E. Aarts · J. L. Encarnação (Eds.)

True Visions

The Emergence of Ambient Intelligence

Springer

10

Broadband Communication

P. Lagasse and I. Moerman

"I do not fear computers. I fear the lack of them."

Isaac Asimov

10.1 Vision

Mobile connectivity is a prerequisite for the deployment of true AmI systems. The concept of Ambient Intelligence requires the interaction of a variety of devices, appliances, sensors, and processors with persons who should not feel constrained in their movements when using the system. In order to best understand the goals and requirements of the broadband communications that underpin the AmI system let us imagine the existence of a "connectivity ether." Pervasive from a personal to a global scale, present anywhere, anytime, this connectivity ether makes us a part of a global electronic nervous system that connects all people and devices with processing or sensing capabilities.

The essential properties of this connectivity ether are that it is scalable, adaptable to changing circumstances, and self-healing. Above all it must be trustworthy and must keep us secure while being respectful of our privacy. As we do not want to wait, it must have a very low latency for whatever service or application we want to use. Finally it must protect us against unwanted data and against information overload by intelligent filtering of the exponentially growing amount of data sent over the network.

In order to achieve this vision there is no need to start from scratch. A wide variety of performant wireless technologies are already available or are currently being developed. The challenge we face is how to build this connectivity ether from the myriad of ICT devices, systems, and networks that are currently available or under development. A number of issues that need to be resolved are quite obvious:

The various mobile and wireless networks and protocols must seamlessly converge with the fiber-optic feeder and backbone network.

The complexity of interconnecting all those heterogeneous networks must be completely hidden so as to transform the complexity of the technology into simplicity of use.

- The restoration and self-healing capabilities should guarantee 100% availability. In view of the underlying complexity of the network, a self-healing and graceful degradation capacity appears to be the only solution to achieve the availability required by life-critical applications, which are bound to exist in such a pervasive network.
- The usual technical specifications of network bandwidth must be translated into a requirement for low latency of service.
- Regarding trust and security, leaving the user stranded to protect himself against all forms of e-garbage and cyber-crime is simply not an option.

10.2 Ongoing Evolutions in Broadband Communication

Nowadays, a lot of broadband applications are taken for granted in fixed networks. These applications require a high level of Quality of Service (QoS) and are generally characterized by high bandwidth requirements and low latencies, which can currently only be offered by fixed broadband access technologies (such as DSL, cable, and FTTH). The challenge future telecom operators and service providers are facing is to examine how these bandwidth hungry and QoS demanding services can also be provided in wireless access networks.

The number of Internet users is still expanding rapidly (from 688 million users, end 2003 (ITU-D, online) to 889 million users in March 2005 (Internet-World Stats, online)), while the number of fixed broadband Internet users is increasing even more rapidly. End 2004 there were about 150 million broadband users (Point Topic Ltd., online), which is 17% of all Internet users. However, when looking at the mobile subscriber statistics revealing 1,688 million subscribers in December 2004 (GSM World Statistics, online), we may conclude that the mobile terminal is without any doubt the most popular terminal. These numbers clearly indicate the great potential of broadband communications, not only for fixed, but also for mobile broadband services. In future users will take it for granted to have access to the same broadband services from their mobile terminal as they have now from their fixed terminal. The unprecedented growth of mobile communications further triggers the emergence of novel mobile applications and services, of which an overview is given in Chap. 9 (Mobile Computing).

In this chapter we first give an overview of the various wireless technologies, leading to the question whether we really need all of them. Then we will describe the concept of personal networking that is currently under development and that will allow to integrate the variety of wireless technologies into a network that comes close to the realization of the ideal situation of the connectivity ether.

10.3 Too Many W

To address the question v
gies for wireless commun
existing technologies.

10.3.1 Overview of E

During the past decade,
been observed in mobile
of the main wireless tec
characteristics (theoretic
application domains.

The IEEE has develo
in terms of application do
802.15, online) define sho
works (WPAN). A WPA
in which a limited numb
PDAs, etc.) communicat
nections. The different 8
The 802.15.1 standard (
medium data rate (1 Mb
technology to connect re
need for a cable. Current
technology. The 802.15.3
high date rates (up to 48
ter known as the global
platform for low data rat
tion.

The group of 802.11 s
(Wireless Fidelity), aims
cial hotspots and has a
(outdoor). The most pop
nowadays, are 802.11b a
ing at 2.4 GHz and interc
originally only allowed ir
Europe for indoor comm
because of interference v
WLAN standard develop
Institute) (ETSI HIPER
the competing standards
to perform better in tern
a definite lead in the wo
ments. However, more re
had been proposed to s

10.3 Too Many Wireless Technologies Today?

To address the question whether or not there are too many different technologies for wireless communication we first present an overview of some of the existing technologies.

10.3.1 Overview of Existing Wireless Technologies

During the past decade, many technological and experimental advances have been observed in mobile and wireless networks. Table 10.1 gives an overview of the main wireless technology standards available today and their main characteristics (theoretical bit rate, spectral frequency, coverage area), and application domains.

