Communications of the Association for Information Systems

Volume 12 Article 13

August 2003

Delivery of Personalized and Adaptive Content to Mobile Devices:A Framework and Enabling Technology

Dongsong Zhang
University of Maryland, Baltimore County, zhangd@umbc.edu

Communications of the Association for Information Systems

DELIVERY OF PERSONALIZED AND ADAPTIVE CONTENT TO MOBILE DEVICES: A FRAMEWORK AND ENABLING TECHNOLOGY

DONGSONG ZHANG
Department of Information Systems
University of Maryland, Baltimore County
zhangd@umbc.edu

ABSTRACT

Many innovative wireless applications that aim to provide mobile information access are emerging. Since people have different information needs and preferences, one of the challenges for mobile information systems is to take advantage of the convenience of handheld devices and provide personalized information to the right person in a preferred format. However, the unique features of wireless networks and mobile devices pose challenges to personalized mobile content delivery. This paper proposes a generic framework for delivering personalized and adaptive content to mobile users. It introduces a variety of enabling technologies and highlights important issues in this area. The framework can be applied to many applications such as mobile commerce and context-aware mobile services.

Keywords: personalization, adaptation, wireless networks, mobile users, content delivery

I. INTRODUCTION

With the advancement of wireless technology and widespread use of mobile devices such as cell phones, personal digital assistants (PDAs), palms, and pocket PCs, many innovative mobile applications are emerging. Wireless technology refers to the hardware and software that allows transmission of information between devices without using physical connections. It allows people to work at the most suitable location and obtain information at anytime. For example, a sales person can check inventory status and place orders through a PDA; people can view the balance of their bank accounts or receive daily stock information via their mobile devices. In the past few years, mobile applications involving multimedia content (e.g. audio and video) attracted more and more attention. For example, Nokia MMS (Multimedia Messaging Service) Center is a high-capacity service platform targeted at GPRS (General Packet Radio Service)/3G(3rd Generation)[1] and 2G operators who offer new interactive and creative multimedia messaging services.

[1] A list of acronyms explaining each of the abbreviations used in this paper follows the references.

Through the MMS Center, subscribers can compose messages containing multimedia content such as text, photo images, voice, and video clips.

One of the challenges for mobile information systems is to take advantage of the convenience of handheld devices and provide appropriate information to the right person in a preferred format. Because users' information needs and preferences differ, they would like to obtain personalized content to satisfy their individual needs. In addition, some unique constraints of wireless networks and devices such as limited bandwidth, unstable network conditions, small display screens, and limited processing power pose serious challenges to mobile content delivery. Personalization may include:

- removing data that are not of users' interests in downloading,
- converting images to smaller representations,
- determining an appropriate resolution of images or a compression rate of video in order to adapt to current network conditions [Steinberg & Pasquale, 2002].

Although the importance of personalization and adaptation in mobile content delivery is increasingly recognized, research in this area is still at its early stage. This paper presents a generic and extensible framework for personalized and adaptive content delivery to mobile devices, and introduces a number of enabling technologies and approaches.

II. APPLICATIONS OF MOBILE COMPUTING AND CONTENT DELIVERY

MOBILE WORKERS

Consider a manager traveling to a series of meetings with potential customers. While on the road, he would like to be kept informed of news related to his business, as well as background information about people and companies he will be meeting. In addition, he would also like to get information related to his travel, including local maps, car rental and transportation information, and directions to the hotel and meeting place. All this information can be delivered to his mobile device when it is needed. Mobile devices give field professionals access to information and service once limited to office-based workers. It is estimated that nearly half of a company's workforce will be mobile by 2004 [Thomas, 2002]. Wireless access to information by mobile workers helps move office tasks to the field.

MOBILE E-COMMERCE (M-COMMERCE)

The next phase of electronic business growth will be in the area of wireless and mobile e-commerce [Varshney and Vetter, 2002]. Many new business opportunities and trends are emerging. Gartner Group estimates that, by 2004, at least 40% of business-to-consumer (B2C) e-commerce will be initiated from smart phones supported by wireless application protocol (WAP) [Haskin, 1999]. The development and use of mobile e-commerce (m-commerce) enable customers or business partners to conduct business activities and commercial transactions through wireless communication networks that interface with mobile devices [Tarasewich et al., 2002].

Initial mobile commerce applications aim at the business-to-consumer market [Heijden, 2002; Tarasewich et al., 2002]. Mobile B2C applications include financial services (e.g., banking, brokerage, stock quotes and portfolio tracking, payment), advertising, and information-oriented services. Mobile computing also provides an opportunity to exploit mobile devices for routine B2B activities such as purchasing and bidding. Varshney and Vetter [2002] propose a framework for m-commerce, which consists of four layers:

- m-commerce applications,
- user infrastructure,
- middleware, and

- network infrastructure.

Design of mobile applications needs to consider the general capabilities of user infrastructure (mobile devices). Mobile middleware provides a uniform user interface while hiding details of underlying wireless networks. The network infrastructure also plays an important role in m-commerce because the availability of infrastructure resources and capabilities affect the perceived service quality by users.

In general, the benefits of applying wireless technology to m-commerce include:

- providing new channels to help reach customers and obtain greater control over inventories,
- reducing costs in both money and time,
- increasing customer satisfaction and revenues through the faster responsiveness and timely introduction of new products and services to customers, and
- enabling companies to redesign workflows and task assignments previously constrained by wired processes, thus reducing cycle times for activities such as placing orders and delivering services.

M-SERVICE

The Internet is becoming a vehicle of services rather than just a repository of information. Business is becoming characterized by loosely coupled organizational components [Iyer et al., 2003]. This trend lowers the risks associated with a rapidly changing market by modularizing marketing, production, and delivery and support services. The term web services is being used to refer to loosely coupled, reusable software components that semantically encapsulate discrete functionality and are distributed and programmatically accessible over standard Internet protocols. Web services provide a standard way for heterogeneous systems to share and exchange information and for applications not only to describe their function and data, but also to interface with other applications dynamically. M-services refer to those web services that can be consumed using mobile devices via a wireless protocol such as WAP. They aim to develop secure applications for mobile devices in order to bring convenience usage to customers.

Mobile web service requirements include:

- dynamic web service interoperability that adjusts in real time to match web service availability,
- mobile device capabilities,
- personal preferences of users, and
- wireless network bandwidth.

Both Verizon Wireless and AT&T offer mobile web services. In July 2000, Verizon Wireless, launched an Internet service called “Mobile Web” for its more than 25 million subscribers [Duggan, 2000]. The Mobile Web’s services include local information, search engine functions, shopping (through Amazon.com and Barnes & Noble), organizer functions, and e-mail. In April 2002, AT&T Wireless launched its mMode consumer web service in the United States [Chidi, 2002]. The GPRS enables data transfers up to 144 kbps. The mMode service includes features such as instant messaging, content downloads (news stories or sports scores), and the ability to connect with personalized data like address books and calendars.

M-services need to consider features of the wireless devices on which they will be running. Pre-installing services on users’ wireless device is an option. However, finding the common denominator in an open environment with different users and needs is not trivial. Therefore, on-demand delivery of services to wireless device is more appropriate than pre-installing services [Maamar et al., 2002]. It is desirable for a wireless application to consume dynamically composed services.

