Multimedia Nomadic Services on Today's Hardware

Nomadic computing services can be built with currently available hardware. What is required is the integration, at the service and user-interface levels, of multiple media into a cohesive nomadic information infrastructure and a graceful transition from desktop to nomadic locales.

Chris Schmandt

How will subscribers actually use nomadic services? How might such services change the ways computer users connect remotely and which media they employ as a function of their location? We can begin to learn the answers to these questions without building new hardware devices or waiting for digital communications networks of the future, and by doing so help ensure that future nomadic devices will meet real user needs.

This article describes nomadic computing services built with currently available hardware and used on a daily basis by a small community of users. This environment consists of applications accessing pagers, facsimile, answering machines, telephone lines, speech synthesis, and digital recording and playback. Its key contributions are the integration of multiple media into a cohesive nomadic information infrastructure and a graceful transition from desktop to nomadic locales. This integration is at the service and user interface levels; the point of this article is that we can successfully prototype a suite of multimedia[1] nomadic services by cobbling together applications, shell scripts, and disparate hardware components.

Our nomadic services make strong use of speech to provide remote access over ordinary telephone lines to a suitable set of tasks that users normally perform at the desktop. We place particular emphasis on communication, spanning voice as well as text and facsimile messaging. In addition these services provide personal information management, remote database access, and short-term information caching. The link between nomadic and desktop computing also changes office applications — nomadic services utilize demanding new media and 24-hour access changes how subscribers use and rely on traditional desktop databases.

Using the existing telecommunications infrastructure offers several advantages. It allows us to develop services to meet the evolving needs of real users in a wide range of real locales, from on-campus hallways to home to airports and around the country or the world. It also helps to define services that will be attractive and should be supported by future nomadic hardware. Finally, it helps us glimpse how nomadic computing will influence work and communication styles, and its impact on a group's sense of community. Our experience to date has been very encouraging in suggesting that nomadic computing services will have significant impact on our future work styles.

Matching Media to Location

Truly nomadic access allows computing and communication from a wide range of locations and necessarily spans multiple media. Access is provided over the existing communication infrastructure, making heavy use of cellular and conventional telephones through voice and touch tone interfaces. Facsimile is used for printing in the home or on the road at hotels or other offices; we treat the fax machine as a "universal printer," not a device carried in a briefcase. Alphanumeric pagers deliver short text messages such as e-mail and calendar notifications as well as maintaining a pocket-sized small data cache. Microcasette recorders or answering machines combined with car stereos provide commuter-oriented applications. These services are layered on a computational base, providing speech recording and playback, speech synthesis, speech recognition, analog and digital telephone interfaces, and fax modems.

With users in a range of nomadic locales different media are appropriate to each. Conventional workstations are preferable in the office, but mobile access seriously impacts conventional desktop applications. For example, we may need to support voice in a graphical calendar interface if users are also able to record calendar entries over the telephone.

Other locales more closely match the conventional nomadic models. Home is a stable environment that often serves as a secondary workplace; although home computers are less powerful than those at the office, telephone lines, answering machines, and sometimes facsimile machines are readily at hand. The daily commute is a challenging but potentially rewarding environment for mobile computing,

CHRIS SCHMANDT is the director of the Speech Research Group at the Media Laboratory.

[1] There is an unfortunately growing trend to equate "multimedia" with "motion video." This article is not about video, but about text, speech/audio, and graphics.

0890-8044/93/$03.00 1993 © IEEE

in which the available media are constrained by the mode of transportation. Commuting time is a special case of local area mobility, in which a user is disconnected from conventional computing resources for short periods of time with frequent reconnection possibilities and much value for close, near-term coordination and scheduling.

Finally, wide area nomadic access occurs when users travel to more distant destinations, and may be based in hotels, on airplanes, or in even campgrounds while on vacation, and usually involves more erratic connectivity. In the wide area, we can assume that telephones are ubiquitous (or carried in one's pocket) and fax machines are near at hand, but in a rural location access may be limited to pay phones and even pagers are often out of transmitter range.

The flexible use of multiple media, especially speech, is therefore the foundation of our nomadic environment. Different media are employed in various interaction locales as a function of the communication access method, the data itself, and the activity of the user. Text may be displayed on a workstation or sent as a facsimile, but it may also be spoken by a text-to-speech synthesizer during telephone access or while driving. Digitized voice is widely used, found in applications ranging from voice messages to calendar entries to to-do lists; although voice entry is usually employed in locales lacking a keyboard, it must be retrieved in the office as well. Finally, images such as maps with driving directions, can be viewed on screen or transmitted as a fax.

Applications

This section describes MIT's Media Lab nomadic computing environment. As show in Table 1 we use fax, pagers, and telephone access to digitized and synthesized audio to allow subscribers to manage a number of personal information and news databases. Although these will be described in this section as independent applications (and are actually implemented as such), they are highly interdependent; for example, a good portion of our faxing is invoked by users while in the midst of a telephone-based interaction employing speech synthesis. While each of these applications observed in isolation is a rather simple deployment of existing technology, what is novel is their integration, interaction, and impact on their users' daily work routines.

Each of the services described here is in use on a daily basis at the Media Lab and some have been exported to two additional sites. Most services have 15 to 20 subscribers who access the service at least once daily, although several services have only one consistent daily user (synthesized text mail during the morning commute and daily "personal newspaper" fax to hotels while on the road).

Phoneshell

At the heart of our nomadic environment is Phoneshell, a set of applications invoked over ordinary telephone connections. Phoneshell provides remote access to voice mail, e-mail, a user's calendar and name and address database, a lab-wide dial-by-name telephone directory, and a variety of news, weather, and traffic information. Applications are controlled using touch tones (DTMF) and, experimentally, speech recognition, and respond by playing or recording audio and synthesizing text as speech. Phoneshell is heavily used; it receives 20 to 30 calls per day and although most calls last five minutes or less some last up to an hour. Immediate access (no busy signal) and reliability are important to subscribers so Phoneshell runs on multiple phone lines answered by computers of several architectures (three Sun Sparcstations and one old Sun 386i using a separate file server). Although developed primarily for the Media Lab, versions of Phoneshell have run for smaller user communities at three other sites.