The IEEE has developed several standards, which complement each other in terms of application domain and coverage area. The 802.15 standards (IEEE 802.15, online) define short-range technologies for Wireless Personal Area Networks (WPAN). A WPAN, often just called PAN, is a small wireless network, in which a limited number of personal devices (such as PCs, mobile phone, PDAs, etc.) communicate with each other through short-range wireless connections. The different 802.15 standards are targeted at different data rates. The 802.15.1 standard (Bluetooth) (Haartsen 2000; Bluetooth, online) has a medium data rate (1 Mbps) and was originally meant as a low-cost wireless technology to connect relatively cheap electronic devices hereby avoiding the need for a cable. Currently Bluetooth is however gaining interest as a WPAN technology. The 802.15.3 standard (Ultra-Wide Band or UWB) enables very high date rates (up to 480 Mbps), while 802.15.4 (IEEE 802.14.1, online), better known as the global standard for ZigBee (ZigBee, online) technology is a platform for low data rate, low-cost, and power-efficient wireless communication.

The group of 802.11 standards (IEEE 802.11, online), also known as Wi-Fi (Wireless Fidelity), aims at home and business wireless LANs and commercial hotspots and has a medium coverage area from 10 m (indoor) to 500 m (outdoor). The most popular standards, available in most laptops and PDAs nowadays, are 802.11b and its higher bit rate successor 802.11g, both operating at 2.4 GHz and interoperable. The IEEE 802.11a, operating at 5 GHz was originally only allowed in the United States, but is recently also permitted in Europe for indoor communication only (outdoor communication is forbidden because of interference with RADAR). In Europe there is HiperLAN/2, the WLAN standard developed by ETSI (European Telecommunication Standard Institute) (ETSI HIPERLAN/2, online). 802.11a and HiperLAN/2 have been the competing standards for a long time. Although HiperLAN/2 has proven to perform better in terms of throughput, the 802.11a standard currently has a definite lead in the worldwide market as the top choice for WLAN deployments. However, more recently the 802.11h standard (not shown in Table 10.1) had been proposed to satisfy regulatory requirements for operation in the

Table 10.1. Wireless technology standards

technology	theoretical bit rate	frequency	range	main application
HomeRF	1 Mbps (v1.0), 10 Mbps (v2.0)	2.4 GHz	~ 50 m	wireless network for home and small office
IEEE 802.11b	1, 2, 5.5, and 11 Mbps	2.4 GHz	25–100 m (indoor), 100–500 m (outdoor)	WLAN, hotspot
IEEE 802.11g	up to 54 Mbps	2.4 GHz	25–50 m (indoor)	WLAN, hotspot
IEEE 802.11a	6, 9, 12, 24, 36, 49, and 54 Mbps	5 GHz	10–40 m (indoor)	WLAN, hotspot
Bluetooth (IEEE802.15.1)	1 Mbps (v1.1)	2.4 GHz	10 m (up to 100 m)	cable replacement, WPAN
UWB (IEEE 802.15.3)	110–480 Mbps	mostly 3–10 GHz	~ 10 m	digital imaging and multimedia applications in WPAN
IEEE 802.15.4 (e.g., ZigBee)	20, 40, or 250 kbps	868 MHz, 915 MHz, or 2.4 GHz	10–100 m	home automation, remote monitoring and control
HiperLAN2	up to 54 Mbps	5 GHz	30–150 m	WLAN
IrDA	up to 4 Mbps	infrared (850 nm)	max. 1 m (line of sight)	instant transfer of data or digital images
IEEE 802.16	32–134 Mbps	10–66 GHz	2–5 km	last mile fixed broadband access
IEEE 802.16a	up to 75 Mbps	< 11 GHz	7–10 km (max. 50 km)	
IEEE 802.16e (broadband wireless)	up to 15 Mbps	< 6 GHz	2–5 km	

GSM	9.6 kbps	900 MHz 1.8 GHz 1.9 GHz	300 m (urban) 5–10 km (rural)	speech, SMS
GPRS	max. 171.2 kbps (typically 20–50 kbps)	900 MHz, 1.8 GHz, 1.9 GHz	as GSM	data, MMS, video
UMTS	144 kbps (rural) 384 kbps (urban) 2048 kbps (indoor)	2 GHz range	1 km (macro-cell) 100 m (micro-cell) 75 m (pico-cell)	speech, SMS, MMS, data, video

5 GHz band in Europe. 802.11h is essentially 802.11a with additional European features and actually incorporates several HiperLAN/2 functionalities.

The IEEE 802.16 family (IEEE 802.16, online), better known as the wide-range WiMax technology, is designed to provide wireless last-mile broadband access in the Metropolitan Area Network (MAN), delivering performance comparable to traditional cable and DSL. The main advantages of systems based on 802.16 are the cheap installation cost and the ability to quickly provision broadband services in scarcely populated areas, which are difficult to reach with wired infrastructure.

We further mention another medium-range standard, HomeRF (Negus et al. 2000). This is a wireless standard developed mainly for consumer use. Although a technically strong solution, it lacks strong vendor support.

A final (very) short-range wireless technology, shown in Table 10.1, is IrDA (Williams 2000). IrDA is short for Infrared Data Association, a group of device manufacturers that developed a standard for transmitting data via infrared light waves enabling transfer from one device to another without any cables. Many devices such as PCs, PDAs, and printers are equipped with IrDA ports. IrDA ports support roughly the same transmission rates as traditional parallel ports. The only restrictions on their use are that the two devices data must be very close to each other and that there must be a clear line of sight between them.

