16

Technological Innovations Enabling Automatic, Context-Sensitive Ecological Momentary Assessment

Stephen S. Intille

Health-related behavior, subjective states, cognitions, and interpersonal experiences are inextricably linked to context. Context includes information about location, time, past activities, interaction with other people and objects, and mental, physiological, and emotional states. Most real-time data collection methodologies require that subjects self-report information about contextual influences, notwithstanding the difficulty they have identifying the contextual factors that are influencing their behavior and subjective states. Often these assessment methodologies ask subjects to report on their activities or thoughts long after the actual events, thereby relying on retrospective recall and introducing memory biases. The "gold standard" alternative to these self-report instruments is direct observation. Direct observation in a laboratory setting, however, artificially constrains behavior. Direct observation is also typically too costly and invasive for long-term, large-sample-size studies of people in their natural environments.

Technological innovations are creating new opportunities to capture accurate, real-time data with minimal intrusiveness using techniques such as electronic Ecological Momentary Assessment (EMA). Other chapters in this collection discuss the benefits, challenges, and versatility of electronic EMA as it is being used in current research. This chapter, however, looks toward the future.

New technologies will enable two significant extensions to current EMA methodologies. First, most EMA studies to date have used intermittent collection of self-report data. New technologies will enable EMA studies that combine continuous data collection of subject activities and physiological states with intermittent self-report data collection. Second, new technologies will enable EMA studies where a computer automatically triggers context-sensitive intermittent self-reports based upon analysis of the continuous data stream. Intermittent self-reports can be tied to the observation of particular activities or states that are specified by the researcher but automatically detected by the computer.

stone_Chapter16.qxp 12/14/2006 8:19 PM Page 309

The ability of a computer to continuously analyze sensor data to determine when and how to prompt a subject in response to that subject's activities is a significant extension of the EMA methodology, referred to here as context-sensitive EMA, or CS-EMA (Intille, Rondoni, Kukla, et al., 2003).[1]

The CS-EMA methodology for information acquisition is dependent upon the development of CS-EMA instruments that use sensors to measure and record contextual information such as a subject's location, movement, and physiological states. The sensor data is processed in real time so that the CS-EMA instrument can automatically respond to subject behavior with context-sensitive and tailored self-report queries. A diverse set of CS-EMA devices is possible, each exploiting different sensors, context inference algorithms, and output devices. The query can be triggered via physiological sensors, social interactions, or image processing. Self-reports can be obtained at the moment of interest, or data can be collected to help remind the subject of the key activity of interest and a report can be collected later. In all cases, however, the computer triggers some type of real-time data collection based upon processing a single or set of incoming data streams for a particular event. Each of these variations is considered to be an example of a CS-EMA in the remainder of this chapter.

CS-EMA is a methodological extension to EMA that cannot be achieved without computer instruments that perform the automatic context processing. CS-EMA may therefore require the development of new data analysis techniques because the triggering of self-report prompts (and the way prompted questions are structured) is dependent upon algorithms that automatically process continuous data streams to infer key moments in time that are of interest to the researcher. CS-EMA instruments must exploit advanced technologies, but within 10 years most of the input and output capabilities required will be built into common consumer electronic devices, thereby making CS-EMA instruments economical to deploy for large-sample-size research.

The Gold Standard Technology: Paper

Paper is a versatile and robust technology for gathering and conveying information and therefore provides a useful comparison point against which to assess new alternatives. Paper has many desirable attributes that, until recently, have been difficult for any other technology that supports real-time data collection to match. Paper is extremely lightweight and compact and has a flexible form factor (i.e., physical style that can be easily changed). It is inconspicuous when stored and used, inexpensive, and easy to replicate and modify without special expertise. Paper is familiar to people and easy for most to use, and it allows for easy correction if participants make a mistake during a particular recording session or at a later time. It also allows for subject adaptation when a subject wishes to express an idea the researcher may not have anticipated. Finally, paper has been validated: although the content on the paper may be debated, the use of paper as the medium through which information is collected is rarely questioned.

Paper as a real-time data collection technology does have some fundamental limitations because it is a static, not proactive medium that cannot tailor content to a subject's behavior, subjective states, or interpersonal experiences. Paper is also easy to forget and usually requires manual data entry and coding of information written on it. Paper has no automatic time-stamping of entries, and paper cannot be easily used to ensure that questions are answered in a particular order. Although paper can be combined with other technologies so that subjects can report different types of data, paper does not support the passive acquisition of multimodal data, and paper cannot automatically tailor data acquisition to the context.

The Emergence of Computerized EMA/ESM Tools

The fundamental problem with paper as a technology for real-time data collection is that it is a static medium that has no capability to be proactive. Paper is used in two ways: retrospectively or non-retrospectively. Retrospective instruments using paper require that a person be able to remember information and record that information at the end of each day, week, or some other time interval. Retrospective recall of mundane, regular, or frequent activity, however, is known to be unreliable (Gorin & Stone, 2001). Non-retrospective paper instruments, known as diaries, ask subjects to record information at the time of the activity of interest. Non-retrospective instruments limit recall bias and therefore should be better at reliably capturing the influence of contextual factors on behavior, subject states, cognitions, and interpersonal experiences. A recent study by Stone and colleagues, however, indicated that although subjects reported high compliance with a paper diary instrument, actual compliance was low; "hoarding," where subjects faked diary entries using retrospective recall, was common (Stone, Shiffman, Schwartz, et al., 2002).

One way to improve paper is to combine it with the use of an electronic beeper to trigger use of the paper technology (Csikszentmihalyi & Larson, 1987). Although such auditory signaling improves subject compliance, the result is "still unsatisfactory" (Broderick, Schwartz, Shiffman, et al., 2003). A better solution is to entirely replace the paper technology with an electronic device that can prompt for data entry and also time-stamp entries so that hoarding is not possible (Stone et al., 2002). When electronic devices are used as the medium for conveying and gathering information, diary compliance can be improved (Hyland, Kenyon, Allen, & Howarth, 1993; Stone et al., 2002). An additional benefit is that electronic entry devices can reduce researchers' data entry time and the likelihood of encoding errors.

Electronic EMA (Stone & Shiffman, 1994), also known as electronic experience sampling (Barrett & Barrett, 2001), is a technique where a small mobile computer such as a personal digital assistant (PDA) is programmed to prompt a subject with an audio or visual reminder on a schedule determined by the researcher. The other chapters in this collection describe how the electronic

stone_Chapter16.qxp 12/14/2006 8:19 PM Page 311

sampling methodology is being applied in health-related research fields. As more electronic sampling studies are conducted, recommendations for new experimental methods (Larson & Delespaul, 1990) and statistical models for data analysis (Schwartz & Stone, 1998) are being proposed. Electronic sampling does introduce some new potential problems, such as heightened reactivity, where the sampling may influence the measurement. Recent results, however, indicate that the magnitude of reactivity to electronic sampling may be smaller than previously thought (Hufford, Shields, Shiffman, et al., 2002). The cost of the enabling technologies is also a problem that is slowing widespread adoption of the technique.

The New Opportunity: Context-Sensitive Data Collection and Intervention

Most electronic sampling for real-time data collection in use today simply replaces the paper medium with an electronic device. The diary assessment content is adjusted for display on small screens, and responses are changed to allow for stylus or touch-screen input. The computer operates as an input, output, and time-stamping storage mechanism. These typical electronic EMA devices may improve the quality of data collected from paper instruments, but most are not yet fully exploiting the capabilities of the electronic devices. New technologies will support three key advances:

- **Continuous and passive collection of multimodal data.** New instruments will use emerging electronic devices to comfortably, continuously, and passively collect data. Sensors worn on the body and placed in the environment will record information about a subject's location, physiological states, and activities without requiring proactive self-report.
- **Context-sensitive prompting at appropriate times and places.** Emerging electronic devices will make context-sensitive prompting possible, where questions are automatically triggered based on the subject's behavior, location, physiological states, past responses, and/or activities.
- **Data collection tailored to the individual and to the situation.** Emerging electronic devices will allow an unprecedented amount of longitudinal data to be collected and stored locally on the sampling device. Algorithms that analyze this data will be able to tailor questions based upon a subject's prior behavior or reported subjective states, cognitions, or interpersonal experiences.

Each of these opportunities is described in the sections that follow.

An Example: Technology to Support CS-EMA

Technological innovations will support CS-EMA instruments that automatically compute an appropriate time and place to collect data based on a subject's past and current activity and then tailor the questions asked to the situation (Intille, Rondoni, Kukla, et al., 2003). Later in this chapter the full potential of CS-EMA is discussed, based on technologies likely to emerge in the next 10 years. However, current technologies can be used to create CS-EMA instruments.

First-generation software developed at MIT using present-day mobile computing devices is being used to test the capabilities of the CS-EMA methodology. Software written for PocketPC (Microsoft, Seattle, WA) PDA devices implements the basic functionality commonly found in commercial (e.g., DiaryPRO [invivodata, Pittsburgh, PA]) and open-source (Barrett & Barrett, 2001; Brodsky, Consolvo, & Walker, 2002; Rafaeli, 2003; Weiss & Beal, 2003) EMA programs. The MIT software, referred to here as CAES, includes options for (1) chaining complex sequences of questions based on particular question responses, (2) multiple-choice and multiple-response questions, (3) flexible question recurrence patterns (by weeks, days, hours, minutes), and (4) bounded randomization (min/max time to next query). The researcher "programs" a diary assessment strategy by entering the questions in a comma-delimited file.

