1 The mobile Web landscape

Mobile Web services are introduced in this chapter. We describe the type of services that mobile users are likely to invoke when connected to the wireless Internet and the browser-based access that characterizes how these services are most often used. We proceed with a description of the associated service delivery models that explain how these services can be developed and deployed. There is significant evidence from the NTT DoCoMo network operator experience that shows the success of an "open" model where the mobile operator provides a framework and environment in which third party content developers can deploy their services. The major stakeholders in the wireless Internet ecosystem are examined next. They include the mobile user, the enterprise, the mobile network operator, the content provider, and finally the mobile network and handset manufacturer.

To support the "open" service development model, the underlying networks need to provide a set of services that are easily accessed by wireless service developers. We describe the concept of a mobile network operating system and the value that Web services technology can bring to bear in the realization of this model. We proceed with an introduction of the concept of context, elaborate on its potential to create personalized services, review the pertinent elements of context, and describe how context is collected and stored in the network. Many standard bodies are contributing specifications that will enable the realization of the mobile Web, and we list the major organizations and their stated objectives. Finally, we provide an introduction to XML and XML Schema, the languages used in most of the example code listings in this book.

1.1 Mobile Web services

The currently deployed new generations of mobile networks (2.5G and 3G) that support data services provide an always-on connect capability, with typical data transfer rates of about 50 kilobits per second (kbps) for 2.5G and about 144 kbps and higher for 3G, and, most importantly, their user pricing models are either based on volume of sent packets or else on a flat monthly fee. These mobile networks extend the wired Internet and allow users to have access "anytime and anywhere" to the same information that they are relying on in their home or office. This is a significant change from previous generation mobile networks that were circuit switched with dial-in access, provided low data transfer speeds (typically 9.6 or 14 kbps), and required users to pay a "per minute" connect time fee.

With the wireless Internet, unique information services can be offered which were not available before. Defined user groups with very specific information needs can be easily accessed in real time. For example, music buffs can receive on their cell phone notification from their music store about the current availability of their favorite music CDs, they can reserve a copy of a CD, and they can pay for them with their phone. Similarly, there are surfer sites (for example, Surfline [1]) that can provide up-to-date information on surf conditions along the coast and enable surfers to head to the right beach. The Surfline site provides streaming video of various coasts in the USA and other places in the world. These streams could be sent to mobile users equipped with phones capable of displaying video so that they can decide for themselves whether the surf conditions are optimal.

Mobile Web services address a wide gamut of user interests. These services can be partitioned into distinct categories, some examples of which are:

- General information
  - news feeds;
  - stock quotes;
  - weather reports;
  - horoscopes;
  - online dictionary.
- Travel
  - car navigation;
  - traffic reports;
  - train connections check;
  - airline ticket reservation;
  - flight status check.
- Entertainment
  - restaurant reservation;
  - movie theater reservation;
  - games.
- ecommerce
  - mobile shopping;
  - online banking.
- Connectivity
  - email;
  - finding friends.

Access to these services and their associated content is typically performed through a Web browser. On the wired Internet, information access is performed via Web browsers such as Microsoft's Internet Explorer and Netscape's Navigator that are able to display content that has been marked up with formatting tags. Hypertext Markup Language (HTML) is the primary markup language used to annotate Internet content for browser display. On the wireless Internet, two major browsers used in mobile terminals include the Wireless Application Protocol (WAP) browser specified by the Open Mobile Alliance (OMA) and the i-mode browser specified by NTT DoCoMo. The WAP browser can display mobile content in Wireless Markup Language (WML) and Extensible HTML (XHTML), and the i-mode browser, used mainly in Japan, can display content in Compact HTML (cHTML), a subset of HTML. Dual mode browsers enable access to both WAP and i-mode content.

Over 400 million handsets support the Openwave Systems WAP browser as of July 2003 [2]. Openwave shipped the world's first WAP 2.0 browser during 2001 in Japan, where it is the second most popular browser. Successful mobile data services such as Sprint PCS Vision, KDDI "au", and Vodafone live! are based on WAP technology. Over 20 million Vodaphone live! subscribers, as of July 2004, were using a WAP browser while Japan's DoCoMo had over 40 million subscribers, as of October 2003, accessing the Internet via an i-mode browser [3]. Besides the Japanese handset manufacturers such as NEC, Panasonic, and Mitsubishi, as of 2003 other manufacturers such as Samsung (DoCoMo phones), Siemens (S55 mobile phone), and Nokia (3650 handset), include an i-mode browser in select mobile phones.