The technologies discussed so far are rather fixed wireless communication or access technologies (offering wireless communication or wireless access, but

not really supporting mobility or only low mobility), while the last three technologies presented in Table 10.1 are wireless technologies supporting high user mobility.

GSM or Global System for Mobile communications represents the second generation (2G) mobile phone standard and is world's most widely used mobile system (GSM World (online); Rahnema 1993). GSM was originally developed for Europe, but has now excess of 75% of the world market. GSM was initially designed for operation in the 900 MHz band and subsequently modified for the 1,800 and 1,900 MHz bands. The 900 and 1,800 MHz frequency bands are used in Europe, Asia, and Australia, while the 1,900 MHz frequency band is primarily deployed in America. GSM is designed for voice and is operated in circuit switched mode.

GPRS (General Packet Radio Service) is a packet-based communication service for mobile devices that allow data to be sent and received across the GSM network (Samjani 2002). GPRS enables always-on wireless Internet access over the GSM network without the need for a dial-up modem. GPRS can theoretically achieve 171.2 kbps using multi-slot techniques, but the available bit rate today is generally much less (typically between 20 and 50 kbps). GPRS is an upgrade to the existing GSM network. The GSM network still provides voice and the GPRS extension handles data, hereby allowing voice and data to be sent and received at the same time. GPRS is a step toward 3G and is therefore often referred to as 2.5G.

UMTS or Universal Mobile Telecommunications System is the European implementation of the third generation (3G) mobile telephone standard (UMTS, online). The 3G mobile system is standardized under 3GPP (Third Generation Partnership Project), a collaboration between ETSI and other regional telecommunication standard bodies. The goal of UMTS is to enable networks that offer true global roaming and can support a wide range of voice, data, and multimedia services on rapidly moving wireless devices. UMTS provides multimedia services in the 2 GHz band and offers data rates of 144 kbps at vehicular speed in rural areas (macro-cell), 384 kbps at pedestrian speed in urban areas (micro-cell) and 2 Mbps indoor (pico-cell). Unlike GPRS, UMTS is based on a completely different technology than GSM (UMTS is based on Wideband Code Division Multiple Access or WCDMA), while GSM is based on Time Division Multiple Access or TDMA, implying that a completely different access network has to be built from scratch. This explains why UMTS shows a much more costly and incremental deployment than GPRS.

10.3.2 Do We Really Need All Those Technologies?

From the overview in previous section, it is clear that there exist many wireless technologies. Unfortunately none of the wireless technologies is powerful enough to support all professional and private services of a mobile user anytime and anywhere. Each wireless technology is designed for some specific application or user context. When the user is moving at vehicular speed, then

a wireless technology that supports high user mobility (such as GSM, GPRS, or UMTS) is required. However such technologies are bandwidth constrained, and hence, when the user mobility is low, other technologies, such as a high-bandwidth WLAN may be preferred. Low bit rate power-constrained wireless devices, such as low-cost battery-empowered sensors or small home appliances, do not require a relatively expensive, high bit rate, high power-consuming Wi-Fi radio interface, but will benefit from cheap wireless technologies like ZigBee. It is obvious that the coexistence of different heterogeneous wireless technologies is indispensable for future ubiquitous wireless communications. In future even more advanced, higher bit rate and more power-efficient wireless technologies will emerge, and will have to coexist with current wireless standards. Interworking between heterogeneous wireless networks is hence getting increasingly important.

Some interworking mechanisms are already being developed in the framework of "beyond 3G" systems. "Beyond 3G" systems are considered to be heterogeneous networks with multiple Radio Access Technologies (RATs) as well as reconfigurable user terminals in order to allow mobile users to enjoy seamless wireless services irrespective of their location, speed, or time of the day. They will allow users to choose their access technology according to his or her needs. This concept of "Always Best Connected" is shared by the 3GPP and it has laid the groundwork for a UMTS/WLAN interworking specification (Gustafsson and Jonsson 2003). While WLANs can offer high bandwidth, cellular networks provide a (nearly) full coverage. Cellular operators no longer see WLAN networks as a competitive RAT, but consider WLAN as a true complementary technology to the 3G cellular networks. Cellular operators embrace the benefits of WLAN to build Internet access hotspots and offer roaming agreements for their mobile customers. UMTS/WLAN integration is not only advantageous for the mobile users, but also for the mobile network. As soon as WLAN coverage is available (typically in town or business centers), a mobile user should switch seamlessly from the ubiquitous low bandwidth UMTS network to a high bandwidth WLAN. This will not only improve the QoS, but will also increase the capacity in the cellular network.

WLAN/UMTS interworking can be achieved by means of different degrees of coupling (Apostolis et al. 2002; Pinto et al. 2004): from very tight coupling, where the WLAN access network is considered as a generic UMTS radio access network, to loose coupling, where WLAN access network is connected to the UTRAN (UMTS Terrestrial Radio Access Network) core network via the Internet. The higher the degree of coupling, the faster the vertical UMTS/WLAN handover can be achieved, but the larger the technical (and economical) challenges.

