Protocol Architecture for Universal Personal Computing

Yalun Li, *Member, IEEE*, and Victor C. M. Leung, *Member, IEEE*

***Abstract*— This paper presents a new paradigm for network computing over the Internet called universal personal computing, where mobile users can access computing resources, network services, and personalized computing environments anywhere using any available terminals. The user and system requirements are defined, and an agent-based protocol architectural required to manage different mobile objects, i.e., users and terminals, in this computing environment is presented. Modifications of connection setup procedures between user application programs to enable addressing based on a global user identity are considered. The use of personal agents to facilitate interworking and management functions is proposed.**

***Index Terms*—Internet, mobile computing, personal communications.**

I. Introduction

THE Internet is emerging as the major global network for data communications, providing the means to interconnect a large number of heterogeneous computer networks. A large group of applications such as e-mail, Telnet, FTP, Gopher, World Wide Web (WWW), etc., have been developed to take advantage of Internet communications, and their increasing popularity has led to an exponential growth in the number of hosts and users connected to the Internet. The future version of the Internet protocol (IP) will provide more IP addresses and support real-time traffic, flexible congestion control schemes, and security features [1].

The rapid expansion of wireless communications technologies is making it possible for mobile users with portable computers or personal digital assistants (PDA's) to access information anywhere and at any time. The resulting computing environment, which is often called *mobile* or *nomadic computing* [2], no longer requires a user to maintain a fixed position in the network, and enables almost unrestricted user mobility. With increasing processing power available in PDA's and notebook computers, and technical developments in the fields of radio channel access method, traffic control, and distributed processing, it is inevitable that mobile computing should be integrated with the Internet so that mobile users can continue to access Internet applications while roaming. However, the Internet was originally designed for interconnecting networks in which hosts have static associations with specific networks,

Manuscript received September 12, 1996; revised May 2, 1997. This work was supported by a grant from Motorola Wireless Data Group, Richmond, BC., Canada. This paper was presented in part at IEEE ICUPC'96, Boston, MA.

The authors are with the Department of Electrical Engineering, University of British Columbia, Vancouver, B.C., V6T 1Z4 Canada.

Publisher Item Identifier S 0733-8716(97)06096-4.

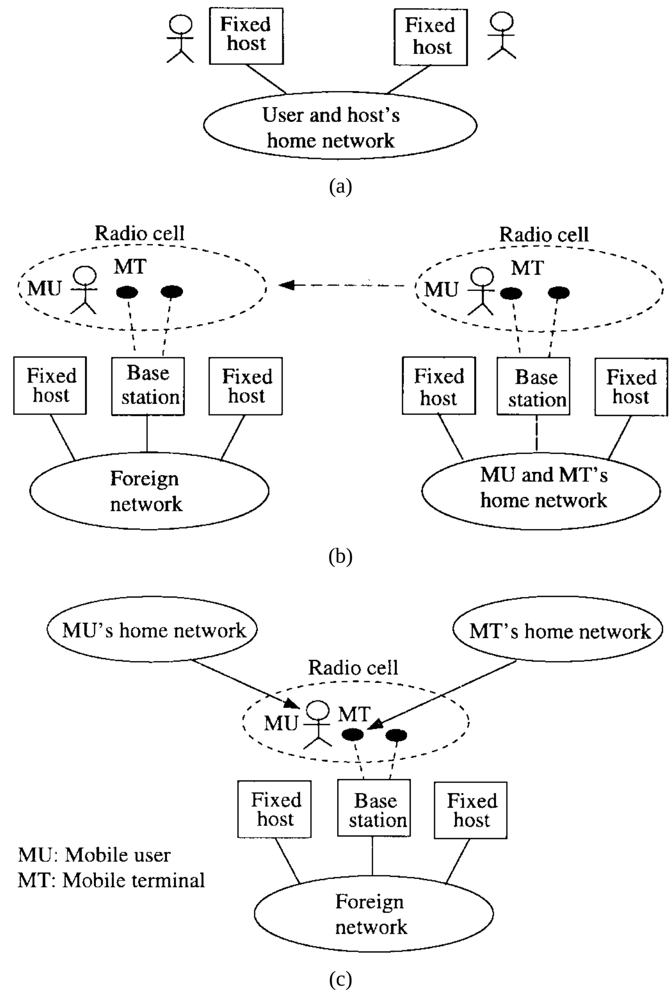


Fig. 1. Expanding dimensions of mobility in computing systems.

as shown in Fig. 1(a), where both users and hosts remain at the home network domain. The IP address of each host serves as the host's unique network identity (ID), and provides the location information for routing purposes. The network part of the recipient host's IP address in each packet is employed to route the packet to the correct network and further to the recipient host. Any host migration to a different network will result in a loss of network service unless the host acquires a new ID associated with the new network. A user who has formally registered and defined her service profile in a network domain enjoys personal mobility only within the home domain that permits the user to access network services from any fixed host in her home network.

0733–8716/97$10.00 © 1997 IEEE

Network computing under the paradigm of nomadic computing is currently centered around the support of host or terminal mobility. The Internet host mobility protocol or *mobile IP* [3], which enhances IP by allowing transparent routing of IP datagrams to mobile nodes over the Internet, has been studied extensively in the literature [4], [5]. The current IETF standard and relating proposals each retain the home IP address of a mobile terminal (MT) for use in identifying it at the network level, but also associate a second IP address (commonly referred to as the care-of address) with the MT to indicate the MT's current location. Incoming datagrams are first routed to the MT's home network, where a home agent redirects the datagrams to a foreign agent at the visited network by tunneling to the care-of address, and the foreign agent in turn delivers the datagrams to the MT.

Mobile IP does not take into account the associations between mobile users (MU's) and MT's. For all practical purposes, these associations can be assumed to be fixed, i.e., each MT can be assumed to be owned by a single MU, which moves together with the MT when roaming away from the home network, as illustrated in Fig. 1(b). The MU continues to enjoy network resources such as personal file space in the home network since the MT is still logically attached to the home network while roaming. However, in reality, not all computer users own notebook computers or PDA's, but they still have the needs to access Internet services from any available (stationary or mobile) terminal while traveling due to business or private needs. With the emergence of low-cost *network computers* (NC's) [6] being promoted by companies such as Oracle and SUN, it is foreseeable that travelers will be able to access network computing terminals at airports, in hotels, or even in bars and cafes. The provision of advanced personalized network applications and services to travelers to enable nomadic computing that is independent of the terminal location and identity could represent the next major advancement in network computing, and is the subject of the present paper.