Medium	On demand (via Phoneshell)	Scheduled transmission
Fax	e-mail, rolodex links, calendar	e-mail summary, calendar, weather (home or hotel)
Pager	Rolodex (for cache)	e-mail, voice mail, reminders
Text-to-speech	e-mail, calendar, rolodex, weather	Commute audio tape (Gannett news)
Digital audio	Voice mail, calendar, radio news Internet Talk Radio	Commute audio tape (BBC news)

■ **Table 1.** *The role of different media for synchronous vs. scheduled information access.*

Phoneshell is described in detail in [11]; a summary of its features follows and is shown in Fig. 1. Phoneshell can play voice mail, record memos, send voice or text replies to voice mail or e-mail, and compose voice, text, or fax messages to persons found in one's rolodex. Phoneshell reads e-mail using text-to-speech synthesis, performs some simple filtering of messages into priority classes, can send a message as a fax, and can generate voice or text replies.[2]

A Phoneshell subscriber can access his or her rolodex, and having selected a "card" can hear selected fields (address, phone numbers, etc.), can access (by voice or fax) data files attached to that card,[3] place a call to that person or send voice or text e-mail or a fax, or send the telephone numbers on the card to his or her pager. A calendar interface speaks calendar entries, including summarization modes for week and month at-a-glance, and allows new entries to be added as voice or text. The calendar can be faxed, and one can remotely enter a default fax number for a range of dates; such an entry triggers another application to automatically send morning announcement faxes (see below). Finally, a news interface presents digitized radio news broadcasts and traffic reports, synthesizes NOAA weather information from around the country (acquired from the University of Michigan public weather server) and text news clips from wire services, and plays Internet Talk Radio programs.

Phoneshell uses a touchtone interface much like conventional interactive voice response systems, but geared towards frequent, expert users. More recent experimental work replaces tones with speech recognition input, and attempts to learn about user interactions and make suggestions [7]. Similar work by Yankelovich [14] provides a speech recognition interface to Sun's Calendar Manager, a group scheduling tool with an Open Windows graphical user interface.

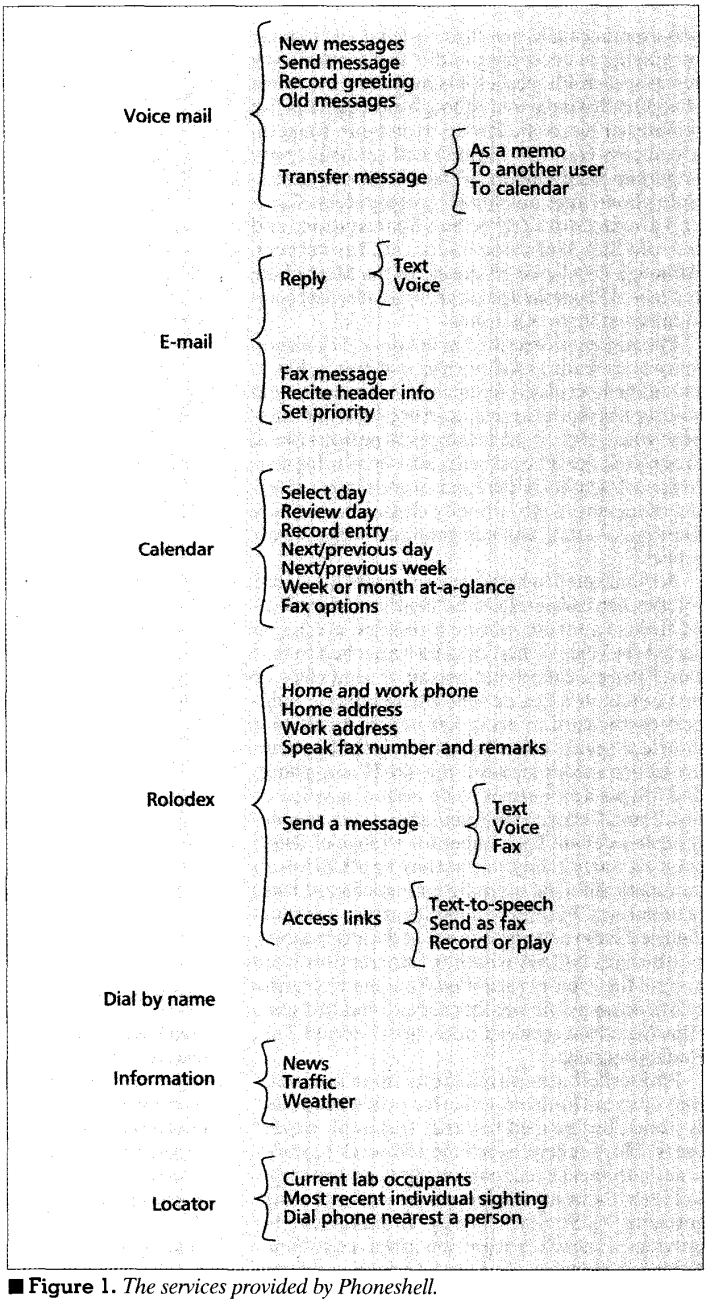
[2] Text is entered as two touchtones per letter. The author can "type" 10 words per minute using this technique. While not very fast, this is comparable to current handwriting recognition products.

[3] Attachments may be text, Postscript, or digitized audio.

Paging

Alphanumeric pagers with nationwide metropolitan service are used for services for which short text messages suffice. E-mail is filtered based on regular expression matching of sender, subject, and recipient fields before being sent to the pager.[4] Since the service provider limits pages to 240 characters, one rarely sees the entire message but often gets its gist. To retrieve the rest of the message, or to reply, the subscriber calls in via Phoneshell (while on the road) or heads for the nearest workstation (while in the Lab). While listening over the phone, a user can mark the sender or subject as "important," so that future messages will be filtered to the pager; this provides a small amount of dynamic control over mail routing while off site.

Every morning our calendars are sent to our pagers, and appointments in the calendar can trigger pages at the time of the appointment if the word "page" appears in the calendar entry. Some users have a weather forecast sent to their pagers in the morning. Subscribers are notified by page of new voice mail, including its origin, length, and return phone number, and can also get return receipts when others in the group hear voice mail sent to them. Users can join a news clipping service that sends to the pager stories about people and companies mentioned in their calendars.