Interworking between heterogeneous wireless access technologies enables a mobile user to enjoy seamless Internet connectivity irrespective of location, speed, or time of the day. However, wireless Internet access anytime and anywhere is not enough to support all a person's professional and private activities. Nowadays an average user has several devices at several distant

locations, such as the mobile devices he carries with him, devices at home, in the office, in the car, etc. A user wants to have a trusted communication between his many local and remote devices, hereby avoiding complex configuration and authentication steps. A more global approach will be needed to provide a global seamless connectivity of heterogeneous devices over heterogeneous communication networks in all circumstances.

10.4 How to Deal with All These Wireless Technologies?

10.4.1 The Concept of Personal Networking

The many technological advances and market demands for mobile and wireless networks have triggered the introduction of a diversity of terminals and wireless devices (notebook, PDA, mobile phone, GPS-terminal, camera, sensors, home appliances, etc.) and the development of various new multimedia services (text and multimedia messaging, real-time video broadcast/multicast, video-conferencing, video-on-demand, wireless payment, gaming, chatting, remote monitoring, etc.). Nowadays an average user carries several personal wireless devices with him, while having more devices at home, in the office, and probably also in his car. The number of personal devices is expected to increase during the next years. Although many devices, wireless technologies, and advanced services are already available today, they are not yet able to adapt autonomously to the user context and the user's preferences. Trusted communication between heterogeneous personal devices may be possible today, but is very complex as a lot of manual configuration and authentication steps are involved. Many newly introduced services have a great potential, but are often perceived as a burden as they do not take into account the user's expectations, the ease-of-use, the limitations of the user's terminal, or the user's network environment.

From PAN to PN

The introduction of the Personal Area Network (PAN) by the IEEE 802.15 working group (see previous section on wireless technologies) is already a first step toward communication between personal devices. However, a PAN only consists of personal devices in the close vicinity of the person and hence only offers a solution at the local scale, while the resources or services needed by a person are not necessarily in the close vicinity of the person. The concept of personal networking, which is studied for example in the IST Integrated Project MAGNET (IST MAGNET, online) (My personal Adaptive Global NET), can bring a global solution to satisfy a (mobile) person's professional and private needs, by providing trusted communication between the many local and remote personal devices. A Personal Network (PN) is a dynamic

collection of interconnected heterogeneous personal devices, not only the local devices centered around the person, but also personal devices on remote locations such as devices in the home network, the office network, and the car network (Niemegeers and Heemstra de Groot 2003). The PN is organized in ad hoc interconnected clusters, consisting of heterogeneous devices, and linked by various suitable interconnection mechanisms (e.g., the Internet, WLAN, GSM, UMTS, PSTN, etc.), as shown in Fig. 10.1. The different clusters are interconnected by secure tunnels between devices with gateway capabilities. The different devices in a cluster may be heterogeneous in terms of processing power, data storage capacity, radio interface(s), battery resources, display size, resolution, etc.

The PN concept can be very well illustrated with the virtual home truck scenario presented in Fig. 10.2. Transportation and logistics represent a major business industry employing millions of truck drivers. Each day, these people spend hours in their vehicle while driving, waiting, or sleeping, and they are often several days away from home. By creating a virtual home environment, these people can be offered the ability to stay in touch with their family, to stay connected with their company and clients, or the possibility to contact their colleague truck drivers, which could have great commercial potential taking into account the large number of truck drivers worldwide.

Consider a truck equipped with a mobile phone, broadband Internet access, LCD display, headset, etc., forming a cluster of cooperating devices. When finished working, a truck driver could set up an Internet connection to his home. At home, a cluster of cooperating cameras, speakers, headsets, provides the truck driver with a virtual home environment. Through this environment, he can virtually walk around, seeing his family, talking with them, watching together a movie, playing games, etc.

When driving, the truck driver can listen to his digital music collection by streaming it from a server in his network at home. When the truck driver stops at a parking, he can read his e-mail, search for colleagues, play a game with other truck drivers, etc. When the truck driver arrives at a client, his

