Pervasive Computing: A Paradigm for the 21st Century

Debashis Saha
Indian Institute of Management-Calcutta

Amitava Mukherjee
IBM Global Services, Calcutta

Although pervasive computing still falls short of Mark Weiser's "walk in the woods," advances in hardware, software, and networking over the past 12 years have brought his vision close to technical and economic viability.

In 1991, Mark Weiser, then chief technology officer for Xerox's Palo Alto Research Center, described a vision for 21st century computing that countered the ubiquity of personal computers. "The most profound technologies are those that disappear," he wrote. "They weave themselves into the fabric of everyday life until they are indistinguishable from it."[1]

Computing has since mobilized itself beyond the desktop PC. Significant hardware developments—as well as advances in location sensors, wireless communications, and global networking—have advanced Weiser's vision toward technical and economic viability. Moreover, the Web has diffused some of the psychological barriers that he also thought would have to disappear.

However, the integration of information technology into our lives still falls short of Weiser's concluding vision:

> There is more information available at our fingertips during a walk in the woods than in any computer system, yet people find a walk among trees relaxing and computers frustrating. Machines that fit the human environment instead of forcing humans to enter theirs will make using a computer as refreshing as taking a walk in the woods.

We still view computers primarily as machines that run programs in a virtual environment that we enter to perform a task and leave when we finish. Pervasive computing presumes an altogether different vision. A device can be a portal into an application-data space, not a repository of custom software that a user must manage. An application is a means by which a user performs a task, not software written to exploit a device's capabilities. And a computing environment is an information-enhanced physical space, not a virtual environment that exists to store and run software.[2]

The need for perceptual information about the environment further differentiates pervasive computing from traditional computing. Sensing devices provide pervasive systems with information such as the locations of people
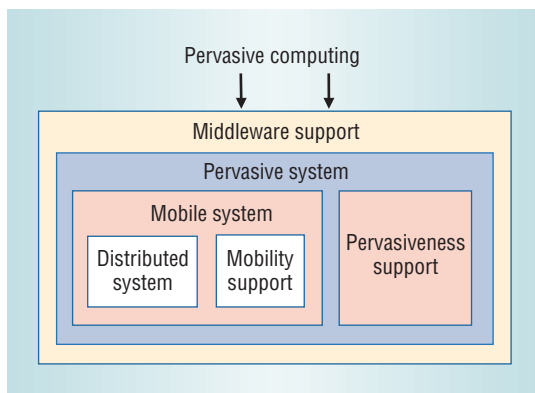
and devices. The system can use this information to interact more naturally with users, moving beyond the desktop legacy of isolated interaction.

EVOLUTION OF PERVASIVE COMPUTING

Pervasive computing defines a major evolutionary step in work that began in the mid 1970s, when the PC first brought computers closer to people. In Weiser's vision, however, the idea of making a computer personal is technologically misplaced. In fact, it keeps computing separate from our daily life.

Although the PC has not delivered the full potential of information technology to users, it certainly took a first step toward making computers (if not computing) popular (if not pervasive). It was also an instrumental factor in the phenomenal growth of hardware components and the development of graphical user interfaces.

Distributed computing

With the advent of networking, personal computing evolved to distributed computing. As computers became connected, they began to share capabilities over the network. Distributed computing marked the next step toward pervasive computing by introducing seamless access to remote information resources and communication with fault tolerance, high availability, and security.[3]

Although the World Wide Web was not designed to be a distributed computing infrastructure, its networking ubiquity has made it an attractive choice for experimenting with distributed computing concepts. It has also created a culture that is substantially more amenable to the deployment of pervasive computing environments than the culture that existed when Weiser first articulated his vision. The ad hoc nature of the Web's growth has proved that we can distribute computing capabilities in a big way without losing scalability. The simple mechanisms for linking resources have provided a means for integrating distributed information bases into a single structure.

Most importantly, the Web has pioneered the creation of a nearly ubiquitous information and communications infrastructure. Many users now routinely refer to their point of presence within the digital world—typically, their homepages, portals, or e-mail addresses. The computer they use to access these "places" has become largely irrelevant. Although the Web does not pervade the real world of physical entities, it is nevertheless a potential starting point for pervasive computing.

Mobile computing

Mobile computing emerged from the integration of cellular technology with the Web.[4] Both the size and price of mobile devices are falling everyday and could eventually support Weiser's vision of pervasive inch-scale computing devices readily available to users in any human environment. Cellular phone systems that separate the handset from the subscriber identity module (SIM) card approximate this model of operation. Subscribers can insert their SIM card and automatically use any handset, placing and receiving calls as if it were their own phone.

