A Small Domain Communications System for Personalized Shopping Assistance

Abhaya Asthana and Paul Krzyzanowski

AT&T Bell Laboratories
Murray Hill, New Jersey, 07922, USA

ABSTRACT

By integrating wireless, video, speech and real-time data access technologies, a unique shopping assistant service can be created that *personalizes* the attention provided to a customer based on individual needs, without limiting his movement, or causing distractions for others in the shopping center. We have developed this idea into a service based on two products: a very high volume hand-held wireless communications device, the PSA (Personal Shopping Assistant), that the customer owns (or may be provided to a customer by the retailer), and a centralized server located in the shopping center to which the customer communicates using the PSA. The centralized server maintains the customer database, the store database and provides audio/visual responses to inquiries from tens of customers in real-time over a *small area* wireless network.

1. INTRODUCTION

Imagine that, as you walk into the shopping center, you take out a device, the size of a "walkman" unit, set it on your shopping cart and say "hello" into its microphone. The unit instantly responds back with a greeting, *"Hello, Jane. Welcome to Kings. Haven't spoken to you since Thursday of last week?"*, and thereon becomes your personal shopping assistant for as long as you are in the shopping center. The device will guide you through the store, providing you with details on products of interest via its built-in screen, pointing out the items that are on sale and doing comparative price analysis, reminding you of things that you were looking for on the previous visit that are now available, allowing you to find the current price of an item by scanning its UPC with a built-in bar code reader, helping you locate an item and, of course, playing your favorite music for *your ears only*.

The small device is only a portal to a complete personal shopping assistant service that can do much more than this, in making your visit to the shop pleasant, convenient and efficient. The scenario becomes even more powerful and intertesting in the mall setting with dozens of big and small stores.

The current methods of advertising in stores, that rely on public address systems or programmable LED displays in the aisles, are not effective due to the noise, distractions and the impersonal nature of the method. Shopping by catalog or from home (such as using Prodigy or TV-based home shopping networks) has great appeal but is not going to take over the world because humans are *nomads and like to roam around*; they enjoy going to malls, touching, feeling and meeting people. What they are frustrated about is where to finds things, determining the right price and the impersonal nature of service.

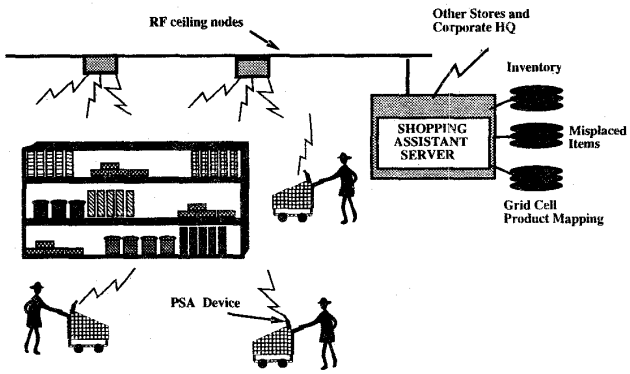

Figure 1. Personal Shopping Assistant Service

By integrating wireless, video, speech and real-time data access technologies, we can create a unique service that personalizes the attention provided to a customer based on individual needs and a shopping profile accumulated over time. This adds not only to customer's convenience and quality of service but can also be a source of direct and indirect revenue for the shopping center.

The service is based on two products. A very high volume hand-held wireless communications device, the PSA (Personal Shopping Assistant), that the customer owns (or may be provided to a customer by the shopping center) and a centralized server located in the shopping center to which the customers communicate using the PSA. The centralized server maintains the product database, the store database, customer profile and provides audio/visual responses to inquiries from tens of customers in real-time over a wireless network.

0-7803-1996-6/94/$4.00 © 1994 IEEE

Amazon/VB Assets
Exhibit 1010
Page 1

We have decoupled the PSA from the PSA server. We prefer to keep the architecture of the handheld PSA device as open as possible. We think of it not as a computing device but an information "portal", an input/output unit. Its functionality is essentially determined by the server with which it communicates. For the present, we will focus on its application for shopping rather than as a personal communication device. In a real product, these functions/services may be integrated. The same device may be used in different service domains by simply changing its operating mode.

The following basic functions and features characterize the PSA and are crucial to its user-interface:

- *Simplicity of Use and Hands-free Operation*: Very few keys (three to four) with a single level menu. Speech is the other input medium.
- *Locate Items*: A primary use of the PSA is to help the customer locate items.
- *Engage the Customer*: Another function of the PSA is to keep the customer's attention by engaging her/him in a somewhat friendly conversation.
- *Direct the Customer*: Help focus the attention of the customer to new product items, or those of particular interest to an individual customer.

