two tier structure: powerful and reliable fixed network with mobile support stations and a large number of mobile hosts, which are roaming within and across multiple heterogeneous networks and are connected by slow and often unreliable wireless links.

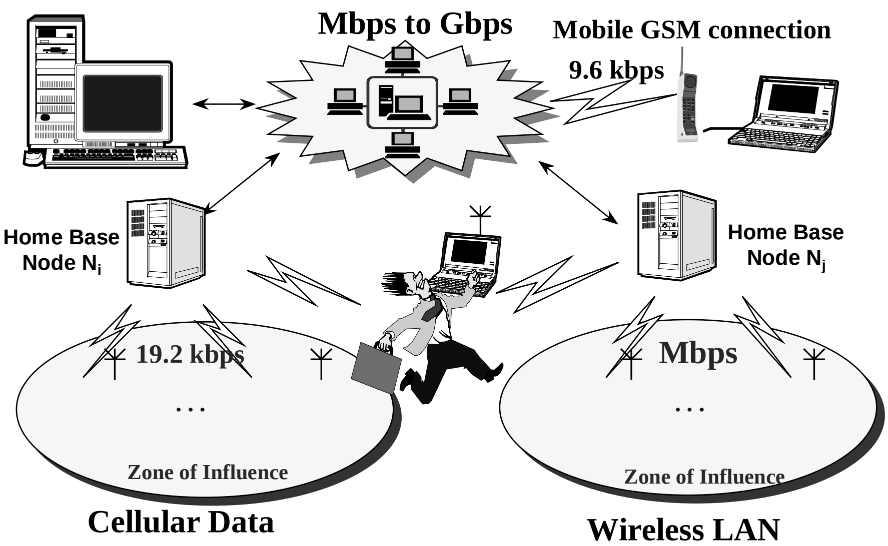

Fig.3. Mobile Computing Environment

Wireless communication services can be grouped into relatively distinct groups (Motorola, 1995; ElWazer & Zaslavsky, 1997a). The grouping is done with respect to the scale of mobility and communication modes and is summarised in Table 1 below.

Table 1. Wireless communications groups

Group	Categorised by;
Cordless Telephones	low mobility, low-power, two-way wireless voice communications, with low mobility applying both to the range and the user's speed.
Cellular Mobile Radio systems	providing high mobility, wide-ranging, two-way wireless communications, with high mobility applying to vehicular speeds and to widespread regional to national coverage.
Wide-Area Wireless Data Systems (Mobile Data Systems)	high mobility, wide-ranging , low-data-rate digital communications to both vehicles and pedestrians.
High-Speed Wireless Local-Area Networks (WLANs)	low-mobility high-data-rate data communications within a confined region, eg. a campus or a large building. An IEEE standards committee, 802.11, has been attempting to put some order into this area.
Paging/Messaging systems	one-way messaging over wide areas
Satellite-Based Mobile Systems	provides two-way (or one-way) limited quality voice and/or very limited data or messaging to very wide-ranging vehicles (or fixed locations)

Wireless networks modulate radio waves or use pulsing infrared light to communicate. Stationary transceivers link the wireless communications to the wired network infrastructure. Wireless communications can be affected by the surrounding physical environment. It interferes with the wireless signal, blocking signal paths and introducing noise and echoes (Forman and Zahorjan, 1994). Wireless communications are characterised by limited bandwidth, high error rates, and frequent spurious connections/disconnections. These factors increase communication latency. This is a result of re-transmissions, re-transmission time-out delays, error control processing, and short disconnections. It might be hard to maintain quality of service (QoS) while the mobile host moves across multiple heterogeneous wireless networks. Mobility can also cause wireless connections to be lost or degraded. Users may travel beyond the coverage of a wireless network or enter areas of high interference. Unlike typical wired networks, the number of devices in a wireless cell varies dynamically and large concentration of mobile users, say, at conventions, hotels and public events may overload network capacity.

Wireless data networks are in operation in most developed countries and are increasingly becoming popular. The current trend in this segment of telecommunications industry is towards open wireless data networks standards and protocols. The proprietary protocols are becoming open to developers and researchers. For example, Motorola opened the RD-LAP protocol (Motorola, 1995). Mobitex protocol (Bellsouth**,** 1996) which was developed by Swedish Telecom and Ericsson has always been open, and the Cellular Digital Packet Data (CDPD) is open to the public too.

The attitude towards wireless networks depends on the type of the wireless network. In a wireless voice service, for example, the user usually appreciates the general characteristics and limitations of radio transmission and is tolerant of occasional signal fades and brief dropouts (Levesque and Pahlavan, 1996). In contrast, it is interesting to notice that data services users are concerned with the accuracy of delivered messages and data, the time-delay, the ability to maintain service while travelling around, and the cost of the service.

Wireless network technologies will continue to develop, sophisticate and offer more services and greater flexibility at lower costs. Users will be able to choose from a wide range of technologies and "mix and match" wireless services with wireline communications in an effort to meet their needs in the most cost-effective manner. An overview of the widely popular network technologies is discussed in (Motorola, 1995; Satyanarayanan, 1993; Levesque and Pahlavan, 1996; McGladdery and Clifford, 1993; Ericsson, 1996; Bellsouth, 1996). Here we summarise the most important features and characteristics with the particular emphasis on wireless technologies available in Australia.