Users can already access the same point in the Web from several different devices—office or home PC, cell phone, personal digital assistant, and so forth. In this sense, for most users, what matters is the view a particular machine provides of the digital world. SIM cards also demonstrate that the end system is becoming less important than the access to the digital world. In this sense, we are well on the way to computers "disappearing," freeing users to focus beyond them.

The "anytime anywhere" goal of mobile computing is essentially a reactive approach to information access, but it prepares the way for pervasive computing's proactive "all the time everywhere" goal. As Figure 1 shows, pervasive computing is a superset of mobile computing. In addition to mobility, pervasive systems require support for interoperability, scalability, smartness, and invisibility to ensure that users have seamless access to computing whenever they need it.
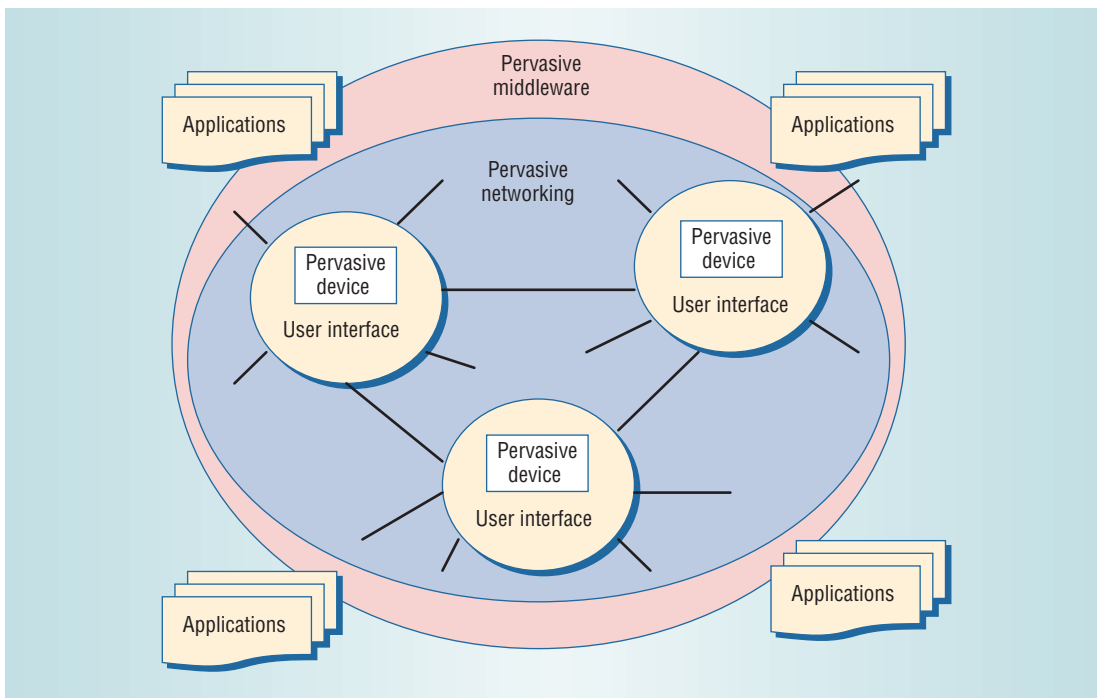
PERVASIVE COMPUTING MODEL

The technological advances necessary to build a pervasive computing environment fall into four broad areas: devices, networking, middleware, and applications. Figure 2 illustrates their relationships.

Devices

An intelligent environment is likely to contain many different device types:

- traditional input devices, such as mice or keyboards, and output devices, such as speakers or light-emitting diodes;
- wireless mobile devices, such as pagers, per-

sonal digital assistants, cell phones, palmtops, and so on; and

• smart devices, such as intelligent appliances, floor tiles with embedded sensors, and bio-sensors.

Ideally, pervasive computing should encompass every device worldwide that has built-in active and passive intelligence. The University of Karlsruhe's MediaCup project[5] is an experimental deployment of everyday objects activated in this sense. The project's guiding principle is to augment objects with a digital presence while preserving their original appearance, purpose, and use.

Sensors that automatically gather information, transfer it, and take actions based on it represent an important subset of pervasive devices. For example, sensors based on the Global Positioning System provide location data that a device can translate into an internal representation of latitude, longitude, and elevation. Stereo camera vision is another effective sensor for tracking location and identity in a pervasive environment. These fast-processing, two-lens digital cameras can record both background images and background shapes. The results are much more robust for tracking motion such as gestures.