The above discussion introduces an additional degree of mobility over terminal mobility that has not been considered in mobile IP. This is referred as personal mobility in the context of personal communication services (PCS) [7]–[9]. Personal mobility enables each user to access a user-defined set of subscribed services, and to initiate and receive any type of communications on the basis of a personal, network-transparent number at any terminal, fixed or mobile, irrespective of geographical location, limited only by terminal and network capabilities and restrictions imposed by the network provider. A global standard for such services, *universal personal telecommunication* (UPT), is being studied by the CCITT [8]. The implementations of UPT services are based on the intelligent network (IN). UPT's network architecture, functional model, and call setup procedure have been discussed extensively in the literature [7]–[11], and distributed control to provide PCS in a broad-band environment has been proposed [12]. These discussions focus on extending point-to-point connection-oriented service provided by the public switched telephone network to mobile users. However, parallel developments in extending network computing services over the Internet to mobile users have been lacking.

We introduce a new paradigm for network computing called *universal personal computing* (UPC). It refers to a computing environment which enables an MU to access computing resources, network services, and personal applications, data files, and environmental configurations through *any terminal*, be it stationary or mobile, anywhere on the Internet, as illustrated in Fig. 1(c). To support such personal freedom in computing, a thorough review of the current implementation of the Internet protocol suite is needed, and new architectures and protocols should be developed to enhance the existing protocol suite to support MU's roaming globally without concerns to each user's location and the terminal used. These developments should be evolutionary in nature so that they are backward compatible with the substantial infrastructure invested in the current Internet. Consequently, we propose an agent-based network architecture that is a compatible natural extension of the network architecture supporting mobile IP [3]. In addition, this architecture is also consistent with the principles of the emerging telecommunications information networking architecture (TINA) [13], and is amenable to implementation using the common object request broker architecture (CORBA) [14], [15]. This facilitates convergence of UPC with UPT in the future.

This paper presents a framework of some of the developments necessary to enable UPC. In the rest of the paper, Section II establishes a number of objectives for the UPC system from the points of view of the users and the system. To manage globally roaming MU's and MT's, logical identifications of users and terminals as mobile objects are discussed. Section III presents the enabling agent-based network architecture with mobile object control entities, registration procedures for mobile users and terminals, and some alternative methods to establish TCP connections between users employing global logical identifiers. Section IV proposes the use of personal agents to facilitate UPC services particularly in a heterogeneous wireless networking environment, in which efficient resource utilization, mobility support, and interworking are important considerations. Section V concludes the paper.

II. Objectives of Universal Personal Computing

A. User Requirements

We assume that each mobile user is associated with a home network which the user normally accesses for network computing services. From mobile users' point of view, a UPC system should satisfy the following user requirements.

- A mobile user can access Internet services from anywhere using any available mobile or fixed terminal.
- Communication with any other person over the Internet is accomplished by specifying the recipient's logical user identifier (LUI) without concern about the other person's real location and the terminal the person is using. The LUI uniquely identifies the user and has global significance.
- A mobile user may be vouched or authenticated by the home network to gain access to services provided by

a visited network, including the ability to communicate with other users or computers anywhere over the Internet, to access services via the home network, and services available locally at the visited network such as printing a document, travel information, reservation services, etc.

- A mobile user has a service profile at her home network which specifies services the user may access through the home network while she is at home or roaming, as well as default parameters for these services and for the mobile user's computing environment in general. Special services to support roaming, e.g., e-mail filtering or digesting, may also be specified.
- A mobile user can access a personalized computing service environment such as personal file system, custom shells, etc., while roaming, so she can access services in an environment that she is accustomed to, as if she was home.
- Services provided to a mobile user are transparent to user movements, i.e., handoffs between radio cells should not disrupt service access or communications in such a way that is noticeable by the user.

B. System Requirements

To meet the above user requirements, the UPC system should satisfy the following requirements.

- Terminals and users should be considered as mobile objects to be managed separately.
- Terminal registration and user registration are required to be coordinated between the network visited by a user and the user's home network, as well as the home network of the terminal in use should it be different from the other two networks. Authentication is needed in the registration process for security reasons, and to facilitate service charging and billing at the respective networks.
- Network services and communications received by a mobile user should be transparent with respect to the logical associations or bindings among user, terminal, and location. These bindings are managed relative to user and terminal mobility.
- Facilities are required to manage the user's local network resources and personal resources at her home network, and to transparently provide a personalized computing environment to the user.
- Adaptation capabilities are provided to harmonize the differences in operating environments, input, output, display, storage formats, etc.
- Certain intelligent agents may be incorporated to optimize the network operation.

The first three system requirements are addressed in the next section by specifying the mobile objects, proposing an agent-based architecture to manage these objects at their respective home or visited networks, and considering enhancements to TCP connection management to transparently set up connections between users' current terminals in use. For the duration of a UPC service access, changing locations of user and terminal will be transparently processed by handoff and location updating mechanisms, so that existing connections and corresponding application programs will not be affected. To resolve challenging problems associated with the last three system requirements, a combination of new technologies in the fields of distributed processing, object-oriented design, and intelligent services should be employed. Section IV discusses using personal agents to satisfy these system requirements, and to provide service mobility to mobile users.

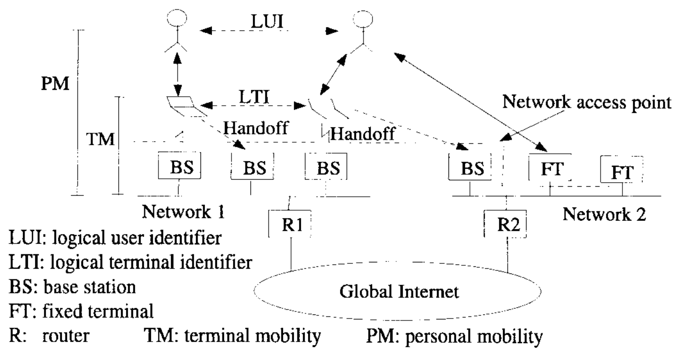


Fig. 2. Dynamic associations of mobile objects.

