Pervasive Computing: Architectural Challenges

Victor Manuel S. Coelho

Dep. Informática, Universidade do Minho
4710-057 Braga, Portugal
vcoelho@portugalmail.pt

Abstract. Computing anytime everywhere - the third wave of computing - is still far from achievement. This communication addresses the on-going research on what is considered to be the two step stones of this third generation: wearable computing and ubiquitous computing, and the efforts to achieve interoperability among these devices. A brief introduction is made to the human-computer interface (HCI), as a method to reduce the problem of user input in wearables, augmented reality (AR), which provides the most common everyday object with sensing, processing and communication capabilities, and context awareness, that allow to choose a certain behaviour at a certain time. Then, based on this analysis, we will explain why these two models will come to be the basics of 'computing everywhere'.

1 Introduction

In the past fifty years of computation there have been two great trends in this relationship: the mainframe relationship, and the PC relationship. Today the internet is carrying us through an era of widespread distributed computing towards the relationship of computing everywhere, characterized by deeply imbedding computation in the world.

The first era we call "mainframe", to recall the relationship people had with computers that were mostly run by experts behind closed doors. Anytime a computer is a scarce resource, and must be negotiated and shared with others, our relationship is that of the mainframe era.

The second great trend is that of the personal computer. The personal computing relationship is personal, even intimate. You have your computer, it contains your stuff, and you interact directly and deeply with it. When doing personal computing you are occupied, you are not doing something else.

The third wave of computing is that of computing everywhere, whose crossover point with personal computing will be around 2005-2020. It will have lots of computers sharing each of us. Some of these computers will be the hundreds we may access in the course of a few minutes of internet browsing. Others will be imbedded in walls, chairs, clothing, light switches, cars - in everything. It is fundamentally characterized by the connection of things in the world with computation, and will take place at a many scales, including the microscopic.

Many everyday items will become ‘smarter’ through integrated processors, memory, sensors, and communication capabilities, and will be able to serve us by tackling additional tasks beyond their normal function. In the long run, we are talking about everyday items like pencils that are capable of digitising everything written with them, clothes that remember the places they have been to, or umbrellas that subscribe to an internet weather information service and will trigger a friendly reminder from the door if your shoes report that you attempt to leave without an umbrella.

All of this leads us to the current topic: information technology becoming omnipresent and entering all aspects of our life, which is exactly what ‘pervasive computing’ or computing everywhere means. This era manifests itself primarily by the migration of classical PC applications into smaller, more specialized ‘information appliances’. These information appliances are being integrated into a common infrastructure that takes care of the synchronization of the distributed and replicated data, guarantees the necessary security, creates interoperability, and offers many other useful services. In the end, it is all about offering adequate access to all kinds of information in any place possible.

We are at the very beginning of this new era of technology, and currently only media artists and scientists are thinking of intelligent umbrellas and similar scenarios that promise to change the physical world into one big interactive computing platform. More concrete are the current efforts to integrate mobile and wearable ‘information appliances’, like mobile phones or other ‘digital assistants’ into Web-based business processes and electronic commerce scenarios, and to enable ubiquitous access to the applications behind them through these devices. However, scalability, flexibility, mobility, security, and heterogeneity are major challenges in the realization of adequate systems and require an appropriate integration into powerful back-end systems as well as a whole suite of modern concepts and standards.

2 Wearable Computing

The integration of mobile and wearable ‘information appliances’ is commonly named wearable computing [1]. The conceptual definition of a wearable computer is that it is always present, always working collection of personal applications. These applications are characterized by being: mobile: where you go, it goes; persistent: your wearable is a constantly functioning resource, and should function across a wide range of physical and social contexts; secondary task: your wearable minimizes the demands on your time, while facilitating the real-world tasks and interaction. It should help you, and not distract you; proactive: wearable applications are agents and can act semi-autonomously, they are capable of taking action to get the user’s attention; context-aware: instead of being deaf and blind like conventional portables, wearables know something about the context of the user and can change its behaviour or take action on the user’s behalf based on context.

The use of wearables will augment and extend the capabilities of the wearer while preserving personal privacy and functioning over a wide range of situations and contexts. It extends the reach and capabilities of a person, providing enhanced communication, memory, sensing and logistical skills. It can for example filter your calls, monitor your health, remind you of a name, or help you catch a bus.