■ **Figure 1.** *The services provided by Phoneshell.*

Fax

Fax is used to deliver messages asynchronously, so they will be available when the recipient is ready (e.g., upon waking in the morning) or to deliver longer or more complicated text data for which speech is inappropriate. While at home, the fax machine may be the most convenient printer. While on the road, faxing also substitutes for other forms of text display such as laptop computers.

Most Phoneshell applications incorporate fax options. While reading e-mail, a long or detailed message may be faxed and, if dealing with a large number of unread messages, a Phoneshell user can send a fax summarizing all new messages (sender, subject, and length of each) and then call back later to listen to a selected few. Any rolodex card can be faxed, or more importantly any file attached to a card, such as a map or driving directions. Phoneshell can fax one's calendar, as well as local or distant weather in short or extended form (which covers several pages with detailed NOAA forecasts and climactic information).

Faxing also occurs as an asynchronous, scheduled event. One of the more popular uses is a morning fax of one's calendar for the remainder of the week. The author has single page fax sent to his home every morning; this fax contains a summary of new e-mail, the calendar for the rest of the week, and the day's weather forecast (Fig. 2). While on the road, a more extended form of this fax is sent to the subscriber's hotel; including in addition a local weather forecast[5] and three pages of text news summaries. The latter is the text of the audio news recorded for listening in the car during the morning commute (see below).

The utility of faxing lies in the universal availability of fax machines and the simple delivery mechanism of having paper available immediately upon waking. While at home preparing for a trip early the next morning, faxing from Phoneshell is the quickest way to get hard copy of phone numbers needed for the trip. For making weekend plans the detailed weather fax provides the required information more quickly than listening to the identical forecast on the NOAA weather radio frequencies. The morning hotel news fax can be read over coffee or later in the day while waiting in a lobby for an appointment.

[4] *SIFT [2] provides a similar service with richer filtering.*

[5] *The forecast is for the nearest weather station, based on the geographic limits of the fax machine's area code.*

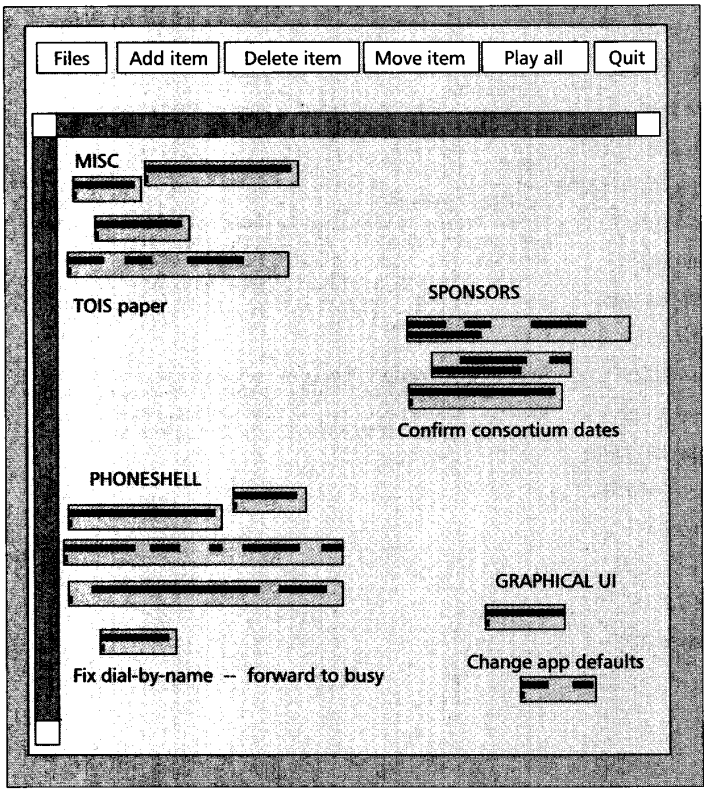
```
/tmp/daily_fax                    Wed Mar 16 08:26:17 1994

587 ackerman@ics.uci.edu          Tue Mar 15 22:18 New 27/1194
 Re: yet another pass
588 geek@media-lab.media.mit.     Tue Mar 15 23:20 New 2/30
 killing phoneshell at moosilauke
589 geek@ media-lab.media.mit.    Tue Mar 15 23:40 New 2/30
 killing phoneshell at beacon
590 iscbarf@BBN.COM               Wed Mar 16 01:03 New 2617/116199
 Revised thesis proposal
591 dwingfotog@aol.com            Wed Mar 16 01:07 New 20/917
 Re: yup, on line!
592 Ben.Stoltz@Eng.Sun.COM        Wed Mar 16 01:51 New 20/902
 Re: Adventure de jour
593 Ben.Stoltz@Eng.Sun.COM        Wed Mar 16 02:13 New 28/1046
 Re: days etc
594 root@media.mit.edu            Wed Mar 16 05:02 New 6/159
 <no subject>
595 mullins@media-lab.media.m     Wed Mar 16 05:33 New 2/225
 New ITR Files
596 jmallory@nas.edu              Wed Mar 16 06:00 New 9/126
 reviewing draft
597 carl@radio.com                Wed Mar 16 06:16 New 30/1217
 Common Ground
598 carl@radio.com                Wed Mar 16 06:21 New 28/1153
 Geek of the Week
599 geek@media-lab.media.mit.     Wed Mar 16 06:53 New 2/11
 page
600 geek@media-lab.media.mit.     Wed Mar 16 07:00 New 11/133
 page
601 reed@interval.com             Wed Mar 16 07:17 New 2/34
 Tricorder??
602 judyb@media-lab.media.mit     Wed Mar 16 08:08 New 7/163
 next percom talk
-------------------------------------------------------------------------------

    Wednesday
    March 16

12:00 - 1:30 DCGS
check with Patty Slusher
check NiF slides
1:30 - 2:30 David McCartney (BT liason)
-------
    Thursday
    March 17

peter's ticket
-------
    Friday
    March 18

3:00 - 4:00 NiF seminar

-------------------------------------------------------------------------------

Greater Boston metrolpolitan area forecast
National Weather Service Boston MA
605 AM Eastern time.  Wednesday MAR 16 1994

..Dense fog advisory until 9am.
Today.  Dense early morning fog.  Then sunshine. Then clouding
up with a chance of afternoon showers.  High in the mid 40s.
West wind 10 to 15 miles per hour.  Chance of showers 50
percent.  Tonight.  Cloudy with light snow likely.  Low 25 to 30.
Northwest
Wind increasing to 10 to 20 miles per hour.  Chance of snow
60 percent.
Thursday.  A chance of morning flurries. Then partly sunny.
Brisk with a high near 40.

CH/PH

Extended forecast.
Friday.  Increasing clouds with a chance of light snow Friday
afternoon and night.  Lows in the 20s.  Highs in the 30s.
Saturday.  Fair.  Lows in the 20s.  Highs 35 to 45.
Sunday.  Clouding up with a chance of rain.  Lows in the 30s.
Highs in the 40s.

RW

The National Weather Service information is provided by the
University of Michigan Weather Underground project and the
National Science Foundation-funded Unidata project, from a
data feed broadcast by Alden/Zephyr Electronics, Inc.
```

