

THE UBIQUITOUS WIRELESS COMMUNICATION THAT ROPES MOBILE DEVICES

CH. V. PHANI KRISHNA¹ & S. V. SWAMY²

¹Professor, Department of CSE, Sri Venkateswara College of Engineering and Technology, Thirupachur,
Tiruvallur, Tamil Nadu, India

²Assistant Professor, Department of CSE, Dr. L. Bullayya College of Engineering for Women, Visakhapatnam,
Andhra Pradesh, India

ABSTRACT

The concept of mobile computing may be described as a computing and technological alchemy that enables people to have access to computing and informational resources irrespective of temporal and spatial circumstances through such devices as laptops and Personal Digital Assistants (PDA). However, the act of simply having a portable computing device available for use does not in itself constitute mobile computing. The essence of mobile computing dictates that there is some network access involved in the process, whereby the user gains access to organizational information, or to some remote “home base”. Recently, the term mobile computing has increasingly been used to refer to the Personal Communication Systems (PCSs), which enables people to access private information such as bank accounts and email correspondence.

KEYWORDS: Mobile Computing, Networking Infrastructure, Global Connectivity, Miniaturization, Hybrid Communicator

INTRODUCTION

To provide an indication of the stage at which mobile computing is currently, Scott McNealy, Chairman and Chief Executive Officer of SUN Microsystems, recently in a BBC interview from the World Economic Forum described his vision of the future of mobile computing. In doing so, he indicated the pervasive and progressive direction in which mobile technology is currently heading. In his futuristic vision, McNealy foresees a person being equipped only with a cellular telephone, through which they would have access to the Internet. A further step of simply inserting a credit card or bank card would give that person access to their bank accounts, to credit, enable them to purchase a plane ticket, or even a ticket to an ice-hockey (Jokerit) match [BBC99]. Indeed, part of this vision has already been implemented in Finland by a number of local banks. Customers are able to use a GSM cellular telephone running a WAP [Wireless Application Protocol] application to transact banking operations [Dis99d, MER99].

Although the services mentioned here are grouped under the umbrella of mobile computing, in the last five years there has arisen a noticeable chasm in emerging mobile systems based on the characteristics of the mobile device. Early mobile architectures such as Coda [MES95], and DIANA [ACD94] were developed with an underlying philosophy of attempting to mimic traditional desktop or wired systems. These systems can be categorised as file and workflow systems as their usage is usually aimed at accessing common file and database systems to enact high-powered transactions that are executed on devices which have relatively high storage and processor capacity. In stark contrast, many of the modern mobile architectures, such as WAP are categorised as Personal Communication Systems (PCS) [PRW98]. As the term

“Personal” implies, such systems are concerned with allowing the user access private information such as bank accounts. Moreover, architectures for PCSs are intended for handheld devices whose trademark characteristic is their dearth of computational resources.

In spite of being afforded the prestigious title of “computing paradigm of the future”, the road ahead for mobile computing in providing and guaranteeing a stable and efficient computing solution for a diverse user community is not completely clear. There is still a dire need for further research in a range of sub areas to uncover new techniques that can be applied to the challenges that accrue from the mobile arena. For example, Katz called for deeper research in the ubiquitous wireless network that supports mobile devices [Kat95]. Subsequent to this call being issued, there have been several valuable research initiatives in this direction, such as [FPLV98, PRW98, and SB98].