CONTEXT-AWARE APPLICATIONS

Applications for mobile users should take advantage of contextual information, such as location or orientation, to offer greater services. Enabling mobile devices and applications to adapt automatically to the changing environments will enhance the user experience. The term ‘context’ can be defined as “any information that characterizes a situation related to the interaction between users, applications, and the surrounding environment [Dey and Abowd, 2001].” It is typically the location, identities of nearby people, objects, and changes to the objects.

The use of context in mobile applications is receiving increasing attention [Gellersen et al., 2002]. A typical mobile context-aware application is a personal tour guide. For example, museums can provide mobile devices and allow users to take personalized tours, enabling them to go anywhere they like and receiving information about the location where he/she is. Such systems may require several components: map component (knowledge of the physical surroundings), information component (keeps all the information about the sites that a tourist should visit), positioning component for finding where the tourist is, and messenger for the tourist to send or receive information. Another example is mobile conference information systems that are designed to provide location-based information to conference attendees [Okoli et al., 2002]. It facilitates the conference attendees to access conference proceedings. The program updates can be broadcasted to mobile devices of attendees to reflect last-minute changes. People can also use the system to find conference rooms or set up meetings with others.

In context-aware applications, some information such as maps/directions is available on the Web. Therefore, the wireless application should provide access to the information on the Web, which is delivered on demand. One of the difficulties of using context in mobile applications is that there is no common way to acquire and handle context. Some researchers propose multi-sensor context-awareness approaches [Gellersen et al., 2002]. Each individual sensor just captures a small aspect of an environment. Then, the captured data is combined to generate a bigger picture to characterize a situation.

MOBILE BATTLEFIELD INFORMATION SYSTEMS

Obtaining information about the situation on the battlefield in a timely manner is critical to commanders. A portable, wireless battlefield administration tracking and information system can store and transmit information via hand-held and mobile devices to a centralized data collection point, where the corresponding decisions can be made to respond to the current situation [Impson et al., 2001].

III. A FRAMEWORK FOR PERSONALIZED AND ADAPTIVE CONTENT DELIVERY TO MOBILE DEVICES

WHY PERSONALIZATION AND ADAPTATION?

One definition of personalization is “the use of technology and customer information to tailor electronic commerce interactions between a business and each individual customer” [Personalization Consortium, 2003] . In this paper, we define personalization as

> *the use of technology and user information to customize multimedia content in order to satisfy the needs of individual users.*

It is achieved by using information obtained either previously or in real-time about the user, his/her interest and preference, properties of mobile devices, and wireless network conditions.

In general, the content transmitted to and from mobile devices is unique because of the limitations imposed by wireless networks and mobile devices. A number of operational factors that distinguish wireless networks from wired networks, including [Macker, 2001]:

- Less bandwidth

- Increased likelihood of channel interference
- Low reliability (high data loss rate)
- More frequent topology changes (e.g. due to node mobility and resource failures)
- Higher transmission latency
- Lower physical security of media

These limitations of wireless networks pose challenges to developers of wireless applications. For example, multimedia data transmitted in a wireless network may encounter serious delay or data loss when the available bandwidth is reduced.

Some users may prefer to accept low quality data with short delay when the network traffic is heavy. To serve users with different network capabilities, the applications must be able to adapt to network changes (e.g. throughput). This requirement is critical for mobile multimedia applications because the multimedia content, especially audio and video, demands a higher peak bandwidth. Hence, wireless applications should be network-aware to be able to adapt to network variations.

Most mobile devices are constrained by small screen size, restricted input methods, lack of color, less memory, and limited power. The small screen size renders Web pages confusing and cumbersome to look at, while less memory and limited power suggest that applications should avoid intensive computation on mobile devices. Since device properties and capabilities differ, it is more effective if the content is customized and presented in a manner that best suits a particular type of device.

In addition, users' interests and information needs differ. Some users are interested in the latest stock quotes and foreign exchange rates, while others are interested in weather or sports news. It is desirable for wireless applications to deliver appropriate content to users' mobile devices to satisfy their individual needs. In the meantime, irrelevant information should be either blocked or removed from the content sent to the user. Users may prefer specific media. For example, some users may prefer a video clip or an image rather than a short text message, although it results in longer delay. When network traffic is high, some users may prefer to receive a low-resolution video or even only the sound track of the video instead of original high-resolution video in order to shorten the latency. Therefore, wireless applications must determine how to customize the content based on individuals' preferences and network conditions.

THE FRAMEWORK

The framework is shown in Figure 1. It was developed to define a generic and extensible architecture that enables personalization and adaptation of multimedia content delivery and presentation for mobile users. Each element of the framework is now discussed.

Mobile Browser-Enabled Devices and WAP

Users may, for example, want to view the latest news or scoreboard of a live sports game provided by the CNN Web site on their mobile devices. The WAP enables wireless devices to access the Internet by using proxy technology. The WAP proxy provides functions including: [WAP Forum, 2001]

- WAP gateway
- Content encoders and decoders: the content encoder converts WAP content into a compact format in order to better utilize the limited bandwidth of wireless networks.
- User agent profile management: create and maintain user profiles describing client capabilities and personal preferences.
- Caching proxy: improve perceived performance and network utilization by maintaining a cache of frequently accessed resources.

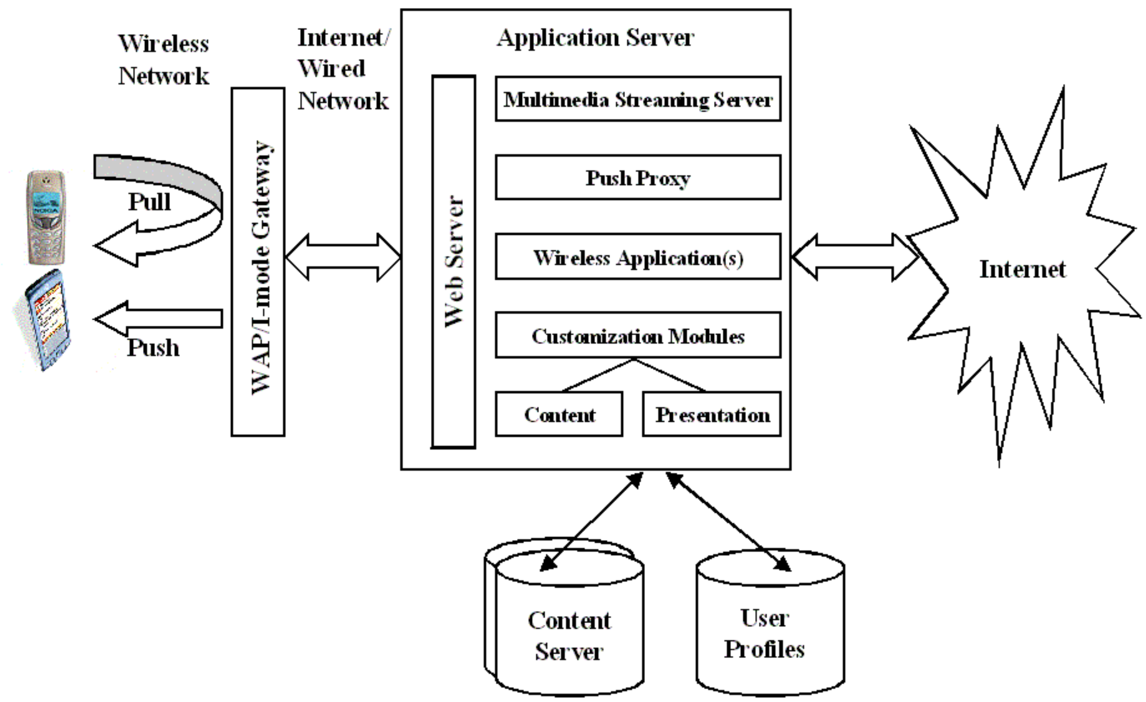

Figure 1. A Framework for Personalization and Adaptation of Multimedia Content in Wireless Environments