■ **Figure 2.** *A sample of the daily fax delivered to the author's home.*

Portable Audio

Analog audiotape is used to provide stored portable audio. Each morning a computer at the office calls home and and records a personalized audio newscast. This news consists of a text part, which is synthesized, and a digitized part. The text source is three Gannett news columns, which provide daily summaries of telecommunications, technology, and travel news. About a dozen categories are currently available. After filtering to correct some pronunciation errors, the text is sent to a speech synthesizer that dials the home answering machine and recites this news (at nearly twice normal speech rate) onto a tape to be played on a car stereo while commuting to work. After the text news, a digitized 15-minute BBC newscast is transmitted in the same fashion; it is time compressed to 10 minutes using the SOLA algorithm [10]. Such a time compression algorithm speeds up playback without changing a sound's pitch, thus better preserving intelligibility.

This answering machine service as well as the home fax is facilitated by a telephone feature provided by the local exchange carrier in the central office. For a small monthly fee, up to two additional



■ **Figure 3.** *A mixed voice and text to-do list.*

numbers can be added to a subscriber's phone line; calls to each cause audibly distinctive ringing patterns. The line terminates in a box that listens to the first ring, and then routes the call to one of three additional telephone jacks, at which point any devices connected to each of the three virtual lines receives subsequent rings. This allows the computer to call the dedicated answering machine early in the morning without ringing any telephone.

The morning commute audiotape is an example of both media conversion — using speech synthesis — as well as audio as a data type of its own — the digitized radio newscast. Audio is also used an a data type for input; Phoneshell users record audio annotations into their calendars, rolodexes, or to-do lists, in addition to recording voice messages which can be sent as multimedia e-mail attachments in MIME [9] or other formats.

Audio is an attractive medium for use with handheld devices as it requires no screen or keyboard, saving size and power. Our earliest work with portable audio devices involved recording on microcasette recorders, digitizing the tape back at the office, then automatically segmenting the audio recordings and appending them to a voice and text project list, as shown in Fig. 3. These audio clips could then be cut and pasted by hand into other applications, such as calendar entries or voice mail messages. Small digital audio recorders, also without any computer interfaces yet, are also appearing on the market. Although occasionally useful, it proved awkward and time consuming to manually connect the recorder to one's office workstation and invoke the digitize and segment processing. But this experience did motivate the addition of a memo feature to the voice mail Phoneshell interface; in voice mail a single keypress triggers recording of a memo, which is treated as a voice mail message for screen display only. This proves to be a convenient means of getting an idea (usually something that one forgot and must deal with in the morning) into the office via the workstation.

To improve the utility of audio reminders on portable devices requires a more structured and interactive environment for recording and playing them back. To this end we designed VoiceNotes, a user interface to a prototype handheld digital audio recorder with speech recognition [12]. VoiceNotes allows its user to configure and maintain a number of lists of notes, such as "shopping," "travel," and "phone calls," and record notes into each; while listening to a list individual notes can be skipped or deleted, and the list can be skimmed. Notes can be assigned attributes such as *important* or *reminder* with an alarm time. Subsequent VoiceNotes interfaces have provided a more conversational means of browsing and recording notes. This work reveals the strengths of a well-designed speech-user interface, but is not part our practical experiences with nomadic computing because the current hardware is in fact tethered to a base station computer and hence not portable.

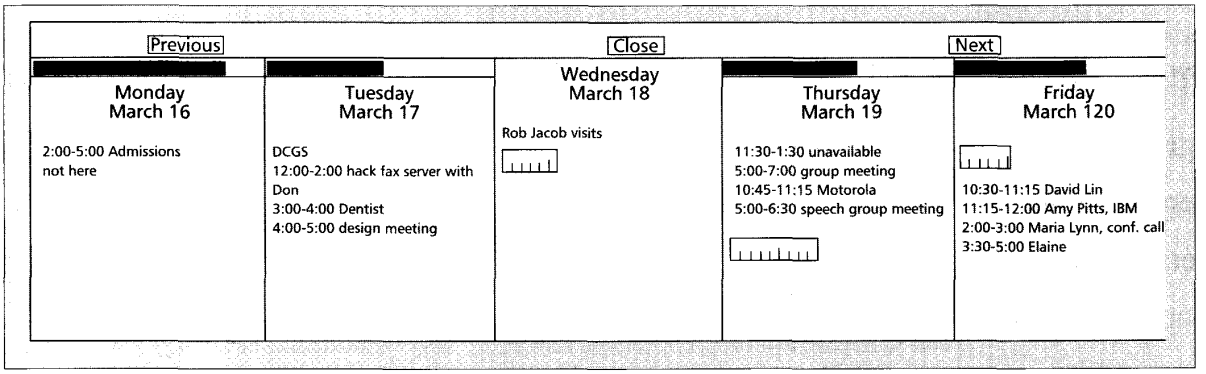
Active Badges

Active Badges are small infrared transceivers built by Olivetti Research Laboratory [13]. The badges periodically transmit their ID to sensors around the lab, and a host computer polls the sensors to track badge location. Badges also have several buttons and can be made to emit a variable series of tones while they are within infrared range of a sensor.

The primary use of Active Badges at the Media Lab is tracking people; location information is available via Phoneshell, a graphical user interface, or a command line program. Location information is used to disable e-mail forwarding to the pager while in one's own office. Badges have also been used for routing telephone calls. If one's phone rings while one is not in the office but is in the building, the badge can emit a tone (different tones can be used depending on the caller's identity); by pressing one of the badge buttons the call can be dynamically forwarded to the nearest telephone. This arrangement is similar to that of [6], but meant to be faster by virtue of running only on campus.