III. Mobile Object Management

A. Mobile Object Identifications

To enable personal mobility, an LUI is needed to uniquely and directly identify a user irrespective of the terminal used. The LUI is used as the basis for sending and receiving messages and for charging a user for services. A mobile user could access network computing and communication services using a fixed or mobile terminal attached to a UPC service supporting network by specifying her LUI and providing additional authentication information. In this paper, we consider mainly the most general situation involving MU's employing MT's belonging to different home networks. Each MT is identified by a *logical terminal identifier* (LTI) to enable management of the terminal's communications in the presence of terminal mobility, i.e., possible changes in network attachment points in the midst of data transfers. Fig. 2 illustrates the situation of dynamic associations among mobile objects and networks. An MT could hand off to a neighboring base station within the same network, or hand off to a base station that belongs to another network.

Although the LUI could be in many forms such as a personal number or even the photo of a user, a user's LUI should be a unique name on the global network, independent of the current location, consistent with the current naming method in Internet applications, and provide a ready association with the user's home network. Table I lists a few examples of naming a person or a server in Internet applications. A simple LUI could be composed of the *user ID* of a user at her home network concatenated with the user's *home domain address*, such as yalun@ee.ubc.ca. This obviates the need for global user-name servers, and is compatible with the current Internet architecture.

Since IP was designed before host mobility was entertained, an IP address both identifies a host and specifies its location for the purpose of packet routing. In the mobile IP, a host's

TABLE I
EXAMPLES OF ADDRESSES USED IN DIFFERENT APPLICATIONS

Applications	Server or personal addresses
Teleconference	yalun@ee.ubc.ca
e-mail	yalun@ee.ubc.ca
www	http://www.ee.ubc.ca/~yalun
ftp	ee.ubc.ca

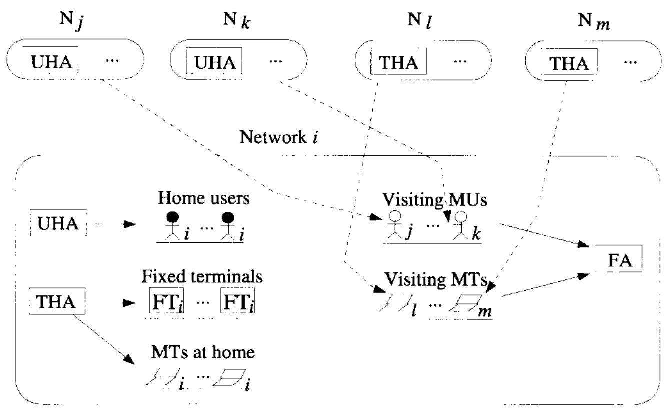


Fig. 3. Multinetwork architecture for mobile objects.

home IP address is used as its LTI, and its physical location while roaming is given by a forwarding (care-of) address. A tunneling mechanism is designed to redirect packets addressed to an LTI to the terminal's current location. Thus host mobility is transparent to the transport layer, and the protocol extension is backward compatible with the existing Internet protocol suite. For compatibility reasons, UPC should continue to employ these methods of terminal and location identification.

B. Network Architecture

Fig. 3 presents a conceptual multinetwork architecture to support personal and terminal mobility in UPC. Each network is a separate administrative domain which has the necessary facilities called agents to manage visiting users and terminals, as well as the users and terminals calling the network home. Since MU's are independent of the MT's, user-related information and terminal-related information must be contained in different logical databases in these agents. Their control functions should also be separated as follows.

1) Terminal's Home Agent (THA): Each network has a THA which maintains a home list or database identifying all MT's that the network is configured to serve. An MT is a normal terminal with additional software that allows it to move through the Internet in a manner transparent to software above the network routing layer within the host. Terminal-related information such as terminal identity (LTI), terminal profile, terminal authentication key, and current terminal location (care-of address) are stored in the THA database. The terminal profile identifies the terminal's capabilities such as resident operating system, file format, graphical user interface, display mode and resolution, etc., so that the serving network can coordinate with the terminal's software to create the computing environment required by the user, or a close approximation of such an environment.

TABLE II
TYPES OF MOBILE USER INFORMATION

Service Profile Information	Location Information
• Authentication information • Service access information • Access domain information • Usage accounting and billing information	• Location binding • Type of terminal • Type of services

2) User's Home Agent (UHA): Each MU belongs to a home network, the administrative domain where the user is registered on a long-term basis. The UHA in each network maintains a list of users associated with this network as their home domain. The UHA includes a user database that records the LUI's of all users calling the network home, and the pertinent information for each user, including the service profile and location information. Table II lists the possible items in a user's home database entry. In particular, the service profile contains the user's authentication information, defines the entitlement of the user to access services and resources, and includes the user's service preferences and environmental defaults. Service customizing for an MU could also be defined in the service profile by specifying a number of selections to deal with different situations, i.e., when the user is not at home and/or is accessing a different kind of terminals; e.g., only important e-mails identified from the topic or source address fields will be forwarded to the user when she is accessing a wireless network via an MT. This database information is accessed whenever the MU roams into a new network and accesses services and resources. Database management operations on the service profile can be performed only by authorized network administrators or by the user herself. The location information provides the mobility binding among the LUI, the LTI of the terminal the user currently uses, and the location of the terminal (e.g., in the form of a care-of IP address). This information is updated whenever the user moves or changes her association with a terminal.

3) Foreign Agent (FA): As an MU migrates over the Internet, she needs to access network computing resources and services from different networks connected to the Internet. To enable service access, the MU must first establish temporary residence in the foreign domain. Each network serving mobile users has a foreign agent (FA) which enables the user and her terminal to be temporarily associated with the network. Each foreign agent maintains a visitor list identifying all MU's currently visiting this network. The entry for each MU mirrors the user's mobility binding stored in the user's UHA. The FA may serve as a default router terminating tunnels [4] extended from UHA's or THA's to enable packet redirection to an MU's terminal. To provide more efficient service to an MU, the FA may cache pertinent parts of the user's service profile.