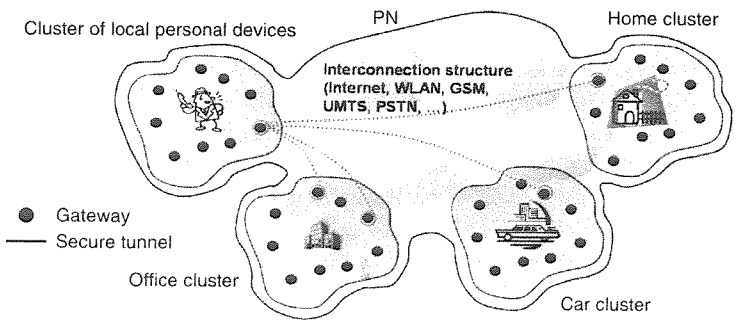


Fig. 10.1. Personal Network concept

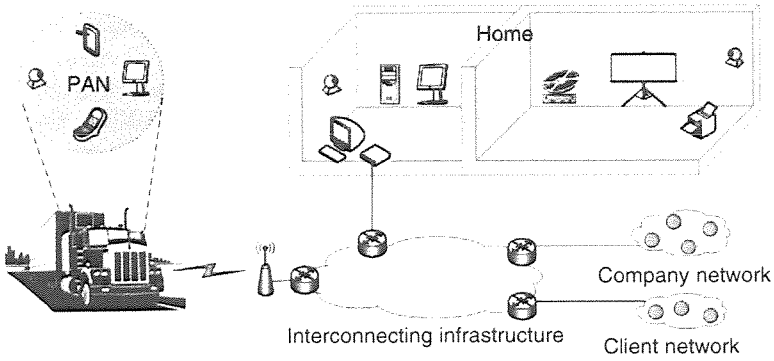


Fig. 10.2. Virtual home truck

PAN can connect to the client's company network and download the necessary documents. The documents can be digitally signed, handed over to the client and a copy can be uploaded to the truck driver's company, reducing the administrative burden. It should be noted here that the client network is not part of the PN of the truck driver. A PN should hence not only be able to support its own personal services, but also services offered by foreign devices or foreign networks.

From PN to Multi-PN

A PN has been defined earlier as a distributed networking solution to support a person's professional and private services by integrating all of a person's devices capable of network connectivity, whether in his or her (wireless) vicinity or at remote locations such as office and home. However, interaction and communication between multiple persons may also be of prime importance in some private or professional scenarios and should therefore also be supported by the PN concept. Multi-PN communication is actually a special case of communication of a PN with foreign nodes, where the foreign nodes belong to another person's PN. The multi-PN communication has already been touched in the previous example of the virtual home truck, when addressing the interactive gaming between truck drivers. The conference/meeting scenario in Fig. 10.3 presents another, more illustrative example of multi-PN.

Consider the case of a conference (or meeting) where multiple persons, each representing a single PN, are attending a conference, taking place in different locations. Each room is equipped with a wireless network printer, an access point providing wireless Internet access, a large display, and a file server to temporary store relevant files. A number of participants also have Internet access through the use of their mobile phone. As not everyone is directly within each other's send range, multi-hop routing is used in order to enable person-to-person connections. If someone has interesting information for other

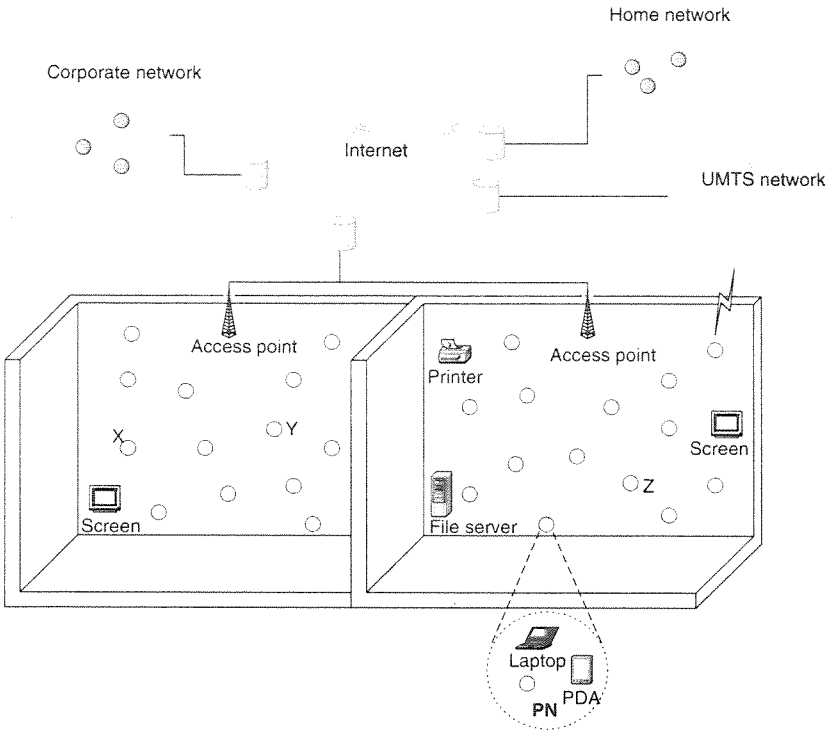


Fig. 10.3. Conference/meeting scenario

persons, he can upload his files to the file server, hereby setting the necessary protection mechanism. The subgroup of persons for whom the information is destined, is automatically informed and can download the information at their convenience. Upon download, some authentication step is required in order to prevent misuse of information. If needed, files can be printed on the network printer. At a given moment person X would like to give person Y some information about another interesting project. Unfortunately this information is stored on the laptop of a colleague Z attending a meeting in a remote room. The colleague is searched for over the multi-PN and the requested file is transferred through multi-hop routing. If needed, a person can access his personal or corporate files by using one of the access points (directly if the person is within the range of an access point or via multiple hops when the access point is out of reach) or even by borrowing the uplink (e.g., a UMTS link) of another person, if this person grants access to the first person. If someone quickly wants to present some extra information to the other attendees in the meeting, it can be sent to the large display. During a meeting multiple persons can work on the same file, which is displayed on the large screen.

For more examples of potential PN and multi-PN scenarios we refer to Chap. 9 (Mobile Computing).