At the Desktop

Nomadic services are tightly related to the ordinary desktop computing environment and will certainly change it. Initially, nomadic services (including most of those described in this article) will provide remote and possibly reduced functionality for performing the same tasks and accessing the same databases currently used in the office. But nomadic access will likely change the desktop as well. Impoverished interfaces may force users to be more organized (e.g., linking files to rolodex cards instead of relying on searching for them by name or keyword). Knowing that the recipient will receive mail on a pager causes the sender to be terse. Composing text messages with a telephone keypad also results in very short messages!



■ **Figure 4.** *A calendar that supports voice entries.*

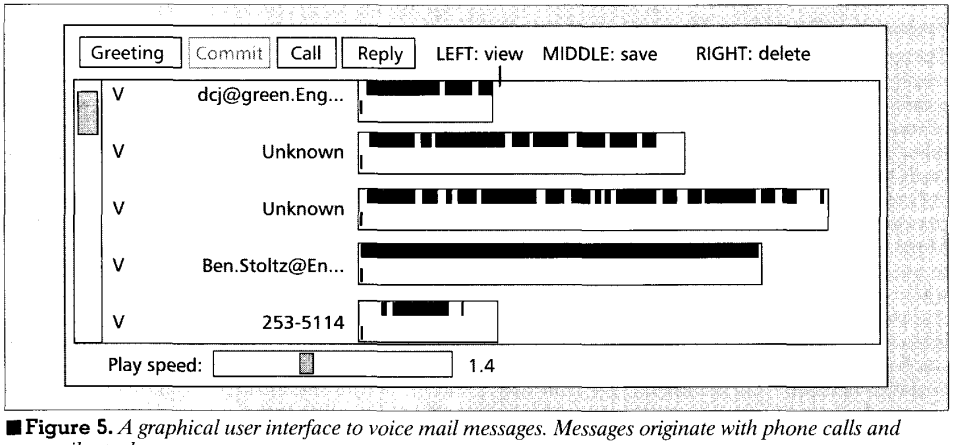
However, some changes will be deeper. The most telling example in our environment is the new need to support audio as a datatype across a range of applications, because telephone access is much more conducive to recording than typing. But applications such as e-mail readers and calendars are text-based, perhaps with graphical user interfaces; what happens when they must support audio as well? The increased use of fax for mobile access will necessitate support for graphics, and perhaps more important is the need to integrate "ink" (hand drawings which are not interpreted as characters) drawn on small portable screens.

Our approach for audio is to store it in files, and use references to filenames in the underlying databases. Applications retrieve entries from the database and, depending on the medium, present it appropriately, synthesizing text over the telephone, or presenting a graphical media controller on screen.[6] A consistent graphical user interface is provided by our SoundViewer widget [3], which displays the sound as speech and silence intervals and allows random access to playback using the mouse. Sound and text widgets can be freely intermixed, as shown in Fig. 5. A graphical user interface provides flexible access to voice messages, either recorded from the telephone or received as audio attachments to e-mail; a user can control playback speed, reply, and forward audio to others or move it into other applications, as shown in Fig. 5.

Service Categories

Our work uses a variety of technologies to provide a number of services to real users. This is not a demonstration system, but rather a serious means of working outside the office with expectation of a high degree of reliability. The various components have been developed incrementally to serve specific needs of our subscribers, and have significantly changed the work habits of some of them. The result is increased communication within several work groups, with decreased latency of responses to important messages and the ability to keep in touch from a much wider range of environments than previously possible.

It is important to develop new and more highly integrated nomadic computing platforms capable of employing emerging digital wireless communication networks, but it is also important to learn from experiments such as ours how these devices might be used. To the end user, the technology is less important than the service it provides (although hopefully future devices will be powerful enough to support more sophisticated user interfaces and customizable service profiles than those described here). This section categorizes the needs filled by the nomadic services described here and discusses the impact of the diversity of media they employ. Some totally new services will evolve around nomadic computing, such as finding the nearest gas station or getting directions to one's hotel while traveling.



■ **Figure 5.** *A graphical user interface to voice mail messages. Messages originate with phone calls and e-mail attachments.*

[6] *A terminal-based application simply displays the text of the filename; the user cannot hear it without calling in, but knows it is there.*

■ ■ ■ ■ ■

Two factors greatly enhance the communication role of a nomadic environment: filtering and the asynchronous notification of messages provided by a pager.

But the core services will most likely reflect and be based on the kinds of activity we perform at our desktop workstations, and this is the focus of our application suite.

Communication

Clearly communication is the dominant theme of the services provided in this environment. Remote access to voice and e-mail is the most common activity across all users. The rolodex is most often accessed as the means to dial a phone number by name or generate a short e-mail message from an inconvenient location. It is an important advantage to be able to respond to messages by voice or text; the send—receive—send cycle of a series of messages plays an important role in the ability to work together. Although awkward to do over the telephone, features such as editing the Cc list of an e-mail reply have turned out to be very valuable for forwarding messages and sharing information.

Two factors greatly enhance the communication role of a nomadic environment. The first is filtering; all pager users and all steady Phoneshell users employ at least some filtering to prioritize messages and determine which are allowed through to the pager. Filtering is not necessary for a low-volume user who might receive five or ten messages between non-workstation mail reading sessions, but for higher volume or extended absences the speech interface is simply too slow without filtering. For example, the author receives approximately 75 messages a day, of which several may be "very important," ten or so "important," and 15 "personal," he or she always reads the first two categories, sometimes the "personal" category, and rarely messages of lower priority. It is encouraging that effective message management can be provided by a very simple filtering scheme.

The second factor is the asynchronous notification of messages provided by a pager. The short time delay between transmission and receipt at the pager often allows for the interchange of a series of responses (the e-mail "conversation") or sometimes triggers a phone call. When users who know eachother have pagers, this leads to more intimate communication, such as an invitation to "call me at home now if you want to talk." Phoneshell is an important adjunct to the asynchronous notification because it allows access to the complete message and facilitates a reply.