In a similar manner, the software infrastructural or architecture aspect of mobile computing is also in need of further research and development. This need does not arise only out of the intrinsic relevance and importance of developing efficient software architectures, but in addition, this need for research is further heightened by the recent and ongoing technological development within the mobile domain. The appearance of such developments as WAP, and UMTS has injected additional urgency in the need for new approaches and techniques for confronting data management, as well as a host of other important issues. In light of the above, the primary importance of this thesis lies in its intended focus on structural concepts that are independent of the logic and semantics of specific applications, but which can be moulded to accommodate implementation of current and emerging technologies. Ultimately, the realisation of these generalised software approaches and techniques will assist in strengthening the mobile computing paradigm. Furthermore, the increasing demand for streaming data applications, such as multimedia applications, on the mobile platform will inevitably result in the building of heavier and more resource consuming applications. In the author’s opinion this development inevitably introduces questions pertaining to levels of guaranteed service that must be provided by mobile applications. In addition, object oriented technologies, such as Java are generating a significant impact on the design of mobile applications. All of these developments must be reflected in the software architecture of future mobile applications. Therefore, developing an understanding and appreciation of software architectures that are pertinent to mobile computing, and of the valuable role that they play in the entire mobile computing schema, will undoubtedly contribute towards the development of richer applications. Such architectures are an essential medium for confronting the data management complexities in the highly distributed environment, as well as accommodating the influx of new technologies.

PREVIOUS RESEARCH

Research on mobile computing originating from both academia and industry has witnessed a dramatic increase in the last five years. A testament to this rise in research interest has been the establishment of a number of dedicated journals, such as MONET (Mobile Networking and applications, Mobile Networks), as well as Andrew Seybold's Mobile Communications Outlook. There is also a tremendous wealth of information covering this domain that is located on the WWW. For example, Aline Baggio's [AB99] provides an excellent repository of bibliographic information covering all aspects of mobile computing.

A common feature of all published research findings on mobile computing is the general acceptance of the multitude of environmental and device ailments that afflict the domain. These ailments include insufficient bandwidth, predictable and unpredictable disconnection, and small resource deficient mobile devices [AloK93, FoZ94, Kat95, PiSa97, SaP98, Bahl98].

In fact, it has been the drive to alleviate these ailments that has been the primary focus of academic and industrial research over the last five years. For example, [BaP97] on supporting mobile clients in accessing client server databases, and [Ray94] on the issues involved in implementing mobile computing over the ATM network. Despite the valuable contributions that these and other similar research initiatives have presented, there has nevertheless been a lack of attention given to the development of software architectures for mobile computing that are infrastructural in character. The value of such structural and high-level approaches is evident in the independence that such architectures retain of the different applications and their semantics, while simultaneously permitting customization of individual applications. In this vein, [SaP98] has presented an insightful taxonomy of mobile system architectures. These computational models may be grouped into mobile client-server models and mobile agent models. The client server models may be further sub-divided into the client-agent-server model, the client- intercept-server model, and the peer-to-peer model. These generic formulations largely comprise the building blocks, or meta-meta level resources that more concrete software architectures are built on.

Although, as previously stated, research on providing software architectures has not been the subject of extensive research efforts, there have been some highly commendable research endeavors in this domain, which have resulted in imaginative and workable architectures. Such endeavors have included the previously mentioned DIANA [ACD94], and the L²IMBO architectures [DFWB98].

The QoS issues that are raised in the L²IMBO approach represent features that are becoming increasingly common place in mobile computing. For example, Kassler and Schulthess adopted the approach of introducing QoS at both the application and the transport levels in order to handle short-term fluctuations on the wireless link [KaS99]. It is certain that QoS issues will become more important as continuous data, and all the complications that are introduced by routing such data over the wireless medium, becomes more pronounced in the mobile computing domain. QoS guarantees are essential requirements for continuous data due to the time critical nature of the data exchange. Such QoS guarantees will include minimum bandwidth reservation, as well as power and processing thresholds, and must be given due consideration in protocol composition of the mobile system.

In addition, there are several on-going industry-led standardisation initiatives that will play a crucial role in the future of mobile computing, and which will impact significantly on software architectural issues. Chief among these standardization initiatives is the WAP Forum's WAP stack specifications, and development environment that consists of WMLScript, and WML. In addition, the emergence of Symbian's EPOC OS for ROM based computing will provide the stability and support that is needed to drive the mission critical mobile applications of the future.