C. Registration and Authentication

UPC allows mobile users to use any terminal to access network computing resources and services at any network connected to the Internet and supporting such service capabilities. Since network access points are not necessarily

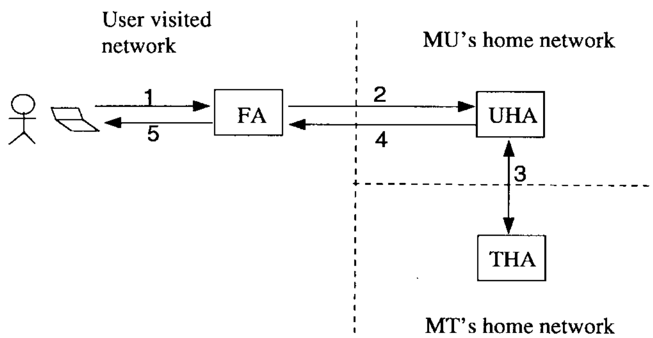


Fig. 4. User registration procedures.

under the control of the same administrative authority, a new set of interdomain mechanisms is needed to allow users to perform security operations in visited domains. To enable these capabilities, the serving network must verify the ID's of both user and terminal, and ascertain the capabilities and service preference of both, and bindings must be established among the MU, the MT, and the current location of both. The necessary information exchange constitutes a set of registration procedures, in which databases in the UHA and THA are accessed, and the successful establishment of these bindings results in the respective databases at the UHA, THA, and FA being updated. Authentication is a necessary part of the registration procedures to verify user and terminal ID's, and to prevent possible attacks on the exchange of registration information over the Internet.

The two-step registration procedure [16], involving terminal registration and user registration, seems appropriate. Terminal registration, using the procedures defined in mobile IP [17], is performed first to register the MT with its FA in the visited network and with the MT's THA at its home network, informing it of the MT's new care-of address. To provide network services locally, the terminal profile should be transferred to the FA before or at the same time that the THA returns a *registration reply* message to the FA.

The user registration illustrated in Fig. 4 follows similar procedures as terminal registration. The numbers in the figure correspond to these steps.

1) The MU sends a registration request message to the FA through the MT. The request includes the MU's LUI and the address of her UHA.
2) The FA relays the request to the UHA, attaching the address of the MT's THA.
3) The UHA copies the necessary terminal configuration information from the MT's THA in order to evaluate suitable services to the mobile user.
4) The UHA confirms the registration by sending a *registration reply* message to the FA. To optimize services to the MU, a user's personal agent (PA) may be created by the FA using inputs provided by the UHA. This topic will be further discussed in the next section.
5) The FA forward the registration reply message to the MU.

To illustrate the most general situation, we consider MU's and MT's from different home networks visiting a foreign network. If a user is accessing a fixed host or an MT at its home network, the terminal registration process is abolished. In whatever situations (i.e., whether the MU is at home or not), user registration should always be carried out. The only difference is that the MU at home registers directly with her UHA without involving any FA.

Since user registrations will be performed on a per-service access basis, and connections between the end application programs are mapped to TCP connections between two real hosts, changing terminals during a service session will not be transparent to the user, i.e., process migration between different terminals is not supported during a service access. A user has to release her association with the old terminal, close or suspend existing applications, then start the application from a new terminal after the user registration process. As different terminals may provide different capabilities required by a user, she may register on more than one terminal to access different services using the same LUI. Messages for individual services will be routed to the corresponding terminals, based on the different bindings established in the different registration instances. For example, during a multimedia conference, a user may use one terminal to share an editing space with other users, a different terminal for video communications, and another terminal for audio communications with other conference participants. Moreover, a user may have two LUI's, e.g., one for business use, and the other for personal use. Different service profiles will be defined for the two LUI's, and services for different purposes will be billed to the corresponding accounts.

In order to prevent third-party invaders, both terminal registration and user registration messages have to be authenticated. The former authentication happens between the MT and its THA, and the latter happens between the MU and her UHA. Terminal authentication guards against malicious registrations that cause an MT's THA to misdirect packets destined for the MT. The user authentication process not only protects users' personal resources from being accessed or tampered with by unauthenticated users, but also safeguards visiting users' access to local network resources, and ensures that services are properly charged to the users. While mechanisms such as "keyed MD5" [17] and user passwords are commonly used to authenticate registration messages from MT's and MU's, respectively, appropriate encryption mechanisms should also be available to protect other sensitive data.

In the current mobile IP, every time an MT crosses the boundary between two radio cells, a location updating at its THA [4] is necessitated if the cells are associated with different FA's. Considering that intranetwork handoffs will happen more frequently than internetwork handoffs, and personal mobility will require care-of address updating at both UHA and THA, location updating of intranetwork handoffs should be handled within the visited network to reduce the signaling across the networks (see Fig. 5). Since the network ID will not change in the same network, only handoffs between different networks should require updating of care-of addresses at the UHA's and THA's. This calls for using a single FA in each wireless network, or coordinating multiple FA's so that they function as a single logical entity.

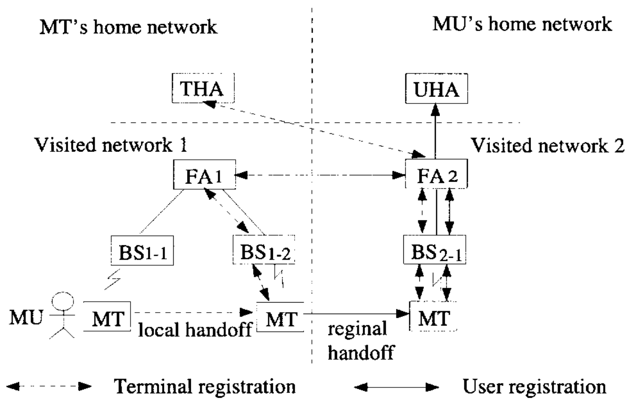


Fig. 5. Location updating of intranetwork and internetwork handoffs.

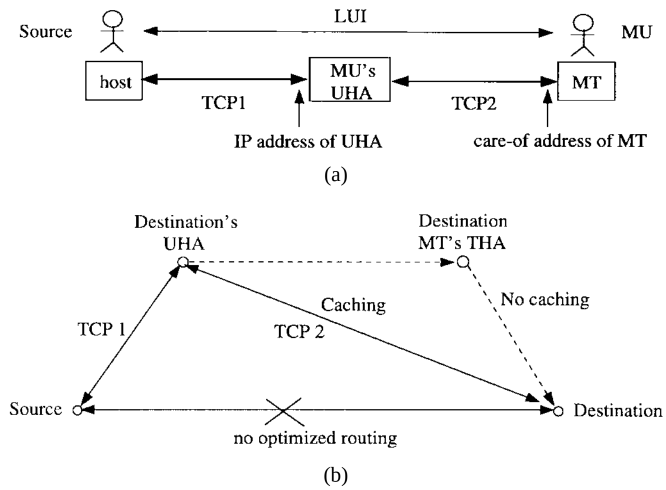


Fig. 6. Transparent solution.

