Do you see this success transferring to other markets like the USA and Europe?

Cell phone usage is spreading worldwide. At the same time, cell phones are becoming smaller and more portable. By enabling these devices to process e-mail and access the Internet, as well as facilitate the instant exchange of information, we are allowing ourselves to participate in the new information society.

Why did you choose extended HTML rather than a more specialized language optimized to a radio link?

HTML is a de-facto Internet language that allows anyone to become an information provider. That is why we decided to adopt it.

Wireless Application Protocol (WAP)

Mjumo Mzyece, University of Strathclyde

The second of our articles on mobile data in current cellular systems provides an overview of the Wireless Application Protocol (WAP). The article gives the background and motivation for WAP, before going on to outline some of the basic technical details of WAP technology. Several examples of WAP applications and services that have either already been deployed or could be deployed are described. Finally, there is a brief description of a few factors that could impact WAP's bid to become the global standard for wireless applications and services.

Introduction

Since the early 1990s, there has been phenomenal growth in the global wireless communications industry. Alongside this rapid numerical growth have been continual technological developments to meet emerging requirements. One such requirement is for wireless networks to provide services other than the traditional voice, fax and low bit rate data services. The Wireless Application Protocol (WAP) is designed to provide advanced information and telephony applications to wireless devices such as cellular phones and personal digital assistants (PDAs). The first release of WAP (version 1.0) was in April 1998 and there are currently more than 8 million WAP subscribers worldwide [1].

This article gives an overview of WAP, beginning with the background and motivation for the technology. The technical aspects of WAP are then outlined followed by some examples of existing and possible WAP applications. Finally, the future prospects and technical developments of WAP are briefly discussed.

Background and Motivation

There are a number of background and motivating factors that have led to the development of WAP:

Growth in Internet, Wireless and Handheld Device Technologies

During the last decade or so, there has been tremendous growth in three high-technology sectors: the Internet, wireless communications and smart handheld devices.

The Internet has grown explosively in terms of both the number of Internet users and the number of Internet hosts. An Internet host is any computer system that is connected to the Internet, whether by permanent or temporary, dial-up or direct connection, and has an associated domain name and Internet protocol (IP) address. According to Internet pioneer Vint Cerf, at the end of 1991 the Internet had over 700 000 hosts used by over 4 million people [2]. In March 2001, it is estimated that there are more than 116 million Internet hosts and more than 428 million Internet users [3]. The International Telecommunication Union (ITU) projects that by the end of 2002, there will be approximately 600 million Internet users [4].

Running almost concurrently with this growth in the Internet has been the equally extraordinary growth in the number of mobile cellular networks and subscribers (Figure 1). The number of mobile cellular subscribers worldwide grew from a mere 11 million in 1990 to 472 million in 1999 [4].