Pervasive networking

The number of pervasive devices is expected to multiply rapidly over the next few years. IDC, a market analysis firm, has predicted that, by the end of 2003, the number of pervasive devices will exceed the estimated worldwide population of 6 billion people. Specifically, there will be more than 300 million PDAs; two billion consumer electronic devices, such as wireless phones, pagers, and set-top boxes; and five billion additional everyday devices, such as vending machines, refrigerators, and washing machines embedded with chips and connected to a pervasive network.

As a consequence of this proliferation, many current technologies must be revamped. In addition to extending the backbone infrastructure to meet the anticipated demand, global networks like the Internet also must modify existing applications to completely integrate these pervasive computing devices into existing social systems.[4]

Pervasive middleware

Like distributed computing and mobile computing, pervasive computing requires a middleware "shell" to interface between the networking kernel and the end-user applications running on pervasive devices. As Figure 2 shows, this pervasive middleware will mediate interactions with the networking kernel on the user's behalf and will keep users immersed in the pervasive computing space. The middleware will consist mostly of firmware and software bundles executing in either client-server or peer-to-peer mode.

User interfaces are another aspect of middleware. Standard Web browsers represent the high end of interface sophistication. They use more color, graphics, and controls than users typically expect on pervasive devices. Mobile computing has already introduced microbrowsers. For example, phone.com's UP.Browser is implemented on several commercially available digital phones.

Pervasive applications

Pervasive computing is more environment-centric than either Web-based or mobile computing. This

means that applications will guide the middleware and networking issues to a large extent. Consider a heart patient wearing an implanted monitor that communicates wirelessly with computers trained to detect and report abnormalities. The monitor should know when to raise the alarm, based on its knowledge about the environment. So this is much more than simple wireless communication.

EliteCare (www.elite-care.com) is an assisted living complex that applies similar technologies to improve the quality of life for elderly residents.[6]

ISSUES AND CHALLENGES

As a superset of mobile computing, pervasive computing subsumes mobile computing's research issues while opening up new ones unique to itself. In all cases, pervasive applications should disappear into the environment.

Scalability

Future pervasive computing environments will likely face a proliferation of users, applications, networked devices, and their interactions on a scale never experienced before. As environmental smartness grows, so will the number of devices connected to the environment and the intensity of human-machine interactions.

Traditional development requires recreating the application for each new device. Even if an enterprise could generate new applications as fast as it adds new devices, writing application logic only once—independent of devices—would have tremendous value in solving the applications scalability problem.

Furthermore, applications typically are distributed and installed separately for each device class and processor family. As the number of devices grows, explicitly distributing and installing applications for each class and family will become unmanageable, especially across a wide geographic area.

Heterogeneity

Conversion from one domain to another is integral to computing and communication. Assuming that uniform and compatible implementations of smart environments are not achievable, pervasive computing must find ways to mask this heterogeneity—or uneven conditioning,[3] as it has been called—from users. For instance, a sophisticated laboratory and a department store may always differ in their infrastructural smartness. Pervasive computing must fill this gap at some level, say middleware, to smooth out "smartness jitter" in the user's experience.

For networking, developers have faced protocol mismatch problems and learned how to tackle the large dynamic range of architectural incompatibilities to ensure transnetwork interoperability. Mobile computing has already achieved disconnected operation, thereby hiding the absence of wireless coverage from the user. Middleware may borrow similar concepts to dynamically compensate for less smart or dumb environments so that the change is transparent to users.

But the real difficulty lies at the application front. Today, applications are typically developed for specific device classes or system platforms, leading to separate versions of the same application for handhelds, desktops, and cluster-based servers. As heterogeneity increases, developing applications that run across all platforms will become exceedingly difficult.

Integration

Though pervasive computing components are already deployed in many environments, integrating them into a single platform is still a research problem. The problem is similar to what researchers in distributed computing face, but the scale is bigger. As the number of devices and applications increases, integration becomes more complex. For example, servers must handle thousands of concurrent client connections, and the influx of pervasive devices would quickly approach the host's capacities. We need a confederation of autonomous servers cooperating to provide user services.

Integrating pervasive computing components has severe reliability, quality of service, invisibility, and security implications for pervasive networking. The need for useful coordination between confederation components is obvious. This coordination might range from traditional areas such as message routing or arbitrating screen usage to new challenges such as deciding which application can use a room's light intensity to communicate with the user. For a wide area federation, message access is the primary requirement. Routing between servers introduces the possibility of messages from a single producer using multiple paths and, hence, arriving at a consumer out of order or duplicated.

Invisibility

A system that requires minimal human intervention offers a reasonable approximation of invisibility. Humans can intervene to tune smart environments when they fail to meet user expectations automatically. Such intervention might also be part of a continuous learning cycle for the envi-

ronment. To meet user expectations continuously, however, the environment and the objects in it must be able to tune themselves without distracting users at a conscious level.