Public Packet Data Networks (PPDN) are provided to the public by service providers that offer wireless data communications. Private networks, used by fleet operators and support services such as emergency services, also use this type of networks (Motorola, 1995). Data transmission is only provided on these networks and use an infrastructure of base stations, network control centers, and switches to transmit the data. Enterprise systems and third-party service providers can connect host data systems to the network via wireline communications. Charges are based on the amount of data transmitted, not the connect time.

Packet-switched data networks are more economical to operate than similar circuit-switched networks. They allow many devices to share few communication channels. Transmission speeds vary from 4800 bps to 19.2 kbps. However, the actual transmission time and throughput is determined by the network load and overheads, and cannot be precisely specified. Packet data networks are well suited for short data transmissions where the overhead of setting up a circuit is not warranted for the transmission of data bursts lasting only seconds or less. Two widely used packet data network worldwide are the Motorola's DataTac (Motorola, 1995) and Ericsson's Mobitex networks (Ericsson, 1996; Bellsouth, 1996).

Cellular digital packet data (CDPD) is a new technology that transmits data packets over existing analog cellular networks. It is ideally suited for established voice cellular analog network operators who wish to add wireless data communication to their existing services. CDPD has the same in-building coverage as the current voice cellular analog networks. CDPD is widespread in the USA and does not have much future in Australia as analog cellular networks are currently being phased out and will be completely replaced by digital networks in the next 2-3 years. By the same token, it is unlikely that CDPD will become popular in South East Asia, as digital wireless networks play the dominant role there.

There are only three established wireless telecommunication carriers in Australia so far: Telstra, Optus and Vodafone (provides GSM service only). The networks owned and operated by these carriers cover most of the available worldwide wireless telecommunication technologies. Telstra, meanwhile, is the major provider in the telecommunications industry in Australia. The discussion here focuses on public wide-area wireless networks. Sound analysis, references and resources related to wireless LANs can be found at the wireless LAN alliance (WLANA) home page (http://www.wlana.com//index.html).

There are two types of wireless data networks in Australia: circuit-switched and packet-switched networks. In circuit switched networks data is carried over a voice system such as AMPS, GSM, some Satellite Services or trunked mobile radio. Australia uses Advanced Mobile Phone Service (AMPS) for its analogue network which uses the 800 MHz band. The AMPS network is solely operated by Telstra. However, under a certain agreement, Optus uses the same AMPS network for its customers (Budde, 1995; AUSTEL, 1996). Under the Australian Telecommunications Authority (AUSTEL)'s guidelines, Australia will move completely to the new digital standard GSM (Global System for Mobiles), which uses the Time Division Multiple Access (TMDA) standard and operates using the 900 MHz band, by the year 2000.

There are currently three GSM networks in Australia. Telstra operates MobileNet GSM network supplied by Ericsson. Optus Communications operates a second GSM network supplied by Nokia. Vodafone, which provides only GSM services in Australia, but not data services, operates the third GSM network supplied by Ericsson.

There are three major two-way packet-switched networks in Australia: MobileData operated by Telstra based on Motorola's DataTac 5000, Mobitex operated by United Wireless based on Ericsson's Mobitex (Ericsson, 1996), and the satellite service Inmarsat-C which is operated by Telstra.

Telstra MobileData network is based on Motorola's DataTac 5000. It uses the Motorola open protocol RD-LAP 19.2. The raw data rate is 19200 bps on 800MHz, 25kHz channel. Net data rate is estimated at 14,400bps. Wireless modem are half-duplex, ie, radio cannot receive packets while transmitting. Telstra MobileData is believed to cover 70% of the Australian population, mostly in metropolitan areas. Telstra MobileData supports fleet connectivity where one host application communicates with up to thousands of wireless modems (terminals). On Telstra MobileData there are three protocols involved: Standard Context routing (SCR), Native Control Language (NCL), and RD-LAP 19.2. MobileData users do not have to deal at all with the latter protocol. This protocol communicates between the wireless modem terminal and the MobileData base station. Figure 4 (Motorola, 1995) depicts the locations of the above mentioned protocols.

