INVITED COMMENTARY

KEY CHALLENGES IN COMMUNICATION FOR UBIQUITOUS COMPUTING

GAETANO BORRIELLO
DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
UNIVERSITY OF WASHINGTON
AND INTEL RESEARCH SEATTLE

What was novel about the outlook espoused by Weiser is that it firmly placed computing in the background and not as an end in itself. The focus is shifted from the technology to the users. The key message is that users' attention is the valuable resource and not the computing devices.

Ten years ago Mark Weiser heralded a future of ubiquitous and invisible computing in a now often referenced article in *Scientific American* [1] in which he stated:

"The most profound technologies are those that disappear. They weave themselves into the fabric of everyday life until they are indistinguishable from it."

What was novel about the outlook espoused by Weiser is that it firmly placed computing in the background and not as an end in itself. The focus is shifted from the technology to the users. The key message is that users' attention is the valuable resource and not the computing devices. Cheap and abundant computing can be used to connect the virtual and physical worlds so that more can be done for the user with less direct manipulation that requires the user's undivided attention [2, 3].

A simple example can serve to make the message clearer. In traditional computing, a user wanting to know how long it will take them to drive to their next appointment would check a real-time traffic Web site and enter their current location and destination. Although this is a powerful capability, it nonetheless requires the user to remember to check in a timely manner to adjust their route or leave earlier than forecast and to enter mundane data into the user interface. In ubiquitous computing, an autonomous agent would constantly be checking the user's current location (available from GPS or in-building wireless infrastructure) and their calendar to determine if an adjustment in schedule should be made. If traffic conditions warrant it, the user can be signaled and their PDA can show a brief reminder with the minutes left before they should head off.

Another example uses a biologist in a laboratory that would like to record all the steps in their experimental protocol so that others can faithfully reproduce them. Traditional computing would provide the biologist with a PDA or laptop in which to enter each step but requiring the biologist to interrupt the flow of their work to do so. This kind of mental switching can lead to notes being missed or recorded incorrectly. Ubiquitous computing would provide the user with an up to date view of the experiment at one or more displays available throughout their lab. Each of their actions would be sensed and automatically inserted into the experimental flow along with important parameters and settings of instruments.

We are finally at the point that we can truly realize ubiquitous computing. Two important trends in embedded computing have made this possible: Moore's Law and increasingly lower-power wireless communication. Embedded processing is cheap enough that it can be a part of virtually every manufactured product. The processors are more sophisticated than ever, take up a very small volume, consume ever-decreasing energy, and are integrated with memory and communication units. This development, the ability to add computing power to most anything, is in itself phenomenal and has led to sales of billions of embedded processors per year. But what makes it really interesting is that these devices are now communicating [4]. Advances in radio communication, along with its miniaturization, make it also possible to wirelessly link all these devices. It is only through the collaboration made possible by communication between devices that the interesting applications envisioned by Weiser can be fully realized.

Thus, communication and networking are at the heart of ubiquitous computing technology. However, if we look at today's computing systems and how they communicate, they are far from achieving the fluidity necessary for the examples above. We are required to become network and system administrators to get even the simplest wireless networks up and running. We have to consider addressing and naming as well as security and authentication, among many other aspects. Furthermore, we must often personally interact with almost every device involved in the communication in order to adjust settings and enter passwords. As we increase the numbers of devices and the different communication methods they use, we will quickly collapse under the burden of fielding and maintaining what are potentially thousands of devices per person: embedded in our homes and cars; present in our streets, buildings, and infrastructure; and carried on our persons. Current methods and approaches just don't scale.

There are many challenges that must be faced and surmounted before we can fully realize computing applications that truly weave themselves into the fabric of our lives [3]. In this article, we will focus on four key ones: heterogeneous networks; geographic vs. network topology; short-lived and intermittent connectivity; and evolution of long-lived systems.

Heterogeneous Networks

Thousands of task-specific embedded devices will not all communicate using the same method. There are several reasons for this. First, they will have different bandwidth requirements with corresponding power consumption. At one extreme we will have embedded sensors that can harvest enough power from their environment to communicate a few bytes (e.g., a switch that can send a radio packet with the energy from being flipped on) while at the other we will have two-way video cell phones [5]. Second, they will be designed for a wide range of form-factors and will be constrained by the volume available. Third, they will need to coexist with many other devices, and it will not be possible to have them all use the same portion of the spectrum. We are already seeing work on dealing with the crowding of the spectrum in the Bluetooth/WiFi interference debate [6].

This heterogeneity will have important implications. Of course, we will need transport protocols that transfer packets from one band to another. But more importantly we will need different approaches to how connections are established and maintained. After all, devices will try to shut down as often as possible to save power, and data will have to take hops onto many different frequencies to get to its destination [6]. Will we be able to continue to base our protocols on the end-to-end principle as in TCP/IP? More likely, we will not. The bounds on round-trip times will have increasingly wider variance and it will be difficult to efficiently communicate using current acknowledgement and timeout approaches. Some ideas in this direction look at making the data more autonomous and using a series of asynchronous acknowledgements to connect the segments of the network over which the data will travel [7].

Geographic vs. Network Topology

As our communication networks have grown we have already seen how different the path between two devices can be in the virtual and physical worlds. A wireless PDA owned by a visitor to a company may be very physically close to the PCs in the employees' offices but it can be simultaneously very far in the virtual world of communication networks. The PDA has to first communicate with a wireless base station on a nearby building, go through that carrier's switching network to an Internet gateway that connects it back to the company through a series of routers. Once at the company's door, the packets will most likely have to be admitted through a series of firewalls (which may require authentication and more round-trip communication). Finally, the packet will be routed to that PC that was only a foot away from the originating PDA through a series of internal routers that interconnect all of the company's offices. This is quite a waste of network resources, not to mention being a power drain on the PDA. In addition, we had to know the address of the nearby PC in order to properly address the data.