The most important aspect of the MIT CAES software is that it can record and respond to sensor data, such as a subject's position (from Global Positioning System [GPS]), the subject's posture (from accelerometers), and the subject's heart rate (from a wireless chest strap monitor). Typically a researcher using EMA has three options: (1) sampling on a time-based, regular-interval schedule, such as every 30 minutes, (2) sampling on a time-based variable random schedule, such as on average once every 30 minutes or sometime randomly within every 2-hour window, and (3) event-based sampling in response to subject initiative, where the subject is told to make a data entry whenever performing a particular activity (Stone & Shiffman, 1994). Devices with sensors allow a fourth option: triggering self-report based on sensor readings of the context. With such a device, contextual factors such as a subject's location, physical activity, proximity to others, or physiological state can trigger self-report assessments. Further, the assessments can be tailored to the context that has been detected from sensors. Any contextual situation that the device can automatically detect can be used to trigger, tailor, or trigger **and** tailor a self-report.

The MIT CAES software has been developed in a modular fashion that allows new context-sensing sensors and software to be plugged in. These sensors permit researchers to use context-sensitive sampling where questions are asked only when a subject does a specific thing. For example, in one study CAES software was used to implement a CS-EMA protocol where a subject was prompted to respond based upon changes in heart rate. Heart rate was monitored using a wireless chest strap that sent data to a receiver attached to a PDA (Rondoni, 2003). In another study, CAES software was used to implement a CS-EMA protocol where a subject was prompted based on changes in posture. Posture was detected by a CAES software module that analyzed data sent to a receiver plugged into the PDA from two wireless accelerometers worn on the body (Ho & Intille, 2004). In a pilot study, the CAES software has been used to implement a CS-EMA protocol where self-reports are triggered when someone is near a

stone_Chapter16.qxp 12/14/2006 8:19 PM Page 313

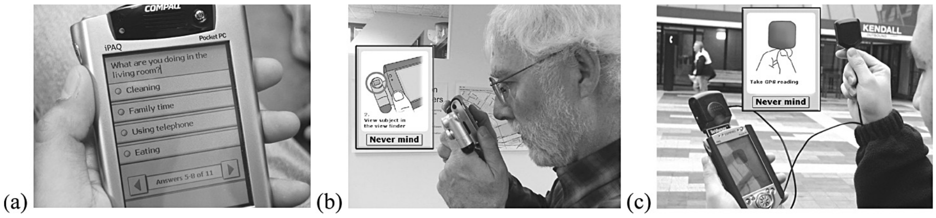

Figure 16-1. Screen shots of MIT CAES software.

particular part of the community, as determined by a GPS receiver plugged into the PDA and a GIS database.

In each of these experiments, the CS-EMA methodology was run using CAES software on a PDA. The researcher scheduled questions to occur during the time or just after the activities of most interest occurred, thereby minimizing the interruption burden of the EMA technique on a subject while maximizing the amount of data acquired about the target behavior. Figure 16-1 shows screen shots of the MIT CAES software.

Table 16-1 lists the properties of paper and current mobile computing devices (e.g., PDAs) that enable real-time data collection. The delivery technology dramatically affects the implementation of EMA. For instance, most current EMA options are implemented on PDA devices. PDAs are getting smaller, lighter, less expensive, and more convenient to use each year, but paper can still be carried much more inconspicuously and comfortably. PDAs offer the significant advantage of audio prompting with time-stamped and time-sensitive electronic data entry, but the prompting can become irritating for the participant in longer experiments or for those with aggressive prompting schedules. The interruption burden of traditional electronic sampling is one of its primary drawbacks.

By using new technology to prompt subjects only during activities of interest, CS-EMA may increase the volume and quality of data acquired about the target behaviors while minimizing the burden of sampling for the subject. CS-EMA, however, requires that subjects use computing devices that can receive data from sensors. Many sensors today are bulky, uncomfortable, and expensive, but this situation is about to change.

Driving Trends

There are four trends that will facilitate the creation of new instruments to support CS-EMA: miniaturization, plummeting costs, ubiquitous wireless networks, and accurate activity recognition algorithms.

Table 16-1. Properties of paper when used for a diary assessment, current PDAs when used for EMA, current PDAs with sensors when used for CS-EMA today, and mobile devices with sensors (in 10 years) when used for CS-EMA

Property	Paper for diary	PDA for EMA (today)	PDA with sensors for CS-EMA (today)	Mobile device with sensors (in 10 years) for CS-EMA
Lightweight	✓	✓-	✓-	✓
Compact	✓	✓-	✓-	✓
Inconspicuous	✓	✓-		✓
Inexpensive	✓			✓
Flexible physical form factor	✓			✓
Easy to replicate (technology readily available)	✓			✓
Easy to modify assessment strategies	✓			?
No special expertise required to implement	✓			?
Familiar to subject	✓	✓-		✓
Subject can adapt if desired	✓			
Validated	✓	✓		?
Not easy to forget/leave behind		✓-	✓	✓
Low interruption burden				
Proactive reminders to enter self-reports		✓	✓	✓
Automatic data entry		✓	✓	✓
Automatic time stamps		✓	✓	✓
Questions can be asked sequentially based on prior responses		✓	✓	✓
Multimodal continuous data collection			✓-	✓
Auto tailor prompts and questions to context			✓-	✓

A "✓" indicates the author believes the technique satisfies the property, with "✓- indicating the property is satisfied but to a somewhat lesser degree. A "?" indicates that the author is uncertain.

stone_Chapter16.qxp 12/14/2006 8:19 PM Page 315

Device and Sensor Miniaturization

The computer industry has a phenomenal history of miniaturizing technologies—a trend expected to continue. Many of the portable phones, PDAs, and music players on the market today are actually 400-MHz+ computers with the capability to store gigabytes of information and transmit and receive data wirelessly. They can operate for a full day on a single night's charge. For output, these devices have bright, color video displays and audio. For input, the devices have touch screens, buttons, keyboards, handwriting recognition, and audio. These tiny computers can fit in a pocket, and they have enough speed and memory to collect, store, and process data from multiple sensors on wireless networks. The devices are essentially wearable computers, computing devices with small and ergonomic physical styles that permit users to comfortably wear or carry them at all times (Starner, 2002). The devices can perform nonstop data collection, processing, and archiving—functionality that could be exploited to create a device to implement CS-EMA protocols.

Miniaturization of wireless sensor technologies is also ongoing. Small, low-power sensors that can be worn on the body or placed in the environment are in development in labs today. Sophisticated sensors are already migrating into common consumer devices. New watches, for example, are capable of displaying and collecting information from wireless sensors worn on other parts of the body, such as a chest strap heart rate monitor (FitSense Technology, Southborough, MA). New mobile phones, PDAs, and watches have sensors such as GPS-based position locators, voice recognizers, fingerprint recognizers, and cameras. In labs, prototypes of jewelry that measure blood pressure volume, galvanic skin response, blood oxygen levels, and hand gestures have been created (Asada, Shaltis, Reisner, et al., 2003; Gandy, Starner, Auxier, & Ashbrook, 2000).

Plummeting Costs

Both the size and the cost of these devices are decreasing, even as their capabilities increase. Two barriers to widespread use of CS-EMA are the cost of the hardware required to collect and analyze sensor data in real time and the cost of developing, customizing, and maintaining the technology.

Intel, a manufacturer of computer, networking, and communications products, predicts that Moore's Law (Moore, 1965)—a doubling of computer power every 2 to 3 years—will hold for another decade. Therefore, the functionality-to-price ratio of mobile computing devices that can be used for computerized sampling will continue to improve. Within 10 years, mobile devices such as phones will have the computational, memory, and networking capacities of today's desktop computers but with "anytime, anywhere" access. These devices will have sensors such as miniature cameras, microphones, GPS, and accelerometers built in. They will have local area networking capabilities such as Bluetooth so that they can collect data such as heart rate and motion measurement devices from sensors or "sensor jewelry" worn on the body. Most significantly for researchers, many of the devices will be available "free" with the purchase of cellular phone contracts. As the technology matures,

stone_Chapter16.qxp 12/14/2006 8:19 PM Page 316

researchers will be able to exploit powerful computing infrastructures that have been purchased by people for their own entertainment, communication, and personal health care needs (Intille, Larson, & Kukla, 2002).

The cost of the technology will be spread across the consumer base. For example, the FCC has mandated that all new mobile phones must have GPS position-finding capability that can pinpoint the phone's position within 300 feet (Federal Communications Commission, 2004). Phone manufacturers are competing with each other to keep the cost of this technology, initially estimated to be $15 to $20 a phone, as low as possible. The ability to track a person in his or her community has benefits in nearly every subfield of preventive health research, and the capability to determine this information will exist in nearly all mobile phones in use by 2010.