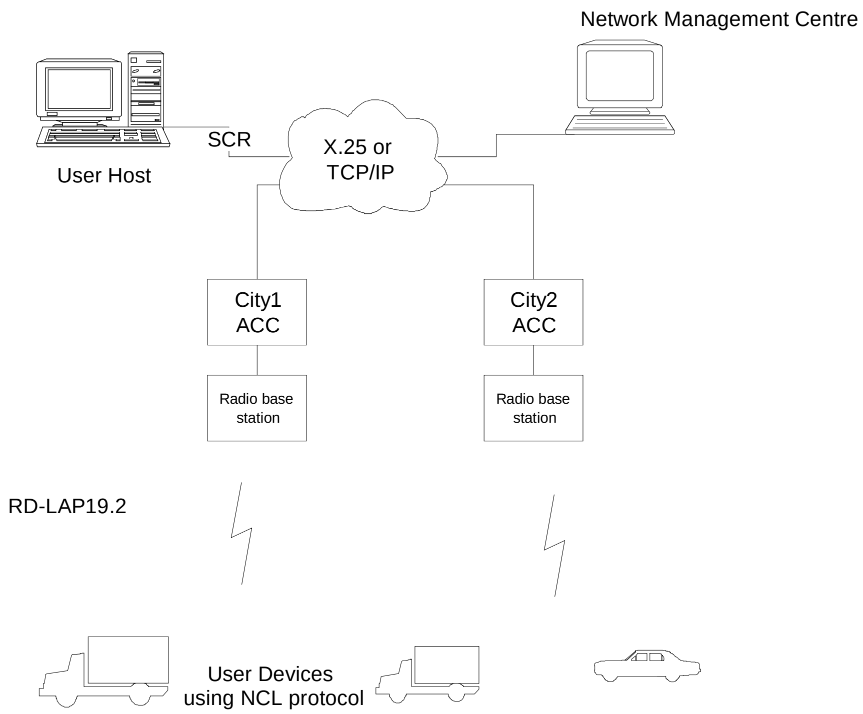

Fig.4. DataTac 5000 architecture

Currently, the GSM networks in Australia can transmit fax and data at 9600bps. Packets over GSM networks can be transmitted via the SMS (short message service) service, with a limited packet size of 140 Octets or 160 characters. There are plans to enhance the GSM technology. A packet data service similar in functionality to CDPD is being designed for GSM networks, but will not likely be commercially available until 1999 at the earliest. GSM will be able to provide packet data capacity now being designed, known as General Packet Radio Services (GPRS). This is a packet radio network service which will provide higher-speed data services for mobile users up to 200 kbps (Ericsson, 1996). GSM phones will soon be able to handle video-conferences, multi-media applications. A new part of the GSM standard is being developed. This is known as High Speed Circuit Switched Data (HSCSD) (Ericsson, 1996) and will boost user capacity from 9600 bps to 64 kbit/s and higher.

These future enhancements to the GSM packet and voice/data services will put a lot of pressure on the existing two-way dedicated packet-switched data in Australia. In two years time, circuit-switched future GSM will make packet-switched DataTac and Mobitex obsolete in terms of speed, capacity, and efficiency. Wireless customers might prefer to have only one handset for voice, data, fax, and packet-switching, which will be a GSM handset.

Australia is a vast continent and has different types of wireless customers. For instance, GSM suits customers in the GSM coverage area. Customers in rural areas of Australia need satellite coverage. Therefore, having too many standards in Australia proves to be a realistic assumption in the wireless networks industry. Also, different wireless telecommunication carriers will have different coverage maps.

Concluding this subsection, it must be said that many research projects in wireless telecommunications are ongoing. For instance, connecting different carriers' networks supplied by different vendors will prove to be very challenging research from a technical and marketing point of view. Internetworking different packet networks such as DataTac and GPRS, by building multi-protocol gateways (ElWazer & Zaslavsky, 1997b), is also another interesting research area. Although the forthcoming proposed 64kbps circuit-switched GSM network may not be enough for a real-time mobile multimedia application, it will prove to be the long-awaited seed to start research in this area. GPRS is expected to host fast packet-switched type of applications such as surfing the Internet.

Wide-area wireless networks (WANs) pose a series of limitations that most current information systems are not equipped to deal with. For example, the net data rate for wide-area wireless networks is limited to a maximum of 14,400 bps in the DataTAC RD-LAP 19.2 kbps capable protocol. The proposed 200 kbps GPRS over GSM networks will also prove to be insufficient for efficient information systems. The convergence of wireless LANs and wireline WANs may become an answer for efficient wireless information systems. Researchers might need to look at this type of convergence as it means existing efficient information systems, designed on wireline networks, will need minimal modifications to become wireless. For instance, wireline TCP/IP based applications can be ported to the wireless computing environment with little modifications to host and client programs. Wireline WANs with 100 Mbps fiber optic backbone combined with efficient wireless LANs using data compression techniques will support real time wireless information systems.

1.2 System Level Support for Mobile Computing

The second layer of figure 2 involves system level support for mobile computing. This support includes database management systems (DBMS), distributed operating systems, algorithms and models that enables mobility in distributed computations across heterogeneous computer networks.

Changes and enhancements to operating systems, as well as other system level additions are discussed in Imielinski and Korth (1996), Imielinski and Badrinath (1994), Marsh, Douglis and Caceres (93). The major issues typically involved here, include disconnection management, file system design, concurrency, query processing, transparent delegation of tasks and responsibilities, data replication, and transaction models.

Support for disconnected operations is very important for mobile computing systems as the moving host can frequently disconnect from the network while dropping out of the coverage areas. Unlike traditional distributed systems, this should not be treated as a failure. Implicit support via the OS, or explicit support embedded in the application should enable continued processing on a local file system. Such disconnected processing can also be based on advance caching (Chung and Cho, 1998), when the application is mobility aware and downloads enough data to survive between two successive connection/disconnection cycles. The most known example of disconnected operation capable system is Coda file system (Kistler and Satyanarayanan, 1992, http://www.cs.cmu.edu/afs/cs.cmu.edu/project/coda/Web/docs-coda.html). After reconnection, the changes to files are propagated to the file server and client's updates are synchronised with updates at the server side. The Ficus file system (Heidemann et al, 1992, http://ficus-www.cs.ucla.edu/ficus/) also supports disconnected operation with certain limitations related to read/write ratios.