10.4.2 Main Challenges of Personal Networks

Some existing technologies may offer solutions to a specific aspect of the PN scenario. However, in order to realize the full PN scenario meeting a person's communication needs, many (existing and new) technologies need to be combined into an integrated solution. The main PN research challenges are listed below.

Self-organization

A PN needs to be easy to use, setup, configure, and maintain as well as fast and secure. The different heterogeneous personal devices have to automatically form clusters. Once a personal device is switched on, it should be immediately recognized by the other local personal devices and automatically integrated in the cluster, without user intervention. A cluster has to be capable to deal with different link technologies. Not all devices necessarily have the same radio interface; some devices may have multiple radio interfaces and will have to forward data between devices with different radio interfaces.

The clusters have to automatically discover potential gateways, which can give access to interconnection structures. If a device in a cluster needs access to a device or service in another cluster, the most appropriate gateway has to be selected in each cluster and a secure tunnel has to be created between the gateways.

Support of Dynamics and Mobility

A PN consisting of several clusters interconnected with each other over fixed infrastructure can be very dynamic: devices and clusters may only intermittently be accessible due to mobility, energy constraints, and radio link characteristics. Clusters may change their point of attachment to the fixed infrastructure (e.g., when the user together with the devices he carries is moving.), clusters may merge and split (e.g., the cluster around the user can merge with or split from the home or office network), devices may enter or leave (e.g., by switching on or off, or because of a bad radio link). This dynamic behavior will strongly influence the communication in PNs, and one of the challenges is to provide connectivity and to allow service continuity in a dynamic PN environment. The user mobility and network dynamics may have a serious impact on many other mechanisms such as addressing, routing, gateway discovery, service discovery, and context discovery.

Naming, Addressing, and Routing

Each service, device, cluster, and PN should have a unique name and all network interfaces a unique address. The reason to have names is to hide irrelevant information from users and to facilitate user access to devices and services. The user does not need to be aware of the actual location/directory or any other detailed information of the subject. As the clusters and the PN are continuously dynamically changing, the addressing of different devices and interfaces may be altered. The user should not worry which dynamically assigned address corresponds to which device, but only use the same name for the same device. The addressing should further allow efficient routing.

Ad hoc Networking

Multi-hop aspects and ad hoc networking techniques can play an important role in the realization of the PN concept, not only in scenarios where multiple PANs (clusters) come together and form an ad hoc network (such as in the conference scenario when no fixed infrastructure is involved), but also for PN-wide routing (cluster to cluster over fixed infrastructures). A PN may not be seen as a pure ad hoc network as defined in the IETF MANET context: "massive network of hundreds or thousands of possibly highly mobile nodes in which routing plays a leading role." In the PN concept, the focus will be much broader and much more challenging than what is called "MANET routing." Context awareness can be exploited to incorporate additional intelligence in existing ad hoc routing protocols. Further, the possible existence of multiple radio interfaces within one cluster (Bluetooth, 802.11b, 802.11a, Zigbee, etc.) imposes additional challenges on the ad hoc routing protocols. In the PN vision, an ad hoc network of multiple clusters will normally not operate as a stand-alone network (as opposed to pure ad hoc networks). Clusters will interact with other clusters or foreign nodes via Internet access points over infrastructure-based networks.

Context Awareness and Context Discovery

Context awareness is a prerequisite for realizing the AmI vision. A PN has to adapt as quickly as possible to context. Hereby context means time information (day, hour), location (geographical position and environment), the person's agenda and schedules, the person's activities, network resources (e.g., available bandwidth), network conditions (e.g., quality of the link, latency), availability of personal devices, device properties (e.g., type of display, resolution, battery capacity), presence of other persons and their available devices and services, user profile (e.g., preferences, subscriptions for Internet access, and charging policies), etc.

A context discovery framework is required for collecting and updating context information and for offering relevant context information to any context-aware application or service. If a context discovery framework is available, a context-aware application or service does not have to care about gathering of the context information, but can just rely on the context discovery framework. Below we present a few examples of context-aware services.

- Gateway discovery and gateway selection: if multiple gateways are available, the gateway selection may be based on the bandwidth and latency requirements of the required service, on battery capacity, on available Internet subscriptions and charging limits, etc.
- Ad hoc routing: routing can be more efficient if constraints like battery capacity and quality of the links are taken into account. Routing can further benefit from the presence of other person's devices (provided that the other person grants access to his devices and the necessary security precautions are taken).
- Service discovery: the result of a service request may depend on the user preferences, the location, the time information, the devices properties, the available bandwidth, etc.
- QoS support: through interaction with the context (and service) discovery framework, the best network, terminal, and service parameters can be automatically selected for a particular service and a given context. As the context discovery frameworks not only keep information about the network conditions and terminal capabilities, but also are aware about the user preferences, a PN is offering more than just QoS. A PN is able to translate the user's (subjective) Quality of Experience requirements into (objective) QoS parameters.

Service Discovery

A PN contains many services that must be discovered and used. A service discovery mechanism is needed to locate the place, where the service can be found, to specify what the service provides and to describe the procedure how the service can be accessed and controlled. The service discovery mechanism should not be limited to local service discovery in the vicinity of the user, but should also be able to detect remote services in remote clusters of the PN or services offered by other persons or service providers. The service discovery mechanism has to further take into account the dynamics of the PN and has to be aware that there is no permanent connectivity between nodes and between clusters. It can hence not (always) rely on infrastructure-based servers.