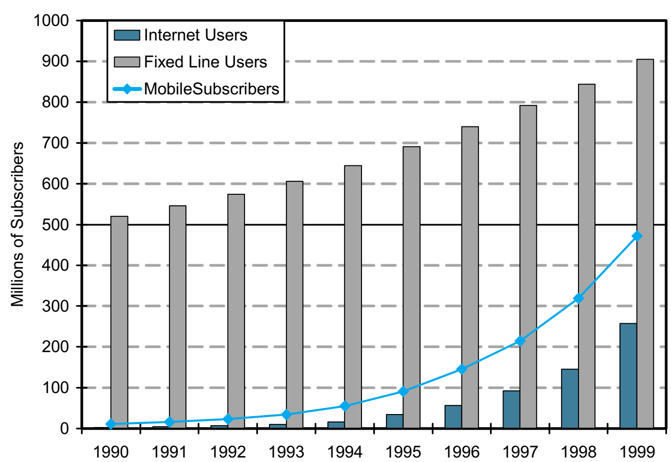

Figure 1 Worldwide Growth in Mobile and Internet Subscribers

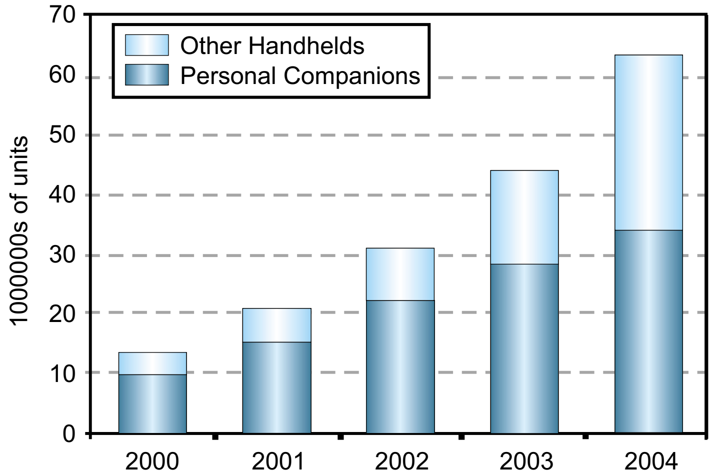

Figure 2 Unit Shipments of smart handheld devices (Source: IDC Research, 2001)

The third and most recent area of rapid growth has been in the smart handheld devices market. Some examples of such devices are PDAs, palmtop PCs, gaming consoles and smart phones. Figure 2 shows an analysis and forecast for the smart handheld device market from 2000 to 2004.

Convergence of Internet and Wireless Communications

The Internet and wireless communications have conventionally been regarded as separate technologies. This is because originally, the Internet was designed to carry mainly data traffic, whilst wireless networks were designed to carry mainly voice traffic. This boundary has become increasingly blurred in recent years.

According to one estimate, the Voice-over-IP (VoIP) market was equivalent to 2.7 billion minutes of traffic in 1999 and will grow to about 35 billion minutes, with revenues of US$19 billion, by 2004. Other sources estimate that by 2004 as much as 25% or 40% of all international calls will be carried over IP [5].

Mobile telephony, on the other hand, has seen some remarkable growth in the carriage of data. The huge popularity of the GSM-based Short Message Service (SMS) vividly illustrates this strong trend. The GSM Association forecasts that by December 2001, global SMS traffic will soar to 25 billion text messages per month, up from 15 billion in December 2000 [6].

These are just two examples of the convergence of traditional Internet/IP/data communications (which are data-centric) and traditional wireless communications (which are voice-centric). WAP delivers both data and voice services to wireless devices.

Need for Standardisation

Given the extremely competitive global telecommunications market and the resultant desire to provide value-added services on top of voice telephony, there is a need for a standard for the delivery of wireless applications and services. Such a standard has to be open, global and enjoy wide industry support. The alternative is a fragmented wireless services market with many proprietary standards. WAP fulfils the role of being a single standard platform for wireless services and applications.

What is WAP?

WAP is an optimised communications protocol stack *and* an application environment designed for the deployment of advanced information and telephony services for wireless devices. It has been *optimised* to cope with the constraints of the wireless operating environment, and the limitations of its targeted wireless devices.

A Brief History of WAP

WAP is the creation of the WAP Forum [7], a consortium of wireless operators, Internet companies, terminal manufacturers, network infrastructure manufacturers, software companies and other interested parties. The WAP Forum was founded in June 1997 by Phone.com (formerly Unwired Planet and now part of Openwave Systems), together with Ericsson, Motorola and Nokia (the world's three largest cellular handset manufacturers). As of January 2001, the WAP Forum membership stood at over 630 [1].

The latest approved version of WAP at the time of writing (March 2001) is version 1.2.1 which was released in June 2000. The following discussion is based upon the specifications for WAP 1.2.1.

WAP Specifications

The WAP specifications comprise almost 40 documents that in total run into the hundreds of pages [8]. Given the volume of the specifications, it is not possible (or even desirable, for that matter) to discuss all the details of WAP in an article of this length and nature. Therefore only the basic features of WAP will be highlighted.

Conceptually, WAP can be considered as comprising two complimentary parts:

- A *lightweight, optimised communications protocol stack* closely based on the Internet protocol stack;
- An *application environment* that provides a platform for the delivery of advanced information and telephony services to wireless devices.

Most Internet applications have been designed for desktop and other powerful computers, and for medium-to-high bandwidth available from fairly reliable wired communications networks.