Replication is necessary in a distributed mobile file system because of the very nature of a mobile host. Mobile hosts can not remain connected to the rest of the system all the time. This is mainly due to two reasons, firstly, to save resources on a mobile host and, secondly, a mobile host can go out of a zone of influence thus causing disconnection. Replication in mobile database systems is discussed in Zaslavsky et al (1996) and in a number of papers from the ECOOP'95 and ECOOP'96 Workshops on Mobility and Replication (http://www.di.uminho.pt/~cbm/wmr-programme.html and http://www.diku.dk/distlab/ workshops/wmr96/pgm.html).

Chrysanthis (1993) provides an axiomatic definition for two transaction types, namely reporting transactions and co-transactions to model the interaction between a mobile host and its base node, but provides no support for operation during disconnection. Pitoura and Bhargava (1995) revise the transaction concepts for mobile computing and take into account the vulnerability of the computation performed at mobile hosts. They also introduce the concept of *transaction proxies* to support database recovery. For each transaction executed at a mobile host a dual transaction, called proxy, will be executed at the base node of that mobile host. The proxy transaction only include the updates of the original transaction. Thus, any time a transaction is submitted to a mobile host, its proxy is submitted to its base node. In that sense proxy transactions correspond to taking periodic backup of the computation, which is performed at a mobile host. The operation during disconnection is not discussed.

The twin-transaction model was introduced by Rasheed and Zaslavsky (1997). This model suggests replicating not only data items from mobile hosts, but also the process of executing the transaction at both the mobile host and its MSS. The twin-transaction model is application sensitive since it achieves high availability at the expense of slight ambiguity introduced as a result of executing a transaction based on previous patterns of behaviour and history. The proxy-transactions and twin-transactions are different in their essence. The former is generated when the mobile is connected and a transaction is submitted to it, whereas the latter is generated no matter whether the mobile host is connected or disconnected. Also proxy-transactions are used to take periodical backups of the computation done on an MWS whereas twin-transactions are more in the lines of emulating the transaction process of an MWS and stationary network and are used for synchronisation as well on reconnection.

Certain work has been done in the area of maintaining data consistency in mobile distributed environments. There are two type of transactions namely weak and strong transactions. The weak transactions are left to have inconsistency thereby not adhering to the ACID properties of traditional transactions. Moreover, clusters over a group of sites are defined and inconsistencies which are present in the cluster are not propagated outside the cluster. An slight problem with the method proposed by Pitoura and Bhargava (1995) is that each data item has two versions on the mobile host. This may not be a viable option as storage capacity is limited and disk space being a valuable resource cannot be used to store two versions. The essential tradeoff to the disconnection problem is between the cost of storing two versions versus communication and bandwidth cost.

Chu et al (1992) discuss the issue of data availability and use data inference to reduce the degree of replication in the system. Maintenance of a knowledge base is required when data is not available, possibly due to some site failures. In the model it is proposed to allocate data objects. The rules, which are based on application semantics, are required to reduce the search space. This model presents an extension to traditional approach of data replication to increase reliability and availability without requiring a higher degree of replication.

Many interesting issues of data management in mobile computing environment still are very much open research issues. For example, database recovery under partially available information becomes hard to resolve. Power efficient query optimisation is another example of an open research issue in mobile databases.

1.3 Applications Over Wireless Data Networks

The combination of networking and mobility will engender new types of information systems. These systems can support collaborative environments for impromptu meetings, electronic bulletin boards whose contents adapt to the current viewers, lighting and heating that adjust to the needs of those present, Internet browsing, hotel and flight reservation, navigation software to guide users in unfamiliar places and tours, wireless security systems, wireless electronic fund transfer point of sale (EFTPOS) systems, remote E-mail, enhanced paging, wireless facsimile transmission, remote access to host computers and office LANs, information broadcast services, and law-enforcement agencies, to name just a few applications (Ji and Zaslavsky, 1996). Some of possible characteristic scenrios are discussed below.

The field sales application moves the trading pit to the site of the customer. It applies to the situation where all the salespeople travel to sell. The system requirements are as follows. Firstly, the terminals will likely to be portable computers with radio modems. They have a large amount of storage with sufficient processing power. Secondly, the data access pattern is two-way communication. Order entries, and database access and updates are required. Thirdly, as for performance requirement, instant response is desirable though the transactions are not massive. Fourthly, the price is negotiable with the wireless network operator. Finally, the transmission medium implemented in the wireless network is transparent to the user. The preferred one will be radio frequency (RF) communication as the satellite network does not provide in-building coverage.

Field service is different from the field sales application though they sound similar. In addition, GPS (Global Positioning System) uses the wireless network so that the positions of mobile users/vehicles are reported to the central/regional office. The system requirements are as follows. Firstly, the terminals used for this type of applications are, in most cases, vehicle based. Secondly, the data access patten is two-way communication: GPS data to the central office, and remote access, updates of the home databases, and order entries into the databases. Thirdly, the performance requirement is that instant response is desired, but the system is not transaction intensive. Fourthly, besides the normal charges, there is likely a cost for GPS equipment. Finally, the transmission medium would either be radio or satellite.