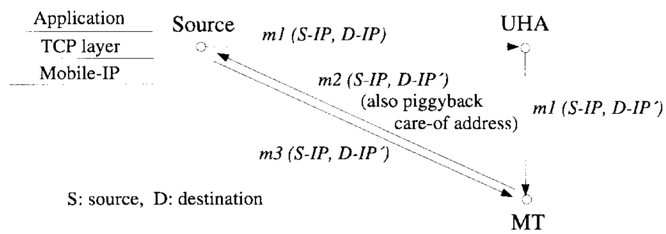


Fig. 7. Modified handshake sequence.

D. TCP Connection Management

To reliably communicate with each other, two application programs on different hosts have to set up a virtual circuit connection, which requires knowledge of the host addresses at the two end points. Since UPC users communicate with each others using LUI's without necessarily knowing other user's current location or host address, the current TCP protocol cannot directly set up a connection with the other user's MT. Therefore, the TCP connection protocol should be enhanced to utilize the dynamic binding among the LUI, LTI, and care-of address.

In the Internet protocol suits, TCP uses the *connection* as its fundamental abstraction. A connection is identified by a pair of endpoints. TCP defines an endpoint to be a pair of integers (*host*, *port*), where *host* is the IP address for a host and *port* is a TCP port on that host. For convenience in the following discussion, the port numbers in endpoint identities are ignored because port numbers will be kept unchanged in the enhanced protocol. Thus, a connection from machine 19.26.0.36 to machine 129.20.2.3 may be defined by the endpoints (19.26.0.36, 129.20.2.3). To establish a connection, TCP uses a three-way handshake [18] to ensure that both sides are ready to transfer data, and to allow both sides to agree on initial sequence numbers. The TCP software at each endpoint uses a *finite-state machine* to trace the state of a connection.

To enhance TCP for connection establishment between two MU's in the UPC system, three alternative methods are possible.

1) The Transparent Solution: As mentioned before, a source communicates with a UPC user by her LUI such as yalun@ee.ubc.ca. The name server may translate this name into the IP address of the UHA at her home network domain. Whether or not the other user is at home, the only choice at the source is to set up a connection with the other user's UHA. If the user is not at home, and she is registered on a terminal at a remote network, the UHA could set up a separate TCP connection with the user's current host since the UHA knows the user's current LTI. The UHA plays the role of a gateway that conveys network messages between the source and the MU. Fig. 6(a) illustrates the sequence of messages in the two separate TCP connections. Two finite-state machines are used in the UHA to record the connections at either sides. The TCP connection between the UHA and MU's terminal is supported by mobile IP. As the MT's care-of address is updated at the UHA whenever the user and MT move to a new location, the UHA can always tunnel packets directly to the MT's care-of address. Thus, triangular routing from the UHA to the MT's THA, then to the MT's care-of address can be avoided.

This method does not require any protocol changes at the calling host, so it is completely transparent to the calling host. But it prevents the calling host from communicating directly with the mobile user. Every packet must pass through the UHA. Thus, it has the same triangular routing problem as mobile IP, and caching technology cannot be used to provide an optimized routing since the source and the destination transmit all of the packets to the MU's UHA address, as shown in Fig. 6(b).

2) The TCP Layer Solution: The second method needs a slight modification of the TCP layer at the source. As shown in Fig. 7, the source requests to set up a connection with a mobile user by sending message 1 (*m*1) to the UHA using IP address D-IP. The UHA transfers the message to the user's MT by substituting the endpoint IP address with the MT's IP address, and the message is tunneled to the MT's care-of address. Then the MT acknowledges the setup request directly to the source by message 2 (*m*2), which also piggybacks the MT's care-of address, thus enabling the source to tunnel future messages directly to the mobile user. The TCP layer at the source is modified to recognize that this message acknowledges the original request message 1. After the source replies to the MT with message 3, the TCP connection is successfully set up between the source and the mobile user.

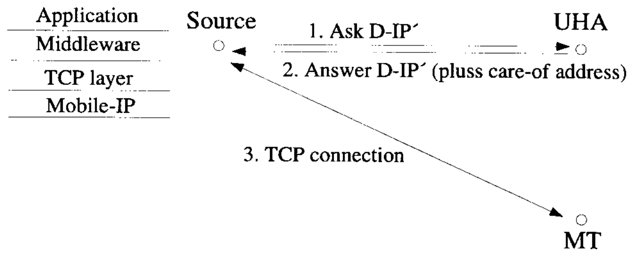


Fig. 8. Middleware layer solution.

The advantage of this method is that, by means of mobile IP, the source can communicate directly with the mobile user. Route optimization employing cache agents is applicable. The problem is that personal mobility is no longer transparent to the TCP layer. Should universal personal computing become an important feature of the Internet in the future, this modification at the TCP layer should be considered in future TCP/IP standard activities.

3) The Middleware Layer Solution: Besides personal mobility, future nomadic computing systems should support a variety of other features and capabilities. It has recently been suggested [19] that these functions are best incorporated in the middleware level of the commonly accepted layered protocol architecture. This motivates us to consider a solution to support personal mobility at the middleware layer. By means of an appropriate directory service, the middleware protocol of the source could check the IP address of the mobile user's MT at the UHA (see Fig. 8) before setting up a connection with the user's MT. When the UHA returns the MT's IP address to the source, the MT's care-of address could be piggybacked on the message, so that the source could tunnel packets directly to the MT.

This method employs middleware layer capabilities to manage mobility so that the TCP layer need not be modified. Of course, this method is not transparent to the source in the sense that the source needs to engage a UPC-compatible middleware. Nevertheless, it presents a clean solution with a clear layered architecture in alignment with OSI principles.

In summary, the purpose of all three methods is to make the source party know the destination party's current associated host address (LTI), thus enabling a UPC system to provide personal mobility functionality on top of host mobility provided by mobile IP. The piggybacked care-of address is just a byproduct of the protocol. After the connection is set up, future updating of care-of addresses caused by changing locations is carried out by mobile IP. The transparent solution gives the simplest but least efficient method. The TCP layer solution is efficient, but not backward compatible with the present network. The middleware layer solution seems the best candidate since it offers the flexibility to incorporate a rich set of UPC functions and is backward compatible. However, given the very large number of hosts already deployed within the Internet, it seems quite likely that most will not be upgraded to support personal mobility for some time. To provide a scalable transformation, the transparent solution may be employed by conventional hosts before they are upgraded.