As in the wired Internet, the wireless user can leverage a search engine to find relevant services and content within the available wireless Web pages. For example, Google has a wireless search service [4] that is offered by wireless operators that include Sprint, AT&T Wireless, Cingular, Qwest, Nextel, Vodafone, and others. Google's search technology can be accessed from any number of terminals, including mobile phones, PDAs, and the Pocket PC. The Google service searches through the over five million WML pages created for wireless WAP terminals. It can also expand its search to the over three billion pages of the World Wide Web. Every viewed page is translated on the fly to text and all search results are displayed in text-only format that fits the terminal screen. The i-mode service has similar search engines, for example, to answer user queries the Oh!New? engine will search Compact HTML pages in i-mode sites.

1.2 Mobile Web service delivery

With the breadth of information and service types that wireless users wish to access, no mobile operator can provide all the required solutions. In the old telecom model, the operator was responsible for providing network communication features and these were rolled out in the network in successive versions of the network software. In voice networks, these features included facilities such as call forwarding, and voice mail.

This was a network-centric model where intelligence was provided in the operator's network servers.

In the Internet model, intelligence has migrated to the edge of the network, and is provided mainly by third parties that specialize in specific content offerings. Network operators could adopt the "transport system" model, where they don't own or manage content and only provide transport services. The drawback of this model is that communication prices become the only tool for customer retention, putting significant pressure on operator revenues. An alternative model is for the operator to take responsibility to develop, aggregate, and deliver mobile content. The operator's portal site links to a select set of third party content providers with whom the operator has established licensing agreements. In this "closed" environment, the operator has sole responsibility for deciding what content is offered on its portal, and the content providers' customers are the operators, not the end-user subscribers. This model was adopted until recently by most operators outside of Japan [5]. The evident drawback of this model is that no incentives were in place for content providers to develop new and exciting services, as they had no way to charge the end user for any improvements. In some cases, the content providers had to pay the operator for an entry on its wireless portal site. Mobile subscribers complained about the lack of variety in services, which in turn led to limited service use, so that ultimately the operators experienced a limited growth in wireless Internet usage.

Mobile operators such as NTT DoCoMo, have made conscious decisions to become a coordinator and facilitator for service delivery and not to perform Web content development or even not to purchase content from Internet companies [6]. DoCoMo chose to become a portal site operator and facilitate user navigation to find useful content. The i-mode portal menus contained 3600 sites as of June 2003. DoCoMo established a very elaborate procedure for evaluating content providers and deciding which would appear on the official i-mode portal. This role is not unlike the one of wired Internet providers such as AOL. AOL too, does not create content but rather provides a platform for content delivery which is accessed via its portal. In this latter service model, the operator provides a platform, the wireless network, for service delivery and service management functions (see Figure 1.1). Content providers can rely on the billing services provided by the operator to track usage and bill the users. A corresponding fee is paid to the operator; for example, a nine percent fee is levied in the case of DoCoMo from subscriber fees paid to content providers.

While mobile operators have chosen to provide official site menus, they allow users to access the broader Internet as well. The user has to key in the Universal Resource Locator (URL) of a specific site he/she wishes to reach: not a very convenient task on a phone device. Nevertheless, it is interesting to note that the number of sites that are accessed outside the operator's portal far exceeds the number of operator provided sites. DoCoMo counted 66 400 i-mode sites, as of June 2003, which were not listed in its menus. These sites usually provide dedicated content for a small number of users that would not justify a slot on i-mode's official menu. The mobile service provider's