Security and Privacy

Secure communication must be guaranteed in the PN through adequate key establishment, encryption, and authentication mechanisms. Security is required at different layers: at the link layer for secure connectivity between

devices within the same radio domain, at network layer for secure routing and secure tunneling over insecure network infrastructures (such as the Internet), at the service level for secure service and context discovery and services access. All data needs to be secured: not only application data, but also control and management information. Different levels of trust relationships between communication entities should be possible in terms of duration (from ephemeral to permanent trust relations) and in terms of cryptographic keys (from lightweight keys to highly robust keys). Keys and/or passwords used in authentication and encryption procedures must be installed in each personal device so that they are not accessible to unauthorized people. The PN must provide a simple management tool to the end-user for establishing and managing trust relationships.

10.4.3 Some Personal Network Solutions

PN Architecture

Figure 10.4 shows the PN architecture, proposed by the IST MAGNET project. The architecture is composed of three abstraction levels (ALs): the connectivity, the network, and the service AL.

The connectivity AL consists of various wired and wireless link layer technologies, organized in radio domains, including infrastructure links.

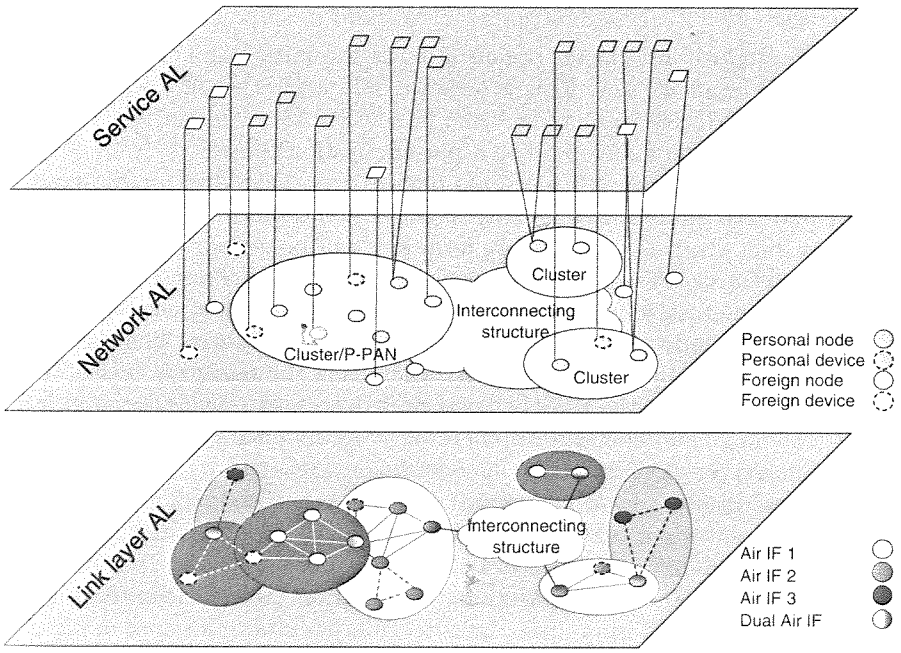


Fig. 10.4. PN architecture: view of three abstraction levels

A distinction is made between node and device, where a node is a device that implements IP. The link layer will allow two nodes or devices implementing the same radio technology to communicate if they are within radio range.

To allow any two devices within a PN to communicate, a network AL is needed. This level divides the nodes and devices into personal and foreign nodes and devices, based on the trust relationship between each other. Only devices that are able to establish a permanent or long term trust relation can be part of the user's PN. A long term trust relation implies that sharing and borrowing devices is also taken into account (e.g., family devices, devices from work). Personal devices that have such a long term common trust relation form clusters and these clusters can communicate with other clusters via interconnection structures. The P-PAN or Private Personal Area Network is a special cluster and represents the collection of personal devices around the person.

The highest level in this architecture is the service AL, which incorporates two types of services; public and private services. Public services are offered to anyone, while private services are restricted to the owner or trusted persons by means of access control and authentication.

The rest of this chapter will mainly focus on architectural concepts at the network AL.

Cluster Organization

The first step toward cluster organization is the initial assignment of a long term (or permanent) key. This procedure only occurs when a new personal device is introduced for the first time in the P-PAN. The long term key establishment can be understood as a personalization process.

At the connectivity AL, personal devices that have direct radio connectivity and that share the same trust relation will establish a trusted communication link. To this end, a short term key will be generated starting from the long term key. When more personal devices are discovered with which a trusted communication link can be established, a cluster of personal devices is formed that only consists of trusted communication links. The short term key establishment happens frequently, each time the cluster is formed (e.g., every morning when the user switches on his personal devices around him), while the long term key establishment only occurs once (when a new device is introduced for the first time).

Once secure link layer connectivity has been realized, communication at the network level can take place between personal devices, without using devices that have different or no ownership relation. To this end, the network layer needs access to the connectivity information at each node. Not all devices in a cluster will have direct links to all other nodes, because of different link layer technologies or radio-range limitations. This implies that a cluster might be a multi-hop network.