IV. Personal Agents in UPC Systems

MT's in future UPC systems could be as simple as a diskless and palm-sized NC or as powerful as a full-featured multimedia desktop or notebook computer. Mobile users will use the terminals that are available and best suited to their needs. Because all the user-related information is unavailable at an arbitrary MT in a UPC system, users will depend heavily on network connectivity to access this information at their home networks. Although a user may access different terminals, she would prefer to operate in a familiar computing environment, and delegate the tasks of managing this environment to the system. To meet this objective, we propose to employ an agent-based architecture (see Fig. 9) where a personal agent (PA) within the network manages a user's computing environment based on the capability profile of the user's terminal and the user's service profile. Other researchers have also proposed to employ mobile agents in mobile computing systems to efficiently handle tasks such as handoff management [19]–[23].

In the proposed architecture, when an MU registers with her current MT at the FA at a visited network, a PA is created by the FA for this user. The PA copies user-related and terminal-related information, such as the user's service profile and the terminal profile, from the UHA and THA, respectively, in order to perform the following functions on behalf of the MU.

A. Interworking

The PA provides local access to the user's service profile and the terminal's capability profile to enable fast service access and harmonize functional differences between heterogeneous systems. The service profile records all of the services and applications to which the user subscribes or has access. Besides conventional services/applications such as e-mail, FTP, file editing, WWW browsing, etc., traveling mobile users may wish to access weather reports, traffic information, stock prices, bank accounts, maps, and other information that resides on the Internet or in their home network. Each service/application may have specific requirements on the processing host and the surrounding resources. Therefore, not all subscribed services/applications are accessible from any MT at any visited network. The PA could perform a mapping function to check the system capability, as shown in Fig. 10, comparing application requirements, terminal capabilities, and local network resources to determine if the MT can launch a subscribed service/application. In the case of a negative result, the alternative choice is to launch the service/application at a local host or server, or remotely at the user's home network, and transfer the results to the MT. The MT is then reduced functionally to a simple terminal. In some cases, a user may have a favorite application which needs to be accessed at home while roaming. If the MT is compatible with this application, it can simply be downloaded from the home network and launched at the MT. However, if the MT is not compatible, the user may run the application remotely at the home location, which may be too slow and expensive in communications costs. Another possibility is that, if the application is available in source code, it may be downloaded

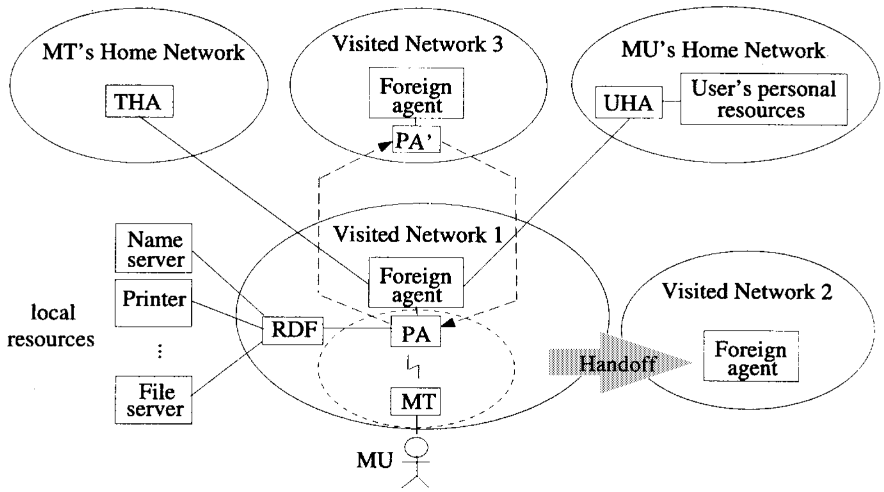
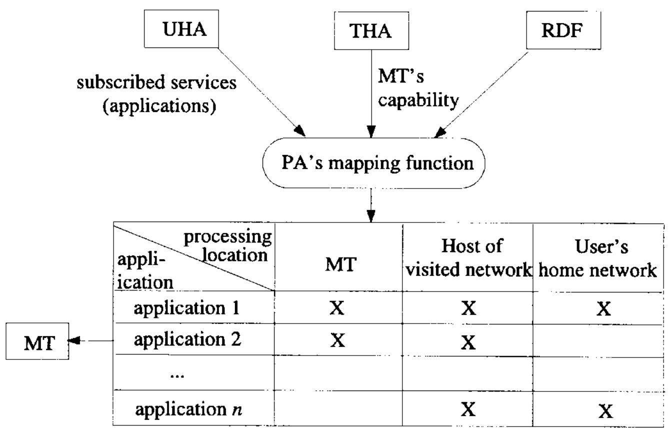


Fig. 9. Agent-based personal computing architecture.



appli-ication \ processing location	MT	Host of visited network	User's home network
application 1	X	X	X
application 2	X	X	
...			
application *n*		X	X

Fig. 10. PA's mapping function to determine available services.

and recompiled or interpreted at the visited location (e.g., applications programmed with languages such as Java or Telescript that could be run on any platforms). If several alternatives are all viable, some performance metric may need to be developed so that the most suitable alternative can be selected under different conditions.

B. Signaling

The PA can perform complex signaling functions on behalf of the mobile user. In wireless systems, congestions over the shared radio link, transmission errors, and handoffs can seriously degrade the performance of users' data connections. It is therefore beneficial to avoid complex signaling procedures that consume valuable radio bandwidth and are susceptible to errors. By implementing an efficient protocol to enable the PA to keep track of the current state of the terminal over a wireless channel, the PA can act on behalf of the terminal to perform any required complex signaling procedures over the network. One example of such a signaling function is incorporated in the recently proposed indirect transport layer protocol (I-TCP) [24], intended to overcome the adverse effects of transmission errors and handoffs in a wireless network to transport protocol performance [25], [26]. With I-TCP, the TCP connection between an MT and a fixed host is broken up at the edge of the fixed and wireless network, and a special network entity is assigned to coordinate between the two connections. In the UPC system, the user's PA can perform the function of this coordinating entity.