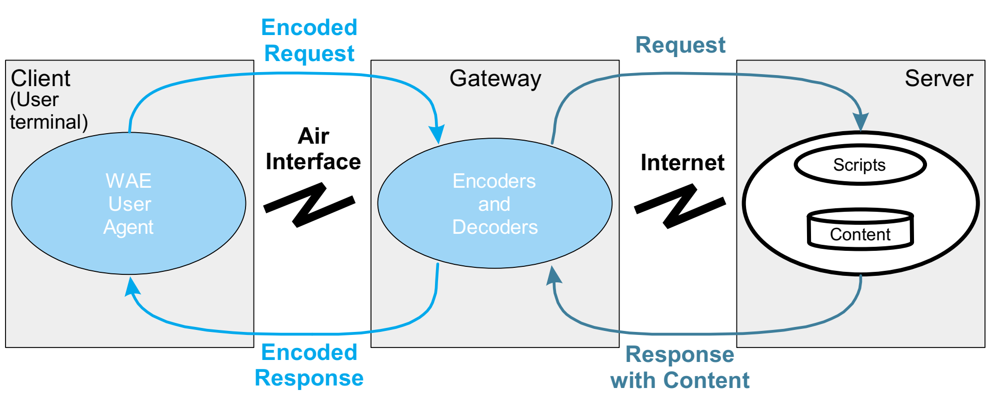

Figure 3 The WAP Architecture

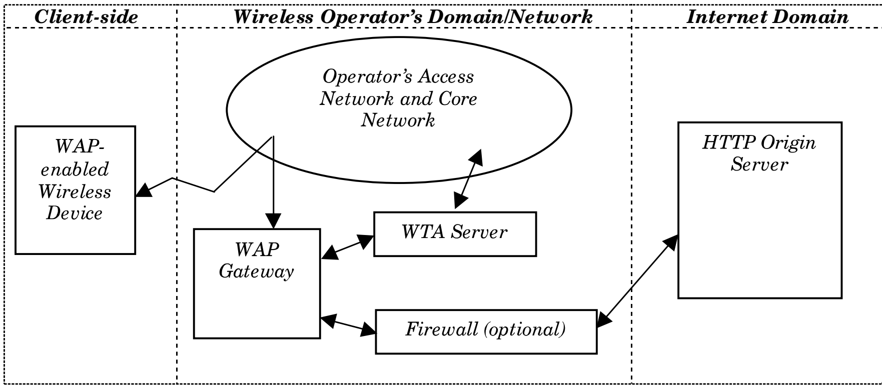

Figure 4 Interconnection of key network components for WAP service

In contrast, the types of devices targeted by WAP have considerable limitations compared to desktop computers. These limitations include:

- Limited processing power i.e. less powerful CPUs
- Limited memory (ROM and RAM)
- Limited power supplies
- Limited displays
- Limited input devices (e.g., a phone keypad).

In addition, wireless communications networks present a number of constraints not found in their wired counterparts. Due to limited radio spectrum, the peculiarities of radio propagation and the mobility of users, wireless communications networks tend to suffer from:

- Limited bandwidth
- Interference and fading effects
- Unpredictable and variable delays
- More unstable connections
- Less predictable availability.

The WAP specifications are specifically designed to address these limitations through a number of optimisations which will be discussed shortly.

The WAP specifications are also designed to be both *bearer-independent* and *device-independent*. Therefore, WAP is not tied to any particular wireless bearer technology or any particular type of wireless device. WAP can therefore be used with any bearer from GSM to CDMA, and on any device from a pager to a PDA.

The WAP Architecture

The WAP Architecture Specification acts as the starting point for understanding the WAP technologies and specifications. It provides an overview of how the WAP technologies fit and work together. Figure 3 shows a schematic of the WAP architecture.

As already mentioned, the WAP architecture is closely modelled on the Internet architecture. The Internet architecture works on the *client/server model* in which content and applications stored on an *origin server* (a web server, for instance) are accessed by a *client* (such as a web browser).

Similarly, WAP content and applications are stored on origin servers using existing server technology such as the World Wide Web's Hypertext Transfer Protocol (HTTP). These applications and content are then accessed by the WAP client using a *microbrowser*.