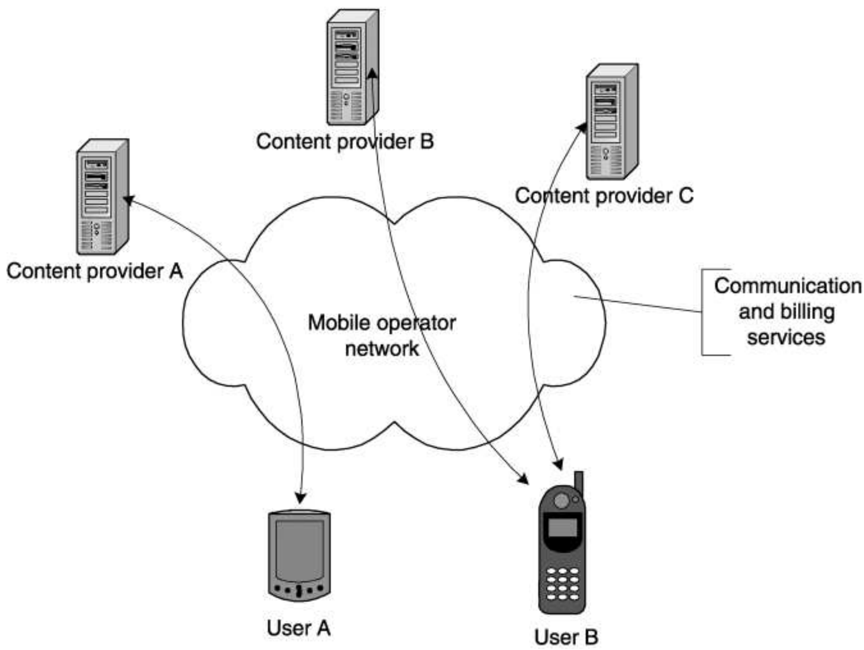

Figure 1.1 Mobile operator: bridging between content providers and mobile users.

revenue from wireless communication charges to these unofficial sites can be very significant and even exceed the average revenue per user (ARPU) from official sites, as in the case of DoCoMo [5].

Early success of wireless Internet use occurred in Japan. The penetration of cellular phones is higher than that of personal computers, one of the reasons being the high access cost to wired Internet services. The cost of DoCoMo's i-mode wireless Internet access is by contrast a small fee (about $3) per month, a content service subscription fee (between $1 and $3, depending on the service), and then there is a low per packet cost (a packet is 128 bytes, that is, 1 kilobit). As a result, i-mode is a major service of Internet email, online banking, ticket purchasing, game-playing, and peer-to-peer video messaging and gaming, with over 40 million users since 2003. There is a variety of Internet access plans available to mobile subscribers in the USA. For example, the Cingular GSM operator charges a fee of ($0.01 per 1 kilobyte), with the cost decreasing with prepaid monthly plans.

1.3 The mobile Web stakeholders

The parties that will benefit from the growth of the mobile Web include, first, the end users, and then, the enterprises that will be able to provide new connectivity capabilities to their employees, the mobile network operators that will deploy the enabling networks,

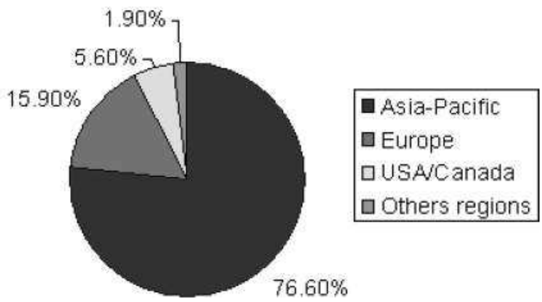

Figure 1.2 Mobile data subscribers by region.
(Source: © 2003 EMC. All rights reserved.)

the content providers that will reach their customers through the wireless Internet channel, and finally the mobile network and handset manufacturers that will build the support network elements and deliver enhanced handsets that will allow for an enjoyable end-user experience.

The mobile user

Future visions for the wireless Internet user have been articulated by software companies such as Sun and Microsoft, and adopted by mobile operators such as NTT DoCoMo. Significant growth is expected in specialized mobile services such as driving directions, traffic reports, tour guides, and any location-based services such as mobile shopping. In related user scenarios, the mobile terminal becomes the "network computer" and the "wallet PC". It is both an information device and a tool for enacting transactions based on the retrieved information. For example, users will use their mobile terminal to search and order products to be delivered at a nearby store, and, once at the store, they will use emoney to pay for the products with their mobile terminal.

As of September 2003 there were over 100 million wireless Internet users, with the majority in Japan and Korea, while fast growth rates were experienced in Europe. Figure 1.2 shows the distribution of mobile data users in 2003.

The enterprise

In the quest to enable employees, for example, sales staff, to be as productive when out of the office as when they are in the office, many enterprises support technologies and services that allow for "anywhere and anytime" connectivity to office information sources. This type of connectivity is made possible through wireless Internet solutions that leverage Web services [7] in the enterprise's domain. Web services provide a standards-based framework for integrating an enterprise's applications. They can supplant any existing middleware, enable programmers to interweave a company's applications as well as applications provided by external services, and reduce the reliance on proprietary solutions to implement application integration solutions.