The purpose of the PSA is to help the customer navigate within a store and inquire about items. In some cases, it may choose to use information derived from the customer's shopping history to provide relevant information to the customer, but it does not attempt to take on any personal attributes. It's goal is to be helpful and, as such, the services are to be removed from those of blasting commercials at the user. Since stores may maintain customer profiles, the privacy issues related to maintaining such customer profiles may also be a concern to some people. This is beyond the scope of the personal shopping assistant, but, to address this issue in this context, we divide the set of customers into two classes: *regular customers* and *store customers*. A *regular customer* is an anonymous one, a customer with no identity within the store. Such a customer will receive basic PSA services such as generic sales information, location tracking, product identification and inquiry services, and miscellaneous background services (such as music). A *store customer*, on the other hand, is one who signed up with a store. This is analogous to store discount cards popular in a variety of stores today. Customers will typically get additional discounts and the PSA server will maintain and use a shopping profile on the customer.

The PSA is a *shopping assistant* not a personal assistant. The term *personal* in the PSA refers only to the fact that the device allows communication between the store server and the customer on an individual basis. For a store customer such service can be personalized on a mutually agreeable basis.

2. ARCHITECTURE

The Personal Shopping Assistant (PSA) consists of two components: the PSA unit itself and the PSA server — a computer responsible for controlling all of the PSAs in its domain.

Several considerations dictate the design. There are constraints on the size, weight, power consumption and the frequency bandwidth needed for the handheld PSA device. At a minimum, the PSA must include functions for: analog radio frequency transceiver, baseband processing for communication (equalization, coding, packetization), image decompression, display unit, speech codec and a key pad.

Since the service is in a local domain many problems are simplified. Moreover, the nature of the application itself offers several opportunities for optimization. The information transfer in this application is asymmetric in nature. While the down link from the server supports high quality digital images (full motion video in the future), the bandwidth requirement for the reverse link from the PSA is considerably less.

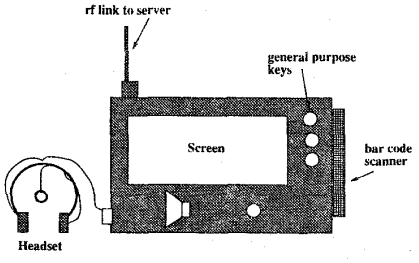

Figure 2. PSA Handheld Device

By design there is no computation performed in the portable unit; all such tasks are done in the server. This implies that data that is sensitive to noise will not be transmitted over wireless. So the system can operate even in an error prone transmission environment.

2.1 PSA

The PSA is truly a terminal for multimedia personal communication and not simply a notebook computer attached to a wireless LAN modem. The PSA unit is a

small, low-cost device about the size of a Walkman-type cassette player. In fact, it even resembles one in that it may be clipped onto the belt (or a shopping cart) and has a set of headphones with a directional microphone attached to it. It differentiates itself from PDAs and other small computers not only in its set of input/output interfaces but in that its functionality is determined by the server with which it communicates. The PSA box strives to present the user with the minimal possible interface. On the front of the box is a relatively large dot-matrix LCD display along with three control buttons (whose functions are specified dynamically by the PSA server). The side of the box has a laser LED which is used for scanning UPC codes. Another button on the front activates the scanner.

Thus, the only user outputs of the PSA are the LCD display and audio through the headphones. The user inputs are voice (via the microphone), bar code scans, and presses of the few control buttons. An additional, albeit hidden, input on the PSA is a receiver which listens to "location beacons". The PSA uses triangulation on these signals to obtain an approximate measure of its geographic location.

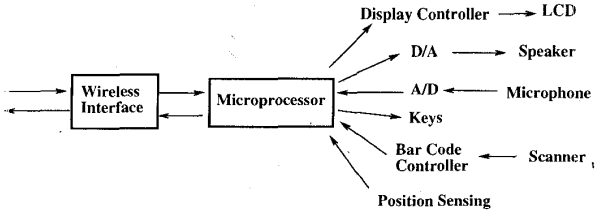

Figure 3. Block Diagram of the PSA

To keep costs low and functionality extremely flexible, the PSA by itself does nothing except relay information to and from the PSA server. Since the wearer may wander around untethered, the sole means of communication is wireless. A cellular system is employed, with each PSA server acting as a cell controller. If the wearer of the PSA wanders away from one controller such that the signal strength of another is stronger, control is switched to the new controller.