Sitting in between the WAP origin server in the *Internet domain*, and the WAP client in the *wireless domain* is a *WAP gateway* or *proxy*. The WAP gateway is the translator between the Internet domain and the wireless domain and comprises the following functionality:

- Protocol Gateway – The protocol gateway translates requests and responses from the WAP protocol stack to the Internet protocol stack, and vice-versa.
- Content Encoders and Decoders – The content encoders translate WAP content into compact encoded formats to reduce the size of transmitted data over the wireless interface (also known as the air or radio interface).

Figure 4 depicts the interrelationship of the key network components required for a WAP service.

The WAP Protocol Stack

WAP provides an optimized five-layer protocol stack. The layered structure enables services and applications to utilize the features of the WAP stack through a set of well-defined interfaces. The five layers are:

- **Application Layer** Wireless Application Environment (WAE)
- **Session Layer** Wireless Session Protocol (WSP)
- **Transaction Layer** Wireless Transaction Protocol (WTP)
- **Security Layer** Wireless Transport Layer Security (WTLS)
- **Transport Layer** Wireless Datagram Protocol (WDP)

Figure 5 shows the WAP protocol stack in relation to the Internet protocol stack.

It is worth noting that the *bearer services* at the base of the stack are not part of the WAP specifications.

External applications and services (in the application layer) may access any of the lower layers directly. Each layer of the WAP protocol stack will now be defined briefly.

Internet	WAP
HTML/JavaScript	Wireless Application Environment (WAE)
HTTP	Wireless Session Protocol (WSP)
	Wireless Transaction Protocol (WTP)
TLS-SLL	Wireless Transport Layer Security (WTLS)
TCP/IP UDP/IP	Wireless Datagram Protocol (WDP)
	BEARERS CURRENTLY SUPPORTED BY WAP: **GSM SMS, USSD, CSD, GPRS**; **IS-136 R-DATA, CSD, PACKET**; **CDMA SMS, CSD**; **PDC CSD, PACKET**; **PHS CSD**; **CDPD**; **IDEN SMS, CSD, PACKET**; **FLEX & REFLEX, DATATAC**

Figure 5 WAP protocol stack in relation to the Internet protocol stack

Wireless Application Environment (WAE)

WAE is the uppermost layer in the WAP protocol stack, and it provides a general-purpose application environment combining elements of WWW and mobile telephony technologies. The key components of WAE are as follows:

- **WAE User Agents** – client-side in-device software that provides specific functionality (e.g., display content) to the end-user. Micro-browsers and phonebooks are examples of WAE user agents.

- **Wireless Markup Language (WML)** – a lightweight markup language that has been implemented using the eXtensible Markup Language (XML). XML is a meta-language used for defining markup languages. Unlike HTML which is rigidly defined for all applications environments, WML has been *specifically* designed and optimized for use on limited capability wireless terminals and for transmission in wireless channels. WML uses a *card and deck metaphor*. WML documents are made up of multiple *cards*. Each user interaction and navigation through an application is described by a set of cards called a *deck*. Decks are downloaded from the WAP origin server as and when required. Each card in a deck defines a single unit of interaction with the user.
- **WMLScript** – a lightweight scripting language, based on JavaScript. WMLScript adds procedural logic and intelligence to WML decks to allow for capabilities such as validation of user input and access to device facilities and peripherals.
- **Content Generators** – Applications or services on origin servers that produce standard content formats in response to requests from user agents in the mobile terminal. WAE does not specify any standard content generators, but allows for any content running on a typical HTTP server commonly used on the World Wide Web. WAP content may be static (already resident on the server) or dynamic (produced on-the-fly in response to some user input or behaviour e.g. content from CGI scripts). To specify content and resources, WAP uses Uniform Resource Locators (URLs), the same addressing scheme used on the Internet. WAP also uses Uniform Resource Identifiers (URIs) to address resources that are not accessible via commonly used protocols e.g. a URI is used for local access to a wireless device's telephony functions.
- **Wireless Telephony Applications (WTA)** – WTA provides a collection of telephony based extensions to WAP that allow for the deployment of services such as initiating phone calls, call-management, handling of text messages and phonebook control.
- **Content Formats** – a set of well-defined data formats with optimized compression for transmission on wireless networks. These include encoded WML format, WMLScript bytecode format, Wireless Bitmap format (WBMP), electronic business cards, and electronic calendars and scheduling exchange format.