Wireless Markup Language (WML) acts as a page description language within WAP [Tarasewich et al., 2002]. It is intended for use in specifying content and user interface for narrowband devices. The WAP architecture also includes supporting servers, which provide services to proxies, mobile devices, and applications as needed. For example, a Public Key Infrastructure (PKI) portal allows devices to initiate the creation of new public key certificates. Another successful service offering wireless web browsing and e-mail from mobile phones called i-mode, produced by NTT DoCoMo, works with specifically designed Web sites.

A mobile device contains a micro browser and gains access to the Internet through access points (AP). Recent major advances in mobile devices became driving forces for a variety of new wireless applications. These advanced features include:

- Ready-to-go wireless solutions

  Many new models of mobile devices, such as Compaq iPAQ h5555 pocket pc, Sony CLIE NZ90, and Tungsten™ C from Palm, Inc., enable users to access Internet, email (e.g. using Microsoft Outlook), and corporate data via the integrated Bluetooth™ technology and WLAN (Wireless Local Area Network) 802.11b Wi-Fi wireless capability. The Wi-Fi technology that operates in the 2.4 GHz range can offer data speeds up to 11 megabits per second. It connects mobile users virtually anywhere with Wi-Fi access and enables them to download emails or browse websites with efficiency.

- Multimedia capability

  Despite the limited bandwidth of wireless networks, some new mobile devices are equipped with advanced multimedia capability, which can potentially provide the power of digital video and photography. For example, the Sony PEG-NZ90 CLIÉ handheld features a built-in two mega-pixel camera in support of still pictures. In addition to a built-in MP3 audio player, it also allows users to record and playback video in MPEG-4 format and store it on the Memory Stick®.

- Enhanced Business-Critical Security

  Many business services are being extended to mobile users. As a result, securing data and communications on mobile devices becomes a critical issue. Some mobile devices provide enhanced business-critical security features. For example, Compaq iPAQ h5555 pocket pc integrates a biometric fingerprint reader to protect data via fingerprint-recognition, VPN (Virtual Private Network) software, F-Secure FileCrypto data encryption, and other mechanisms.

Pull and Push

Two models for wireless applications provide information/services to the user with a mobile device: pull and push. Delivery of content can be done using SMS, email, or WML documents. In the pull model, the mobile client requests specific content from the server. The server finds the relevant content either from local content servers or from other sources including the Web [Cohen et al., 2002]. For example, applications may use Internet search engines to search for information. Finally, the identified content is sent back to the client.

In the push model, a server delivers content to the mobile client without receiving a request from him/her. In other words, this model allows content to be automatically sent or “pushed” to mobile devices by applications via a push proxy[2]. Push functionality is extremely relevant to real-time applications that send notifications to their users, such as news, stock price, new product information, and traffic updates. Without this functionality, clients must inquire application servers for new information, which increases the network traffic and delays just-in-time information acquisition.

The push model can be used for multicasting information. Rather than sending a message to every user (i.e. broadcasting), multicasting sends messages to a selective group of mobile users who share common interests. For example, a wireless application can use intelligent agent technology to collect the latest information about the stock market automatically from websites of the Financial Times or the Wall Street Journal under agreement, and then process and multicast it to a specific group of users. Providing support for wireless multicast communications is one of the challenges m-commerce presents to service providers and application developers [Varshney, 2002]. To achieve this goal, it is necessary to manage group membership. Applications need to create user profiles and a group list, and update them as participants change their interests or move from one service area to another due to the user mobility.

WAP/I-mode Gateway

The gateway mainly translates requests from a wireless protocol stack (e.g. the WAP 1.x stack) to the WWW protocols (e.g. HTTP and TCP/IP). It also performs DNS (Domain Name Service) lookups of the servers specified by the client in the request URLs.

---

[2] See above for a discussion of proxies

Application Server

The application server is the core of the framework. In addition to wireless applications, the computer that serves as the application server may also contain several other servers including the Web server and multimedia streaming server, although those servers can reside on other server machines. In addition, two customization modules perform personalization and adaptation to the content and its presentation.

Multimedia Streaming Server

Some multimedia information, such as a video clip of a new movie or an image of a product, is more attractive and informative to the customer than a text. Such multimedia information is especially helpful in mobile business applications such as sales promotion and product introduction. Streaming media technology is widely used in Web-based application systems. In comparison to traditional downloading method, streaming technology enables users to play an audio or video file while it is downloading rather than to play it after the entire file is downloaded to the local client, which significantly shortens the latency.

Streaming services require sufficient bandwidth to ensure the bit-rate needed by each media stream and the strict delay variation needed to avoid buffer underflow at streaming clients. Studies on streaming media in wireless environments include Curran and Parr [2002], Goose et al. [2002], and Yoshimura et al. [2002]. Curran and Parr [2002] propose a framework for the transmission of streaming media to mobile devices. Their solution involves exploiting mobility prediction to predict a mobile terminal's future location based on its previous history. Goose et al. [2002] propose a 3D audio extension to SMIL (Synchronized Multimedia Interaction Language) and a server-based framework able to receive a request and process such an SMIL file, create multiple simultaneous audio objects, then to multiplex them into a single audio, and prepare the result for transmission to a mobile device. In general, the selection of the streaming server partly depends on the multimedia format(s) that mobile devices support. Yoshimura et al. [2002] present a mobile CDN (Content Delivery Network) architecture for streaming media, in which content segmentation, request routing, pre-fetch scheduling, and session handoff are controlled by SMIL modification.

Content Server and User Profiles

The content server can be a database server that holds enterprise data such as business transaction data, customer information, and inventory data. It may also contain the metadata of multimedia content (e.g. metadata of video, audio, and images). The user profiles contain well-organized information about individual users and their mobile devices, which is then used for content personalization and adaptation.

Network Infrastructure

Network infrastructure includes the wireless network (e.g. WLAN), wired network, and the Internet communication. The first-generation wireless networks, which emerged in the early 1980s, were analog, voice-centric mobile systems. The second generation appeared in the 1990s, featuring digital mobile, circuit-switched networks. Time Division Multiple Access (TDMA), Global System for Mobile (GSM), and Code Division Multiple Access (CDMA) are the basic second-generation digital wireless telephone technologies. The 3G wireless networks offer broadband, packet-based transmission of multimedia providing date rates of 2 Mbps.