A smart environment can implement tuning at different system levels. For example, network-level devices will require autoconfiguration. Current manual techniques for configuring a device with addresses, subnet masks, default gateways, and so on are too cumbersome and time-consuming for pervasive computing.

Automated techniques to dynamically reconfigure the network when required are also crucial to realizing the pervasive computing vision.

Perception: Context awareness

Most computing systems and devices today cannot sense their environments and therefore cannot make timely, context-sensitive decisions. Pervasive computing, however, requires systems and devices that perceive context. Mobile computing addresses location- and mobility-management issues but in a reactive context—responding to discrete events. Pervasive computing is more complex because it is proactive. Intelligent environments are a prerequisite to pervasive computing.

Perception, or context-awareness, is an intrinsic characteristic of intelligent environments. Implementing perception introduces significant complications: location monitoring, uncertainty modeling, real-time information processing, and merging data from multiple and possibly disagreeing sensors. The information that defines context awareness must be accurate; otherwise, it can confuse or intrude on the user experience.

ComMotion, a location-aware computing environment that addresses these issues for mobile users, is under development at the MIT Media Lab (www.media.mit.edu/~nmarmas/comMotion.html). Microsoft Research is investigating Radar,[7] an in-building location-aware system.

Smartness: Context management

Once a pervasive computing system can perceive the current context, it must have the means of using its perceptions effectively. Richer interactions with users will require a deeper understanding of the physical space.

Smartness involves accurate sensing (input) followed by intelligent control or action (output) between two worlds, namely, machine and human. For example, a pervasive computing system that automatically adjusts heating, cooling, and lighting levels in a room depending on an occupant's

electronic profile must have some form of perception to track the person and also some form of control to adjust the ventilation and lighting systems.[3]

PERVASIVE INITIATIVES

Both academia and industry have recently advanced pervasive computing projects. Although our selection is far from exhaustive, it suggests the current state of the art in pervasive computing.

Aura

Carnegie Mellon University characterizes its Aura project (www-2.cs.cmu.edu/~aura/) as "distraction-free ubiquitous computing." The project aims to design, implement, deploy, and evaluate a large-scale computing system demonstrating a "personal information aura" that spans wearable, handheld, desktop, and infrastructure computers.

Aura is a large umbrella project with many individual research thrusts. Darwin is an intelligent network at Aura's core. Coda is a distributed file management system that supports nomadic file access, and Odyssey provides operating system support for resource adaptation.

These products and others are evolving within the Aura project, which emphasizes pervasive middleware and application design.

Endeavour

The University of California at Berkeley's Endeavour project (endeavour.cs.berkeley.edu/) is an academic effort that focuses on the specification, design, and prototype implementation of a planet-scale, self-organizing, and adaptive "information utility." This smart environment is pervasive—everywhere and always there—with components that flow through the infrastructure, shape themselves to adapt to their usage, and cooperate on tasks.

Endeavour's key innovative technological capability is its pervasive support for fluid software. It includes processing, storage, and data management functionality to arbitrarily and automatically distribute itself among pervasive devices and along paths through scalable computing platforms that are integrated with the pervasive networking infrastructure. The system can compose itself from pre-existing hardware and software components to satisfy a service request while advertising the services it can provide to others.

Oxygen

The Oxygen project (www.oxygen.lcs.mit.edu/),

> **Pervasive computing requires systems and devices that perceive context.**

an MIT initiative, envisions a future in which computation will be freely available everywhere, like oxygen in the air we breathe. The project rests on an infrastructure of mobile and stationary devices connected by a self-configuring network. This infrastructure supplies abundant computation and communication, which are harnessed through system, perceptual, and software technologies to meet user needs.

The Oxygen project is focusing on eight environment-enabling technologies. Its emphasis is on understanding what turns an otherwise dormant environment into an empowered one to which users shift parts of their tasks.

Portolano

In its Portolano project (portolano.cs.washington.edu/), the University of Washington seeks to create a testbed for investigating pervasive computing. The project emphasizes invisible, intent-based computing, which infers users' intentions via their actions in the environment and their interactions with everyday objects.

Project devices are highly optimized to particular tasks so that they blend into the world and require little technical knowledge on the user's part. In short, Portolano proposes an infrastructure based on mobile agents that interact with applications and users. Data-centric routing automatically migrates data among applications on the user's behalf. Data thus becomes "smart," and serves as an interaction mechanism within the environment.