Wireless Session Protocol (WSP)

WSP provides the upper WAE with a consistent interface for two types of session services. First, a *connection-oriented service* that operates above the transaction layer protocol (WTP) and the datagram layer protocol (WDP). And second, a *connectionless service* that operates above WDP only, bypassing WTP. Both types of session services may operate in secure or non-secure modes (with and without WTLS respectively).

WSP facilitates organized exchange of content between co-operating client/server applications. Specifically, it allows applications to:

- establish and release sessions between clients and servers;
- agree on a common level of protocol functionality using capability negotiation;
- exchange content between client and server using compact encoding;
- suspend and resume sessions.

WSP is based on HTTP/1.1.

Wireless Transaction Protocol (WTP)

WTP provides the services necessary for interactive "browsing" applications (i.e. request/response applications). During a browsing session, a client requests information from a server, and the server responds with that information. This request/response pair of events is called the *transaction*. The purpose of WTP is to reliably complete a transaction taking into account both the degree of reliability required and the cost of achieving that reliability. WTP runs on top of a datagram service (WDP) and optionally a security service (WTLS).

WTP specifies three classes of transaction service:

- Unreliable invoke message with no result message;
- Reliable invoke message with no result message;
- Reliable invoke message with one reliable result message.

Other key features of WTP include:

- Acknowledgements, duplicate removal and re-transmissions;
- Optional user-to-user reliability whereby the WTP user confirms every received message;
- Concatenation and separation of multiple Protocol Data Units (PDUs);
- Message orientation so that the basic unit of interchange is an entire message and not a stream of bytes;
- Asynchronous transactions by which a responder sends back a result as data become available.

Wireless Transport Layer Security (WTLS)

WTLS is the security layer protocol in the WAP protocol stack. The WTLS layer operates above the transport protocol layer (WDP). The WTLS layer is an optional layer, and whether or not it is implemented depends on the required level of security for a given application. The purpose of the WTLS layer is to provide privacy, data integrity and authentication between two communicating applications.

WTLS is based on the Internet standard Transport Layer Security or TLS protocol (previously called Secure Sockets Layer or SSL). WTLS provides the following features:

- Data integrity
- Privacy
- Authentication
- Denial-of-service protection

Wireless Datagram Protocol (WDP)

WDP is the transport layer in the WAP protocol stack. The WDP layer operates above the data capable bearer services supported by the various wireless network types. It provides a common interface to the upper layer WAP protocols and enables them to function independently of the underlying wireless network.

This bearer independence is achieved by adapting the WDP layer to specific features of the underlying bearer. Figure 6 depicts WDP bearer adaptation for four different bearers.

WDP offers transparent service to the upper layers of the WAP protocol at the Transport Service Access Point (TSAP) interface. This allows applications to operate transparently over different underlying bearer services.

The varying heights of each of the bearer services and their respective bearer adaptations indicates the differences in operation of the various bearers, and the different WDP adaptations required to maintain consistent service to the upper layers above the TSAP interface.

Each individual bearer service that is supported by WAP has its own WDP profile. This is called the *WDP Bearer Dependent Profile*. It defines the operation between a wireless