Currently, the most widely available and implemented WLANs comply with the 802.11b standard. The recent availability of 802.11a and access points, however, introduces what is becoming another mainstream wireless LAN solution. In contrast to 802.11b radios, which can transmit at 2.4 GHz and send data up to 11 Mbps using direct sequence spread spectrum modulation, 802.11a radios transmit at 5 GHz and send data up to 54 Mbps using orthogonal frequency division multiplexing with shorter-range coverage.

A few companies started developing practical 4G technologies and global standard. For example, on May 28, 2003, NTT DoCoMo, Inc. announced plans for a field trial of a fourth-generation (4G) mobile communication system. In indoor experiments announced in October 2002, DoCoMo's 4G system demonstrated maximum information bit rates of 100 Mbps for the downlink and 20 Mbps for the uplink. It can support much faster transmission of multimedia content than previous generations.

IV. APPLICATION SERVER: CONTENT PERSONALIZATION AND ADAPTATION

In the proposed framework, in addition to wireless applications, the application server may also hold a push proxy that initiates the push operation, the Web server, multimedia streaming server(s), and two customization modules for personalization and adaptation.

In the pull model, after receiving a request from the client, the applications first determines the potential sources (e.g. enterprise database or Web) of the requested information. In the push model, the applications can use intelligent agent technology to collect certain types of information automatically. The personalization and adaptation are performed based on user profiles and network conditions before delivering the relevant content to the mobile devices.

CONTENT PERSONALIZATION BASED ON USER PROFILES

A user profile is a record of user-specific data. In addition to personal information interest and preference, the user can also define different situations and specify behavior desired in those situations in the user profile. Therefore, a user profile may include, but not limited to, the following information:

- User information

  User information includes user ID, background information, personal interest represented by either keywords or information/service categories, preferences (e.g. media preference, summarization method, and priorities among data items).

- Target device information

  The limitations of mobile devices exercise constraints on the display of information. For example, the small screen size can only display a few lines of text. The device information can specify, for example, how many lines of text are in the display, how many characters per line in the display, and the resolution of the screen. Another reason for maintaining profile information of a type of device is to manage the session of disconnected clients, since disconnection can occur during any stage of a business process for reasons such as network time-out or network failure, low battery, and local memory constraints [Sairamesh et al., 2002].

- Service profile: service restrictions and user availability.

- Wireless network information: network ID, topology, and configuration.

Knowing user profiles allows wireless applications to manage users' needs for information, deliver information in a predictive manner (profiting from the WAP push mechanism), and track clusters of information dynamically. Push-based information delivery gained popularity as a way to relieve the information-hunting burden of users [Cetintemel et al., 2000]. Push depends on knowing user interests for making scheduling, content customization, bandwidth allocation, and routing decisions. Managing user profiles and providing the right information in an appropriate format is one of the core issues on the application server side [Sairamesh et al., 2002].

A number of users that share the same interests can form a user group. A user can be a member of several groups. In our framework, a central database server is used to manage all user profiles and can be accessed by all requesting applications. In reality, user profiles can be located at other sites [Bougant et al., 2003]:

1. User profiles are stored on a single centralized server – the only profile server. Whenever there is a data transfer request, each application needs to access the central profile server to retrieve information from the database. The advantage is that, because there is

only one central server, the profiles are easy to maintain. The disadvantage is that the workload of the profile server increases when a large number of applications try to access the profiles simultaneously.

2. User profiles are duplicated on several servers at different locations. The advantage is that it is manageable for different network architectures. The drawback is that frequent updates may bring network traffic congestion because all profile databases must be synchronized after the updates.
3. User profiles are divided into multiple subsets and stored on several databases. The profiles stored in each database are different. The advantage is that specific services can be managed for certain users. The problem is that it is difficult to separate user profiles into different sub-databases.
4. Each user profile is stored in the database on his/her client device locally. There is no profile server. The advantage is that it is easy to maintain the profile and profiles do not create workload on the server side. However, the wireless applications have to broadcast, instead of multicast, information to all registered mobile clients. Then, the information filtering process is performed through accessing local user profiles on the client device to determine if the received information is relevant or not. Because of the limitations of mobile devices, this approach is not a good choice.
5. A main server contains a root for user profiles, but the details of profiles are stored in several sub-databases. The basic idea is that the root provides a form of index for user profiles, which leads to a particular sub-database for further, detailed processing. In this method, the profiles are well organized and stored in different databases based on their semantics. The difficulty lies in how to create such a root.

User profiles for text-based data were extensively investigated in the context of information filtering and selective dissemination of information. Those systems use user profiles and information retrieval techniques to filter unstructured textual documents. Similarly, the push proxy can assess the overlap or similarity between a document and a user profile to determine if the document is of interest to the user. In the Boolean model, a user profile is constructed by combining keywords with Boolean operators. An “exact match” semantics is used. In the similarity-based model, the profiles and documents are compared with one another and the similarity is calculated based on the vector space model [Rijsbergen, 1979]. A document with a high similarity score is considered relevant to the user’s interest; thus it should be delivered automatically to the user. In a content-based multicast approach [Shah et al., 2002], filters are used to sift the irrelevant information. Only relevant information of a user’s interest is delivered. Filters are located at some nodes of a multicast tree. Each filter residing in a node contains personalized information (user profiles) of clients. Although filters bring exact information to the client, the computational process in the filter to sift out the pertinent information creates a delay. Therefore, locating the filters depends on two principles:

- minimizing the total traffic in a multicast path through calculating the estimated traffic load for every node in which the filter is located. The node with the least traffic load is selected.
- determining the number of filters that minimizes the delay of transmission.

User profiles can also be used for adapting the multimedia content. Because of bandwidth limitations, the user may experience a long delay, especially for audio or video data. The purpose of QoS (Quality of Service) management is to keep track of the ongoing QoS, compare it with expected performance, detect possible QoS degradation, and then tune network resources accordingly to sustain the QoS at a satisfactory level [Chen et al., 2000]. Therefore, some minimum QoS must be provided to support smooth audio/video playback in a wireless environment. The multimedia content should be adjusted dynamically based on the network condition (e.g. traffic) and media preference of the user. For example, when the network traffic is high, a video with a high-compression rate or even only the sound track of the video may be preferred to be delivered to mobile devices to avoid long latency. As a result, users can specify their preferences such as desired image resolution, video compression rate and frame rate, and preferred medium type under different situations in their profiles. iMobile developed by AT&T is a mobile service platform that authenticates users who send requests from various mobile devices [Chen et al., 2002]. The service platform transcodes video content based on user and device profiles. It addresses various issues like authentication, authorization, transcoding, adaptation, and deployment. It integrates 'iVideo' server, which manages actual video content delivery in the wireless network. The platform consists of a client side proxy and a video proxy. The client proxy decodes the incoming video stream and forwards it to the display units. The application layer monitors the decoded bit stream. The client maintains a running average of the streamed bytes and measures the download throughput. It then notifies the video proxy to adjust the frame rate accordingly. The client proxy is also responsible for gathering profile information and sending it to the server.

Some wireless applications or services aim to provide precise and concise information to users based on their interests at different times. The time dimension is involved because users frequently change interest or preference. For example, consider a business trip: a person who travels to another city for a conference needs a variety of information, such as direction to the hotel, flight information, and conference schedule, during her trip. However, it is usually not necessary to view all the information simultaneously. Before she leaves office for the airport, she wants to check the flight information; after arriving at the destination airport, she needs the directions to the hotel; when she attends the conference, she wants to know the conference schedule. For such time-dependent information acquisition, the user profile should not only specify the user's information interest, but also show the priorities and preferences of information [Cherniack et al., 2001], together with the time stamps.