The application scenario for emergency response is a special one. Features for such an application area are hard to define since they vary from one situation to another. For example, bushfire control is different from flood control. For illustration, we take the example of bushfire control. Firstly, the wired network setup is almost impossible although the wired network can be set up in the command centre. The terminals needed would be the ones found in the field service applications, terminals installed in the fire fighting vehicles. Secondly, the data access is two-way communication as both access and updates of the distributed databases are necessary. Thirdly, the system performance requirements are strict: instant response and high volume of data per time unit as maps are accessed and updated frequently. Fourthly, the cost of using the network is not of priority. Fifthly, the transmission medium in this case has to be satellite(s) because other media do not cover less populated areas unless a backbone network is also wireless, eg, a wireless ATM. Sixthly, the network setup should be fast as time is critical. Finally, the GPS functionality is also required for this situation as the positions of all vehicles are reported to the commander centre, which can be used to reflect the global picture of the burning forests. The eventual internetworking scenario is illustrated in Figure 5 (Direct Broadcast Satellite Technology, 1996). Three satellite networks are interconnected: LEO (low earth orbit) for the communication among fire fighting groups and the communication between the control centre and the fire brigades; GEO (geo-stationary orbit) for the communication between remote databases and the command centre; and a special satellite network for GPS. An alternative scenario may involve a heterogeneous wireless network, which may include a satellite link and several wireless LANs.

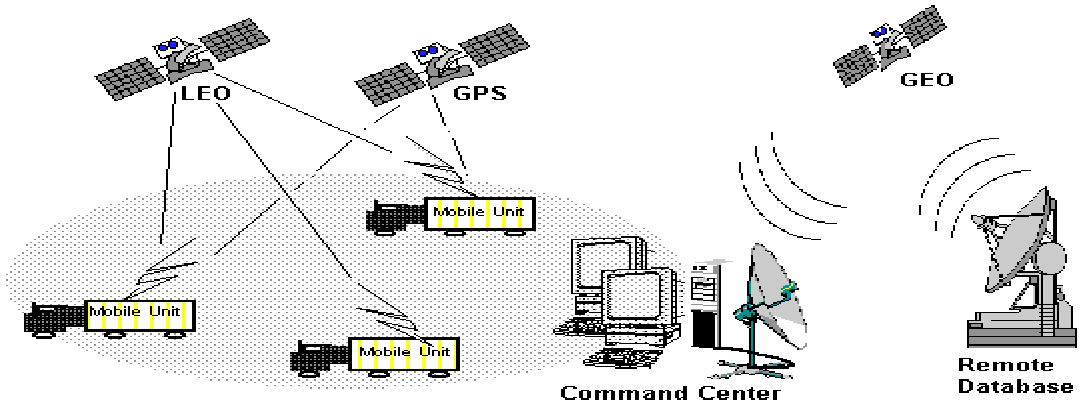

Fig.5. Internetworking for Bushfire Control Application

2. Current Status of Mobile Computing Research

Mobile computing is a rapidly emerging research and development area. Examples of applications which are in a great need of mobile access to computing resources include agriculture, bushfire control, mass disasters, wildlife monitoring, non-stop business activity and many others. There has been a considerable research effort going on around the world with respect to mobility. A number of research groups have been set up all around the world to deal with the various aspects of mobility. Table 1 gives a condensed summary of research groups world-wide, which are looking into various aspects of mobile computing. Figure 6 depicts the geographical distribution of research efforts in mobile computing world-wide.

Table 1. Summary of Ongoing Research Projects in Mobile Computing

GROUP NAME	WHO	WHERE	WHAT
DATAMAN (http://paul.rutgers.edu/~acharya/dataman.html)	T. Imielinski, B.R. Badrinath	Rutgers Uni, NJ, U.S.A.	Mobile Computing Issues with respect to data management Distributed algorithms and services Data Broadcasting Indirect Protocols Data Replication Issues
INFOPAD (http://infopad.eecs.berkeley.edu/)	EECS Dept.	Uni of California, Berkeley,USA	Developing a platform and prototype for providing wireless access to multimedia data.