The mobile network operator

Mobile carriers often have competed on subscription prices and offers of more minutes or free minutes during off-peak hours. The carriers differentiating themselves on service offerings will drive the future. In the voice world, many operators are now attempting to replicate Nextel's success with its unique Direct Connect voice service that establishes, at the touch of a button, a walkie-talkie like radio connection with one or more users for the exchange of short voice messages. This service simplifies connectivity for workgroups that often need to be in touch with each other, and can be used by groups of non-work related users, for example, a family on a vacation trip.

In the wireless Internet space, operators can differentiate themselves by the data services that they provide. Operators may choose to develop their own unique content, purchase specialized content from third party providers, or else provide a platform environment, referred to as a mobile ecosystem, for delivery of services offered by external parties. NTT DoCoMo, for example, chose the latter option and is providing billing services as well as a portal site for mobile service access. On the other hand, virtual network operators or operators of special purpose networks that operate a vertical market system may choose to have closer control of the content they provide to their subscribers.

In the "open" mobile ecosystem environment, the network operator can provide service management facilities such as service usage tracking and customer billing. Some research studies estimate that an operator's revenues from data services could be at least double those that are achieved in a "closed" model [5]. User ownership is perceived to be a key ARPU driver. To prevent erosion of value through service openness, the network operator can own the subscriber by managing the user's identity, the user's personal information (for example, content preferences), network presence, and user location.

NTT DoCoMo has pursued its goal to expand i-mode's "open" model in markets outside Japan. Through investments in non-Japanese operator shares, DoCoMo has formed alliances with operators to offer the i-mode service in markets that include Taiwan (KG Telecommunications), the Netherlands (KPN Mobile), the USA (AT&T Wireless), Spain (Telefonica Mobiles), and others, and as of July 2004 there were over three million i-mode users outside Japan [3]. DoCoMo sets up licensing agreements with mobile operators to provide them with patents, technologies and the know-how necessary to launch the i-mode service. DoCoMo has also played a major role in helping operators establish close ties with content providers and equipment vendors, which is essential in order to build a strong foundation for the i-mode service.

The content provider

A mobile user accesses business functionalities via Web services offered by content providers. A Web application may often combine multiple Web services, so that the application appears as one integrated business function. For example, a number of Web services used in a supply-chain architecture can be presented as a single application such as in the case of a product ordering system. Similarly, a restaurant reservation service can operate in conjunction with a restaurant review service, and a credit card processing system. Content providers could provide the Web services components or else take on their integration and offer the resulting comprehensive applications.

The mobile network and handset manufacturer

Networks that enable mobile terminals to access Web content require network elements with dedicated functionality to support the data transmissions, screen content generation, and contextual awareness of user situations, for example, their location. Corresponding support is needed in the handsets. Prior generations of mobile terminals provided small text-oriented displays with a few short lines of text that did not present attractive interfaces for information display via the installed micro-browsers. The newer mobile terminals have larger graphics-oriented screens, many with color, which enhance the wireless experience and entice users to access the Internet more often. Some operators come up with the specifications of the phones they wish to offer their subscribers. For example, NTT DoCoMo placed orders for phones with specific requirements with Japanese handset manufacturers. These phones include rich feature sets that support the goal of making the wireless Internet access experience effortless and enjoyable.

1.4 Mobile Web networks

The future of the mobile Internet will depend very largely on the ability of the wireless industry to deliver valuable services that are easy to use. Some of the major challenges for this success include the establishment of a standard user identification method, financial transaction standards for mobile commerce, and easy inclusion of third party service offerings in mobile networks. The ubiquitous availability of operating systems such as Windows and Mac OS, have facilitated the tremendous growth of application offerings on the desktop computer. A similar growth in mobile applications will occur as network operators open their networks to third party service providers.

1.4.1 The mobile network operating system

Mobile network operators could consider themselves as a component of the "mobile network operating system" where they provide services such as user location and user identification that are leveraged by the mobile applications. The other components of this mobile network operating system include the user mobile terminal and the application server that hosts the mobile applications, each providing its respective operating system services. For example, the mobile terminal provides a browser user interface, and the application server provides services for generating markup that is ultimately displayed by the mobile terminal's browser. There are additional service components as well, and these will be described in future chapters.

Growth of the mobile applications market will be assured when application developers will have straightforward access to all service components of the mobile network operating system: the user terminal capabilities, the network operator services, and the application server features. Operators will be able to differentiate themselves by the type of services that they provide to application developers. For example, if an operator has unique services that track the mobile user's context (e.g., whether the user is currently driving a vehicle), then the application provider will be able to better adapt an application's behavior to the user's changing context. This ability can directly affect the "ease of use" factor that determines whether users will adopt wireless Internet services.