2.1 Context-Awareness

Perhaps the most important feature of a wearable, and what differs it from a normal PDA, is context-awareness [2]. The environment and the context surrounding the user can give cues to enhance usability. The devices can sense what the user does need next, what he or she is probably going to do, and what actions can be taken to ease the task of the user. All the sensors carried by the user are responsible for sensing what is going on in our world. The statement “context is more important than position” becomes central for future applications. It is important to clarify that by context we implicitly mean the people around the user, the situation (in a meeting, making a phone call, having a coffee, walking), the environment (location, temperature, time, light), how the user feels (pulse, body temperature). All this additional information about the context can make the computing device act and react more sophisticated and less obtrusive.

2.2 Implicit Human-Computer Interface

At this point, we are ready to introduce a new concept: implicit human-computer interface (HCI) [3], which consist of using user activity in the real world as input to computers. It is an action performed by the user that is not primarily aimed to interact with a computerized system but which such a system understands as input. Having in mind current and upcoming technologies, such as increased processing power even on mobile devices, availability of sensors ranging from temperature to cameras, and the resulting perceptional capabilities, devices will start the shift from explicit HCI towards a more implicit interaction based on situational context. As an example we can take a look at the garbage bin, that scans in the bar code of products and reproduces the information for a shopping list. The action performed by the user is the same as with any other garbage bin. The recognition of the system and the built-in interpretation of the system make use of the action performed by the user. The user does not explicitly interact with the computer, thus the process describes an implicit interaction.

3 Ubiquitous computing

Ubiquitous computing [4] is a paradigm in which networked computing resources are present anywhere. Users augment their computing and communicating capabilities with lots of nearby computing devices, and the network achieves an infinitesimal capillarity reaching everywhere. Resources are mobile and have wireless network connectivity. In such scenario services make use of information, processing capabilities and storage and output resources obtained anywhere. We can characterize ubiquitous computing by two main attributes: ubiquity, where interactions are channelled through multiple interfaces rather than a single workstation and transparency, where the technology becomes so embedded in everyday life that it essentially becomes invisible to the users.

Whether it's with refrigerators, in cars, around the office or on the high seas, powerful new systems that you can access through words and maybe even gestures are promising to friendly-up the world. The ultimate aim is to seamlessly blend the analog human world with all things digital. That way, by accessing it through an infrastructure as widespread as electric power is today, you will tap into this world on your terms and in your language, not a machine's.

Less than a decade ago, such dreams were confined to far-out future factories such as SRI, Xerox Corporation's Palo Alto Research Center (PARC) and MIT's Media Lab. But recent advances in computing power, storage, speech recognition and especially wired and wireless networking, coupled with the rise of the World Wide Web, are bringing the dream within grasp of the real world. Researchers agree more uniformly than ever on where technology is headed, allowing what was previously a mass of visions and predictions about the future to now be classified into three broad technological frameworks: 24/7/360; who, what, when, where; and the digital companion. These categories translate the importance of pervasiveness, awareness and personalization.

3.1 “24/7/360“

Today, it is a fundamental principal of ubiquitous computing that computational power and services will be available whenever they're needed - that is the 24/7 part. And not just throughout a building, but everywhere - that's the 360, in degrees around the globe. Under the 24/7/360, however, lie two radically different approaches. One continues the drive to push computational power into objects with ever-smaller footprints, via souped-up laptops, handhelds and wearables. The other holds that tomorrow's computing resources will not be carried on specific devices. Instead, they will live on networks. The network becomes the computer.

3.2 Who, What, When, Where

Before even a few folks have the benefit of truly ubiquitous computing, great steps must be made toward creating technology that serves people rather than the other way around. That means objects and services must sense and respond to what is going on around them, so that they can automatically do the right thing, like holding a routine call if you're busy, letting you know if your flight is delayed, or inform you of a traffic jam and suggest a better route. Such feats are known as context-aware computing, as we have seen before in this communication. However, to do this job to the utmost, networks must know something about the people using them, often including their identity and location. This will force a choice: do people want to periodically cede privacy in exchange for better service? Tracking systems for people have been implemented successfully with devices ranging from infrared-emitting active badges, ultrasound transmitters known as BATS and radio-frequency identification tags. Yet, although a world populated by electronic tags promises to extend computing to almost anything, it does not address one of the biggest hopes for ubiquitous computing: that sensors, effectors and actuators can also be incorporated into devices, making systems capable of both processing information and responding to it.