LITTLE WORK (http://www.citi.umich.edu/mobile.html)	CITI(Centre for Information Technology Integration)	Uni of Michigan, Ann Arbor, USA.	Investigating O.S. requirements for mobile computers.
Mobile Communications Research Group (http://www.ee.surrey.ac.uk/EE/CSER/csertext.html)	EE Dept CSER (centre for satellite engineering)	University of Surrey U.K.	Satellite and personal communications system, Universal mobile telecom systems (UMTS)
Mobile Computing Lab (http://www.cs.columbia.edu/mil)	Dan Duchamp	Columbia University, U.S.A.	All aspects of mobile computing
Mobile SIG (http://www.cl.cam.ac.uk/Research/SRG/msig.html)	University of Cambridge, Computer Laboratory	Cambridge University, UK.	From Media access layer to O.S. support for mobile computing, Distributed file systems
MONET (Mobile Internet) (http://fury.nosc.mil/)	Military	Navy, U.S.A.	Networks, Demonstrate the technologies required to develop a department of defence network that is interoperable with the future public-carrier.
Wireless/ Mobile Networks , Mobile IP (http://netlab.itd.nrl.navy.mil/wireless.html)	Navy	Naval Research Lab , Washington D.C., USA	Networks, Distributed computing, Super Computing, Optical Computing.
Pathfinder, BBN (http://malachite.bbn.com/Departments/DistributedSystems/mobile.html)	Bolt Beranek and Newman INC	U.S.A.	Communication Systems
Walkstation II Project (http://www.it.kth.se/Tslab/ws/ws.html)	KTH, Teleinformatik	Stockholm, Sweden	Mobile Multimedia communication via a global system approach.
WAMIS (http://esto.sysplan.com/ESTO/WAMIS/)	Wireless Adaptive and Mobile Information System	U.S.A.	Information Technology support to mobile systems. Achieveing revolutionary improvements in mobile communicaions technology.
Mobile Computing at the Uni of Washington (http://www.cs.washington.edu/reserach/mobicomp/mobile.html)	CSE dept	U.S.A	Mobisaic - An information system for a mobile and wireless computing environment (http://www.cs.washington.edu/homes/voelken/mobisaic/mobisaic.html) Wit - A system infrastructure for mobile handheld computing (http://snapple.cs.washington.edu:600/wit/wit.html)
MITL (ftp://mitl.research.panasonic.com/pub/tr)	Matsushita Information technology lab	U.S.A.	Mobile issues
SONY CSL (http://www.csl.sony.co.jp)	Sony Corp	Japan, Tokyo	Protocols
Lancaster Univ (ftp://ftp.comp.lancs.ac.uk/pub/mpg/README)	Distributed Multimedia Research Group	U.K.	Multimedia Research Group
Active Badge (http://www.cam-orl.co.uk/ab.html)	Olivetti	U.K.	Locating Persons

Mobile-IP (file://playground.sun.com/pub/mobile-ip/mobile-ip.html)	IETF Mobile IP , Working Group (Internet Enginrng Task Force)	U.S.A.	Adapt protocols to support mobility within the Internet. Mobile-Networks
Shoshin Mobile Computing Project (http://ccnga.uwaterloo.ca.wireless/)	Maths Faculty, Dept Of Comp Science	Univ of Waterloo , Canada	Quality of service and mobility management, traffic modelling, Security etc.
Mobile Computing Research at Purdue University (http://www.cs.purdue.edu/research/cse/mobile)	Dept of Comp Sciences & School of Elec. Engineering	U.S.A	CrossPoint Project Data Management in Mobile Distributed Environments High Speed ATM/Broadband Integrated Networks Mobile Environments in Telemedicine Sciencepad Project
Mobile Computing group at ANU	Comp.Sci	Australian National University	Data caching Mobile TCP/IP performance
Mobile Computing group at Macquarie University	D.Skellern, N.Weste, School of MPCE	Macquarie University, Sydney	Mobile IP Wireless LANs Mobile databases
MOBIDOTS (http://www.ct.monash.edu.au/~mobidick/)	School of Comp.Sci and Software Eng.	Monash Univ, Australia	All aspects of Mobile Computing