All of the above work have some bearing on issues pertaining to the type of structural concepts that will be discussed in this thesis. They serve as a starting point in task of developing and ultimately proposing a novel software architecture that can be used in building robust and reliable applications. Another interesting feature pertaining to existing software architectures for the mobile domain is that the great majority of them were developed over five years ago. Given the pace of technological advancement in the mobile computing world, this time period represents an eternity. Consequently, architectures such as DIANA and L²IMBO were developed without thought afforded to such crucial developments as UMTS, or the now pervasive stampede to funnel multimedia over the wireless link. This work of this thesis will be among the pioneering efforts that will explore these new topics.

RESEARCH METHODOLOGY

The research strategy that will be applied in deriving answers to the research questions posed in this thesis is two-fold in character. In that, there are both theoretical and constructive aspects that are involved. To this end, a number of research methodologies were considered in an attempt to identify a research methodology that would afford maximum effect in the tasks of developing and presenting the argued case in the richest and clearest possible manner. For example, Jenkins [Jen85] has developed an insightful taxonomy of research methodologies that can be applied to the IS (Information Systems) field. However, from the standpoint of this thesis, Jenkins' research schema is seen as being too restrictive, as it compartmentalizes and isolates the various research strategies.

Therefore, a more pragmatic and cohesive methodological approach will be adopted in deriving and elucidating the aims of this thesis. To this end, an extensive literature review will provide the basis for developing an appreciation of the prevailing issues within the mobile computing domain. This review will encompass literature gathered from traditional printed sources, as in industry and academic Journals, as well as information that is drawn from the WWW. The knowledge that is gleaned from these sources will then be used to develop and construct a theoretical software architecture.

LIMITATIONS OF RESEARCH

As was stated earlier in the research methodology section, this thesis will enact a version of the systems development methodology. The intent of the research is not to focus on implementation issues such as the coding of any specific application. Rather the intent lies in developing a software methodology that defines how the structure of a particular set of mobile applications should evolve. To this end, this methodology will focus on issues such as how these applications handle their data, how and where the computing is actually performed, and the how results are made available to the mobile device.

In terms of proposing a practical software architecture the emphasis will be on the first two stages of the systems development systems stage of the research methodology that is being applied. Those two stages are constructing the conceptual framework, and developing the system architecture. In this vein, the motivation behind the proposed architecture will exhibit a predisposition towards emerging technologies such as WAP and UMTS, and to the processing of multimedia data on the wireless platform.

CONTRIBUTION TO KNOWLEDGE

The results of this will research will extend the knowledge base of the mobile computing domain in its treatment of issues such as computing models, and data and resource management strategies. Furthermore, these issues will be discussed in light of emerging technologies and standards such as WAP, and UMTS. To this extent, this research can be of value to telecommunication industry players such as Nokia, who are heavily engaged in developing cutting edge mobile computing solutions.

Moreover, the results of this research, particularly the proposed software architecture, can also be utilized by software developers of mobile computing solutions. A clear and definitive taxonomy of viable software architectures that are pertinent to current as well as emerging mobile computing trends can help these organizations to more aptly design and build applications that are mindful of both technological trends as well as the restrictions of the mobile domain.

CHALLENGES OF MOBILE COMPUTING

As can be witnessed from Table 1.1, the resource levels of laptop computers and handheld are quite divergent. When this resource divergence is accepted, the next step is to appreciate that such divergence does influence the software architectural makeup that underlies the applications that run on these devices

The second implement, which meshes with the actual hardware device to bring the concept of mobile computing into existence, is the wireless medium. However, unlike the situation that prevails in the device domain where many of the challenges are mitigated across the scope of different devices, the challenges relating to the wireless medium remain implacable across the entire spectrum of mobile computing. In addition, new challenges are being introduced into the mobile computing alchemy because of the increasing requirement by mobile applications to process multimedia data.