The mobile network operating system will consist of a collection of the above services that are implemented as Web services and can be called by application developers. Web services were promoted as a set of shared protocols that enable disparate systems to talk to each other. These protocols include a Web Services Description Language (WSDL) that describes a service's interfaces, a Simple Object Access Protocol (SOAP) for transporting XML messages, and a Universal Description, Discovery, and Integration (UDDI) directory for storing information about Web service offerings. Web services can replace less flexible methods for information exchange, such as EDI (Electronic Data Interchange), that are used for the exchange of specific transaction data. With Web services multiple WSDL interfaces can be defined for accessing a service, and multiple clients can make use of the provided access methods.

1.4.2 Web services proposition

It is worthwhile revisiting the original motivation for Electronic Data Interchange (EDI) [8] to better understand how its Web services [7] successor enables building mobile network operating systems. EDI became a standard under the X12 American National Standards Institute (ANSI) committee. By December 2000, this standard defined over 300 transactions for the exchange of electronic documents, such as purchase orders, invoices, health care claims, or requests for proposals. Each EDI transaction consists of the transmission of a set of data segments, which are related data elements. For example, date and time are expressed in a separate data segment. With the definition of the electronic data forms exchanged between trading partners came an improvement in the overall business processes. The improvement was mainly due to the reduced intervention by humans. However, EDI is focused on the data forms being exchanged, and whenever these forms change, updates are required to the supporting software. Furthermore, the EDI transactions are not directly integrated with a trading partner's host applications. Separate software is needed to extract business data from an EDI transaction and make it available for use by other business applications.

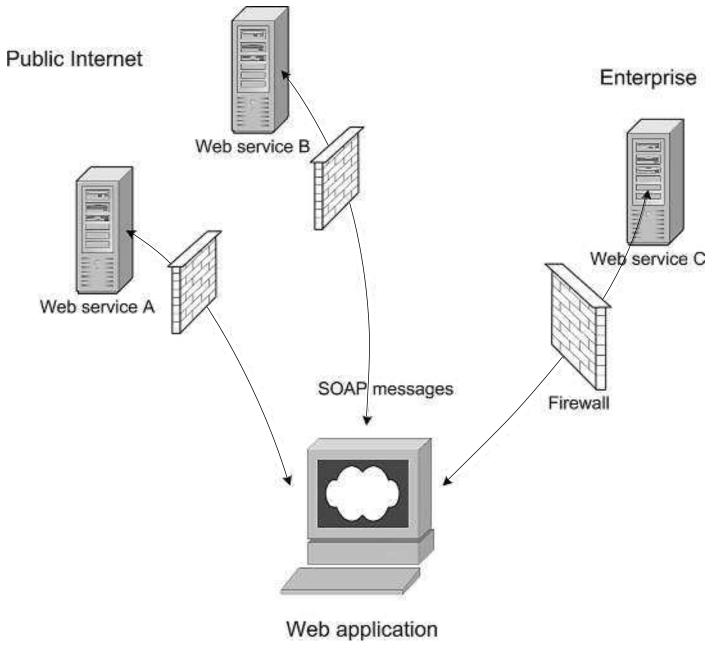

Figure 1.3 A Web application: integrating Web services with SOAP messages.

Unlike EDI, Web services provide a procedure call-like interface where services can be invoked in a way similar to the remote procedure call (RPC) of programming languages. Web services enable the building of software applications that execute on the Internet and use the same software paradigms that were successfully applied in the development of enterprise applications. For example, object-oriented development approaches provide for the reuse of software components, the objects that are the application's building blocks. A Web service is like an object that can be reused by multiple Web applications. A Web application is therefore considered to consist of collaborating Web services. To collaborate, these services rely on standard interfaces in the form of SOAP messages that convey requested operations and are transported on top of the standard HTTP protocol (see Figure 1.3). This interface is both a strength, as messages can then cross enterprise firewalls, and a weakness, as security needs to be enhanced to disallow malicious SOAP messages.

The reusability of Web services components and the relatively easy integration of Web services through standard interfaces are a very potent proposition for Application Service Providers (ASPs). ASPs were not very successful in the past as they had to spend large amounts of their resources on customizing their applications to specific customer needs and integrating them with pre-existing enterprise applications in the customer's premises. The license fees charged by ASPs covered only 20 percent of the application's deployment cost [9]. The ASPs did not realize that customization and integration could amount to as much as 80 percent of an application's implementation effort.