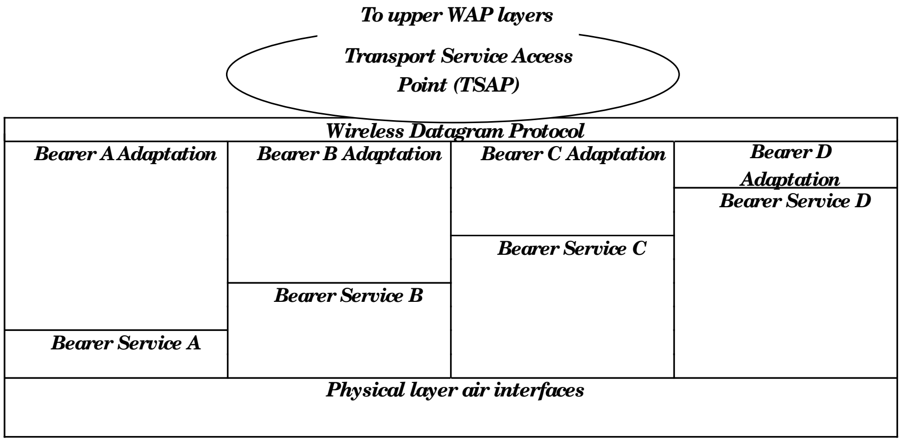

Figure 6 Wireless Datagram Protocol Architecture

terminal and the WAP Gateway over that particular bearer. For example, the current WDP specifications define the following protocol profiles for GSM: GSM SMS, GSM USSD, GSM CSD, GSM GPRS and GSM Cell Broadcast. Similar WDP profiles are defined for each bearer service supported by WAP.

Since the WDP layer provides the interface between the bearer service and the rest of the WAP protocol stack, the WDP specification specifies the bearers that are supported together with their adaptation profiles. As more bearer services such as Wideband-CDMA are developed and deployed, it is expected that they will be incorporated into the WDP specifications.

WAP Applications and Services

There are a variety of WAP applications and services that are existing or feasible, including:

Content-based Services

In this context, content could be current affairs, stock market news, movie clips, sports scores/highlights, cartoons and downloadable jingles. This is the sort of content that has proved very successful for Japanese wireless operator NTT DoCoMo's i-mode service.

Examples of existing content-based WAP applications include: StockSmart (http://agsub.stocksmart.com) which provides up-to-the-minute stock market news and the FT.com WAP site (http://wap.ft.com) which provides business news and analysis. Figure 7 shows the opening screens of the FT.com WAP site viewed with one of the numerous free WAP emulators available on the Internet.

Location-based Applications

Location-based or *position-dependent services* are services delivered to a wireless terminal on the basis of its current location or position. The UK-based National Geospatial Data Framework (NGDF) estimates that as much as 80% of information has some relationship to the earth's surface or is, in other words, geospatially referenced [9]. This gives some indication of just how valuable information about position or location is in general.

There are a wide variety of location-sensitive applications which can be implemented with WAP including:

- **Emergency/Safety Services.** These types of services could include roadside assistance/rescue services and emergency medical services.
- **Localised Information.** This category of services could include targeted advertising, WAP-enabled tour guides, Yellow Pages, weather reports, and car parking information.
- **Traffic/Routing Information.** This is a prime example of the value of timely position-dependent information. A WAP-enabled traffic and routing application would be able to locate the user and only relay information that is useful in their current position. The route planner could dynamically update the planned route to avoid areas of congestion.
- **Tracking Services.** These services use tracking devices such as GPS receivers to locate and monitor moving objects. Obvious areas of applications for individual consumers would be tracking of property, pets or children. Commercial applications could be fleet management and despatch services.

A good example of location-based WAP services is from FedEx. A FedEx customer can use their WAP-enabled device to find the closest FedEx drop-off location complete with driving directions. They can also track the progress of their deliveries via WAP [10].

Financial Services

There are a wide variety of WAP-enabled financial applications and services available. In the United Kingdom, for instance, banks such as the National Westminster Bank and Egg, offer their customers a number of WAP-based banking services.

Gaming/ Entertainment

Wireless gaming and entertainment is a potential "killer app" for WAP services, particularly for younger mobile users. There are already numerous WAP sites dedicated to gaming and entertainment e.g. Wirelessgames.com.

M-commerce Applications

M-commerce or mobile commerce involves buying and selling of goods and services through wireless handheld devices. It is unclear how much real commercial activity there is via WAP-based terminals. Nevertheless, there are many WAP-based m-commerce applications that are being developed, particularly in countries like Finland, which has the highest mobile penetration rate in the world (66.7% according to the ITU) [4]. In Finland, Computer Sciences Corporation (CSC) and Nokia have collaborated with a Finnish fashion retailer to use WAP to send clothing offers direct to mobile telephones.