The structure of the PSA is shown in Figure 3. The PSA's microprocessor has the following tasks:

- processing the protocols involved in setting up a virtual circuit to the PSA server.
- performing compression/decompression on audio I/O.
- directing messages to the appropriate output devices.
- tagging messages from input devices properly before transmitting them to the wireless interface.

The PSAs have an open architecture and we envision several brand names would exist. A customer may own a PSA or it may be provided to a customer by the shopping center or mall. There are number of alternatives for handling the initial sign-on of a customer. It could be based on the PSA identifier, a customer identifier scanned via the UPC scanner, a speaker identification system, or smart card authentication.

2.2 PSA Server

The PSA server is the heart of the PSA system. It bears the responsibility for providing functionality to the hand-held PSA unit. As technology advances and PSA servers can be updated, the same PSA's power and usefulness will increase. As mentioned earlier, the communications structure of the PSA server mimics that of a cellular phone system. Establishing communications with the PSA server may involve speaking a phrase into the microphone, entering a customer identifier code through the UPC scanner or a combination. Once established, the PSA system may use that identifier to retrieve a "customer profile" from its database. The profile may be used as an aid to guiding the customer, for targeted marketing, or for extracting parameters for speech recognition. The customer cannot directly access the profile either for reviewing or altering. Only the PSA system can do that. We refer to a registered customer that has signed on as an *active* customer.

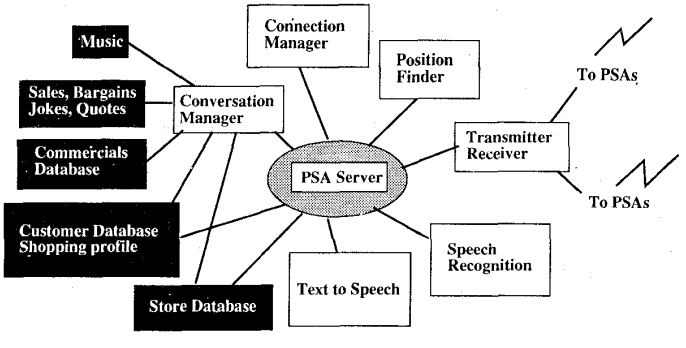

Figure 4. PSA Server Components

Figure 4 illustrates the functional architecture of the PSA server. The boxes are hardware or software functions that collectively provide the PSA service. The

Transmitter and Receiver represent the wireless network subsystem that sends and receives voice and images to and from the PSA units.

The *position finder* function keeps track of where an *active* customer is located (by means of the PSA's position broadcast). The *Speech Recognition* unit provides a limited vocabulary recognition words and phrases in near real time. The server uses *word spotting* to communicate with the customer. In this application the recognition accuracy is not an issue since no critical record updates are to be done. The idea is to engage the customer in a somewhat friendly conversation that is mostly controlled by the server with occasional responses or queries from the customer. If an error is made in recognition, it does not lead to catastrophe such as in process control, a financial loss as in a stock market service or an incorrect schedule placement such as in airline reservation system. The recovery here could be simply to repeat the question or change the subject or use a menu and soft button system to clarify the situation. We believe this conversational property of the proposed application can be exploited to our advantage in being able to initially work with a less than perfect speech recognition technology. The *Text to Speech* plays back the audio streams composed by the conversation manager for individual customers.

The system includes a number of databases. The *Customer Database* maintains the profile of every *registered customer*. This database is modified by the system based on interactions with the customer. The *Commercial Video, Audio Database* is a repository of audio and video strips that can be played on demand or using a directed sales schema for product promotion. The *Store Database* has the price and inventory of the items in the store. Based on customer inputs an order placement record can be created to facilitate ordering of items that are out-of-stock, high demand items, or items requested that are currently not stocked by the store. The *Sales, Bargains, Quotes, Jokes* database contains video and audio clips that the customer can scan or can be played by the system for his amusement or to promote sales. A *Music* unit pipes music into the customer's audio channel.

The *Connection Manager* controls the allocation of all video and audio channel bandwidth. The *Conversation Manager* is responsible for the entire session with a customer. It is the main process that invokes the services of other subsystems as necessary to serve a customer whose session it is controlling.

2.3 PSA Service Components

Because the growth in functionality is essentially unbounded, we will examine a few functions which the PSA and PSA server can reasonably handle with current technology.

2.3.1 Universal Product Code (UPC) Scan

The PSA is expected to be small enough to fit in a normal size palm. Thus, it may be picked up at any time to scan a UPC bar code. Such a scan causes the PSA to send a UPC query to the PSA server. The PSA server will query store's sales database and responds with the price for the scanned item along with other information deemed appropriate (quantity in stock, manufactures suggested price, product information). The response may appear on the display as well as through the headphone. The PSA controller can also keep a running total of all items purchased and query on whether the scanned item should be added to the tally.