Personal Information Management

The second service category is personal information management, of one's rolodex, calendar, and reminder list. The morning calendar fax was one of the most popular features at one site at which this service was offered. The ability to link rolodex cards to arbitrary files which can then be spoken or faxed adds a further dimension to the range of personalized information which can be retrieved. More than once the author has phoned to listen to an attachment containing driving directions while visiting a friend across the country. Although secondary to message retrieval, the importance of personal information management can be gauged by the moderate success of the portable electronic pocket organizers from several vendors, and how such applications seem to be the first to appear on each of the latest generation of PDAs.

Personal information is not read-only; callers can add text or voice entries to their calendars, record memos, record annotations for rolodex cards, and even create new rolodex cards using the touchtone keypad.[7] Our connection model provides for single copies of the various personal databases, which are accessed and modified in real time during a session. This avoids problems with reconciling home and portable copies of databases, and is important where multiple parties may be sharing these databases, such as group scheduling. Storing data at the central location and changing it in real time avoids the problem of later reconciliation between possibly different versions of the data, although recent research into disconnected file systems is beginning to address these concerns [4, 5]. A more robust nomadic product will cache whatever can be stored locally and interact with distant databases as demanded and when connected.

Database Access

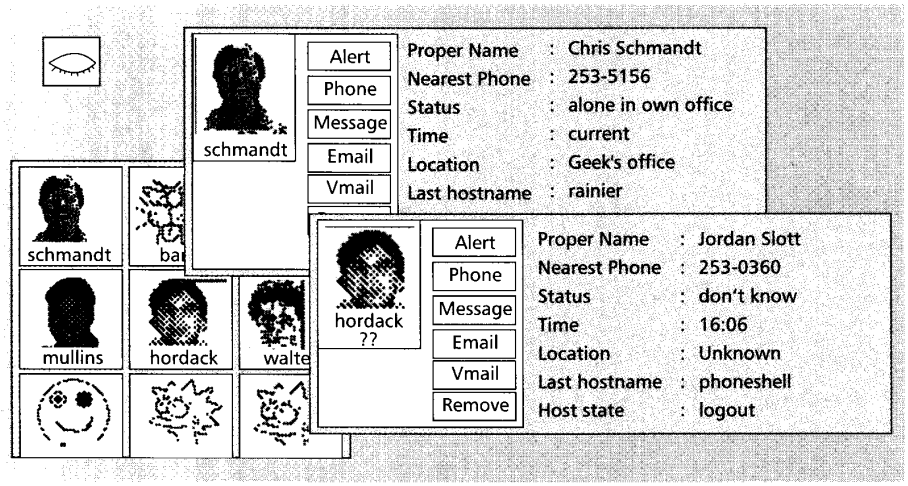
While working at the desktop we routinely gain access to a variety of data from remote locations, ranging from news and weather to travel and traffic services, and increasingly browsing databases through user interfaces such as Mosaic. These are important for nomadic access as well. Phoneshell allows users to type in a city name with touch tones and then fetches and synthesizes a local weather forecast. Text news selected from a user's profile is transmitted each morning as audio or a fax, depending on whether the recipient is at home or on the road. The morning audio tape recaps broadcast international news, and Phoneshell provides access to recordings of the most recent BBC world news, the hourly National Public Radio newscasts, and local traffic reports from another radio station. A morning search of several wire services provides news clips of stories containing matches to the proper noun strings in a user's calendar; these are indicated by a page and can then be synthesized over the telephone.

Data Caching

Although our databases are stored at a central location, various technologies provide a means of caching that data in a form more convenient for the nomadic user. Faxing a rolodex card home just before leaving on a trip is a simple example, where convenient access (over the phone) to hard copy allows it to be stuffed into one's baggage on an as-needed basis. The pager provides a better example. Phoneshell users may send a rolodex card (name and telephone numbers) as a page, which can then be cached in the pager's "personal mailbox," which is non-volatile when the pager's memory fills. Because the pager is small enough to always be in one's pocket, it provides a convenient cache during a trip for flight information, reservation confirmations, and local telephone numbers.

Both the news fax to the hotel and the news audiotape home provide caching, allowing retrieval to be disassociated from connect time to the database. Caching is used less for data entry, although portable digital audio recorders will provide a convenient form for short notes when properly integrated to office computing environments so voice notes can be uploaded to the appropriate multimedia applications.

[7] *In the same Phoneshell session one can create a rolodex card and then send e-mail or telephone the person associated with the card.*



■ **Figure 6.** *The Watcher graphical interface. Faces, on the left, can be expanded; the detail boxes on the right show that one user is active in his office, while another has most recently been active via Phoneshell.*

Maintaining a Sense of Community

Although these services have all been developed to allow users to work while away from the conventional office computing environment, they have the curious effect of helping to maintain a sense of an electronic community, and increase rather than decrease ties with co-workers. How do the nomadic computing applications described in this article aid help to give a sense of community, knowing who else is around and being able to perhaps interact with them at a distance?

Although communication has already been cited as the primary motivation for our nomadic services, the speed of a response is also a good indicator of a sense of community. Receiving e-mail messages on a pager indicates who is working, and perhaps what they are working on; if this triggers a return phone call, even in the middle of the night, contact is made and issues blocking progress can be promptly resolved. A small but very important sense of closure is gained from the e-mail return receipt sent when voice mail messages are heard by the recipient. This receipt can be sent to the pager like any other e-mail as well. These forms or short-delay asynchronous communication are especially useful for work teams at different sites or on differing schedules, such as at a university.

From the point of view of those working regular hours in the lab, one of the problems with flexible work time and increasing use of nomadic access to office resources is that we lose our sense of who has been around and how far away (i.e, disconnected) they might be. A graphical interface, Watcher (Fig. 7), keeps track of users via Active Badge sightings as well as the Unix *finger* command, and displays who is present, where they are, and for those not present, when they were last seen.

Console logins, i.e., in the lab, are distinguished from remote login, either via dialup or from some remote site such as a sponsor company. This same information is available over the phone using speech output via Phoneshell. Watcher can also track when one is the office and on the phone and, more importantly, when one has logged in to read mail via Phoneshell. In short, it provides a concise representation as to each person's most recent contact with the Lab, which in turn implies when one last checked messages, whether one is coming in to work early or late, or how well one is keeping in touch while out of town. Out of concern for privacy, this information is not logged and no history is available beyond the most recent sighting.