The drive towards ubiquitous computing gives rise to smart artefacts, which are objects of our everyday lives augmented with information technology to obtain general context information, available to any application within a given environment. These artefacts will retain their original use and appearance while computing is expected to provide added value in the background. In particular, added value is expected to arise from meaningful interconnection of smart artefacts. This computer augmentation of artefacts is called Augmented Reality (AR) [5]. One of the first experiments in computer-augmented objects is the MediaCup [6], which consists of a common everyday cup with a small computer attached to it. It as sensors, processing and communication capabilities that enable it to collect and broadcast context information. One can see it as a building block for ubiquitous computing.

3.3 The Digital Companion

Enter a third major aspect of ubiquitous computing: software agents, also called bots, that root around behind the scenes to find services and generally get things done without bothering humans with the details. Among the first bots to hit the market could be context-aware applications that seek to prevent information overload by filtering e-mail, phone calls and news alerts. For example, at Microsoft, software agents under development make these decisions based on such factors as message content, the kinds of messages users read first or delete without opening, and the message writer's relationship with the reader or position in a company organization chart. Agents can then determine whether to interrupt or not by correlating that information, with the help of desktop sensors such as microphones and cameras, with whether the person is on the phone, busy at the keyboard or meeting with someone. If the person is out, the agents can even decide whether to track him or her down via pager or cell phone.

What puts the technology into the futuristic agent class, however, is that it employs procedures based on statistical reasoning techniques in order to draw inferences from users behaviour and make its judgments. The same techniques enable the system to learn from past experiences to get better at its job. Under this Open Agent Architecture (OAA) framework, humans take no direct hand in controlling the fleet of servants sent scurrying to do their bidding. They merely express their desires to the OAA through a microphone or keyboard, by drawing on a display screen or even speaking over the telephone, and things get done. Agents will adapt to human needs and will stay with people for years or even decades. Just as a good secretary learns a boss's preferences and even comes to anticipate his or her needs, so will a digital companion serve its human masters. It will be a person's "universal remote for the world."

3.4 Wireless Networking Between Devices

Ubiquitous computing can be distinguished from computing in general by its emphasis on ubiquity, interconnectedness and dynamism. It is intended to be ubiquitous; the goal is to create low-cost, embedded, distributed and non-intrusive computing technology. Networking via both traditional wired and newer wireless technologies plays a central role. Its dynamic nature is a result of mobile and adaptive applications that are able to automatically discover and use remote services. Sharing information requires a suitable communication technology, which preferably should be wireless in order to be in line with the unobtrusive nature of the devices. It must be robust, scale well, and must efficiently use the limited energy of the autonomous device. Finally, the communication technology employed should adhere to a broadly-used standard to leverage from existing communication services in the environment.

Bluetooth [7] is an emerging communication standard that provides ad hoc configuration of master/slave piconets including eight active units at most. It supports spontaneous connections between devices without requiring a priori knowledge of each other. Bluetooth allows data transfers between units over distances of nominally up to 10 meters. The gross data rates of 1Mbps is shared among participants of a piconet. Bluetooth operates in the licence-free 2.4GHs spectrum (2.400-2.484 GHz) and uses frequency hopping spread spectrum (FHSS) to minimize interference problems. The technology is geared toward low energy consumption, and targets the consumer mass market with world-wide availability and low price.

The firsts experiences with Bluetooth were made in the Smart-Its project [8], and show that the future for this standard is very optimistic. It greatly simplifies ad hoc networking using the piconet communication paradigm. By using the freely available ISM band, devices can be used world-wide without alterations. Its frequency hopping technology makes transmissions robust against narrow-band interfaces. With data rates of up to 1 Mbps, Bluetooth also offers more than enough bandwidth for ubiquitous computing applications such as simple sensor networks. Even though Bluetooth modules are currently rather expensive, prices are expected to drop to about USD 5 per unit once mass-production is running full-scale. We expect that when this happens, there will be a boost in ubiquitous devices manufacturing. Ubiquitous computing implemented with Bluetooth is right at the corner.

