Is the WAP Window Closing?

WAP, the Wireless Application Protocol, has been heralded as a major enabler of m-commerce (e-commerce over mobile phones) because it was designed to suit the limitations of the mobile display and the low data transfer rates. We are already seeing substantial advances in display capabilities, however, and this will be matched before long by significant improvements in data rates. WAP sounds great, but will it be enough when mobile terminals can access data faster than today’s PCs and match their display performance?

Introduction

While it is perfectly possible to transmit or receive data using a digital mobile phone such as GSM, it is slow, and the phone itself is not the end device (unless the message is a purely textual Short Message Service (SMS) limited to 161 characters per message). A laptop or PDA would usually be employed as the originator or receiver of such data. This necessity to interconnect two devices makes data transmission less convenient on the move, and the major application of internet access is impractical due to the large quantities of data combined with impractically low speeds.

The premise underlying WAP is that terminals are limited in their functionality and WAP/WML seeks to exploit or overcome (depending upon your point of view) these limitations.

The need for WAP

It is worth reflecting on the circumstances that led to the creation of WAP so that we can better understand the impact of the new technology.

Digital mobile phones such as GSM and CDMA (Code Division Multiple Access - the US 2nd generation digital cellular system) offer digital speech communication from and to the handset. With a digital infrastructure established it made sense to include the capability of carrying data (supplied by some external device) through the phone and into the switched network, providing users with the opportunity of sending and receiving files even when away from the office. In addition, SMS offered the capability of transmitting short textual messages in the same way as pagers except that SMS offered two way messaging.

While these data capable mobile networks were being built, a large growth in the use of the Internet was experienced, with almost all firms and many individuals “surfing the net” for information and setting up websites to announce their presence in the new global market. The PC, which is the prime terminal for connecting to the Internet, is usually connected via a modem providing at least 28.8kbps. With a big screen and full keyboard, the PC can display the full glory of the most artistically creative web pages including graphics, animated images and photographs.

Today’s mobile phones are seen as a logical device for delivering data services for several reasons.

- they have an established data connection to telecommunications networks which are the first point of connection to the Internet for most people, and therefore provide the logical route for data services
- they are mobile, enabling access to data services from anywhere
- many people already have a mobile phone, and will not want to carry a separate device around with them to acquire data services

However, handset design currently centres on the primary purpose for mobile networks, namely speech. By reducing both the quantity and the complexity of the data, it is possible to utilise the limited functionality of existing handsets and give the user access to some parts of the Internet.
WAP is the protocol by which the mobile phone accesses specific sites on the Internet, and handsets need specific "micro-browser" software installed to enable them to search for WAP enabled sites. The downside of creating WAP content on a web site is the need to create two versions of the same thing - HTML for normal browsing and WML, the cut down version for WAP.

The handset limitations that created the need for WAP are set to be reduced or to disappear with increasing data rates delivered to the handset (first from GPRS - General Packet Radio System, and then UMTS - Universal Mobile Telephone System) and considerably improved displays. With the dramatic increase in the rate at which data can be delivered to the terminal, is it reasonable to expect the terminal to remain constrained by its display or keyboard? With PDAs (Personal Data Assistants) already utilising colour, touch sensitive screens capable of displaying full VGA width, it won't be long before they are equipped with wireless communications. In the not too distant future, a protocol that deliberately limits terminal functionality will surely be inappropriate when terminal functionality exceeds today's PCs.

Where does all this leave WAP?

Is current handset design good enough?

While data rates are limited to 9.6kbps there is little point in increasing the handset display or keyboard capability beyond what they can already do. In Japan a new service called i-mode has been introduced which offers a similar service to WAP i.e. a method of connecting to websites from a mobile handset. The display provided on i-mode handsets is larger and clearer than typical 2G handsets, but this is largely to accommodate the graphical text characters that constitute the Japanese written language (they can only display 8 Japanese characters at a time!) The major difference between WAP and i-mode is the ability of the i-mode handset to handle a compressed version of normal HTML pages rather than the special WML pages. However, to be useful for i-mode, these pages must in practice have restricted content, to fit in with the limited displays and slow data speeds.

As long as the data rates remain slow, there will be little incentive to embellish mobiles with high-quality displays. This would imply that apart perhaps from styling, handset design is more than good enough for today's technology.

But for how long?

However, increase the handset’s data rate capability significantly, and the need for a better display arises.

The bottleneck of delivering data is reduced or removed, and suddenly the handset can now manage large graphical images or send large text files. Now the bottleneck is the display and keyboard. Great progress has been made with PDAs where colour touch screens and small but fully functional keyboards have been developed (the Psion 7 already offers a full width VGA screen and full “qwerty” keyboard). These could easily be combined with handset technology to create the personal terminal.

When the communication data rate matches this display and keyboard capability, the limitations of the handset that drove the development of WAP will have disappeared - and so therefore will the need for WAP.
When personal terminals can just as easily handle HTML, the same language the PC uses to communicate with Internet WebPages, why have a cut down version limiting the functionality? Developments in Markup languages will improve the data latency (and hence reduce download requirements) and perhaps allow terminals to display data in the way best suited to them. xML (eXtensible Markup Language) is going some way to achieving this.

So WAP will only be needed until greater bandwidth in the communication channel and increased functionality in the personal terminal are widely available. But when is this likely to happen?

GPRS will give GSM the ability to deliver at least 28.8kbps, probably more eventually, and is being launched now (though there are few handsets available yet). It is quite likely that GPRS equipped handsets will not have greater display and keyboard capability, and WAP will benefit from the increased data rates in the short term. However, PDAs have become very sophisticated and when the first PDA equipped with a GSM/GPRS chipset becomes available, which cannot be far away, will WAP's days be numbered?

Things only get worse for WAP after that. UMTS (third generation digital mobile telephony) will be with us in two years time. While coverage will be patchy at first (though by the end of 2007, 80% of the UK population will have access to UMTS) these phones will deliver much higher data rates (typically 144kbps in most areas, considerably higher than this in some areas). While again they may initially be handsets with simple displays, it will only be a matter of time before the PDA, and the UMTS handset join forces. This may occur in two ways. A direct union as described above with GPRS, or perhaps more likely, through the use of Bluetooth. Bluetooth is a short range digital radio connection, designed to be cheap enough to be built into almost anything and providing sufficient bandwidth to nicely complement UMTS (see our companion White Paper, "Bluetooth as a 3G enabler"). The PDA will be separate from the UMTS terminal, but in direct contact with it through a Bluetooth link. The delivered bandwidth will be considerably better than that currently achievable via landlines and the display and keyboard as good as a PC. Using today's PC software fast surfing of the Internet will be possible without the need for WAP middleware!

What may keep the window open?

It will be noticed that the argument presented above relies on the ubiquitous use of high-end handsets or the common deployment of Bluetooth enabled PDAs. One of the consequences of the high prices paid for UMTS licences however is that operators will be very keen to broaden their market reach and we can expect a very rich variety of services targeted at different market segments.

Some of these segments will want a cheap and cheerful small single device to meet their information service needs. These segments will continue to benefit from a WAP solution and, although xML can be expected to replace WAP in the fullness of time, it could well be many years before that happens. Unless these segments are large and offer significant revenue opportunities for the operators however, the impetus to maintain the WAP oriented services will wane. Once the decline begins the mainstream WAP decks will disappear quickly, and WAP will be remembered as an important stepping stone in the path of personal communications.