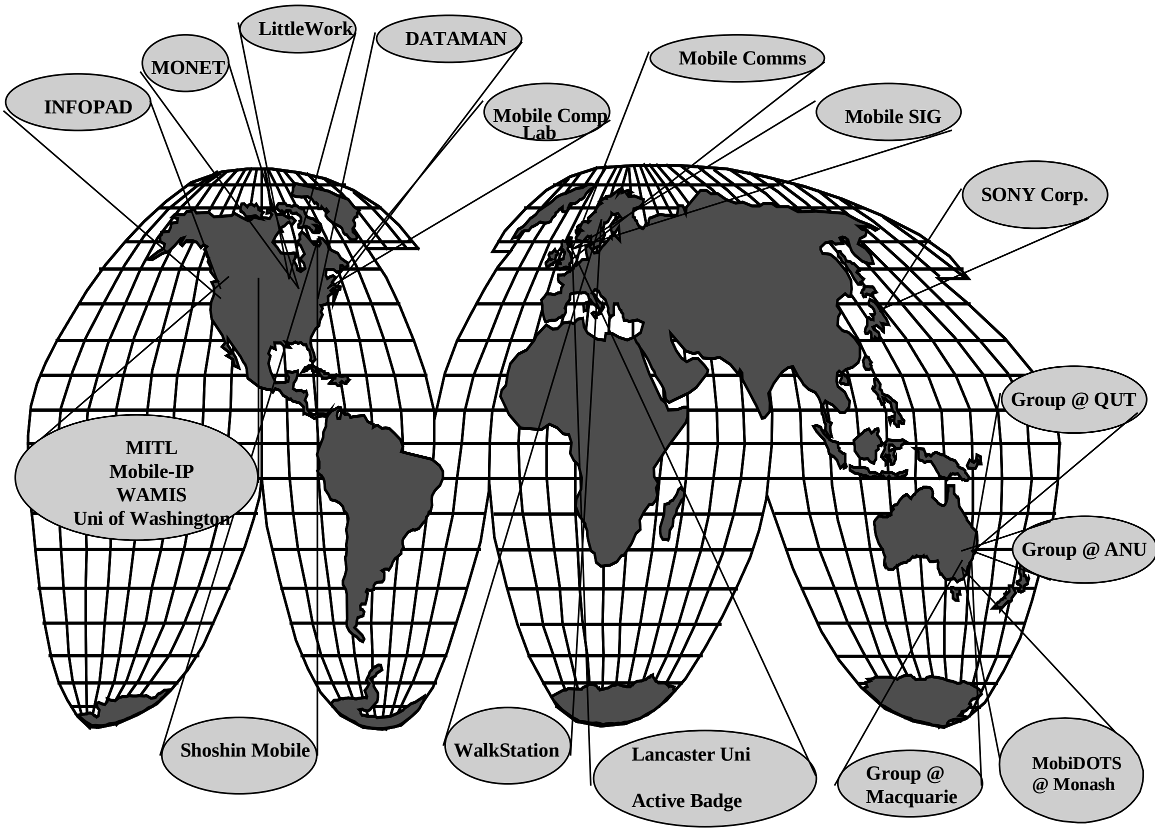

Fig.6. Geographical distribution of research in mobile computing

Mobile computing is enjoying rapidly growing interest and popularity. This is indicated by the fact that almost any international conference in the areas of networking, distributed systems and databases includes now at least one session to discuss mobility issues and impact. Recent conferences that have addressed mobile computing include:

- 16th, 17th, 18th International Conferences on Distributed Computing Systems
- 30th Hawaii International Conference on System Sciences
- 24 International Conference on Very Large Databases (VLDB'98)
- 1st (1995, USA), 2nd (1996, USA), 3rd (1997, Hungary) ACM/IEEE Annual International Conference on Mobile Computing and Networking
- IEEE International Conference on Computer Communications and Networks
- International Conferences on Data Engineering and many others.

Mobile computing research in Australia can be viewed from two perspectives: *engineering* and *computing*. Engineering type research normally involves telecommunications, low-level networking (first three layers of the OSI conceptual model), hardware design and development. Computing direction, on the other hand, addresses systems support for mobile computing, mobility issues in applications, networking and protocols, programming environments for developing mobile information systems. The first two Australian Workshops on Mobile Computing, Databases and Applications provided a forum to discuss various aspects of mobile computing systems. Both these workshops were organised in Melbourne in 1996 and 1997, respectively. The topics covered included: mobile databases, mobile computing and networking, OS and software support for mobile computing, mobile applications and information systems. These workshops attracted mostly Australian academics, researchers and developers. The published papers (Zaslavsky and Srinivasan, 1996; Zaslavsky et al, 1997) can be accessed through the WWW site: http://www.ct.monash.edu.au/~mobidick/. It is planned to continue with this forum and organise an International Conference on Mobile Computing and Distributed Applications (MCDA'98) in 1998. The most up-to-date information can be found on the URL above.

3. Conclusions

This paper provides an overview of concepts, achievements, research issues and challenges in mobile computing. Mobile computing, as it stands today, offers many exciting opportunities. However, the challenges that the research community faces are quite significant. These challenges include mobility aspects, power, frequent connections/disconnections, bandwidth limitations, cost factors, resource scheduling and management, advanced concurrency, replication and synchronisation algorithms. Stronger support and attention from Australian government, academia and business would also be helpful in addressing these challenges by Australian universities and IT industry.

REFERENCES

ALONSO, R. and KORTH, H.F. (1993): Database System Issues in Nomadic Computing, MITL, International Conference on Management of Data, *ACM SIGMOD RECORD*, Vol.22: 388-392.

ASOKAN, N.(1994): Anonymity in Mobile Computing Environment, *IEEE Workshop on Mobile Computing Systems and Applications*, Santa Cruz, CA, USA. Also is available electronically: http://snapple.cs.washington.edu:600/library/mcsa94/asokan.ps

AUSTEL (1996): "Fact Sheet 1", available on http://www.austel.gov.au/info/factsheets/ fs01.htm

BADRINATH, B.R., ACHARYA, A., IMIELINSKI, T. (1993): Impact of Mobility on Distributed Computations, *ACM Operating System Review*, Vol.27, No 2: 15-20.

BELLSOUTH (1996): Mobitex Network Architecture, URL http://www.data-mobile.com/ unshocked/bmd02010.html

BUDDE, P. (1996) "Information Technology Management", *Paul Budde Communication*, Vol. 7: 77-95.

CHRYSANTHIS P.K. (1993): Transaction Processing in Mobile Computing Environment, *IEEE Workshop on Advances in Parallel and Distributed Systems*, Princeton: 77-83.

CHU, W.W., RIBEIRO, B. A. N. AND NGAI,. P.H. (1992): Object Allocation in Distributed Systems with Virtual Replication. *Proceedings of Eighth International Conference on Data Engineering*: 238-245.

CHUNG, H. and CHO, H. (1998): Data caching with incremental update propagation in mobile computing environments, *Australian Computer Journal*, current issue.