The more significant barrier to the adoption of technology able to implement CS-EMA protocols may be the cost of developing, customizing, and maintaining the technology. Every study is different. Therefore, these costs will not drop unless tools are created that allow researchers to generate custom-designed EMA applications. Fortunately, the programming environments for mobile computing devices are improving. For example, many new mobile phones and PDAs can be programmed using platform-independent languages such as Java. Creating flexible tools for researchers so they can take advantage of CS-EMA and other new opportunities is a worthy and necessary goal toward achieving the vision of the ultimate mobile computing technology for running CS-EMA protocols described later in this paper.

Wireless Networks

The third force driving the adoption of CS-EMA will be "anytime, anywhere" wireless network capabilities. Three types of wireless networks will play a role. Wide-area wireless networks such as those used by the mobile phone carriers (e.g., GSM) can be used to send data long distances intermittently. High-speed local area wireless networks (e.g., 802.11b, or WiFi) that are common in workplaces and homes can be used for fast, low-cost transfer of video, audio, and other high-bandwidth data streams. On-body personal wireless networks such as Bluetooth can collect data from miniature, wearable on-body sensors. When combined, these three networks create new real-time data collection and analysis possibilities. For example, one lab prototype used a wireless heart rate chest strap monitor to send data via a personal wireless network to a receiver connected to a common mobile phone; the mobile phone sent the data via the wide-area network to a computer that processed the signal in real time and sent messages about heart condition back to the phone (Qi, 2003).

Activity Recognition Algorithms

The fourth trend that will lead to the development of new technologies that can support CS-EMA protocols is the creation and validation of accurate algorithms

stone_Chapter16.qxp 12/14/2006 8:19 PM Page 317

that use sensor data to automatically detect information, such as the subject's activity, about the subject's context.

The next three sections further discuss the three areas of innovation created by these technology trends: (1) continuous and passive collection of multimodal data, (2) context-sensitive prompting at appropriate times and places, and (3) data collection tailored to the individual and the situation.

Continuous and Passive Collection of Multimodal Data

Technological innovations will create new opportunities for continuous and passive collection of different types of data from people as they engage in free-living behavior. Manual data recording of certain data types, such as physiological states, is known to be difficult and error-prone (Hollenberg, Pirraglia, Williams-Russo, et al., 1997). Further, the more data that are being collected, the more burdensome and less practical self-report data collection becomes. The sensors for passive, multimodal data collection can be classified into two types: those worn or carried on the body and those placed in the environment.

Sensors Worn on the Body

Sensors worn by a person may passively collect data on three contextual factors: the person's motion, physiological states, and the state of the surrounding environment.

Motion Sensors

The actigraph may be the most widely used passive sensing device in use for studies of free-living subjects. An actigraph consists of one to three accelerometers in a watch-like casing that can be worn for long periods and used to estimate gross motion and energy expenditure (Melanson & Freedson, 1995). The actigraph has proven itself to be a useful tool to validate other instruments that measure physical activity (Welk, 2002). Unfortunately, commercial actigraphs do not permit data to be output in real time to a mobile computing device. They also do not sample motion data at sufficiently high rates so that the output from the devices can be used to detect specific activities.

New "real-time" actigraph extensions are being developed in laboratories. Figure 162 shows some "mobile MITes" (MIT environmental sensors): three-axis wireless accelerometers developed at MIT that transmit data to a PDA. Multiple mobile MITes can be worn simultaneously on different parts of the body. The PDA can either save the raw data for later analysis as a traditional actigraph does or process the multiple streams of acceleration data in real time to automatically infer ambulatory activities (e.g., walking, cycling) and the posture (standing, sitting, lying down) of the person. These motion states can be detected reliably in real time from raw accelerometer data, particularly when one accelerometer is placed on the upper body and one on the lower body and correlations between the two are detected (e.g., see Bao & Intille, 2004). Software on the PDA can then respond to the

stone_Chapter16.qxp 12/14/2006 8:19 PM Page 318

person's physical activity with a targeted question during or just after the behavior of interest. The MITes are designed to operate for a full waking day on a single coin cell battery. Participants can wear them continuously if they make one battery change each morning—a burden, but not an unacceptable one for many studies.

In Figure 16-2, mobile MITes transmit three-axis acceleration data in real time to a nearby PDA with a special sleeve that contains the receiver. The PDA can process the data to detect activities in real-time activity to support CS-EMA. MITes can be quickly installed in homes or other environments to detect use of specific devices and to trigger a context-sensitive self-report. The data, which are used to answer context-specific questions, can be transmitted to a PDA carried by the subject.

Some wireless motion sensors have already been commercialized. A wireless FootPod pedometer from FitSense can occasionally send distance traveled and speed data to a mobile computing device, such as a PDA or watch (FitSense Technology, Southborough, MA). The device is easy to use and easy to wear and can even transmit data to some Motorola mobile phones.

Accelerometers are not the only mechanisms that can measure body motion. Gyroscopes can also provide information about body motion, and they are small enough to include on the board shown in Figure 16-2. Gyroscopes, particularly when combined with accelerometers, could provide software running on a PDA or telephone with the relative orientation of the subject's limbs. Gyroscopes or accelerometers placed on both sides of a joint can detect the approximate joint angle. In laboratories, researchers are experimenting with up to 30 accelerometers placed simultaneously on a single person to determine the best number and placement for recognition of particular activities (Kern, Schiele, & Schmidt, 2003; Van Laerhoven, Schmidt, & Gellersen, 2002). Activity detection from these devices could enable CS-EMA protocols with activity-triggered self-reports.

Physiological Sensors

Dramatic improvements in the size, performance, and comfort of on-body physiological sensors for ambulatory monitoring are being made in laboratories, and

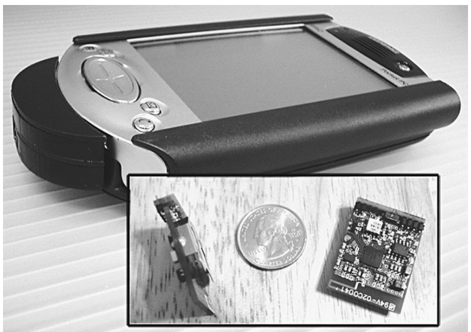

Figure 16-2. Mobile MITes.

stone_Chapter16.qxp 12/14/2006 8:19 PM Page 319

these advances are migrating to consumer devices. The HealthWear armband is worn on the upper arm and continuously collects data on arm movement, skin temperature, heart rate, and galvanic skin response (BodyMedia, Pittsburgh, PA). The LifeShirt is a washable shirt that contains embedded sensors that monitor physiological states such as heart rate, blood pressure, and respiration (VivoMetrics, Princeton, NJ). Although these devices cannot send data in real time to a CS-EMA device, emerging technologies will provide new opportunities for context-sensitive sampling that is triggered by physiological state. For instance, heart rate can be measured with a commercial wireless heart rate chest strap (Polar Electro Oy, Kempele, Finland) (Goodie, Larkin, & Schauss, 2000). Monitoring can take place in real time from a PDA when the appropriate receiver is attached (Rondoni, 2003). Systems that are more comfortable to wear for extended periods are in development. For instance, the SmartShirt is a prototype shirt that sends EKG, heart rate, respiration, temperature, and pulse oximetry data wirelessly to a receiver (Sensatex, Bethesda, MD).

Other sensors enable similar physiological triggering. Galvanic skin response measures perspiration, which is known to correlate with stress. The galvanic skin response can be passively measured using a modified biker's glove (Healey & Picard, 1998). Blood pressure can be measured with an arm or finger cuff that automatically inflates at researcher-specified intervals of time (Kario, Yasui, & Yokoi, 2003). Finger touch can be measured without encumbering the fingertip using a tiny sensor that measures color change in the fingernail (Mascaro & Asada, 2004). A small finger ring can measure blood oxygen levels and wirelessly transmit the signal to a mobile computing device (Asada et al., 2003). "Sensor pills" (Given Imaging, Yoqneam, Israel) that are swallowed can transmit data wirelessly to an on-body receiver, measuring the body's internal state (Mow, Lo, Targan, et al., 2004).

Real-time analysis of multiple data streams from diverse sensor types will create the most powerful context detectors. For example, suppose a researcher wishes to use CS-EMA to trigger subject self-reports only during situations when the subject is undergoing psychological stress. A psychological stress detection algorithm could monitor three sensors for the following states: high heart rate, high galvanic skin response, and low body motion (to rule out excessive physical activity). Only when all three conditions are met (or just after they subside) would the subject receive the stress diary assessment.

Unfortunately, some ambulatory monitoring devices are intrusive and not easy to wear. Blood pressure measurement, for instance, will startle the person because a cuff must inflate, causing interruption and discomfort. Commercialization of the technologies, however, is bringing miniaturization (Bonato, 2003). Most of these sensor devices will soon be wireless and available in physical styles that permit them to be worn by participants inconspicuously, under clothing, for days or weeks at a time without discomfort (Pentland, 1996). Some of these sensors provide continuous streams of data (e.g., motion sensors on the body). Others provide intermittent signals because of the invasive nature of the sensing itself (e.g., blood pressure).

stone_Chapter16.qxp 12/14/2006 8:19 PM Page 320

Wearable Sensors that Measure the Environment

Wearable sensors can measure not only motion or physiological states but also properties of the environment. For example, small ultra-violet photodiodes attached to a wireless transmitter worn on the wrist or neck can be used to send exposure data to software running on a PDA, allowing for sun-exposure context-sensitive queries. GPS systems in phones or PDAs can be used to identify a person's position outside, and software with a GIS database can trigger context-sensitive questions tailored to the subject's movement and the type of environment he or she is in (e.g. park, business district, or near a particular retail establishment).