This chapter examines the technical challenges that are faced by mobile computing, which arise from the characteristics of the mobile hosts and the wireless medium, as well as the newer challenges that are bred as a result of the processing of multimedia data. An understanding of the issues to be discussed here is critical, as these are the issues that must be confronted in the design of mobile systems. The goal of the designer of such systems is to discern a software architecture that exhibits strengths in the areas that are crucial to the implicit philosophy behind the application, and which are reflected in the explicit requirements. The challenges facing mobile computing can be broadly categorised into communication, mobility, and portability associated challenges.

WIRELESS COMMUNICATION

Network failure is one of the great concerns that afflict mobile communication. This section presents the most perplexing of these communication problems, which include network disconnection, and bandwidth starvation and variability.

Disconnection

Disconnection can occur in the mobile environment due to network failure, or as a result of the (communication) c-autonomous nature of the mobile device. A computer system is c-autonomous within the environment if it cannot be compelled by external systems to start a communication session, continuing that session, or be prohibited from initiating communication with an external system. In addition, mobile devices must contend with lower bandwidth, higher error rates, and more spurious disconnections that are caused by the frailty of the wireless link. This environment forces the designer of mobile systems to grapple with the vexing dilemma of spending resources on the network trying to prevent disconnections, or alternatively on spending those resources on enabling the system to cope with disconnections more gracefully by working around them where possible.

It may be advanced that the more autonomous a mobile device is, the better it is able to withstand the trauma of disconnection. For example, applications that are built on the DIANA architecture reduce communications by running entirely locally on the mobile device when the chances of network disconnection are great [ACD94]. Thus, in environments that are beset by frequent disconnections it is highly advantageous that the mobile device is able to operate as a stand-alone device as opposed to as a mobile terminal. However, this presupposes some minimum levels of resources that the mobile device must be equipped with, which cannot necessarily be met by many of the handheld devices that are the focus of this work.

In some cases, round-trip latency and short disconnections can be disguised by operating asynchronously [FoZ94]. This operation mode is counter to the synchronous remote procedure call paradigm in which the client waits for a reply after each request. In contrast, in asynchronous operation a client can send multiple requests before asking for any acknowledgement. In addition, perfecting and lazy write-back are also viable strategies that have been developed to decouple the act of communicating communication from the actual time a program consumes or produces data. Such strategies permit the mobile device to make progress during network disconnections. In terms of multimedia data, this would imply that it would be more advantageous for the mobile terminal to operate in asynchronous mode, since the delays involved in invoking the reliability standards in synchronous mode could be harmful to the time sensitive nature of such data during transmission.

In situations where the mobile terminal is accessing a remote database the problem of disconnection becomes more heightened. This is because the traditional commit protocols are not amenable to the unpredictable disconnections that are typical of the mobile environment. The simplest, and a widely used, commit protocol is the basic two-phase commit protocol (B2PC), and its derivative, the optimistic two-phase commit protocol (O2PC) [LeC98]. There are two main deficiencies involved in using these commit protocols in the mobile environment. In the first place, if the mobile client commences a transaction, and is shortly after disconnected, the locks that were placed on data items at sites involved in that transaction will cause such data items to be inaccessible to other users. Secondly, communication failures can violate the transaction atomicity rule. Thus, in B2PC, the adopted solution is to abort the entire transaction. A high number of such aborts can have a debilitating impact on a system's throughput and fault tolerance [BoDe9]

The DocMan document management system [BuB96] provides a good example of a system that handles network disconnection. It achieves this by allowing the user to download important information onto the mobile device. During those periods when the user is disconnected from the network, work is allowed to progress, and any changes that are made to documents can be reconciled with the master repository when network connection is re-established. Here the problems of blocking and site inaccessibility do not arise. This strategy is justified on the basis that in distributed systems less than 1% of all writes are followed by a write by a different user [KiS92]. However, this statistic has yet to be verified in a mobile environment.