C. Resource Management

The PA may optimize resource usage at the MT, the visited network, and the user's home network in order to maximize the efficiency or minimize the cost of service.

- The PA allows the mobile user to access network services and resources at the visited network without knowledge about the local network configuration. By means of a resource discovery facility (RDF) provided in each network, the PA enables the mobile user to access local resources such as print servers, compute servers, mail servers, etc. In some cases, a user's PA does not have the capability to execute an application locally because of insufficient resources, but a remote server has direct access to the resources needed, or the PA should run a heavyweight application over several remote servers. The PA could create another personal agent (PA′ in Fig. 9), which is similar to an itinerant agent in [20]. PA′ roams among a set of networked servers to seek out ones that can help with the user's task or perhaps collect information from these servers until it accomplishes its task, and then returns the results to the original PA.
- The PA helps to manage the user's personal resources in a transparent way. These resources, including configuration, data, and text files, are usually stored in the mobile user's home domain. For example, the PA may enable the user to manage her home file directories as if they were local. The PA may fetch the user's bookmark file from her home directory when the user invokes the service of a WWW browser. If the user wishes to edit a file, the PA may start the user's favorite file editor, fetch and create a temporary copy of the file, and update the file at the home directory periodically or at the command of the user. If the user accesses e-mail service, the PA may

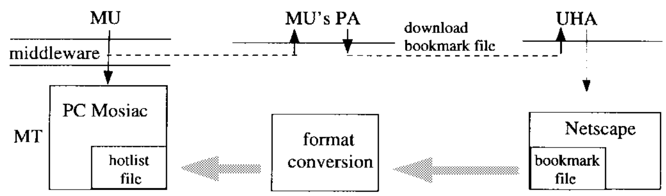


Fig. 11. Example of WWW browsing in a visited network.

present the user's mail box, mailing lists, aliases, etc., from her home network. Transportation of data from the user's home network to the application program at the MT can be managed by the PA in a transparent way on an as-needed basis.

- Given the service/application the user wishes to access or launch, the PA may also help the MT to optimize the usage of battery power or communication bandwidth. The PA could migrate processing-, memory-, or communications-intensive tasks to fixed hosts or servers in the visited network in order to compensate for the lack of MT resources or capabilities.
- Some wireless networks may be quite unreliable, with MT's experiencing frequent loss of connectivity. The PA may serve as a proxy for the MT when connectivity is lost, resynchronizing with the MT when connectivity is once again restored.

D. Mobility and Handoff Management

Mobility management refers to tracking the location of a mobile user by the network, whereas handoff management refers to the maintenance of connectivity when a mobile user hands off between radio cells. A fundamental way in which UPC service differs from conventional operation is that a user's PA may be in charge of the user's connectivity with the rest of her computing environment. Sessions may be created between user and (local or remote) servers during service access. As the user moves, the server interacting with the user may need to change. Thus, physical mobility on the part of the user may result in the virtual mobility of the server(s). However, unlike physical mobility, in the case of virtual mobility, the new server needs context information from the old server to pick up and continue the session seamlessly. To function properly, PA's need to react to mobility, handoff management, and context information transfer events in the network related to the corresponding MT's. In a distributed network environment, in order not to restrict the mobility of MT's, the corresponding PA's may also need to be mobile so that they track the mobility of the MT's.

Fig. 11 shows an example of how the PA facilitates WWW browsing by the mobile user. Suppose the mobile user's service profile identifies a bookmark file associated with the Netscape browser employed by the user at her home network, and the user is registered on a notebook PC in which the installed browsing tool is Mosaic. When presenting the home directory to the mobile user, the PA first converts the file name of the Netscape bookmark file into a format recognizable as a Mosaic hotlist file. If the mobile user runs Mosaic and accesses the hotlist file, the request is intercepted and sent to the PA. The PA then downloads the user's bookmark file from her home network and converts it into the hotlist file format, and forwards this file to the MT for presentation to the user. This enables the user to access important personal information while traveling. In this example, the PA acts as a virtual file server with file name and file format conversion capabilities.

V. Conclusions

We have proposed a new paradigm for future mobile and nomadic computing, called universal personal computing or UPC, which allows the mobile users total freedom to access network computing services and resources and personalized computing environments from any available fixed or mobile terminals. We have presented an agent-based protocol architecture for universal personal computing by identifying the functional requirements, and by specifying the architectural elements necessary to realize universal personal computing. Agents to separately manage mobile user and terminal objects in corresponding networks have been proposed. Registration procedures for updating user's location information to facilitate mobility and handoff management have been described. Connection management enabling a mobile user to be reached using her logical user identifier has been discussed. From the protocol description of terminal and user registration, authentication procedures, and handoff management, we notice that terminal mobility is a special case of personal mobility where both the user and associated terminal belong to the same home network. Thus, UPC system research can take advantage of mobile-IP research efforts, and the mobile-IP standard should be made upward compatible with UPC systems.

We have also proposed the use of personal agents on the fixed network to facilitate interworking, signaling, and resource, mobility, and handoff management functions. This computing architecture will be composed of distributed components that execute in fixed local or remote hosts or servers (instead of in MT's) in cooperation with each other to provide services to mobile users and administrate the interconnected networks. Design of such distributed applications is intrinsically difficult. Issues such as communications, process migration, concurrency, consistency, interworking heterogeneous systems, and fault tolerance must be addressed in the design process.

Universal personal computing is a powerful concept, and many issues remain to be resolved before the full capabilities of this concept can be realized. Work is in progress to address some of these issues, and to implement and demonstrate universal personal computing. For practical reasons, the first implementation will limit "universality" to computing systems with a high level of compatibility. We welcome the participation and contributions of the research community to make universal personal computing a reality.

References

[1] W. Stallings, "IPv6: The new internet protocol," *IEEE Commun. Mag.*, pp. 96–108, July 1996.
[2] L. Kleinrock, "Nomadic computing—An opportunity," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 25, pp. 36–40, Jan. 1995.
[3] C. Perkins, Ed., RFC 2002, *IP Mobility Support,* Oct. 1996.