Ubiquitous computing requires a focus on geographic proximity rather than network topology. Nearby devices should communicate as directly as possible while maintaining security. The classic example of this is the ability to print to the nearest printer. As we increase the number of devices in the environment, there will be more and more types of resources where we will want to interact with specific, geographically close instances [8]. Infrared technology was a step in this direction but never flourished. Will it make a comeback as a way to communicate short-range, within the confines of a room, and with intention (due to having to align the two devices). Bluetooth is the newest short-range RF technology, followed close behind by 802.11 wireless Ethernet. But Bluetooth was designed as a cable replacement technology. How will multiple devices determine with which others they should communicate? How can we be sure our data is secure when other devices owned by others are in close proximity [9]? Too much security will limit what we can do and how seamlessly we can do it. Too little and our privacy may be compromised.

Ubiquitous computing requires a focus on geographic proximity rather than network topology. Nearby devices should communicate as directly as possible while maintaining security. The classic example of this is the ability to print to the nearest printer.

Short-Lived and Intermittent Connectivity

No one will want to change batteries on thousands of devices. They are going to try to limit their power consumption as much as possible, and one of the biggest items in the power budget is communication. Thus, for the foreseeable future it is likely we will be facing the prospect of devices trying to keep their communications short-range, short in duration, and infrequent. However, the more we interconnect devices, the more applications and their users will come to rely on that capability. Tradeoffs between power consumption and connectivity will be at the core of many devices and applications.

Devices will have to be designed to be smarter about intermittent connectivity. They will need to push data to powered data stores in the infrastructure where it can be gathered by collaborating applications. They will need to cache new data until connections can be reestablished and provide reconciled views once they are reconnected. It may even be necessary to move pieces of computation between devices to reach the best balance of communication and power consumption.

Evolution of Long-Lived Systems

Finally, as we place more and more devices all around us, into our clothing, furniture, walls, and roadways, it is highly unlikely we will be in a position to upgrade them all simultaneously. Systems, and the applications they support, will need to evolve how they communicate over time. Devices may begin in relatively quiet surroundings but soon enough find themselves surrounded by others in the same spectrum. Today it is easy to deploy devices because the spectrum is relatively empty and has not yet been partitioned into fine-grain cells.

Evolution means increasing heterogeneity as legacy communication methods continue to be supported as new ones are introduced. But the radio spectrum is limited. We need to plan ahead

As embedded computing penetrates all manufactured goods, we will have to design our communications to cope with rapidly increasing complexity and the emergent behavior and new vulnerabilities it will bring.

for dynamic reallocation of bandwidth [10]. The future will bring many more devices communicating at shorter range but also being dynamically rearranged and locally crowded. We will need adaptive techniques for bandwidth allocation because of numbers and arrangement. This can entail master coordination but it is likely we will require negotiations between devices and applications for scarce bandwidth. How will we protect our systems from denial-of-service attacks that can pop-up anywhere through local jamming?

Summary

Realizing truly ubiquitous computing presents many challenges for communication and networking. This article has provided a brief introduction to four of them. There are many others on the horizon. While our information systems will become more useful and attuned to our needs, they will also become correspondingly more complex through their interconnectedness. As embedded computing penetrates all manufactured goods, we will have to design our communications to cope with rapidly increasing complexity and the emergent behavior and new vulnerabilities it will bring.

References

[1] M. Weiser, The Computer for the 21st Century, *Scientific American*, Sept. 1991, vol. 265, no. 3, pp. 94–104.

[2] D. Tennenhouse, "Proactive Computing," *Communications of the ACM*, May 2000, vol. 43, no. 5, pp. 59–66.

[3] G. Borriello, "The Challenges to Invisible Computing," *IEEE Computer*, Integrated Engineering column, Nov. 2000, vol. 33, no. 11, pp. 123–25.

[4] D. Estrin *et. al.*, "Embedded, Everywhere: A Research Agenda for Networked Systems of Embedded Computers," Committee on Networked Systems of Embedded Computers, Computer Science and Telecommunications Board, National Research Council, Washington, DC, 2001.

[5] J. Paradiso and M. Feldmeier, A Compact, Wireless, Self-Powered Pushbutton Controller, *Proceedings of Ubicomp2001: Ubiquitous Computing*, Sept. 2001, pp. 299–304.

[6] Hewlett-Packard White Paper, Wi-Fi(tm) and Bluetooth(tm): Interference Issues, Jan. 2002, http://www.hp.com/notebooks/us/eng/docs/WiFiBlue.pdf

[7] M. Esler *et al.*, "Next Century Challenges: Data-Centric Networking for Invisible Computing," *ACM Mobile Computing and Networking (Mobicom)*, Aug. 1999.

[8] M. Addlesee *et al.*, "Implementing a Sentient Computing System," *IEEE Computer*, Aug. 2001, vol. 34, no. 8, pp. 50–56.

[9] A. Perrig *et al.*, "SPINS: Security Protocols for Sensor Networks," *ACM Mobile Computing and Networking (Mobicom)*, July 2001.

[10] V. Bose, D. Wetherall, and J. Guttag, "Next Century Challenges: RadioActive Networks," *ACM Mobile Computing and Networking (Mobicom)*, Aug. 1999.