Disparate Media

One important aspect of the applications described here is their reliance on a disparate set of media: text, voice, and some graphics. This is important because the choice of medium is motivated not only by the means by which the subscriber is accessing the service, but also because of the communication role envisioned at the core of these services. Many of the databases we access could be made available as text, but it is not currently possible to automatically transcribe voice mail messages. Even if it were feasible, there are many people and communication situations for which e-mail is impossible or an inadequate substitute for a synchronous voice conversation.

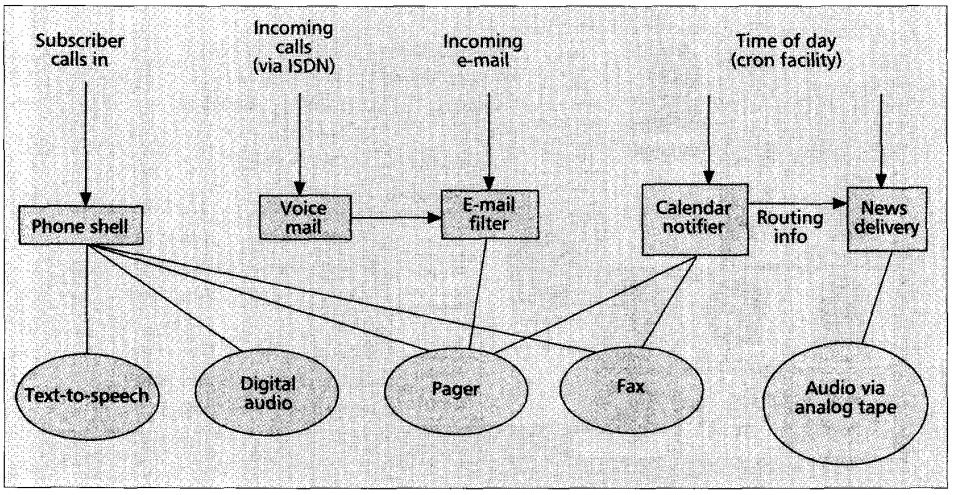
Because of the research goals of our group, our nomadic services make heavy use of voice. Speech has many advantages in addition to its expressiveness and richness as a means for human communication. We can utilize the in-place network and the ubiquitous telephone as a terminal; we can purchase consumer products such as pocket cellular phones and answering machines; and we can listen to speech while our hands and eyes are busy with other tasks, such as driving or going for a walk. At the same time, speech has severe limitations, being both serial and slow. We make heavy use of time compression in an attempt to compensate; both synthetic and digitized speech playback speed is under user control. In fact subjects accommodate quickly to sped up speech, and find it annoying to have to listen to normal rate voice messages after exposure[1].[8] Still it is important to have fallback mechanisms such as the ability to send a long portion of text as a fax, and to recognize that most users would prefer to get most information as text if it were possible.

The choice of voice because it facilitates certain means of nomadic access impacts both databases and local storage formats as well as the user

One important aspect of the applications described here is their reliance on a disparate set of media: text, voice, and some graphics.

[8] *Recently, Sony introduced a consumer answering machine with built-in time scaling.*

It is important to consider how emerging non-textual media will relate to other tasks being routinely performed on the desktop.



■ **Figure 7.** *Our nomadic environment uses many application programs (shown in boxes), which respond to various inputs, coming in from the top, and utilizing multiple communication media (shown in circles).*

interfaces for desktop access to these databases. It is therefore important to consider how emerging non-textual media will relate to other tasks being routinely performed on the desktop. For example, in our case the most popular use of voice on the desktop is the graphical user interface to voice mail, which is even more important to users than using audio as a data type in other applications. In meeting this need, we developed the tools (the SoundViewer widget and some database abstractions) to easily integrate voice into the remaining applications, and allow free interchange of voice between applications. The resulting whole is in some ways greater than the sum of its parts due to this interoperability.

Limitations and Related Work

It is important to recognize a number of limitations in the nomadic computing approach presented here, and to understand its role with respect to the emerging nomadic computing hardware and communication architectures. Our work spans personal information management, handling asynchronous communication, and voice as a data type. The first generation of pocket information managers (such as Wizard and Boss) are portable and emphasize personal information management, but reconciling their data with office workstations and making them accessible to other users is difficult. They also cannot receive asynchronous information. More recent palmtop computers and PDAs incorporate paging options (due to the growing use of the PCMCIA slot); this allows hardware such as a Newton to receive text messages. Work at Bellcore utilizes two-way radio portable radio modems, in conjunction with cellular telephones, to manage calls in real time at a distant location, though a caller must deal with voice menus and await a response [6].

Our approach encompasses both synchronous (on demand) access, as well as asynchronous communication. In order to access voice or manage phone calls, a telephone interface is necessary. The first portable computers to include cellular telephones (such as the Eo) did so awkwardly. But with improving micro-electronics cell phone components will require less space, making this marriage a happier one. An interesting recent development is Simon, from BellSouth and IBM; Simon is a medium-size handheld phone with a touch-sensitive display, graphical user interface, and abilities such as paging, fax receipt, and personal directory lookup. Such a "smart phone" is valuable to the large number of users who would carry a telephone before they would carry a computer.

Filtering is an essential aspect of our nomadic environment and it will be for users of similar services. Although we desire to always be able to reach others ourselves, we value our private lives and would not dream of allowing ourselves to be interrupted by every phone call to our offices or e-mail message sent to our accounts. We did not invent e-mail filtering. Some of the earliest e-mail filtering work was Information Lens [8], which used filters and semi-structured text tags to categorize e-mail messages. And Bellcore's SIFT project uses rule-based filtering to direct e-mail delivery dynamically as well [2].

Because databases are stored in a central location, with only a little distant memory, it is necessary to place a phone call for many interactions, especially via Phoneshell, and call completion time is surprisingly slow. Even with a locally dialed call and use of speed dialing features, 10 seconds or more are required to login to Phoneshell.[9] Using credit cards or other than the default long distance carrier from a pay phone adds to this time. Placing a cellular call usually takes an additional 15 to 20 seconds to ring through to the distant party, and if roaming even longer. This is excessive overhead if all one wishes is to record a three-second reminder for tomorrow morning. In addition, touchtone interfaces with hierarchical menus become tedious and while navigating quickly even expert users sometimes get lost.