[4] A. Myles, D. B. Johnson, and C. Perkins, "A mobile host protocol supporting route optimization and authentication," *IEEE J. Select. Areas Commun.*, vol. 13, pp. 839–849, June 1995.
[5] J. Ioannidis, D. Duchamp, and G. Q. Maguire, Jr., "IP-based protocols for mobile internetworking," in *Proc. SIGCOMM'91 Conf.: Commun. Architectures, Protocols*, Sept. 1991, pp. 235–245.
[6] L. Stuart, "NC specifications revealed," *Network World*, vol. 6, p. 1, June 1996.
[7] G. S. Lauer, "IN architectures for implementing universal personal telecommunications," *IEEE Network Mag.*, vol. 8, pp. 6–16, Mar./Apr. 1994.
[8] CCITT SG 1, *Draft Recommendation F.851 UPT-Service Description, Version 7*, Geneva, 1992.
[9] M. Fujioka, S. Sakai, and H. Yagi, "Hierarchical and distributed information handling for UPT," *IEEE Network Mag.*, vol. 4, pp. 50–60, Nov. 1990.
[10] J. Vandennamaeele, J. B. Thieffry, and H. Decuypere, "UPT, a new dimension in telecommunications provided by IN," in *Proc. Digital Commun. Zurich Seminar*, 1992, pp. 41–54.
[11] ETSI DTR/NA-70201, UPT: General Service Description, Sept. 1991.
[12] M. Veeraraghavan, T. F. La Porta, and R. Ramjee, "A distributed control strategy for wireless ATM networks," *ACM/Baltzer Wireless Networks*, vol. 1, no. 3, 1995.
[13] F. Dupuy, G. Nilsson, and Y. Inoue, "The TINA consortium: Toward networking telecommunications information services," *IEEE Commun. Mag.*, pp. 78–83, Nov. 1995.
[14] S. Vinoski, "CORBA: Integrating diverse applications within distributed heterogenous environments," *IEEE Commun. Mag.*, pp. 46–55, Feb. 1997.
[15] F. J. Carrasco, J. Esteve, M. Felipe, J. C. Garcia, J. C. Moreno, and J. Ruano, "Applying CORBA technology for the implementation of a TINA service," in *Proc. GLOBECOM'96*, 1996, pp. 57–62.
[16] A. Nakajima, M. Kuramoto, and M. Yabusaki, "Intelligent network architecture for personal mobile telecommunication," in *Proc. ICUPC*, 1992, pp. 339–344.
[17] C. Perkins, Ed., Internet Draft, draft-ieft-mobileip-protocol-16.txt, Apr. 22, 1996.
[18] D. E. Comer, *Internetworking with TCP/IP*. Englewood Cliffs, NJ: Prentice-Hall, 1991.
[19] R. Bagrodia, W. W. Chu, L. Kleinrock, and G. Popek, "Vision, issues, and architecture for nomadic computing," *IEEE Personal Commun. Mag.*, pp. 14–27, Dec. 1995.
[20] D. Chess, B. Grosof, C. Harrison, D. Levine, C. Parris, and G. Tsudik, "Itinerant agents for mobile computing," *IEEE Personal Commun. Mag.*, pp. 34–39, Oct. 1995.
[21] I. Iida, T. Nishigaya, and K. Murakami, "DUET: An agent-based personal communications network," *IEEE Commun. Mag.*, pp. 44–49, Nov. 1995.
[22] R. Ramjee, T. F. La Porta, and M. Veeraraghavan, "The use of network-based migrating user agent for personal communication services," *IEEE Personal Commun. Mag.*, pp. 62–68, Dec. 1995.
[23] G. Y. Liu, A. Marlevi, and G. Q. Maguire, Jr., "A mobile virtual-distributed system architecture for supporting wireless mobile computing and communications," *ACM/Baltzer Wireless Networks*, vol. 2, no. 1, 1996.
[24] A. Bakre and B. R. Badrinath, "I-TCP: Indirect TCP for mobile hosts," in *Proc. 15th Conf. Distributed Computing Syst.*, Vancouver, Canada, May 1995, pp. 136–143.
[25] P. Manzoni, D. Ghosal, and G. Serazzi, "Impact of mobility on TCP/IP: An integrated performance study," *IEEE J. Select. Areas Commun.*, vol. 13, pp. 858–867, June 1995.
[26] R. Caceres and L. Iftode, "Improving the performance of reliable transport protocols in mobile computing environments," *IEEE J. Select. Areas Commun.*, vol. 13, pp. 850–857, June 1995.

Yalun Li (S'93–M'95) received the B.S. and M.S. degrees in electrical engineering and computer science from Jilin University of Technology (JUT), P.R. China, in 1982 and 1986, respectively, and the Dr.Ing. degree from the Department of Computer System and Telematics, Norwegian Institute of Technology, in 1995.

He worked as an Engineer in the Microcomputer Laboratory at JUT in 1982 and 1983. From 1986 to 1989, he was a Lecturer in the Department of Computer Science at JUT. Since 1995, he has been a postdoctoral Fellow in the Department of Electrical Engineering, University of British Columbia, where he is responsible for the architecture and protocol design of universal personal computing systems in the "Internetworking Wireless Data Networks for Mobile Computing" project. His research interests include wireless multipoint multimedia communications, personal communications services (PCS), traffic control and analysis, and mobile computing.

Victor C. M. Leung (S'75–M'95) received the B.A.Sc. (Hons.) degree in electrical engineering from the University of British Columbia (U.B.C.) in 1977, and was awarded the APEBC Gold Medal as the head of the graduating class in the Faculty of Applied Science. He attended graduate school at U.B.C. on a Natural Sciences and Engineering Research Council Postgraduate Scholarship, and received the Ph.D. degree in electrical engineering in 1981.

From 1981 to 1987, he was a Senior Member of Technical Staff at Microtel Pacific Research Ltd., specializing in the planning, design, and analysis of satellite communication systems. He also held a part-time position as Visiting Assistant Professor at Simon Fraser University in 1986 and 1987. In 1988, he was a Lecturer in the Department of Electronics, Chinese University of Hong Kong. He joined the Department of Electrical Engineering, U.B.C., in 1989, where he is an Associate Professor and a member of the Centre for Integrated Computer Systems Research. He is also a Project Leader in the Canadian Institute for Telecommunications Research, a Network of Centres of Excellence funded by the Canadian Government. His research interests are in the areas of architectural and protocol design and performance analysis for computer and telecommunication networks, with applications in satellite, mobile, personal communications, and high-speed networks.

Dr. Leung is a member of the ACM.