Another property of the environment is sound, and mobile computing devices now contain sufficient amounts of storage to continuously record audio data. Mehl and colleagues have created the electronically activated recorder (EAR) (Mehl, Pennebaker, Crow, Dabbs, & Price, 2001). Subjects wear the EAR for 2–4 days during which time it records 30s samples of audio once every 12 minutes. The audio is then manually transcribed for analysis. Although the EAR records to tape, new versions of EAR-like devices that use digital devices could not only save the data but process the audio stream in real-time for important keywords or background sounds. If particular words or sounds are spotted, context-sensitive self-report diary assessments could be automatically triggered. In one study, researchers have used automatic speech recognition to determine how and when multiple subjects that work in the same physical environment are interacting with one another (Eagle & Pentland, 2003).

Beepers have been given to family members and simultaneously activated to signal multiple family members to make simultaneous entries on paper diaries (Larson and Richards, 1994). Radio frequency tags worn by subjects can *automatically* detect who is in a particular part of a space and when those people may be interacting with one another. Similarly, the networking capabilities of some mobile computing devices (e.g. Bluetooth) can be used to automatically determine when two people carrying the devices are within a short distance of one another.

Finally, video cameras can now be incorporated into phones or other mobile computing devices to create time-lapse video snapshots of what a person sees. A prototype extension for the MIT CS-EMA software has been used to capture such a record. The subject carries a PDA with a camera plug-in in the front shirt pocket. The device takes one (low-quality) time-stamped picture every 30 seconds that shows roughly what the subject sees. The image stream allows a researcher to more accurately reconstruct the subject's behavior, subjective states, cognitions, and interpersonal experiences throughout the day. The CS-EMA software can also prompt subjects to take pictures of certain objects in the environment. For example, subjects can be asked to take pictures of everything they eat—a new technique in nutrition research (Wang, Kogashiwa, Ohta, & Kira, 2002).

Environmental Sensors (Placed by the Researcher)

In addition to asking participants to wear sensors, other sensors can be installed in their environment. These sensors are also undergoing miniaturization and

stone_Chapter16.qxp 12/14/2006 8:19 PM Page 321

therefore are becoming less of a burden for participants to endure and for researchers to install. Ubiquitous sensor placement is made possible by reduction in sensor size and lengthening of battery life. Work is underway in laboratories to create wireless environmental sensors that are less than a cubic millimeter in size (Kahn, Katz, & Pister, 1999).

Sensors placed on objects in the environment provide information about what people in that environment are doing. If the sensors can wirelessly transmit information to a mobile computer, automatic detection of everyday activities can trigger context-sensitive sampling. For example, at MIT we are conducting experiments where MITes (and their earlier versions) are installed in participants' homes. The sensors are attached to any object that the person may manipulate: doors, cabinets, light switches, dials, appliances, and even large containers such as Tupperware (Intille, Munguia Tapia, Rondoni, et al., 2003). The sensors measure when the objects are moved.

Figure 16-3 shows some early versions of MITes that used magnetic contact switches "installed" in one participant's home. The sensors were literally taped to objects in the home and then used to collect data for 2 weeks. New wireless MITes dramatically cut installation time by using accelerometers instead of contact switches. Algorithms have been developed that use the data from the sensors to automatically detect some activities of daily living using statistical models (Munguia Tapia, Intille, & Larson, 2004).

Other researchers are implementing similar activity-detection algorithms that use different types of environmental sensors. For example, Philipose and colleagues have created a bracelet with an embedded radiofrequency identification tag (RFID) reader. When a subject wears the bracelet and inexpensive RFID stickers are placed on objects in the subject's home, a computer system can detect when the subject touches each of the tagged objects. Algorithms that recognize everyday activities (e.g., making tea) can then be detected automatically, in real time (Philipose, Fishkin, Fox, et al., 2003). In combination with software on a PDA, such systems could be used for studies where the subject's behavior is monitored for specific eating patterns and then context-sensitive questions are asked during or after particular target behaviors.

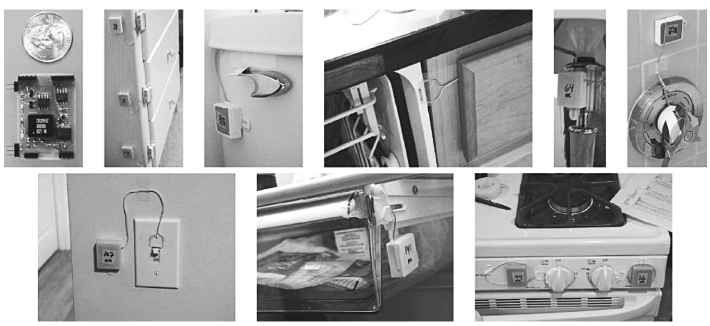

Figure 16-3. Early versions of MITes.

stone_Chapter16.qxp 12/14/2006 8:19 PM Page 322

Other sensors such as microphones or video cameras can easily be placed throughout an environment for research studies. For instance, automatic computer vision analysis of video streams from inexpensive webcams attached to laptops can be used to detect when a person is in a particular room and can transmit his or her location to a PDA running assessment software. The software can implement a CS-EMA protocol that asks for self-reports depending upon how people move about the environment. Air quality properties such as carbon monoxide and dioxide levels, temperature, and even particulate levels can be measured with sensors and used as CS-EMA triggers.

Not all environmental monitoring must lead to CS-EMA-triggered interruptions for the participant. For example, image data can be used as a memory trigger rather than a self-report trigger to minimize the interruption burden of sampling. We call this technique image-based EMA. In a prototype implemented at MIT, a computer monitors a video stream for motion events. When significant motion is detected, a sample is triggered. However, instead of interrupting the subject with a prompt, a static image of the environment and the person is taken, time-stamped, and stored. Later, at a convenient time, such as when riding the bus, waiting in line, or just being idle, the participant can use a PDA device to answer questions. The image-based EMA program displays previously captured images and then presents the researcher's question. In many cases, the image will trigger the subject's memory, helping him or her provide more accurate answers to the assessment questions by providing visual contextual cues about that particular moment in time (Intille, Kukla, & Ma, 2002).

Environmental Sensors Already in Place

Although in many studies investigators may wish to have participants wear sensors or may place them in their living environments, researchers may also be able to exploit the proliferation of sensors in devices and environments. Working with public institutions or corporations, researchers may obtain access to valuable data sources. With appropriate conversion software, it may be possible to send the data via wireless networks to a device enabling CS-EMA protocols. For example, suppose a researcher wanted to explore the relationship between sedentary behavior and television viewing. TiVo (TiVo, Alviso, CA) or another digital recorder company might provide the researcher access to real-time data on the current television station settings. These data could be combined with data from wireless accelerometers worn by participants in the home and enable real-time context-sensitive questioning that is tailored to how people are moving and what content they are watching.

Similarly, suppose a researcher is interested in studying the relationship between sedentary behavior, driving, and the built environment. An automobile manufacturer may be able to provide real-time data from the vehicle itself, and more vehicles have built-in GPS. A PDA will be able to collect data from wireless accelerometers worn by the person, a GPS system in the car, and GIS information obtained over the Internet to implement a CS-EMA protocol where questions are triggered based on driving and walking behavior.

stone_Chapter16.qxp 12/14/2006 8:19 PM Page 323

Obtaining access to data sources from companies or governments can be challenging because these organizations often wish to protect the privacy of their patrons, as well as their trade secrets. Furthermore, implementing the technology required to get access to the data in real time to enable CS-EMA may not always be possible without substantial investment in technology middleware. However, as devices such as home theater systems and automobiles improve, the technologies they use to communicate with other electronic devices are likely to improve, making possible novel data transfer for real-time health assessment research.

Other existing data sources installed in the environment that might interest health researchers are public surveillance cameras, satellite imagery, usage data from home appliances such as stovetops, and radiofrequency tollbooth payment trackers.

In summary, researchers using real-time data collection have new tools at their disposal for passively collecting multimodal data in real time and, most significantly, triggering additional data collection based on automatic, real-time processing of the incoming data streams.

Context-Sensitive Prompting at Appropriate Times and Places

The key technological innovation enabling the CS-EMA methodology is the automatic detection of context. In the simplest case, questions can be triggered in response to a single, prespecified sensor reading—for example, heart rate jumps above a subject-specific threshold. Alternatively, the computer combines many sensor readings simultaneously to infer more complex subject states using pattern recognition algorithms. However a trigger is detected, the key insight is that a subject's context and activities can be used to determine when and where self-report data are collected. Mobile computing devices, including mobile phones, are now sufficiently powerful to perform this "just-in-time" questioning (Intille, 2002).

Conveying Information at the Appropriate Place

Telephones have been used for many years to deliver and collect health information from people in their home settings (Friedman, 1998). More than 95 percent of U.S. homes have telephone service (U.S. Census Bureau, 1990). Although the telephone is a ubiquitously available technology, as a technology for real-time data collection it has two serious limitations. First, it is fixed in the home and most people spend a significant amount of time in other settings. Second, the telephone, without additional technologies, has no information available about when it is opportune to prompt for information. Finally, the technology required to use the traditional phone system to proactively call people is costly and complex to maintain.