Sentient Computing

AT&T Laboratories, Cambridge, UK, is collaborating with the Cambridge University Engineering Department on the Sentient Computing project (www.uk.research.att.com/spirit/). The project explores user interfaces that employ sensors and resource status data to maintain a world model shared by users and applications.

The world model for the Sentient Computing project covers an entire building. Interfaces to programs extend seamlessly throughout the building. Computer desktops follow their owners and reflect real-time updates for object locations. This project has led to some new kinds of applications, like context-aware filing systems and smart posters.

Cooltown

Hewlett-Packard's pervasive computing initiative, Cooltown (www.cooltown.com), focuses on extending Web technology, wireless networks, and portable devices to create a virtual bridge between mobile users and physical entities and electronic services.

Cooltown uses URLs for addressing, physical beaconing and sensing of URLs for discovery, and localized Web servers for directories to create a location-aware system that supports nomadic users. It leverages Internet connectivity on top of this infrastructure to support communications services.

EasyLiving

The EasyLiving project of Microsoft Research's Vision Group (research.microsoft.com/easyliving/) is developing an architecture and related technologies for intelligent environments. The project supports research addressing middleware, geometric world modeling, perception, and service description. Key system features include computer vision, multiple sensor modalities, automatic and semi-automatic sensor calibration, and device-independent communication and data protocols.

WebSphere Everyplace

IBM's pervasive computing work focuses on applications and middleware that extend its WebSphere software platform (www-3.ibm.com/software/pervasive/). The company is spearheading consortia and initiatives for open standards to support pervasive computing applications. It is also working with hardware vendors such as Palm (www.palm.com), Symbol Technologies (www.symbol.com), and Handspring (www.handspring.com) to develop a new generation of devices.

Pervasive computing is about making our lives simpler through digital environments that are sensitive, adaptive, and responsive to human needs. Far more than mobile computing, this technology will fundamentally change the nature of computing, allowing most objects we encounter in daily life to be "aware," interacting with users in both the physical and virtual worlds.

While research challenges remain in all areas of pervasive computing, all the basic component technologies exist today. In hardware, we have mobile devices, sensors, and even smart appliances. Supporting software technologies include digital signal processing and object-oriented programming. Advances in networking support mobility management, ad hoc routing, and global reachability. Hence, in the near future, we can expect our homes to network intelligent devices that transparently support our information and communication needs. ■

References

1. M. Weiser, "The Computer for the 21st Century," *Scientific Am.*, Sept., 1991, pp. 94-104; reprinted in *IEEE Pervasive Computing*, Jan.-Mar. 2002, pp. 19-25.
2. G. Banavar et al., "Challenges: An Application Model for Pervasive Computing," *Proc. 6th Ann. ACM/IEEE Int'l Conf. Mobile Computing and Networking* (Mobicom 2000), ACM Press, 2000, pp. 266-274.
3. M. Satyanarayanan, "Pervasive Computing: Vision and Challenges," *IEEE Personal Communication*, Aug. 2001, pp. 10-17.
4. D. Saha, A. Mukherjee, and S. Bandopadhyay, *Networking Infrastructure for Pervasive Computing: Enabling Technologies and Systems*, Kluwer Academic, 2002.
5. M. Beigl, H.W. Gellersen, and A. Schmidt, "Mediacups: Experience with Design and Use of Computer-Augmented Everyday Objects," *Computer Networks*, Mar. 2001, pp. 401-409.
6. V. Stanford, "Using Pervasive Computing to Deliver Elder Care," *IEEE Pervasive Computing*, Jan.-Mar. 2002, pp. 10-13.
7. P. Bahl and V.N. Padmanabhan, "RADAR: An In-Building RF-Based User Location and Tracking System," *Proc. IEEE Infocom 2000*, vol. 2, IEEE Press, 2000, pp. 775-784.

Debashis Saha is an associate professor in the Management Information Systems and Computer Science Group of the Indian Institute of Management-Calcutta. His research interests include pervasive communication and computing, wireless networking and mobile computing, and wavelength-division-multiplexing optical networking. He received a PhD from the Indian Institute of Technology at Kharagpur, India. He is a senior life member of the Computer Society of India, a senior member of the IEEE, a member of both the IEEE Computer and IEEE Communications Societies, and a member of IFIP WG 6.8. Contact him at ds@iimcal.ac.in.

Amitava Mukherjee is a principal consultant at IBM Global Services, Calcutta. His research interests are in mobile computing and communication, pervasive computing and mobile commerce, optical networks, combinatorial optimization, and distributed systems. Mukherjee received a PhD in computer science and engineering from Jadavpur University, Calcutta. He is a member of the IEEE and the IEEE Communications Society. Contact him at amitava.mukherjee@in.ibm.com.