Low Bandwidth

As with disconnection, mobile computing designs are required to be more concerned about bandwidth consumption and constraints than designs for stationary computing because wireless networks deliver considerably lower bandwidth than stationary networks. Although the situation is changing rapidly – even as this work is in progress Sonera is implementing HSCSD (High Speed Circuit Switched Data) – typical wireless networks achieve only a meagre 9 – 14 Kbps for application data [Dis99c]. Moreover, since the available bandwidth is divided among the users sharing a cell, the effective bandwidth per user is even lower. Indeed, for radio transmissions the error rate is so high that the effective bandwidth is limited to less than 10 Kbps [PiSa97]. Accordingly, in mobile computing bandwidth is very much a premium resource and this is further exacerbated by the financial costs of the little bandwidth that is available.

Bandwidth Variability

Wireless networks vary considerably on the degree of bandwidth and reliability that they provide to mobile devices [Kat95]. As with the problem of bandwidth starvation, this variability in the amount of available bandwidth is

caused mainly by the number of users who are simultaneously sharing the resources of a cell, as well as the interaction of the radio signal with the environment. The fundamental difference between bandwidth starvation and bandwidth variability is that the former is an inherent and largely static characteristic of the mobile domain, whereas the latter is characterized by its unpredictable occurrence.

To effectively deal with the constraint of bandwidth variation, the situation calls for mobile systems to be adaptive in their outlook. The term adaptively entails that systems should be able to detect and to adapt to situations in which there is either an abundance of, or a lack of bandwidth, while maintaining delay sensitive communications. In this vein, [WeBa97] have posited the view that applications should be totally aware of changes in the environment, and should assume sole responsibility for implementing appropriate actions. An opposing extreme of that position entails the system bearing the complete burden of detecting and counteracting all environment changes [NoS95]. However, the most effective adaptation strategy is one that strikes a balance between these polarities. Such a compromise approach would enable individual applications to determine how best to adapt, but at the same time allow the system to centrally monitor the resources, and ensure that they are used effectively [Nos95].

Table provides some insightful information on how mobile communication mediums compare with a number of other communication technologies. This table serves to portray the depth of the bandwidth starvation and variability problems that afflict mobile connections. It also serves to indicate the impact that 3rd generation technologies will have on the future of wireless communications.

Table 1: Connection Mediums

Network Type	Bandwidth	Coverage	Availability	Usage Cost
Fixed LAN [Sta94]	10 M – Gbps	Local	High	Low
Satellite [NoH95]	9.6 Kbps- 3 Mbps	National	Medium	High
GSM				
SMS [Dis99c]	9.6 Bps	International	High	Low
Circuit Switched Data [Dis99c]	14.4 Kbps	Regional	Medium	High
HSCDS [Pere98, Dis99c]	57.6 Kbps	Regional	High	Medium
GPRS [Egg99, Dis99c]	115 Kbps	Regional	High	Low
EEDGE [Dis99c, Pere98]	48 Kbps	Regional	Medium	Low
UMTS				
UMTS (WCDMA) [Pere98]	144 Kbps – 2 Mbps	Global	High	Low
Wireless Broadband				
Wireless LANs [Pere98]	1 – 25 Mbps	N/A	N/A	N/A
Wireless ATM [Pere98]	25 - 155 Mbps	N/A	N/A	N/A

Security Risks

Security risks are higher in mobile computing precisely because it is easier to connect to the wireless link. Therefore, the security of wireless communication can be compromised much more easily than wired communication, and this risk is especially higher if the transmission range stretches over a wide area [Kat95]. This situation heightens the requirement for mobile computing designs to include suitable security measures that would detect and prevent the system from being violated by unauthorised parties.