By comparison, new mobile computing devices are making "anyplace" communication possible at a very low cost. Mobile phones are exploding in usage,

stone_Chapter16.qxp 12/14/2006 8:19 PM Page 324

and many people carry them everywhere they go. There is almost one mobile phone for every two people in the United States (*Trends in Telephone Service*, 2003). The percentage of mobile phone owners will grow because many people are replacing wired telephones with wireless service (Horrigan, 2003). By 2002 in Finland (which the U.S. tends to follow in technology adoption), 86 percent of the male population and 78 percent of the female population over age 10 had a mobile phone (Statistics Finland, 2003). As these devices are further miniaturized and more sensing capability is incorporated into them, they can be exploited by health researchers to get information to or acquire information from subjects (Collins, Kashdan, & Gollnisch, 2003). New phones will have built-in GPS positioning that will automatically provide position information when outdoors. Wireless networking capability may soon provide information about location when the devices are inside buildings (Bahl & Padmanabhan, 2000).

Other emerging technologies will make it easier to get messages to a subject at the appropriate place. Bluetooth headsets such as the FreeSpeak (JABRA Corporation, Ballerup, Denmark) can wirelessly transmit a private audio message to a person's ear from a mobile computing device. Heads-up display devices such as the SV-6 PC Viewer (MicroOptical Corporation, Westwood, MA) that attach to the ear or glasses and provide a video display to a person as he or she moves about the environment are also dropping in price. Even interactive voice response systems in vehicles, such as OnStar (General Motors, Detroit, MI), could be used to collect data from subjects.

Looking 10 to 15 years into the future, new low-cost display devices for homes and workplaces may transform typical home environments into spaces where information can be presented not only in the right room but also on the most appropriate object, such as furniture, walls, floors, overlaying a picture, or placed on a particular device such as an outlet.

Figure 16-4 shows a prototype living room at MIT. In the corner is an Everywhere Displays Projector (EDP) (Pinhanez, 2001) consisting of a computer

Figure 16-4. Prototype living room with an Everywhere Displays Projector in one cabinet; it can display information on most surfaces in a typical-size room.

stone_Chapter16.qxp 12/14/2006 8:19 PM Page 325

projector, a computer-controlled mirror, and a standard desktop computer. After calibrating the device, the computer can use a perspective transformation to warp an image so that, when it is projected onto objects in the environment, it retains its correct size and shape. If the computer is also connected to a computer-controlled camera, it can detect the laser dot and allow the laser pointer to be used to "click" on the image, no matter where it is projecting (Intille, Lee, & Pinhanez, 2003). Such advances in computer hardware and the widespread adoption of consumer technologies for communication and entertainment will create new opportunities to display messages at the appropriate place, possibly allowing prompting strategies that grab a person's attention without audio interruption.

Detecting an Appropriate Time

A greater challenge than presenting or requesting information at the right place may be to develop algorithms that can use sensors to determine an appropriate time.

Detecting the Activities that Matter

Entire fields in engineering are devoted to developing algorithms that can collect noisy sensor readings and infer a person's activity or context. Recent advances in pattern recognition algorithm development and improvements in processor speed have led to systems that can robustly infer events from noisy sensor data. The most promising techniques typically use a supervised learning algorithm, where a set of labeled examples is used to create exemplar models of each target class and then new examples are classified based on similarity functions with the targets (Witten & Frank, 1999).

Using pattern recognition, raw sensor data can be converted into more meaningful labels. For example, data on limb motion collected from accelerometers worn on multiple body parts can be used to automatically determine which of a list of known activities such as "walking," "cycling," "scrubbing," "working at a computer," and "vacuuming" a person may be performing (Bao & Intille, 2004). One of the key challenges for health researchers is to learn how to exploit these pattern recognition algorithms. The algorithms are statistical—they output only a statistical likelihood that an activity is observed given the observed data. Therefore, when developing CS-EMA sampling protocols and performing data analysis, researchers need to account for both the variability among subjects and the uncertainty in the triggering algorithms. The data analysis methods needed to analyze data collected with CS-EMA tools demand future research.

Minimizing Interruption Burden

In most electronic sampling studies, the researcher assumes that, given a full day and random sampling, the sampling will eventually capture a sufficient number of events of interest. However, a well-known limitation of EMA is the burden the technique places on the participant. One way to reduce the burden is to

stone_Chapter16.qxp 12/14/2006 8:19 PM Page 326

keep the measures brief to avoid subject fatigue (Stone & Shiffman, 1994). This strategy, however, may not be sufficient for long-term sampling studies or for studies where the events of interest are short in duration.

The CS-EMA methodology may provide a solution by limiting sampling to the moments of greatest interest to the researcher. A variation on this approach is to also use context-sensitive computing to determine good (or poor) times to interrupt if a specific activity of interest cannot be detected. Because of the proliferation of electronic devices vying for the user's attention, the detection of good times to interrupt using sensors is becoming an active topic of research for computer scientists (Horvitz & Apacible, 2003). Ironically, EMA is a technique being used to conduct these studies. In one study conducted in the workplace setting, researchers found that a few sensors that detect use of the phone, if anyone is talking, or other indicators of conversation could be better than 75 percent effective at predicting interruptability in the office (Hudson, Fogarty, Atkeson, et al., 2003). Another study found that information proactively presented at physical activity transitions (e.g., a transition from sitting to walking) rather than at random times has a higher perceived value (Ho & Intille, 2004). Future versions of CS-EMA might use such sensors to reduce the burden of sampling by triggering prompts during automatically detected activities or during activity transitions when participants are likely to be most receptive to an interruption.

Exploiting the Network

Even with a variety of sensor inputs available, automatic detection of subject activity is challenging for many activities of interest. For example, although galvanic skin response and other physiological measures can provide some information about emotional state, reliable detection of most emotions is still an active research problem (Picard, 2003). Similarly, automatic detection of certain activities or parameters of those activities (*"he's cooking lazily," "she's being passive-aggressive"*) may be far beyond current capabilities of machine perception. In these cases, a researcher interested in using CS-EMA may need to fall back on other EMA designs.

An alternative is to exploit the ubiquitous networking capability of emerging computing devices. For example, suppose two people living in the same home are participating in a study. Instead of sampling the target individual, if PDA software detects that the two people are near one another, the software might silently prompt the second individual (perhaps using vibration) and present questions about the state of the first. The software can attempt to exploit human intelligence of other people as triggers for sampling the subject of greatest interest. Moreover, answers from the second individual might change the sequence of questions asked of the first. Although this network-triggered CS-EMA protocol could occur in a home, it might also occur in a workplace or other environment where large groups of people work and where not all the people participating know each other well.

Text messaging using the network could also trigger data collection among groups. Text messages have already been used as economical interventions for

stone_Chapter16.qxp 12/14/2006 8:19 PM Page 327

diabetes for children (Franklin, Waller, Pagliari, & Greene, 2003), asthma management (Neville, Greene, McLeod, et al., 2002), and bulimia research (Bauer, Percevic, Okon, et al., 2003). Clever use of simultaneous communication between groups of subjects could heighten the sense of perceived value of the experience for subjects by exploiting social norms. For instance, if subjects in a study are receiving prompts at the same time, a real-time update about who has answered the queries might be a useful tool to increase compliance.

Data Collection Tailored to the Individual and to the Situation

The CS-EMA methodology as described thus far involves continuous collection and analysis of data that are used to prompt self-report at appropriate times and places. However, emerging computer technologies will also facilitate opportunities for personalized tailoring based on an individual's past behavior, subjective states, cognitions, and interpersonal experiences.

User-Specific Devices

The mobile computing technologies that will enable the next generation of CS-EMA instruments will be highly personal devices. Already mobile phone users are beginning to use their phones to store detailed personal information and files. As the interfaces between the phones and desktop computers are improved in the next 5 years, phones will become the primary mechanism by which people carry and exchange data. For heavier computer users, phones are likely to contain information about personal contacts, meetings, and movement patterns around the community (from GPS). It is not unreasonable to expect the mobile devices to also contain large amounts of health data. As the physiological and motion sensors reach comfortable sizes, first athletes and then the general public will begin to use them for personal health monitoring.

Longitudinal Data Storage—Lifelong Records

A trend toward personalizing mobile computing devices will take place at the same time as another important technological innovation: "free memory." Storage devices the size of a quarter can now store gigabytes of information, and the cost of memory has plummeted. For all practical purposes, within 10 years digital memory storage capacity will no longer be a consideration for most typical data collection tasks, even those that take place on mobile computing devices. "Free memory" and personalized devices will mean that people will never need to delete the digital data they collect. Suppose a mobile phone user has a wireless pedometer that saves walking data and GPS position data on the phone. That person could easily save a lifetime of walking data on a single phone and never have a reason to delete it. In fact, as discussed in the next section, the

stone_Chapter16.qxp 12/14/2006 8:19 PM Page 328

phone might collect a longitudinal record of nearly everything the person sees, hears, and does.

Looking to the Future: CS-EMA+10

In the next 10 years, new technologies will enable CS-EMA protocols to be widely deployed for health research. How will the properties of new CS-EMA instruments compare with the current dominant technology of paper? Consider what an instrument that implements a CS-EMA protocol might be like 10 years from now (CS-EMA+10), after mobile computing, wireless sensors, ubiquitous networks, and "free memory" have permeated the marketplace.