Figure 7 FT.com WAP site (Source: FT.com,2000)

Speech-recognition Applications

Applications incorporating speech recognition are one of the most exciting prospects for WAP. The ability to use voice input would increase the ease-of-use and convenience of almost any WAP application. Coupled with machine-translation techniques, speech recognition could even enable WAP applications that translate between different languages. In July 2000, Conversa announced that it had successfully ported its speech recognition and text-to-speech engine to Phone.com's WAP-compatible microbrowser [11]. This development enables a user to control the microbrowser through voice commands.

Conclusions

WAP could potentially become *the* standard for the delivery of advanced applications to wireless terminals. However, there are some points of concern:

- **Competition from NTT DoCoMo's i-mode service.** NTT DoCoMo's hugely successful i-mode service [12], its recent aggressive expansion strategy outside the Japanese market and its rapid introduction of higher bitrate third-generation cellular networks may pave the way for i-mode to become the global standard for wireless applications and services.
- **Competition from other Technologies.** In this category are GSM-specific innovations such as the SIM Application toolkit and the Mobile station application execution environment (MExE), and also handheld operating systems such as PalmOS, EPOC and Windows CE. All these technologies are rapidly evolving to incorporate features to enable wireless applications. For example, Microsoft has devoted considerable resources to enabling application development for the Windows CE platform.
- **Limitations in the current release of WAP.** There are some serious limitations in the current release of WAP including lack of support for multimedia and inadequate security features. The next release of WAP (version 2.0) is scheduled for June 2001 [13] and it is supposed to address some of the current shortcomings of WAP. Notably, it will be based on xHTML, with backwards compatibility to XML.

Whether in the long-term WAP will succeed in becoming the standard for wireless applications standard worldwide is very uncertain. What is certain though is that WAP will continue to be, at least in the short-to-medium-term, a leading contender to become that global standard.

References

[1] WAP Forum (11 January 2001). *WAP Update 2001.* http://www.wapforum.org/new/WAP2001Update.ppt

[2] Vinton G. Cerf (1991). *A Brief History of the Internet and Related Networks.* Internet Society, History of the Internet. http://www.isoc.org/internet/history/ cerf.html

[3] Telecordia Technologies NetSizer (March 2001). *Evaluating the size of the Internet.* http://www.netsizer.com

[4] International Telecommunication Union (June 2000). *Telecommunications Indicators Update.* http:/ /www.itu.int/journal/200006/E/html/indicat.htm

[5] Craig McTaggart and Tim Kelly (29 May 2000). *Background Issues Paper.* ITU IP Telephony Workshop, 14-16 June 2000 http://www.itu.int/osg/sec/spu/ni/iptel/ workshop/iptel.pdf

[6] GSM Association (12 February 2001). *More Than 200 Billion GSM text messages forecast for full year 2001.* http://www.gsmworld.com/news/press_ 2001/press_releases_4.html

[7] WAP Forum (2001). *WAP Forum web site.* http://www.wapforum.org

[8] WAP Forum (2001). *WAP Forum Specifications.* http://www.wapforum.org/what/technical.htm

[9] National Geospatial Data Framework (March 2000). *Discovery Metadata Guidelines: Version 1.2.* http:// www.ngdf.org.uk/Metadata/metguide/metaguide12.pdf

[10] FedEx (2001). *FedEx Wireless Solutions.* http:// www.fedex.com/us/ebusiness/wireless

[11] Conversa (25 July 2000). *Conversa speech recognition engine ported to Phone.com WAP-compatible microbrowser* http://www.conversay.com/Company/7_25_00.asp

[12] Kei-ichi Enoki (2001). i-mode, *IEEE VTS News*, 48 (2) (this issue), p. 4–12

[13] Wireless Developer Network (2001). *An Interview with Scott Goldman, CEO of the WAP Forum.* http:// www.wirelessdevnet.com/channels/wap/features/ goldman.phtml

Mjumo Mzyece (mjumo@ieee.org) received a B.Eng. in Electronic and Electrical Engineering from the University of Manchester, England in 1996. From 1996-1999 he worked in the area of networks and communications at a copper mining conglomerate and an Internet service provider. After completing an M.Sc. in the Department of Electronic and Electrical engineering at the University of Strathclyde, Scotland in September 2000, he joined the Department's Mobile Communications Group to pursue a Ph.D. on link quality control techniques for packet-based mobile radio networks.