2.3.2 Voice Recognition

The microphone input on the PSA enables the PSA controller to receive voice inputs. A customer may use the microphone to issue voice queries. In the simple case, these can be used to guide the customer to locating the product. Most stores deal with a limited domain of items (and those items are nouns, limiting the problem of recognition even further). As such, a high level of sophistication in speech recognition is not necessarily required. Moreover, the PSA wearer can use a control button on the PSA to skip to the properly recognized word through the guidance of the controller.

A sample dialogue may be:

Customer:
"Nintendo"
PSA Server responds:
"Nintendo games are in aisle 4.
Game cartridges are in aisle 5".
or
Customer:
"Peas"
PSA Server responds:
"Fresh vegetables are in aisle 15.
Frozen vegetables are in aisle 8.
All Green Giant products are 25% off today."

2.3.3 Audio Output

The headphones may be used for a variety of purposes. In addition to query responses, the PSA controller can offer the wearer a choice of background music (or news, or stock prices) to listen to while

shopping. The audio output can also be used for targeted marketing and providing information based on position.

For a more direct use of non-voice audio, consider the use of a PSA as a shopping assistant in an audio store. Currently some audio stores have a booth to which a shopper can bring a compact disk of interest, scan the bar code, and listen to previews. With a PSA, the process will be far more dynamic — the shopper will listen to previews on the spot (this will also be more convenient for the store: less misplaced CDs, no lines, happier customers). A more advanced PSA server can collect statistics on a person's buying patterns and recommend certain disks for previewing.

2.3.4 Display

The PSA's display dominates surface on the face of the PSA. It is primarily used to supplement the audio data that a customer is receiving and direct the user's choices. For example, when a product is scanned, the PSA can display the product information along with displaying identifiers for the programmable buttons. In this instance the buttons may take on the functions of "add item to tally", "delete item from tally", and "more information". Current technology makes it feasible for the PSA to have a high-resolution bit-mapped monochrome or color LCD display. Hence, in addition to textual information and prompts, the PSA can also display pictures of products or video clips. The CD-store scenario of the previous section can now be extended to a video store. The capabilities become limited by the available bandwidth (which can be expected to grow with the maturing of the involved technologies).

2.3.5 Databases

A customer can never update anything in the PSA server's databases. Thus there are no security issues. Many stores now maintain a rather extensive customer database. For example, food stores offer "savings club" cards which allows them to gather shopping statistics on a per-customer basis. The granularity of these databases makes them useful in highly directed marketing. Suppose you enter a audio/video store. Upon establishing communications with your PSA, the controller queries the customer database and sees you recently purchased a VCR. In greeting you, it can also announce "I thought you might be interested to know that BASF video tapes are on sale today".

Combined with location data from the PSA, the marketing can be even more directed. Let's examine a supermarket scenario. A customer would not want to be incessantly badgered having the entire sales flier recited to him upon entering the store. However, a list of often-purchased items can be sent to the PSA display and the customer may be alerted to examine that. While wandering through the store, the PSA controller may use position data from the PSA to find out which items are being sold in that area. Any of those items which are in common with the list of frequently-purchased items may fire a trigger to alert the customer:

PSA Server:
"Do you need to get milk?"
or
PSA Server:
"All Budget Gourmet entrees are 25% off"
or
PSA Server:
"Watch your step; this floor was just washed"

Since the customer has time to answer questions if asked casually, surveys can be conducted for tracking shopping patterns, dietary queries or product details. Story telling could be provided to keep kids entertained and occupied.

3. IMPLEMENTATION

The service builds upon some of the existing technologies such as wireless networks, image compression and VLSI technology, low power circuits, active memories, database, speech recognition and synthesis.

The wireless system employs a micro-cell architecture that minimizes overall system bandwidth consumption while supporting tens of users accessing simultaneously in a small domain. Through prototyping and actual experimentation we need to refine the architecture and protocol for the wireless network that is asymmetrical in its usage. The traffic in this application is characterized by short, word-oriented queries or responses from the user, and longer bursty audio and image traffic from the server. We expect the link from server to PSA to carry high volume, bursty voice responses for queries, steady image/audio traffic for sales/product information.

A production implementation of the PSA must be practical, low cost and power efficient. Eventually, for a low cost, low power, practical implementation, a VLSI solution for the PSA and the communication interface board on the sever end will have to be developed.

For our experimental prototype we are using off-the-shelf hardware components with a Plessey ARM610 microprocessor as the PSA controller and a network of SPARC 10s as the PSA server.