PRESENTATION PERSONALIZATION

Although mobile devices are portable, their limitations pose challenges to content presentation. Different mobile devices have different screen size, colors, resolution, power, and other specifications. The small screen size can only display a few lines of text. Most pages on the World-wide Web, however, are designed for display on desktop computers with at least 640 * 480 resolution color monitors. As a result, direct presentation of most WWW pages on these small mobile devices is aesthetically unpleasant, un-navigable, and even illegible. Therefore, determining how to present content on mobile devices more effectively is critical to navigating the information and achieving user acceptance [Freire et al., 2001].

Four approaches are available to display Web pages on small screens. They can be classified according to authoring roles [Bickmore and Schilit, 1997]:

- device-specific authoring,
- multiple-device authoring,
- client-side navigation, and
- automatic re-authoring.

Device-Specific Authoring refers to authoring a set of Web pages for a particular type of mobile device. Users of a particular type of device will only have access to a selective set of services. The Web pages for these services can all be designed up-front for the specific device's display.

Multiple-Device Authoring identifies a set of target devices rather than a single device. A Web document is transformed into a number of rendered documents to cover the devices within the set based on a pre-defined mapping mechanism. One example is HTML cascading style sheets (CSS) that define a set of displaying attributes for different structural portions of a document. The user can specify a style sheet for content presentation. Another example of this approach is that portions of the document can be tagged with a level of abstraction measure. Users can specify a preferred level of abstraction they wish to view and are presented with the corresponding detail. Different levels of content abstraction are available for displaying content on mobile devices [Buyukkokten et al., 2001].

Client-Side Navigation. The user is given the ability to navigate a single web page interactively by altering a portion of it at any given time. For example, the user can use the scroll bar on the document display area, or use 2D fish eye visualization technique to display content [Clamage et al., 2003]. Active outlining techniques, in which the user can expand and collapse sections of documents dynamically, are implemented as a client-side navigation approach.

Automatic Re-Authoring: In this approach, an application will re-author any web document designed for the desktop via a series of transformations based on characteristics of the target display device, so that the document can be displayed appropriately. For example, the size of an image on a Web page may be reduced to fit a small screen better. For an original document that contains both text and images, the re-authored document that can display on the mobile device may only include some key text sentences, or only the most important image, or a shortened document that contains both. The user can specify such preferences in his/her profile.

Based on levels of content abstraction presented on the mobile device, approaches to displaying Web content can be categorized as follows [Buyukkokten et al., 2001]:

1. All – systems display the entire information at once;
2. Incremental – systems show only the first line of the content. The user can expand the line to display the rest if he is interested in seeing further details;
3. Keywords – systems first only show the keywords extracted from the content. Users can retrieve the details if those keywords are of their interests;
4. Summary – systems initially display summaries of content, and then show the details if users want;
5. Keyword/Summary - at first, systems show the keywords of content, then users can see the summaries if they are interested. Finally, users can retrieve the entire content in case they want to see all the details. As a result, the user profile can include the following pieces of information for personalized content presentation: the preferred displaying format(s) and/or abstraction level, and properties of his/her device.

The five approaches offer trade-offs. If systems directly display the original content, obviously users must use the scroll bar frequently to view it. It becomes even worse if the user finds that the content is not exactly what he wants after spending a lot of time to browse it. If systems provide a multi-level view of a Web page (e.g. keyword -> summary -> original document), by looking at keywords and/or summaries, users may be able to quickly identify the content of their interests for further details and avoid unnecessary perusing of irrelevant content if the keywords or summaries do not seem interesting to them. However, they need an additional step to retrieve and view the original content, which is stored either on the local device or on the application server. This process would either consume the extra power of the mobile device or increase the network traffic. Therefore, it is important for applications to determine which presentation scheme is more effective in a given circumstance. For example, if there are a large number of documents are returned to the user, then a multi-level view of documents may be a good choice.

CONSTRUCTING AND MAINTAINING USER PROFILES

Constructing and maintaining mobile user profiles manually can be a time-consuming and ineffective process. The purpose of building user profiles automatically is to identify the information that users like to access, so that mobile application systems can self-update user profiles.

User profiles can be managed and updated dynamically. A common approach to updating profiles automatically uses user feedback, which is typically used in information retrieval research. In this approach, the text documents that mobile users point out relevant and interesting are represented as a vector of term and weight pairs (Vector Space Model). Such vectors are merged and grouped to represent user interests. The grouping process can use clustering algorithms. Cetintemel et al. [2000] propose a multi-modal, self-adaptive profiling approach by using a clustering algorithm based on relevance feedback. Each profile vector is given a strength value. Users give feedback to the system when the information is displayed to them. The system then uses this feedback to adjust profiles. If the feedback is positive, the strength is increased; otherwise, it is decreased. When the strength value of a vector is smaller than a predefined threshold, this vector is considered irrelevant to the user's interest and is removed from the user profile. In this way, users do not have to update their profiles by themselves. A potential problem of the user-feedback approach is that not every user likes to provide feedback to systems frequently.

Many researchers also attempted to use data mining techniques to generate user profiles from user log data. The user browsing history is a good source to infer what users want and to derive user profiles. The commonly used data mining techniques for generating user profiles include association rule mining and clustering [Wu et al., 2001]. Another interesting issue is adaptive user clustering and profiling. Instead of responding to users' requests, it would be beneficial if the application system can predict users' future needs or actions, so that the application server can pre-fetch data and pre-process it. Phatak and Mulvaney [2002] employed both user and URL clustering information to improve completeness and relevancy. For each user, the system maintains a list of the closest URL clusters. When a user logs in, the system first retrieves links from the URL clusters he/she is closest to. Data mining techniques are also applied to automatically cluster mobile users into groups [Wu et al., 2001]. The documents that are represented using the vector space model can be clustered into a set of categories. The user's interest is represented as a list of (category name, count) pairs. When the grouped documents and user information are available, the system partitions all users into a number of common-interest groups.

CONTENT ADAPTATION BASED ON NETWORK CONDITIONS

Four critical network parameters affect the performance of multimedia content transmission [Fluckiger, 1995]:

- network throughput (i.e. bandwidth)
- transit delay (i.e. transmission latency)
- delay variation caused by the network congestion, and
- error rate

In this section we primarily discuss content adaptation based on network traffic. Two steps are involved in content adaptation to the network changes:

(1) monitoring the network and detecting changes, and

(2) making necessary adjustments to adapt to changes

As a process of collecting traffic information, network traffic monitoring is regarded as the initial step for a number of mechanisms of network control such as connection admission control and QoS maintenance [Fu and Venkatasubramanian, 2001]. The most common classification of existing network monitoring methods is based on the traffic generated by the method.

In **active monitoring** [Wu et al., 2000], network measurements are obtained by sending additional testing messages, which will inevitably introduce extra traffic to the network. By incurring the cost of extra traffic, active monitoring obtains more control in the monitoring process.