DIRECT BROADCAST SATELLITE TECHNOLOGY, (1996) URL: http://http.cs.Berkeley.edu/~randy/Daedalus/BARWAN/DBS.html

DOUGLIS, F., KAASHOEK, F., LI, K., CACERES, R., MARSH, B., TAUBER, J.A. (1994): Storage Alternatives for Mobile Computers, *First Symp. on Operating Systems Design and Implementation*, Monterey, California, USA: 25-37.

ELWAZER, M. and ZASLAVSKY, A. (1997a): Infrastructure support for mobile information systems in Australia, *Proceedings of the Pacific-Asia Conference on Information Systems (PACIS'97)*, Brisbane: 33-43

EL-WAZER, M. and ZASLAVSKY, A. (1997b): Multi-Protocol Gateway Approach To Packet-Switched Wireless Data Networks Interconnection, *3rd Asia-Pacific Conference on Communications*, APCC'97, 7-10 December, Sydney, Australia

ERICSSON Inc (1996), *GSM: the future*. Available at the URL: http://www.ericsson.se/systems/gsm/future.htm

FORMAN, G.H. and ZAHORJAN, J. (1994): The Challenges of Mobile Computing, *IEEE Computer*, Vol 17(4): 38-47.

HEIDEMANN, J.S., PAGE, T.W., GUY, R.G., POPEK, G.J. (1992): Primarily Disconnected Operation: Experience with Ficus, *2nd Workshop on the Management of Replicated Data*: 9-12.

IMIELINSKI, T. and BADRINATH, B.R. (1994): Mobile Wireless Computing: Challenges in Data Management, *Comms. of the ACM*, Vol.37: 18-29.

IMIELINSKI, T. and KORTH, H. (1996): Introduction to mobile computing, In: *Mobile Computing*, T.Imielinski and H.Korth (Eds), Kluwer Academic Publishers: 1-43.

JI, G., and ZASLAVSKY, A. (1996): Future Database Applications on Mobile Platforms, *7th Australasian Conference on Information Systems*, Tasmania: 275-285.

KISTLER, J. and SATYANARAYANAN, M. (1992): Disconnected Operation In The Coda File System, *ACM Transaction On Computer Systems*, Vol.10, No 1: 3-25

LEVESQUE, A. and PAHLAVAN, K. (1996): Wireless Data, *The Mobile Communications Hamdbook*, IEEE PRESS: 553-567.

MARSH, B., DOUGLIS, F, CACERES, R. (1993): System Issues in Mobile Computing, *Matsushita Information Technology Laboratory*, Technical Report MITL-TR-50-93.

MCGLADDERY, W. and CLIFFORD, R. (1993): Survey of current and emerging wireless data networks", *Proceedings of Canadian Conference on Electrical and Computer Engineering*, IEEE Canada: 1000-1003.

MOTOROLA Inc (1995): *Wireless Data Communications: An Overview*. Also available at the URL: http://www.mot.com/wdg/

NZAMA, M., STANSKI, P., ZASLAVSKY, A. (1997): Philosophy of mobile computing in heterogeneous distributed environment: ET effect in computing world", *Proceedings of the 2nd Australian W/S on Mobile Computing, Databases and Applications*, A.Zaslavsky, B.Srinivasan, H.Schmidt (Eds), 27-28 November, 1997, Melbourne, 37-45.

PITOURA, E. and BHARGAVA, B. (1995): Maintaining Consistency of Data in Mobile Distributed Environments, *15th International Conference on Distributed Computing Systems*: 404-413.

RASHEED, and A. ZASLAVSKY, A., (1997): A Transaction Model to Support Disconnected Operation in a Mobile Computing Environment, *4th International Conference on Object-Oriented Information Systems*, OOIS'97, Brisbane, November, 1997

SATYANARAYANAN, M., KISTLER, J.J., MUMMERT, L.B., EBLING, M.R., KUMAR, P., LU, Q. (1993): Experiences with Disconnected Operation in a Mobile Environment, *USENIX*

Symposium on Mobile and Location Independent Computing, Cambridge, Massachusetts, USA: 11-28.

YEO, L.H. and ZASLAVSKY, A. (1994): Submission of Transactions from Mobile Computers in a Cooperative Multidatabase Processing Environment, *Proceedings IEEE/CS 14th International Conference on Distributed Computing Systems*, Poland: 372-379.

ZASLAVSKY, A., FAIZ, M., SRINIVASAN, B., RASHEED, A., LAI, S.J. (1996): Primary Copy Method and its Modifications for Database Replication in Distributed Mobile Computing Environment, *15th Symposium on Reliable Distributed Systems*, Canada, October 23-25: 178-187.

ZASLAVSKY, A. and SRINIVASAN, B. (Eds) (1996): *Mobile Computing, Databases and Applications*, Proceedings of the Australian Workshop, Monash University, Melbourne, 1 Feb.

ZASLAVSKY, A., SRINIVASAN, B., SCHMIDT, H. (Eds) (1997): *Mobile Computing, Databases and Applications*, Proceedings of the 2nd Australian W/S, 27-28 November, 1997, Melbourne, 119 pages.

ZHOU, X.D., ZASLAVSKY, A., RASHEED, A., PRICE, R. (1998): Efficient object-oriented query optimisation in mobile computing environment, *Australian Computer Journal*, (current issue)