A device implementing a CS-EMA protocol for studying how, when, and where people engage in physical activity might operate as follows. The device would consist of software that runs on all major mobile phones. Ten years from now, well over 150 million mobile phones will be in use daily in the United States. People who wish to become research subjects will obtain the software using the touch screen on the phone to access the project website and download and run the program.

The typical phone will be a powerful, miniature computer with a 0.5- to 1-terrabyte storage card; it will easily synchronize with a computer via low-power wireless networks. Some people will use phones with a pocket-sized physical shape and style; others will have phones with watch or armband physical styles. Nearly all phones will have color screens with touch-input displays.

Each phone will have a local body network used to collect data from built in and add-on sensor devices. Pedometers and some physiological monitors (e.g., heart rate) may be common. These devices will always be activated and collecting data, most likely so phone users can monitor their own health. Researchers using phones for CS-EMA protocols may distribute additional sensors to subjects as well, and the sensors will use the local body network to transmit data to sampling software running on the phone. For physical activity measurement, the sensors will consist of miniature accelerometers embedded in comfortable elastic wrist and ankle bands and a wireless heart rate chest patch. Subjects may wear them for months, charging their batteries at night. Two to five accelerometer bands may be worn on different parts of the limbs. Additional data will be collected on the subject's position from the phone's built-in GPS or indoor tracking system using the local network inside buildings. The researchers may ask subjects to place a few hundred sensors like MITes around their home to improve recognition of certain types of activities and detect events such as eating, opening the refrigerator, cooking, exercising, sleeping, and socializing.

The software for monitoring physical activity will save and then immediately process the incoming signals. The accelerometers will send three-axis accelerometer data to the phone at 20 Hz. The phone will continuously process the sensor data and use statistical learning algorithms to automatically detect certain physical activities specified by the researcher. The algorithms will recognize many activities with 80 to 90 percent accuracy (Bao & Intille, 2004). Using more sensors will typically improve recognition performance. Heart rate in combination with the accelerometer data will determine what activity is being performed and with what intensity.

When a target physical activity is detected, such as "walking," "cycling," "scrubbing," or "vacuuming," the phone will prompt the subject and present context-specific (multiple-choice) questions that the researcher has constructed. The questions may change based on the subject's prior activity, including pedometer data collected 2 years ago but saved on the subject's phone. Subjects will use the phone's touch screen to answer the questions after receiving an audio prompt.

As it gathers the self-report data, the CS-EMA+10 device can create a "digital diary" of the subject's activity by recording what a user sees and hears (Clarkson, 2002; Gemmell, Bell, Lueder, et al., 2002). The researcher can use this diary to interpret surprising results. The value of such an experiential memory device was realized as early as 1945, but only recently have improvements in the size and cost of technology made the vision achievable (Bush, 1945; Gemmell et al., 2002).

The phone will collect all of the following information: a continuous video stream (from an embedded camera), a continuous audio stream of everything the subject hears and says (from the phone's microphone), a continuous accelerometer data stream of the subject's limb motion, a continuous data stream of physiological parameters such as heart rate, a continuous data stream of subject's location in the community, and other miscellaneous data about how the subject is feeling as reported by the user occasionally via a mobile computing device user interface. When compressed, a year of data can be stored with less than 1 terabyte of memory. By 2007, 1 TB of memory may retail for approximately $300 (Gemmell et al., 2002).

In some situations the software running on the telephone may change the question that is asked based on an automatically detected deviation from past baseline performance. For example, a person may be reminded of previous answers from the previous days, weeks, or years (e.g., *"Yesterday at this time you felt tired. Do you feel more or less tired right now?"* or *"You've done this activity at about the same time every day for the last month. Why do you think that is the case?")*. The device can go one step further, however, and use image-based EMA to show the subject video or play the subject audio from periods of time in the past. The computer might play a video clip of an activity that occurred 2 months ago and ask the subject to compare the current situation to the past event.

People using the telephone software might participate in CS-EMA studies for years if researchers ensure that the context-sensitive interruptions do not become a burden. Participation would require no more than telling the telephone how many questions a week are tolerable and letting the program automatically run in the background. In some studies, participants using devices running CS-EMA protocols may start to view themselves less as "subjects" and more as "active volunteers." The software could be structured so that some studies are experienced as long-term participatory conversations rather than impersonal and untailored question-and-answer sessions.

stone_Chapter16.qxp 12/14/2006 8:19 PM Page 330

All the technologies that are needed to make this scenario a reality exist today and prototypes are being tested at MIT and elsewhere. As consumer technologies drop in size and price, such instruments could be deployed on very large scales with millions of telephone users. Open research questions include how to analyze the huge amount of data that such a tool would produce and how to provide researchers with the tools they need to develop context-sensitive sampling protocols without requiring expertise in the sensor technologies.

Privacy, Ethical, and Practical Considerations

As with any research study, the use of CS-EMA protocols, particularly those deployed with tools available 10 years from now, requires that informed consent be obtained from subjects. The investigators must ensure that participants understand what data are being collected, what the data could be used to infer, and how they can opt out of data being collected at particular times. Investigators must take the most care when sensors are being used to continuously collect data without the direct input of participants, because they may not realize how much data a consumer device such as a mobile phone could (or already does) collect. They may also forget that data are being collected when no proactive action is required.

A device implementing a CS-EMA protocol may be capable of recording a person's location, activities, interactions with other people, physiological states, and psychological states. Such descriptive data sets may make it increasingly likely that EMA data collected in one study can be retroactively reused for a different purpose in another study at a later date. However, the density and versatility of the datasets may also make it more likely that datasets will be of value to other parties, such as the government or family members interested in the behavior of particular subjects. The probative power of the data sets may necessitate improved procedures for masking subject identity. This is especially true because the data sets may contain audio and video samples that can be used to identify the participant, and current technologies do not allow researchers to automatically anonymize those data sources without obscuring much potentially relevant data.

The EMA technology that is continuously monitoring raises several other concerns that need to be addressed. One is that investigators may wish to request data collected by consumer devices used by participants prior to their enrollment in the study. A standard mobile phone, for instance, may store detailed information on the person's prior whereabouts, social connections, and physiological state—perhaps from the last several years. A participant may wish to provide only the relevant longitudinal data to the researcher, but in practice it will be more difficult to recall and flag data that he or she does not wish the researcher to see. The participant may therefore turn over the data without fully comprehending the amount of information it contains.

stone_Chapter16.qxp 12/14/2006 8:19 PM Page 331

Another concern raised by EMA continuous monitoring is that some sensors may unintentionally capture data on nonconsenting people who interact with the subject. In situations where data about interaction with others are collected, it may not be possible for the investigator to obtain informed consent from all people with whom the subjects interact, either before or after the study. Such data could be removed once it is identified, but marking instances for deletion may require seeing of the data.

Finally, EMA continuous monitoring may capture data from the participant that is unrelated to the core study goals but that requires a response by the investigator. For example, while analyzing heart-rate–prompted questions, the investigators may discover heart arrhythmias. Should the researcher communicate this information to the subject? Alternatively, the investigators may determine that the subject engaged in an illegal activity. Do they report this information to the authorities? Subjects may also mistakenly feel that they are being monitored in real time and adjust their behavior accordingly. A researcher would not want a subject to defer medical care because he or she mistakenly felt that the system monitoring heart rate would trigger a warning if a serious heart problem was observed.

In addition to privacy and ethical concerns, future CS-EMA studies offer some practical challenges. A device running a CS-EMA protocol on a future computing device could generate orders of magnitude more data than current studies. New tools will be needed to manage these datasets and allow researchers to easily manipulate multimodal data sources with sampling rates that vary from days to milliseconds. Although many studies will be possible using standard telephones and commercial sensors such as electronic pedometers, in other cases a substantial infrastructure investment may required to collect the data desired (e.g., television content from a proprietary digital video recorder/tuner). Once again, special tools may be required for some studies that collect and process data. In most cases, the context detection algorithms will function as "black boxes." This may create some analysis challenges because the researcher may not have access to reliable data on the performance on the algorithms that are analyzing data to trigger self-reports.

Conclusions

Electronic sampling techniques, and particularly context-sensitive sampling techniques that use new sensor technologies, will create exciting new opportunities for health researchers interested in studying behavior, subjective states, cognitions, and interpersonal experiences of free-living subjects in natural settings. The CS-EMA methodology is likely to provide value for any health-related research where behavior and context are intertwined. For example, one can imagine studies such as the following:

- A study using CS-EMA to gather data on the links between television viewing, eating behaviors, and sedentary behaviors. Sensors worn on the

stone_Chapter16.qxp 12/14/2006 8:19 PM Page 332

body, placed in the kitchen, and placed on the television trigger questions when people are sedentary, snacking, and/or watching particular types of television shows.
- A study on ultraviolet exposure where questions about sun exposure are triggered only when a subject has been getting unusually high UV exposure for at least 15 minutes
- A study on walking behavior and the built environment, where subjects are asked specific questions based on their typical walking behavior, current walking speed and duration, and type of space they are currently in
- A study on the impact of social contact on perceptions of psychological well-being, where a group of friends or coworkers are prompted to describe how they feel just after an encounter with one another. Audio snippets from the conversations could be saved and analyzed by the researchers.