Secure communication over insecure wireless channels is one of the main benefits of digital networks. As a result of the binary nature of the data, strategies such as encryption and interleaving can be more readily and effectively introduced into the communication framework, than in an analogue system. The security advantages of digital network is borne out by the reality that up until the present point in time, which represents a period in excess of 10 years, there have been no reported incidents of the GSM encryption algorithms having been violated.

Mobile Multimedia Communication

The term multimedia refers to two or more media – audio, video, or text - that are presented together. As Pekkola noted, all multimedia products can be grouped into one of the following three categories [Pek98]:

- **Multimedia Conferencing:** These systems allow geographically separated users to be able to communicate via audio and video. Some applications also offer means of data collaboration tools such as a shared whiteboards or other common workspaces. Examples of these systems include collaborative virtual environments (CVEs), multi-point videoconference systems and shared workspace applications such as BSCW and the AltaVista Forum.
- **Media-on-Demand:** This is used to solicit some kind of media from the server. For example, if users want to see television-news on their computer, they order news from the media server, which broadcasts the news back to them.
- **Multimedia Mail:** This is like ordinary mail, but it is also consists of some elements of multimedia files which come as attachments. These multimedia files include, for example, MPEG or AVI video or even interactive multimedia applications such as interactive WWW pages. In addition, ordinary files that are transferred with multimedia files through the network can be bundled into this class.

As noted earlier, mobile communication suffers from severe bandwidth restrictions and higher than average bit error rates. It follows from these technical limitations that the response or round trip time is a critical problem that degrades the performance of mobile applications [SaP98]. In terms of multimedia data transmission, as Table 2.2 shows, these limitations serve to aggravate the impact that real-time data transmission constraints impose on network connections.

CONCLUSIONS

This paper provided an overview of the challenges that are endemic of mobile computing. The remainder of this thesis will focus on how solutions to those challenges are reflected in the design of mobile systems. To this extent the tone will assume a more practical outlook as reference will be made to a number of currently existing architectures, the purpose of which is to afford the reader an opportunity to grasp a practical view of the issues that are being discussed. The specific architectures that will serve this purpose are WAP (Wireless Application Protocol), In addition, the JMAM (Jyvässkylä Mobile Architecture for Multimedia Data) architecture, which is one of the primary aims of the thesis, will

enter onto the stage for the first time.

REFERENCES

1. Ahmad T. et al, The DIANA Approach to Mobile Computing in MOBIDATA (Interactive Journal of Mobile Computing) Vol. 1, No. 1, 1994.
2. Swarup A. et al, Balancing Push and Pull for Data Broadcast, SIGMOD Record (ACM Special Interest Group on Management of Data), 26(2), 1997 pp. 183 – 194. 1997
3. Agrawal, P., Energy Efficient Protocols for Wireless Systems in 9th IEEE International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC), 1998, Vol. 2, pp 564 – 569.
4. Alonso R. and Korth H. F., Database System Issues in Nomadic Computing, SIGMON Record, 1998, pp 388 – 392.
5. Bahl P., Supporting Digital Video in a managed Wireless Network, IEEE Communications Magazine, Vol. 36 6, June 1998, pp 94 – 102.
6. British Broadcasting Corporation, Interview with Scott McNealy from the World Economic Forum in Geneva, Switzerland, January 1999.
7. A New Epoch Dawns for Wireless SYMBIAN in DISCOVERY – Nokia's Telecommunications Magazine, Vol. 49, June 1999, pp 24 – 27.
8. Eggleter S., the Mobile Business in Telecommunications, August 1999, Pp 88 – 89.
9. kassler A., Schulthess P., An End-to-End Quality of service Management Architecture for Wireless ATM Networks in Proceedings of the 32nd Hawaii international Conference on System sciences, 1999 pp 10.
10. Papastavrou S., Samaras G., Pitoura E., Mobile Agents for WWW Distributed Database Access, Proceedings of the 15th International Conference on Data Engineering, March 1999, (ICDE 99), pp 228 - 237.