Passive monitoring techniques [Mao et al., 2001], in contrast, rely only on the traffic that applications generate as they communicate with other nodes in the network. The network status information is piggybacked on packets traversing the network, so it will not cause extra traffic in the network. Passive monitoring is more accurate than active monitoring because the actual user packets are measured. However, fast processing is necessary for measuring all the actual user packets securely.

A **hybrid monitoring** scheme is expected to be a better solution by combining the advantages of active and passive monitoring approaches [Landfeldt et al., 2000]. For example, active probing can be used only on demand when passive information is unavailable, such as when there are no open connections and a host wants to know the network conditions for setting up a new connection.

The bandwidth monitoring for multimedia traffic poses challenges because of the bursty and delay sensitive nature of applications.

- First, it is important but difficult to determine the sampling rate. A node in a wireless network obtains the traffic state information by sampling the network and stores the information in its local database. The traffic data collected via less frequent sampling may not reflect the real network condition. A high sampling rate will gather more accurate traffic information, but can also lead to high bandwidth consumption [Fu and Venkatasubramanian, 2001]. Most existing techniques usually take a simplistic approach towards parameter collection by using a monitoring module to frequently sample the residual capacity of a network link at fixed intervals and update the database. Adjustable sampling rates using statistical analysis techniques based on time-series have also been studied [Fu and Venkatasubramanian, 2001]. In this approach, an adaptive range is obtained such that the deviation between the predicted and actual values remains in the range with high confidence. Then, a bound on the sampling rate is determined. The parameter collection process adjusts the range and sampling rate dynamically based on the status of incoming traffic.
- Second, it is necessary to combine different solutions to the problems of adaptive bandwidth reservation and rerouting under traffic situations. Mobile agent technology seems to be a promising solution for wireless network management and QoS control [Manvi and Venkataram, 2002]. The agents can be used to monitor the traffic and link failures, and/or reserve bandwidth on the alternate route by interacting with the nodes in the path through spare capacity and reservation tables.
- Third, the fluctuation of network congestion is the main factor that harshly degrades the multimedia service. Many adaptive mechanisms are proposed to maintain the quality of service until the transmission is finished. Kwon et al., [2003], for example, propose a technique to ensure the quality of data transmission during the network traffic fluctuation over a time period. A QoS parameter, the cell overload probability (the probability of excessive number of data transmission in the network), is redefined. Then, a distributed call admission control (CAC) algorithm and bandwidth adaptation algorithms (BAA) are used to limit the cell overload probability. CAC aims to guarantee the boundary of cell overload probability. It will allow new data transmission if the current transmission can maintain its QoS and the new request can reach its desired QoS. BAA attempts to minimize the number of data transmissions with lower desired bandwidth. This approach mainly focuses on the bandwidth adjustment for existing data transmission if there is a new request.

Adaptation in networks can be implemented in three ways [Bolliger, 2000]:

Receiver-initiated Adaptation [Noble, 2000], where adaptation work is done by the receiver of the data. This strategy has a benefit of scalability, since servers do not need to be changed when there are new applications requesting adaptation. It is also more efficient because it understands the context of the object to be adapted, such as the level of importance of an image in a Web page to be delivered. Therefore, it can make better response time-quality tradeoffs for complex applications.

Proxy-based Adaptation [Mao et al., 2001]: Adaptation functionalities are implemented in a proxy between the sender and the receiver. An advantage of this approach is its transparent design. Neither the receiver nor the sender needs to be changed to support network adaptation. Therefore, it is scalable and cost effective. The cost of this transparency is that proxy has little knowledge about the receiver, the sender, or the context of the object to be adapted. This approach typically provides a few static adaptation policies, which is not very flexible [Bolliger, 2000]. In addition, the proxy adaptation approach must track the network condition of both the server-to-proxy and proxy-to-server connections to make adaptation decisions.

Sender-initiated Adaptation: Senders as the content providers have full control on the content quality and more computational power to make reliable adjustment on content quality. However, sender-initiated adaptation may lack scalability. Both proxy-based and sender-initiated strategies are fine for mobile multimedia applications, while the receiver-initiated adaptation strategy is the least appropriate, since mobile devices such as PDAs usually have very limited power and computing resources.

The key requirements of adaptation for mobile multimedia applications are [Mao et al., 2001]:

- automated data format adaptation without user intervention,
- graceful quality degradation,
- seamless handoffs across networks during roaming, and
- high QoS with low jitter, delay, and guaranteed bandwidth.

Different approaches can be applied to implement adaptation for mobile multimedia applications:

Multiple Encoding: servers can store multiple copies of the same image or audio/video stream with different encoding methods or parameters. Applications choose an appropriate copy dynamically based on current network conditions. Mahajan et al. [2002] propose an adaptive architecture based on user preference to increase the perceptual value of the multimedia stream. After the connection is established with a multimedia server, the mobile client periodically sends feedback about bandwidth availability to the server. The server stores several copies of streaming data (e.g. video) encoded at different fidelity levels. Based on the bandwidth feedback and user preference, the server dynamically determines an appropriate copy of audio/video stream. The user preferences are specified in terms of preferred QoS parameters such as resolution and frame-rate. This approach is the easiest way of adaptation for multimedia applications. However, this approach is limited in that multiple versions of the same multimedia stream with different parameters must be prepared and stored, resulting in much more processing time and storage space.

Transcoding refers to mapping a non-scalable stream to another non-scalable stream with different compression rate (e.g. MPEG to H.263). A higher compression rate results in a smaller data size. Transcoding avoids the necessity of storing multiple files corresponding to various bit rates [Steinberg and Pasquale, 2002]. However, this approach requires decoding and recoding of each media stream. Therefore, it is computation-intensive and requires deep understanding of different encoding and decoding schemes.

Layered Encoding (Transcaling and Multicasting) is the most popular way to provide adaptation for multimedia applications [Radha, 2001; Wu et al., 2000]. This approach usually generates a base-layer and one or more enhancement layers to cover the desired bandwidth range. By applying multicast, no separate delivery of streams is needed. Layers can be added or dropped by joining or leaving a multicast group. For example, if congestion is detected, the highest layer is dropped. Transcaling, on the other hand, is a generalization of transcoding. It derives one or more scalable streams covering different bandwidth ranges from another scalable stream.

Rate Shaping: The application can change the following media encoder parameters: image resolution, video frame rate, quantization parameter, and movement detection threshold.

Considering agility and stability in mobile wireless networking environments, a conservative adaptation policy is desired, since it can lead to a more stable operation [Margaritidis and Polyzos, 2000]. Users normally prefer lower but stable QoS to higher but varied and unpredictable QoS.

OTHER ISSUES IN MOBILE CONTENT DELIVERY

Routing

An ad hoc network is a collection of mobile nodes forming a temporary network without the centralized administration or standard support services regularly available on conventional networks [Curran and Parr, 2002]. Ad hoc routing algorithms include Destination-Sequenced

Distance-Vector Routing and Temporally-Ordered Routing algorithms. Statistical search theory is used by maintaining a history of prior known mobility patterns of users. Based on the prior information, it is possible to compute a vector of probability mass functions concerning the likely location of a target station [Curran and Parr, 2002]. Some mobile wireless routing approaches proactively maintain routes between all source-destination pairs (i.e. table-driven). In on-demand routing approaches, however, routes are established reactively to a given destination when needed (i.e. traffic driven) [Macker, 2001]. In those approaches, it is assumed that it is not necessary to maintain routes between all source-destination pairs at all times in a dynamic topology.