Table 16-1 (presented earlier in the discussion of emerging technologies) compares four technologies for real-time assessment: paper used for a diary, PDA software used for computerized EMA, PDA devices with context-sensitive software and sensors used for CS-EMA, and mobile devices 10 years from now used for CS-EMA+10. The table, which is speculative and makes gross simplifications, nonetheless suggests that researchers interested in studying free-living subjects may have a powerful new assessment instrument at their disposal as the technology that enables widespread adoption of CS-EMA matures. This new tool may be characterized by the following:

- A CS-EMA+10 device will be as lightweight and compact as the typical mobile phone, which most people will carry everywhere they go. The CS-EMA+10 device will be as inconspicuous as the mobile phone, which will have permeated society.
- A CS-EMA+10 device will be inexpensive to deploy for research. Consumers will absorb the cost of the mobile computer and sensor technology for personal entertainment and communication purposes. Hundreds of millions of people will own sufficiently powerful mobile computers, permitting instruments to be deployed in studies with extremely large sample sizes.
- A CS-EMA+10 device will not have a flexible physical form factor but it will have a comfortable one. It will have flexible function; the researcher can structure the way that information is presented to the subject to achieve a particular end.
- Many versions of CS-EMA+10 devices will be easy to replicate and modify without special expertise because tools will be developed so researchers can set up context-sensitive protocols. The investment in such tools will be justified by the enormous sample sizes that can be obtained for tools that are widely deployed.
- The CS-EMA+10 technology will operate on mobile computing devices that are familiar to subjects and easy for most people to use.

- The CS-EMA+10 technology can be developed so that participants can correct and adapt their entries—for example, by leaving audio notes for the researcher to explain their answers.
- The CS-EMA+10 tool may still raise questions about validation, reactivity, and selection bias, as does current electronic EMA.
- Although CS-EMA+10 instruments will be easy to carry everywhere, they will not be easy to forget because they will proactively prompt for information and the technology will be the subject's primary phone.
- Most data collected via CS-EMA+10 tools will not require costly, time-consuming, and error-prone manual data entry and coding.
- CS-EMA+10 technology will have automatic time stamping of entries, prohibiting hoarding. It will also allow the researcher to ensure that questions are answered in a particular order.
- The CS-EMA+10 device will allow acquisition of both self-report data and passive acquisition of multimodal data.
- In ways not possible via any other instrument, the CS-EMA+10 device will permit researchers to automatically tailor data acquisition to the context.

Finally, although this chapter is about the use of new technologies for real-time, context-sensitive assessment, the same technologies could be powerfully deployed for just-in-time intervention (Intille, 2004). A computing device can automatically determine when and how to present information to motivate behavior, attitude, or belief changes at "teachable" moments. Evaluation of such technologies is likely to require the use of the CS-EMA protocols.

ACKNOWLEDGMENTS

Dr. Intille is supported, in part, by National Science Foundation ITR grant #0112900 and the House_n Consortium. He thanks Jennifer Beaudin for comments on a draft of this paper. Emmanuel Munguia Tapia designed the wireless sensors shown in Figure 16-2. John Rondoni implemented the first version of the MIT PDA software for CS-EMA.

Note

1. In past work, this extension to the EMA methodology has been called context-aware experience sampling, or CAES (CAES, 2003; Intille, Rondoni, Kukla, et al., 2003). Context-aware is a term commonly used by ubiquitous computing researchers to indicate systems that automatically detect and respond to context.

References

Asada, H. H., Shaltis, P., Reisner, A., Rhee, S., & Hutchinson, R. C. (2003). Mobile monitoring with wearable photoplethysmographic biosensors. *IEEE Engineering in Medicine and Biology Magazine, 22*(3), 28–40.

Bahl, P., & Padmanabhan, V. N. (2000). RADAR: an in-building RF-based user location and tracking system. In *Proceedings of IEEE Infocom 2000* (Vol. 2, pp. 775–784). IEEE Press.

stone_Chapter16.qxp 12/14/2006 8:19 PM Page 334

Bao, L., & Intille, S. S. (2004). Activity recognition from user-annotated acceleration data. In A. Ferscha & F. Mattern (Eds.), *Proceedings of PERVASIVE 2004* (Vol. LNCS 3001, pp. 1–17). Berlin Heidelberg: Springer-Verlag.

Barrett, L. F., & Barrett, D. J. (2001). An introduction to computerized experience sampling in psychology. *Social Science Computer Review, 19*(2), 175–185.

Bauer, S., Percevic, R., Okon, E., Meermann, R., & Kordy, H. (2003). Use of text messaging in the aftercare of patients with bulimia nervosa. *European Eating Disorders Review, 11*, 279–290.

Bonato, P. (2003). Wearable sensors/systems and their impact on biomedical engineering. *IEEE Engineering in Medicine and Biology Magazine, 22*(3), 18–20.

Broderick, J. E., Schwartz, J. E., Shiffman, S., Hufford, M. R., & Stone, A. A. (2003). Signaling does not adequately improve diary compliance. *Annals of Behavioral Medicine, 26*(2), 139–148.

Brodsky, M. Z., Consolvo, S., & Walker, M. (2002). *IESP Software.* Retrieved April 15, 2004, from http://seattleweb.intel-research.net/projects/ESM/iESP.html

Bush, V. (1945, July). As we may think. *The Atlantic Monthly, 176,* 101–108.

CAES. (2003). *Context-Aware Experience Sampling Website.* Retrieved January 15, 2003, from http://web.media.mit.edu/~intille/caes

Clarkson, B. P. (2002). *Life patterns: Structure from wearable sensors.* Unpublished Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA.

Collins, R. L., Kashdan, T. B., & Gollnisch, G. (2003). The feasibility of using cellular phones to collect ecological momentary assessment data: application to alcohol consumption. *Experimental and Clinical Psychopharmacology, 11*(1), 73–78.

Csikszentmihalyi, M., & Larson, R. (1987). Validity and reliability of the Experience-Sampling Method. *The Journal of Nervous and Mental Disease, 175*(9), 526–536.

Eagle, N., & Pentland, A. (2003). Social network computing. In A. K. Day, A. Schmidt, & J. F. McCarthy (Eds.), *Proceedings of the International Conference on Ubiquitous Computing* (Vol. LNCS 2864, pp. 289–296). Berlin, Heidelberg: Springer-Verlag.

Federal Communications Commission. *Enhanced 911.* Retrieved May 31, 2004, from http://www/fcc/gov/911/enhanced

Franklin, V., Waller, A., Pagliari, C., & Greene, S. (2003). "Sweet Talk:" text messaging support for intensive insulin therapy for young people with diabetes. *Diabetes Technology and Therapeutics, 5*(6), 991–996.

Friedman, R. H. (1998). Automated telephone conversations to assess health behavior and deliver behavioral interventions. *Journal of Medical Systems, 22*(2), 95–102.

Gandy, M., Starner, T., Auxier, J., & Ashbrook, D. (2000). The Gesture Pendant: a self-illuminating, wearable, infrared computer vision system for home automation control and medical monitoring. In *Proceedings of the Fourth International Symposium on Wearable Computers (ISWC'00)* (pp. 87–102). IEEE Press.

Gemmell, J., Bell, G., Lueder, R., Drucker, S., & Wong, C. (2002). MyLifeBits: fulfilling the Memex vision. In *Proceedings of ACM Multimedia '02* (pp. 235–238). New York: ACM Press.

Goodie, J. L., Larkin, K. T., & Schauss, S. (2000). Validation of Polar heart rate monitor. *Journal of Psychophysiology, 14*(3), 159–164.

Gorin, A. A., & Stone, A. A. (2001). Recall biases and cognitive errors in retrospective self reports: A call for momentary assessments. In A. Baum, T. Revenson, & J. Singer (Eds.), *Handbook of health psychology* (pp. 405–413). Mahwah, NJ: Erlbaum.

Healey, J., & Picard, R. (1998). StartleCam: a cybernetic wearable camera. In *Proceedings of the Second International Symposium on Wearable Computers* (pp. 42–49). IEEE Press.

stone_Chapter16.qxp 12/14/2006 8:19 PM Page 335

Ho, J., & Intille, S. S. (2004). *Using context-aware computing to reduce the perceived burden of interruptions from mobile devices* (Technical report). Cambridge, MA: Massachusetts Institute of Technology.

Hollenberg, J. P., Pirraglia, P. A., Williams-Russo, P., Hartman, G. S., Gold, J. P., Yao, F. S., & Thomas, S. J. (1997). Computerized data collection in the operating room during coronary artery bypass surgery: a comparison to the hand-written anesthesia record. *Journal of Cardiothoracic and Vascular Anesthesia, 11*(5), 545–551.

Horrigan, J. B. (2003). *Consumption of information goods and services in the United States.* Washington, DC: Pew Internet and American Life Project.

Horvitz, E., & Apacible, J. (2003). Learning and reasoning about interruption. In *Proceedings of the 5th International Conference on Multimodal Interfaces* (pp. 20–27). New York: ACM Press.