3.5 Interoperability Among Ubiquitous Computing Devices

Most electronic devices can't talk to each other. But they should. Unhappily, today's technology barons can't agree on a standard. Sun's Jini [9] and Microsoft's Universal Plug and Play (UPnP) [10] are battling for dominance in the new connected world. Jini and UPnP are the most prominent examples of ubiquitous computing. Some of those modules will live on computer networks. Others will live inside small devices. And all of them will be able to talk to each other and work together.

Jini. Today, if an application wants to print, it must contain printing code. Tomorrow, it will just need to know how to talk to a printer. The self-containing printing module will do the rest. That's the promise, anyway. And it is Sun Microsystems who is doing most of the promising, under the umbrella of its Jini initiative.

The idea behind Jini is to store a tiny Java program in the smarts of any device. When the Jini device is plugged into a computer network, the Jini program will "announce" itself to the other devices attached to the network and let them know what it can do. Each Jini device also agrees to follow "rules" about how it will interact with other devices. But, and here's a critical difference, there's no central figure of command. Devices swap information with one another. The printer can communicate with the digital camera, which can also send its pictures to a display. In addition, Jini devices don't need a special network or miles of cable to work together. Plug two Jini devices into the electrical outlets of your living room and, in principle, they can send data to one another via the power lines. Other devices may beam data to one another through wireless connections such as those used by cellular phones or TV sets' remote controls. The idea of Jini technology very simply is to use one of the most powerful of the ideas from distributed systems: the idea of agents. Each device or service is an object represented by an agent which moves in the network. These agents are not defined by protocols, they're defined as programming language types. So I don't have to get you to implement a protocol, I can simply pass you an executable agent over the network.

UPnP. Microsoft, in turn, is developing its own set of ideas for information appliances. UPnP is based in the same protocols we use everyday. Their concept is if you want something to print, there's already an internet printing protocol. This technology will enable multiple PCs to share a single IP address, automatically configure the network and allow for on-demand connections, without having to crack open the case. To do so it will use the IEEE 1394 specification, Universal Serial Bus, or a phone or power line. This capability will enable devices on the network to be instantly recognized and share information. Detailed information on each device down to tech support links can be read directly off each device.

Yet, both are under development or about to enter the early-adopter phase. Jini still has not won strong support from some big names, notably IBM, Oracle and HP. Microsoft has to overcome misgivings about its previous Plug and Play effort, which was a mediocre effort at best. Can Microsoft make the necessary philosophical shift? With Jini ubiquitous computing, centralized computers and networks lose control. Small computers, devices and services work as an alliance rather than a hierarchy. Will UPnP fit into that scheme?

We may have two competing systems for a while, meaning this great concept of ubiquitous computing will suffer from slow acceptance until there is a single standard.

4 Problems in Wearables and Ubiquitous Computing

When researchers came up with the phrase ubiquitous computing they envisioned computers embedded in walls, in tabletops, and in everyday objects. In ubiquitous computing, a person might interact with hundreds of computers at a time, each invisibly embedded in the environment and wirelessly communicating with each other.

In the purest form of ubiquitous computing, all computation is contained in the environment rather than on the person. This extreme has several problems.

The first and probably the most important problem is that ubiquitous computing environments pose serious privacy risks. By watching everything a user does these systems have the potential to leak all our actions, preferences, and locations to others unknown to us, now or in the future.

The second problem is that it is often difficult to maintain personalization of ubiquitous computing system. In the worst case, every time a new person joins a work-group or community her personal profile needs to be added to every device. Even if all the devices and environments on a campus share a personal profile database, profiles need to be updated every time she moves to a new site.

The wearable perspective suggests that instead of putting sensors and cameras in the room, put them on the person. In the purest form, the wearable user would do all detection and sensing on her body, requiring no environmental infrastructure at all.

Wearables offer a solution to most of the problems mentioned above. Because the wearable always travels with the wearer, personal profiles never need to be transferred to a new environment. And because a wearable might stay with a user for many years her profile can automatically evolve over time. Furthermore, wearable computers are an inherently more private location for sensitive data because it never needs to leave the person. If the data is never transmitted it becomes much harder to intercept, and the distribution of personal profiles across several wearables (possibly owned by many entities, each with a vested interest in keeping his own data private) makes them a less convenient target for compromise.