Resuming Multimedia Stream Delivery From A Disconnected Point

Compared with wired networks, wireless networks are much more unstable and unreliable. The transmission of a video or audio stream in a wireless network may be interrupted due to the disconnection between the mobile device and the network. If the system can capture the point of a multimedia stream when it is interrupted and continue to send the rest of the stream from that point rather than from the very beginning of the stream after the mobile device is re-connected to the network, it can save a lot of time and increase user satisfaction.

Security

In a mobile environment, anytime, anywhere access to mission-critical data by users with mobile devices brings security challenges to the mobile infrastructure. WLANs are generally more susceptible to eavesdropping than wired networks. Given the nature of WLANs, there can be a number of security exploits such as:

- insertion attacks when an unauthorized wireless user attempts to consume service at no cost;
- interception and unauthorized monitoring; and
- encryption attacks.

As a result, for sensitive business data, proper security mechanisms are required to protect the privacy and ensure the integrity of the data.

WLANs require secure access to an AP in a manner different from wired LANs. The mobile device and the user attempting to connect to an AP must be authenticated. The packets transmitted to an AP can be encrypted via Wired Equivalent Privacy protocol [Boncella, 2002]. PKI, smart card technologies, biometrics, digital certificates, or combinations of the above techniques have been used in security control for mobile applications. In mobile commerce, end-to-end security between the Internet and mobile devices is indispensable.

Security channels between mobile clients and the WAP gateway are supported by WTLS (Wireless Transport Layer Security), which provides server and/or client authentication via certificates similar to X.509 certificates. The channel between the WAP gateway and the Web server is supported by TLS (Transport Layer Security). Translation between WTLS and TLS is executed at the WAP gateway. A new secure layer can be implemented inside the wireless application environment to provide end-to-end security [Soriano and Ponce, 2002]. In mobile database systems, cross-domain authentication mechanisms (roaming agreements) will involve many complicated authentication activities, especially when a roaming path is long.

V. CONCLUSIONS

Given the growing popularity of mobile devices and mobile Internet access, many new wireless applications that aim to enable anytime, anywhere information acquisition emerged. Due to the unique constraints posed by wireless networks and mobile devices, those applications are highly desired to provide users with personalized content. This paper presents a framework for personalized and adaptive mobile content delivery, and discusses a variety of enabling techniques in details. The framework can be applied to many applications such as mobile commerce and context-aware mobile services.

We are currently implementing a prototype system for personalized mobile content delivery using J2ME (Java 2 Platform, Micro Edition) based on the framework, which is helping us understand the problems better and develop new technical solutions. Although it is challenging, we truly believe that this line of research can greatly increase both the effectiveness and efficiency of mobile applications and benefit mobile users.

Editor's Note: This article was received on June 18, 2003. IT was with the author for approximately three weeks for one revision. It was published on August __, 2003

REFERENCES

Bickmore, T.W. and B. Schilit (1997) "Digestor: Device-independent access to the Word Wide Web". *Proceedings of the WWW6 Conference*, Santa-Clara, CA, April 7-12.

Bolliger, J. (2000) "A Framework for Network-Aware Applications", *Ph.D. dissertation*, ETH Zurich.

Boncella, R.J. (2002) "Wireless Security: An Overview", *Communications of the AIS*, (9), pp. 269-282.

Bougant, F., F. Delmond, and C. Pageot-Millet (2003) "The User Profile for the Virtual Home Environment", *IEEE Communication Magazine*, (41)1, pp. 93-98.

Buyukkokten, O., H. Garcia-Molina, and A. Paepcke (2001) "Seeing the Whole in Parts: Text Summarization for Web Browsing on Handheld Devices", *Proceedings of the Tenth international conference on World Wide Web*, Hong Kong, May 1-5.

Cetintemel, U., M.J. Franklin, and C.L. Giles (2000) "Self-Adaptive User Profiles for Large-Scale Data Delivery". *Proceedings of the 16th International Conference on Data Engineering*, San Diego, CA, February 28 - March 3.

Chen, K.T., J. Y., and C.C. Ko (2000) "Monitoring QoS Distribution in Multimedia Networks", *International Journal of Network Management*, (10), pp. 75-90.

Chen, Y.-F. et al. (2002) "Personalized Multimedia Services Using a Mobile Service Platform", *Proceedings of the 2002 IEEE Wireless Communications and Networking Conference*, Orlando, FL, USA, March 17-21.

Chidi, G.A. (2002) "AT&T Wireless Launches Mobile Web Service in US", http://wireless.itworld.com/4256/ 020416attwireless/page_1.html

Clamage, A. et al. (2003) "A Fisheye Calendar Interface for PDAs: Providing Overviews for Small Displays", *Proceedings of the Annual Symposium and Open House (HCIL)*, College Park, MD, May 30.

Cohen, D. et al. (2002) "Personalized Pocket Directories for Mobile Devices", *Proceedings of the 11th International Conference on World Wide Web*, Honolulu, Hawaii, USA, May 7-11.

Curran, K., and G. Parr (2002) "A Framework for the Transmission of Streaming Media to Mobile Devices", *International Journal of Network Management*, (12)1, pp. 41-59.

Dey, A.K., and G.D. Abowd (2001) "A Conceptual Framework and a Toolkit for Supporting the Rapid Prototyping of Context-Aware Applications", *Human-Computer Interaction*, (16)24, pp. 97-166.

Duggan, E. (2000) "Verizon Hits the Horizon at Full Speed", *The Business Journal*, August 28, http://southflorida.bizjournals.com/southflorida/ stories/2000/ 08/28/story8.html

Fluckiger, F. (1995) *Understanding Networked Multimedia.* Englewood Cliffs, NJ: Prentice Hall

Freire, J., B. Kumar, and D. Lieuwen (2001) "WebViews: Accessing Personalized Web Content and Services", *Proceedings of the WWW10 Conference*, Hong Kong, May 1-5.

Fu, Z., and N. Venkatasubramanian (2001) "Adaptive Parameter Collection in Dynamic Distributed Environments", *Proceedings of the 21st International Conference on Distributed Computing Systems*, Mesa, AZ, USA, April 16-19.

Gellersen, H.W., A. Schmidt, and M. Beigl (2002) "Multi-Sensor Context-Awareness in Mobile Devices and Smart Artifacts", *Mobile Networks and Applications*, (7*)*, pp. 341-351.

Goose, S. et al. (2002) "Streaming Speech3: A Framework for Generating and Streaming 3D Text-To-Speech and Audio Presentations to Wireless PDAs as specified Using Extensions to SMIL", *Proceedings of the Eleventh International Conference on World Wide Web*, Honolulu, Hawaii, USA.

Haskin, D. (1999) "Analysts: Smart phones to Lead E-Commerce Explosion", all*NetDevices,* Nov. 3

Heijden, H. (2002) "Mobile Business Processes: Cases from Sweden and the Netherlands", *SSE/EFI working paper series in business administration*.