Hudson, S. E., Fogarty, J., Atkeson, C. G., Avrahami, D., Forlizzi, J., Kiesler, S., Lee, J. C., & Yang, J. (2003). Predicting human interruptability with sensors: A Wizard of Oz feasibility study. In *Proceedings of the Conference on Human Factors and Computing* (pp. 257–264). New York: ACM Press.

Hufford, M. R., Shields, A. L., Shiffman, S., Paty, J., & Balabanis, M. (2002). Reactivity to ecological momentary assessment: an example using undergraduate problem drinkers. *Psychology of Addictive Behaviors: Journal of the Society of Psychologists in Addictive Behaviors, 16*(3), 205–211.

Hyland, M. E., Kenyon, C. A., Allen, R., & Howarth, P. (1993). Diary keeping in asthma: Comparison of written and electronic methods. *British Medical Journal, 306*(6876), 487–489.

Intille, S. S. (April/June, 2002). Designing a home of the future. *IEEE Pervasive Computing,* 80–86.

Intille, S. S. (2004). A new research challenge: persuasive technology to motivate healthy aging. *Transactions on Information Technology in Biomedicine, 8*(3), 235–237.

Intille, S. S., Kukla, C., & Ma, X. (2002). Eliciting user preferences using image-based experience sampling and reflection. In *Proceedings of the CHI '02 Extended Abstracts on Human Factors in Computing Systems Human Interface Short Abstracts* (pp. 738–739). New York: ACM Press.

Intille, S. S., Larson, K., & Kukla, C. (2002). Just-in-time context-sensitive questioning for preventative health care. In *Proceedings of the AAAI 2002 Workshop on Automation as Caregiver: The Role of Intelligent Technology in Elder Care.* Menlo Park, CA: AAAI Press.

Intille, S. S., Lee, V., & Pinhanez, C. (2003). Ubiquitous computing in the living room: concept sketches and an implementation of a persistent user interface. In *Adjunct Proceedings of the Ubicomp 2003 Video Program* (pp. 265–266).

Intille, S. S., Munguia Tapia, E., Rondoni, J., Beaudin, J., Kukla, C., Agarwal, S., et al. (2003). Tools for studying behavior and technology in natural settings. In A. K. Dey, A. Schmidt, & J. F. McCarthy (Eds.), *Proceedings of UbiComp 2003: Ubiquitous Computing,* (Vol. LNCS 2864, pp. 157–174). Berlin, Heidelberg: Springer.

Intille, S. S., Rondoni, J., Kukla, C., Anacona, I., & Bao, L. (2003). A context-aware experience sampling tool. In *Proceedings of the Conference on Human Factors in Computing Systems: Extended Abstracts* (pp. 972–973). New York: ACM Press.

Kahn, J. M., Katz, R. H., & Pister, K. S. J. (1999). Mobile networking for "Smart Dust." In *Proceedings of the 5th Annual ACM/IEEE International Conference on Mobile Computing and Networking* (pp. 271–278). New York: ACM Press.

Kario, K., Yasui, N., & Yokoi, H. (2003). Ambulatory blood pressure monitoring for cardiovascular medicine. *IEEE Engineering in Medicine and Biology Magazine, 22*(3), 81–88.

stone_Chapter16.qxp 12/14/2006 8:19 PM Page 336

Kern, N., Schiele, B., & Schmidt, A. (2003). Multi-sensor activity context detection for wearable computing. In *European Symposium on Ambient Intelligence (EUSAI)* (Vol. LNCS 2875, pp. 220–232). Heidelberg: Springer-Verlag.

Larson, R., & Richards, M. (1994). *Divergent realities: The emotional lives of mothers, fathers, and adolescents.* New York: Basic Books.

Larson, R., & Delespaul, P. (1990). Analyzing experience sampling data: A guidebook for the perplexed. In M. deVries (Ed.), *The experience of psychopathology: Investigating mental disorders in their natural settings* (pp. 58–78). Cambridge, UK: Cambridge University Press.

Mascaro, S. A., & Asada, H. H. (2004). Measurement of finger posture and three-axis fingertip touch force using fingernail sensors. *IEEE Transactions on Robotics and Automation, 20*(1), 26–35.

Mehl, M. R., Pennebaker, J. W., Crow, M. D., Dabbs, J., & Price, J. H. (2001). The Electronically Activated Recorder (EAR): a device for sampling naturalistic daily activities and conversations. *Behavior Research Methods, Instruments, and Computers, 33*(4), 7.

Melanson, Jr., E. L. & Freedson, P. S. (1995). Validity of the Computer Science and Applications, Inc. (CSA) activity monitor. *Medicine and Science in Sports and Exercise, 27*(6), 934–940.

Moore, G. E. (1965). Cramming more components onto integrated circuits. *Electronics, 38*(8).

Mow, W. S., Lo, S. K., Targan, S. R., Dubinsky, M. C., Treyzon, L., Abreu-Martin, M. T., & Papadakis, K. A. (2004). Initial experience with wireless capsule enteroscopy in the diagnosis and management of inflammatory bowel disease. *Clinical Gastroenterology and Hepatology, 2*(1), 31–40.

Munguia Tapia, E., Intille, S. S., & Larson, K. (2004). Activity recognition in the home setting using simple and ubiquitous sensors. In A. Ferscha & F. Mattern (Eds.), *Proceedings of PERVASIVE 2004* (Vol. LNCS 3001, pp. 158–175). Berlin, Heidelberg: Springer-Verlag.

Neville, R., Greene, A., McLeod, J., Tracy, A., & Surie, J. (2002). Mobile phone text messaging can help young people manage asthma. *British Medical Journal, 325*(7364), 600.

Pentland, A. (1996). Smart clothes. *Scientific American, 274*(4), 73.

Philipose, M., Fishkin, K. P., Fox, D., Kautz, H., Patterson, D., & Perkowitz., M. (2003). Guide: Towards understanding daily life via auto-identification and statistical analysis. In *Ubihealth 2003: The 2nd International Workshop on Ubiquitous Computing for Pervasive Healthcare Applications.*

Picard, R. W. (2003). Affective computing: challenges. *International Journal of Human-Computer Studies, 59*(1–2), 55–64.

Pinhanez, C. (2001). The Everywhere Displays Projector: A device to create ubiquitous graphical interfaces. In G. D. Abowd, B. Brumitt, & S. A. N. Shafer (Eds.), *Proceedings of the Conference on Ubiquitous Computing* (pp. 315–331). Berlin, Heidelberg: Springer-Verlag.

Qi, Y. (2003). iDen Heart Rate Demo. Cambridge, MA: Personal communication.

Rafaeli, E. (2003). *The Personality, Motivation, and Cognition (PMC) Diary (Version 2. 0).* Retrieved October 1, 2004, from http://www.personality-project.org/pmc/diary.html

Rondoni, J. (2003). *A Context-Aware Application for Experience Sampling and Machine Learning.* Unpublished M.Eng. Thesis, Massachusetts Institute of Technology, Cambridge, MA.

Schwartz, J. E., & Stone, A. A. (1998). Strategies for analyzing ecological momentary assessment data. *Health Psychology, 17*(1), 6–16.

Starner, T. E. (2002). Wearable computers: no longer science fiction. *IEEE Pervasive Computing, 1*(1), 86–88.

Statistics Finland. (2003). *The Evolution of the Information Society.* Retrieved August 15, 2004, from http://www.stat.fi/index_en.html

stone_Chapter16.qxp 12/14/2006 8:19 PM Page 337

Stone, A. A., & Shiffman, S. (1994). Ecological momentary assessment (EMA) in behavioral medicine. *Annals of Behavioral Medicine, 16*(3), 199–202.

Stone, A. A., Shiffman, S., Schwartz, J. E., Broderick, J. E., & Hufford, M. R. (2002). Patient non-compliance with paper diaries. *British Medical Journal, 324*(7347), 1193–1194.

Trends in Telephone Service. (2003). Retrieved April 15, 2004, from http://www.fcc.gov/Bureaus/Common_Carrier/Reports/FCC-State_Link/IAD/trend803.pdf

U. S. Census Bureau. (1990). *Historical census of housing tables telephones.* Retrieved April 15, 2004, from http://www.census. gov/hhes/www/housing/census/historic/phone. html

Van Laerhoven, K., Schmidt, A., & Gellersen, H. W. (2002). Multi-sensor context aware clothing. In *Proceedings of the 6th IEEE International Symposium on Wearable Computers* (pp. 49–56). IEEE Press.

Wang, D. H., Kogashiwa, M., Ohta, S., & Kira, S. (2002). Validity and reliability of a dietary assessment method: the application of a digital camera with a mobile phone card attachment. *Journal of Nutritional Science and Vitaminology, 48*(6), 498–504.

Weiss, H., & Beal, D. (2003). *PMAT: Purdue Momentary Assessment Tool.* Retrieved October 1, 2004, from http://www.mfri.purdue.edu/pages/PMAT/

Welk, G. J. (2002). Use of accelerometry-based activity monitors to assess physical activity. In G. J. Welk (Ed.), *Physical activity assessments for health-related research* (pp. 125–142). Illinois: Human Kinetics Publishers, Inc.

Witten, I. H., & Frank, E. (1999). *Data mining: Practical machine learning tools and techniques with Java implementations.* San Francisco: Morgan Kaufmann.