Of course, one might still infer a person's location by the fact that a room's resources are being controlled by a particular network address. Traffic analysis is a serious threat to privacy, and the TCP/IP protocols have no features for anonymity. Finally, sometimes privacy leaks are inherent in the application itself. For example, an application that shows where a person is on a map has no choice but to reveal that information; that's its job. Our goal is not to maintain total privacy, but rather to design a system whereby personal data is distributed on a need-to-know basis.

Many wearable systems have been demonstrated that act very similar to smart rooms and ubiquitous computing. For example, wearables have been used to create proximate selection interfaces for room control as well as personalized room controllers for the disabled. Wearable systems have also been used to help create automatic diaries of a user's state over time, both for health monitoring applications and video diaries. Finally, many applications exist that present context-based information such as tour-guides and general notes related to a user's context. In these systems location is sensed on the wearable either by GPS (for outdoors) or indoors by location beacons. The location beacons are essentially the same as the active badges used in ubiquitous computing, except that instead of being worn by a person to broadcast his identity they are placed in rooms to broadcast locations. Similar systems have used paper labels that are recognized through machine vision to recognize a location or object, while still others recognize objects or locations without any tagging or infrastructure at all.

Wearable systems are well suited to providing privacy and personalization, but they tend to lack in other areas.

Wearable computer systems have trouble maintaining localized information. For example, if information about a single location gets updated then every wearable needs to be given the new information. Furthermore, is it often difficult for a wearable system to sense information beyond the user's local area.

If a wearable is used to control a resource off the persons body, such as a stereo, big screen display, or air conditioner, it is often much easier to design the system with the resource-specific drivers in the device itself. When low-level control is left to the wearable it tends to produce higher demands on the wearable's CPU and wireless network and necessitates that the wearable have code to control each kind of device that might be discovered.

Wearables are also not well suited to managing resources among several people. When more than one wearable user wants to use the same stereo, for example, often it is desirable to have a more intelligent system than simply allowing the last request to take precedence.

5 Conclusion

After exposing the problems in these two architectures we can foresee that the future will most likely be a fusion of both. One complements the other. An upgraded wearable working together with computer-augmented devices will be the way to go. Wearables present no problem in the privacy and personalization matters. The problem of localized less important information will persist, although each time the wearable is in context of that information, it gets updated. As for crucial information updates, you'll just have to teach your wearable to update itself. He will learn! Interoperability obtained with Jini will solve the problem of localized control. Your wearable will be able to talk to every device through agents passed over the network, instead of drivers. Resource management can be resolve by an agent that lives in each resource. Priorities can be established according to usage, hierarchy, age, preferences, and even statistical reasoning techniques.
All the technology necessary to implement this new model is being tested almost everywhere, so you should expect world-wide implementation within the next decade.

References

[1] Wearable computing: What is a Wearable, http://agents.www.media.mit.edu/projects/wearables, (1994)

[2] Hans Gellersen, Albrecht Schmidt and Michael Beigl: Multi-Sensor Context-Awareness in Mobile Devices and Smart Artefacts, http://www.smart-its.org/publication/sensors-in-mobile-devices.monet.pdf, (2001)

[3] Albrecht Schmidt: Implicit Human Computer Interaction Through Context, Personal Technologies 4 (June 2000) 191-199

[4] Mark Weiser: The Computer for the Twenty-First Century, Scientific American (September 1991) 94-10

[5] Albrecht Schmidt and Michael Beigl: New Challenges of Ubiquitous Computing and Augmented Reality, 5th CaberNet Radicals Workshop (5-8 July 1998) Valadares, NR. Porto, Portugal

[6] Michael Beigl, Hans Gellersen and Albrecht Schmidt: MediaCups: Experience with Design and Use of Computer-Augmented Everyday Artefacts, Computer Networks, Special Issue on Pervasive Computing, Vol. 35, No. 4 (March 2001) Elsevier 401-9

[7] Bluetooth Special Interest Group: What is Bluetooth, http://bluetooth.com/util/faq1.asp#faq1, (2001)

[8] Oliver Kasten, Marc Langheinrich: First Experiences with Bluetooth in the Smart-Its Distributed Sensor Network, http://wwwtec.informatik.uni-rostock.de/RA/pact2001/ucc-paper4.pdf, (2001)

[9] Jini: How Jini Technology Works, http://www.sun.com/jini/overview, (2002)

[10] UPnP: About Universal Plug and Play Technology, http://www.upnp.org/about/default.asp, (2001)