The main drawback to our approach is the large number of components required; although engineers may be willing to carry several pieces of equipment and transfer devices between home, car, and office, less devoted technophiles will need more integrated solutions. This is also true in terms of our software architecture. Figure 8

[9] *This includes both call setup time, the time it takes from completion of dialing to the first ringing at the called party, and the time to log in via touchtones.*

shows the various software components; changes in one often require changes in another, especially in terms of scheduling or filtering rules. Similarly although our services are very flexible, configuration usually involves programming an application or shell script; we would prefer an environment that observed subscriber behavior and learned to adapt to it. These observations suggest real problems for future products in terms of managing and configuring multiple possibly conflicting services. Packaging will be a factor in making nomadic technology fashionable. Some of us thrill to reading our e-mail and sending replies by flashlight at a roadside pay phone in the middle of the desert at night, but this is the act of a dedicated nerd and must be made more elegant to appeal to a larger population.

On a minor but annoying note, many of our current telecommunications devices are less than optimal. Current analog cellular telephones are infamous for noise and dropping calls. Pagers (and cell phones) may not work in building cores, basements, or underground on public transportation.

Conclusions

This article describes a nomadic computing environment built with current hardware technology. Its focus was on services, uses of the technology, and integration of nomadic access with ordinary desktop office tasks. The nomadic applications described here enable users to keep in closer contact with their stationary (or virtual) work environment and the data required to to be productive away from it. Different applications and media are employed depending on the data being accessed and the degree of distance from the office and current degree of mobility of the subscriber.

This article attempts to demonstrate the power of nomadic computing by describing how a small sample of users have gone increasingly mobile. Identifying the classes of operations they perform helps define a service-oriented approach to designing future mobile computing environments and their impact on stationary desktop applications. There is much to be learned about nomadic computing by building systems based on today's technology. We can explore how nomadic services might be used or classified based on locale, and differences between synchronous and asynchronous nomadic interactions. We have demonstrated the utility of multiple media based on user access requirements, and this leads to better understanding of how users choose appropriate media depending on their locale. Much research remains to be done in effectively retrieving and manipulating non-textual data. Finally, by observing usage patterns and requirements in an active nomadic user population we can better categorize the different needs nomadic computing can solve. In sum, we hope to gain insights into media requirements and potential roles of nomadic services to be developed on future platforms.

Acknowledgments

Many members of the Media Lab Speech Research Group contributed to all as aspects of the applications described or mentioned in this article, including but not limited to: Mark Ackerman, Barry Arons, Sumit Basu, Lorne Berman, Jeff Herman, Debby Hindus, Angie Hinrichs, Chris Horner, Jennifer Lee, Eric Ly, Lorin Jurow, Matt Marx, Atty Mullins, Sheldon Pacotti, Deb Roy, Jordan Slott and Lisa Stifelman. Don Jackson, of Sun Microsystems at the time, helped port Phoneshell to the Sun site and developed some of the paging and faxing scripts. Don Jackson, Lisa Stifelman, and especially Mark Ackerman provided valuable reviews of early paper drafts.

The majority of this work was sponsored by Sun Microsystems with additional support from Apple Computer and AT&T.

References

[1] D. S. Beasley and J. E. Maki, "Time- and Frequency-Altered Speech, N. J. Lass, ed., Contemporary Issues in Experimental Phonetics, Chapter 12, pps. 419-458, (Academic Press, 1976).

[2] W. S. Gifford and D. L. Turock, "The Electronic Receptionist: A Knowledge-Based Approach to Personal Communications," Proc. of the IEEE Conf. on Sel. Topics in Wireless Commun., IEEE, June 1992.

[3] D. Hindus, C. Schmandt, and C. Horner, "Capturing, Structuring, and Representing Ubiquitous Audio, *ACM Trans. on Office Info. Systems*, vol. 11, no.4, pps. 376-400, 1993.

[4] L. B. Huston and P. Honeyman, "Disconnected Operation for AFS, Proc. Mobile and Location-Independent Computing, pps. 1-10. USENIX, August 1993.

[5] J. J. Kistler and M. Satyanarayanan, "Disconnected Operation in the Coda File System," *ACM Trans. Computer Systems*, vol. 10, no. 1, pps. 3-25, Feb. 1992.

[6] M. Kramer *et al.*, "Use of Two-Way Wireless Messaging for Personal Telephone Management, Proc. of ICUPC '93, Oct. 1993.

[7] E. Ly and C. Schmandt, "Chatter: A Conversational Learning Speech Interface, Proceedings of the spring AAAI workshops AAAI, 1994.

[8] T. W. Malone, *et al.*, "Semi-Structured Messages are Surprisingly Useful for Computer-Supported Coordination, *ACM Trans. Office Information Systems*, vol. 5, no. 2, pps. 115-131, 1987.

[9] M. T. Rose, *The Internet Message*, chapter 6, Prentice Hall Series in Innovative Technology. (P T R Prentice Hall, 1993).

[10] S. Roucos and A. M. Wilgus, "High Quality Time-Scale Modifications for Speech," Proc. of the Int'l Conf. on Acoustics, Speech, and Signal Processing, pps. 493-496. IEEE, 1985.

[11] C. Schmandt, "Phoneshell: the Telephone as Computer Terminal," Proc. of ACM Multimedia '93, pps. 373-382, New York, Aug. 1993.

[12] L. J. Stifelman *et. al.*, "VoiceNotes: A Speech Interface for a Hand-Held Voice Notetaker," Proc. of INTERCHI '93, New York, April 1993.

[13] R. Want and A. Hopper, "Active Badges and Personal Interactive Computing Objects," IEEE Trans. on Consumer Electronics, vol. 38, no. 1, pps. 10-20, Feb. 1992.

[14] N. Yankelovich, "Talking vs. Taking: Speech Access to Remote Calendars," Adjuct Proc. of CHI '94, ACM, 1994.

Biography

CHRIS SCHMANDT received the B.S. and M.S. degrees from MIT, where he has been building speech systems since 1979. He is the director of the Speech Research Group at the Media Laboratory, a position he has held since the creation of the Lab. His current research focuses on user interfaces and applications of speech processing technology, voice as a data type on workstations and hand-held computers, and computer-mediated telephony. His e-mail address is geek@media.mit.edu.

The nomadic applications described here enable users to keep in closer contact with their stationary (or virtual) work environment and the data required to to be productive away from it.