Impson, J., S. Kupst, and N. Mehravari (2001) "Proof-of-Concept Prototype of a Portable Wireless Battlefield Administration Tracking System", *Proceedings of the Military Communications Conference (MILCOM 2001)*, McLean, VA, USA, Oct. 28-31.

Iyer, B. et al. (2003) "Web Services: Enabling Dynamic Business Networks", *Communications of the AIS*, (11*)*, pp. 525-554.

Kwon, T. et al. (2003) "QoS Provisioning in Wireless/Mobile Multimedia Networks Using an Adaptive Framework", *Wireless Networks*, (9*)*1, pp. 51-59.

Landfeldt, B., P. Sookavatana, and A. Seneviratne (2000) "The Case for a Hybrid Passive/Active Network Monitoring Scheme in the Wireless Internet", *Proceedings of the ICON 2000*, Singapore, September 5-8.

Maamar, Z., W. Mansoor, and Q.H. Mahmoud (2002) "Software Agents to Support Mobile Services", *Proceedings of the AAMAS'02*, Bologna, Italy, July 15-19.

Macker, J.P., V. D. Park and M.S. Corson (2001) "Mobile and Wireless Internet Services: Putting the Pieces Together", *IEEE Communications Magazine, June*, pp. 148 - 155.

Mahajan, A., P. Mundur, and A. Joshi (2002) "Adaptive Multimedia System Architecture for Improving QoS in Wireless Networks", *Proceedings of the Pacific Rim Conference on Multimedia*, Taiwan.

Manvi, S.S., and P. Venkataram (2002) "Adaptive Bandwidth Reservation Scheme for Multimedia Traffic Using Mobile Agents", *Proceedings of the 5th IEEE International Conference on High Speed Networks and Multimedia Communications*.

Mao, Z.M., et al. (2001) "Network Support for Mobile Multimedia using a Self-adaptive Distributed Proxy", *Proceedings of the 11th International Workshop on Network and Operating Systems Support for Digital Audio and Video (NOSSDAV-2001)*, Port Jefferson, NY, USA, May 12-14.

Margaritidis, M., and G.C. Polyzos (2000) "MobiWeb: Enabling Adaptive Continuous Media Applications over Wireless Links", *Proceeding of the IEEE International Conference on Third Generation Wireless Communications*, San Francisco, California, June.

Noble, B. (2000) "System Support for Mobile, Adaptive Applications", *IEEE Personal Communications Magazine*, pp. 44-49.

Okoli, C. et al. (2002) "The Mobile Conference Information System: Unwiring Academic Conferences with Wireless Mobile Computing", *Communications of the AIS*, (9), pp. 180-206.

Personalization Consortium (accessed in Aug. 2003), "Personalization Information", http://www.personalization.org/personalization.html

Phatak, D.S., and R. Mulvaney (2002). "Clustering for Personalized Mobile Web Usage", *Proceedings of the IEEE International Conference on Fuzzy Systems*, Honolulu, HI, , May 12-17.

Radha, H. (2001) "TranScaling: a Video Coding and Multicasting FRamework for Wireless IP Multimedia Services", *Proceedings of the 4th ACM international workshop on Wireless mobile multimedia*, Rome, Italy.

Rijsbergen, C.J.v. (1979) *Information Retrieval*, London: Butterworths.

Sairamesh, J. et al. (2002) "Self-Managing, Disconnected Processes and Mechanisms for Mobile E-Business", Proceedings of the WMC'02, Atlanta, Georgia, USA.

Shah, R., R. Jain, and F. Anjum (2002), "Efficient Dissemination of Personalized Information Using Content-Based Multicast", *Proceedings of the 21st Annual Joint Conference of the IEEE Computer and Communication Societies (INFOCOM 2002)*, New York, June 23-27.

Soriano, M., and D. Ponce (2002) "A Security and Usability Proposal for Mobile Electronic Commerce" *IEEE Communications Magazine*, pp. 62-67.

Steinberg, J., and J. Pasquale (2002) "A Web Middleware Architecture for Dynamic Customization of Content For Wireless Clients", *Proceedings the Eleventh International Conference on World Wide Web*, Honolulu, Hawaii, USA.

Tarasewich, P., R.C. Nickerson, and M. Warkentin (2002) "Issues in Mobile E-Commerce", *Communications of the AIS*, (8), pp. 41-64.

Thomas, B. (2002) "IT Trends: Global Mobile Wireless Services 2002 to 2005", Giga Information Group, June 27.

Varshney, U. (2002) "Multicast SUpport in Mobile Commerce Applications", *IEEE Computer*, (35)2, pp. 115-117.

Varshney, U. and R. Vetter (2002). "Mobile Commerce: Framework, Applications and Networking Support", *Mobile Networks and Applications*, (7), pp. 185-198.

WAP Forum (2001). "WAP Architecture", http://www.wapforum.org

Wu, D., Y.T. Hou, and Y.-Q. Zhang (2000) "Scalable Video Transport over Wireless IP Networks", *Proceedings of the 11th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, London, UK, September 18-21.

Wu, K.-L., C.C. Aggarwal, and P.S. Yu (2001) "Personalization with Dynamic Profiler", *Proceedings of the Third International Workshop on Advanced Issues of E-commerce and Web-based Information Systems*, June 21 - 22, San Juan, California

Wu, Y.-H., Y.-C. Chen, and A.L.P. Chen (2001) "Enabling Personalized Recommendation on the Web Based on User Interests and Behaviors" *Proceedings of the 11th International Workshop on Research Issues in Data Engineering*, Heidelberg, GERMANY, April 1-2.

Yoshimura, T..et al. (2002) "Mobility and Wireless Access: Mobile Streaming Media CDN Enabled by Dynamic SMIL", *Proceedings of the Eleventh International Conference on World Wide Web*, Honolulu, Hawaii, USA, May 7-11.

LIST OF ACRONYMS

3G	Third Generation
AP	Access Point
BAA	Bandwidth Adaptation Algorithm
CAC	Call Admission Control
CDMA	Code Division Multiple Access
CDN	Content Delivery Network
CSS	Cascading Style Sheets

DNS	Domain Name Service
GPRS	General Packet Radio Service
GSM	Global System for Mobile
J2ME	Java 2 Platform, Micro Edition
MMS	Multimedia Messaging Service
PDA	Personal Digital Assistant
PKI	Public Key Infrastructure
QoS	Quality of Service
SMIL	Synchronized Multimedia Interaction Language
TDM	Time Division Multiple Access
TLS	Transport Layer Security
VPN	Virtual Private Network
WAP	Wireless Application Protocol
WLAN	Wireless Local Area Network
WTLS	Wireless Transport Layer Security
WML	Wireless Markup Language

ABOUT THE AUTHOR

Dongsong Zhang is assistant professor of Information Systems at the University of Maryland, Baltimore County. He received a Ph.D. in Management Information Systems from the University of Arizona. His current research interests include multimedia-based e-Learning, cross-cultural effect on computer-mediated communication, mobile computing, and applications of Web services. His publications have appeared or will be appearing in the *Communications of the ACM, IEEE Transactions on Multimedia, Information Systems Frontier, Journal of the American Society for Information Science and Technology, Information Resource Management Journal,* among others.

Communications of the Association for Information Systems

ISSN: 1529-3181