All personal nodes within one cluster can securely communicate with each other using IP as the common protocol. The use of IP as a common language makes it possible to have a network layer architecture that is independent from the heterogeneity of the underlying link layer. The cluster can also consist of limited devices that have no IP capabilities. As these personal devices can offer important services to the cluster, they should be connected somehow. To this end, a personal node that is able to communicate with this IP-incapable device will serve as its contact point (proxy), making its services accessible to the other nodes.

As the cluster network architecture is an IP-based multi-hop network, it should provide addressing and routing functionalities in order to enable efficient and secure intra-cluster communication. In order to cope with the specific characteristics and dynamics of clusters (nodes can join and leave the cluster, clusters can merge or split, nodes can move, nodes can have multiple wired/wireless interfaces), a distributed architecture is preferred. In such an architecture, all nodes in the cluster will cooperate in an ad hoc manner to provide the necessary network functionality (cluster formation and management, routing, addressing) and security functionality.

Finally, a cluster will not only operate as a stand-alone network, but it will also interact with its immediate environment, such as local foreign nodes or interconnecting structures. Nodes in the cluster that provide connectivity to nodes and devices outside the cluster are called gateway nodes. Gateway nodes will require some special functions such as address translation, set up and maintenance of tunnels, filtering incoming traffic, etc. These tasks might be quite heavy for some personal nodes, so it is useful to select powerful personal nodes as gateway nodes if possible. The process of finding capable gateway nodes with links to foreign nodes or interconnecting structures is another network function, which has to be provided by the cluster.

PN Organization

A PN can have multiple clusters that are geographically dispersed, but that have access via gateway nodes to the interconnecting structure. In order to form a PN and realize inter-cluster communication, two requirements need to be fulfilled. First of all, the clusters need to be capable of locating each other in order to establish tunnels between them. Secondly, once the PN has been formed, it should be able to maintain itself regardless of changes in gateway and node mobility. For these requirements to be fulfilled, we introduce the concept of a PN agent, a management framework that can be either centralized, under the control of a single provider or in a fixed cluster, or distributed over multiple providers or operators. Clusters that have obtained access to the interconnecting infrastructure announce their presence to this PN agent. This presence information should at least include the name of the PN the cluster belongs to, the point of attachment to the interconnecting structure, i.e., the IP address of the gateway through which the cluster can be reached and some

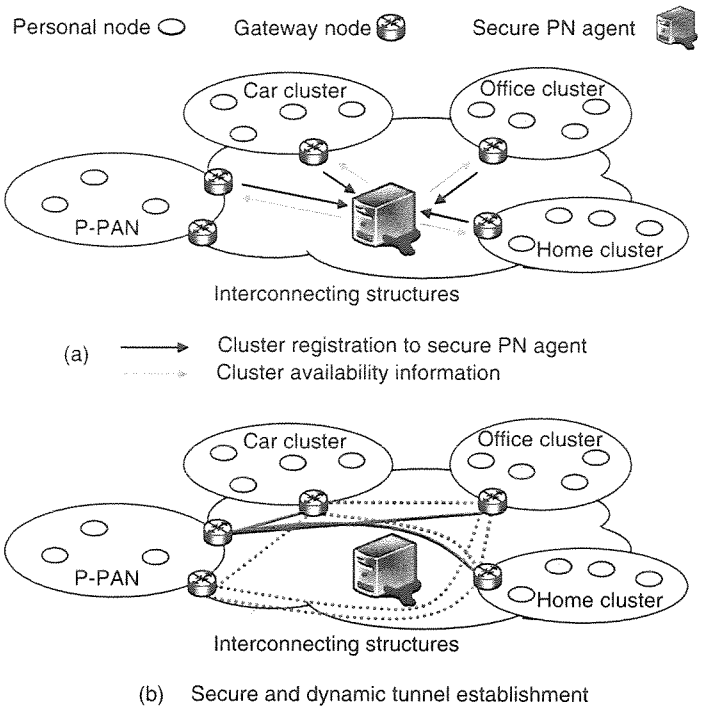
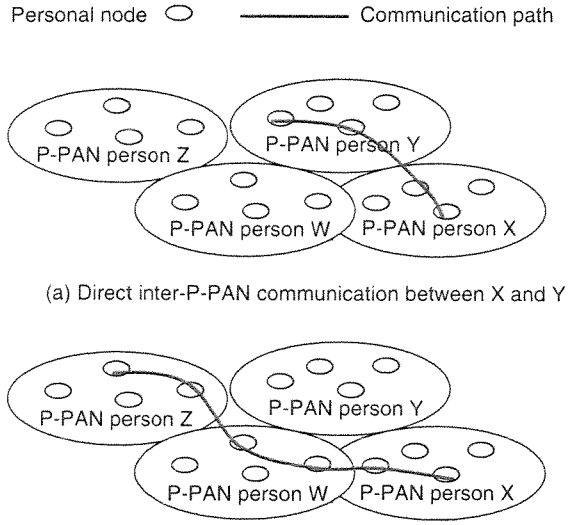


Fig. 10.5. PN formation and maintenance

credentials to verify this information. The PN agent will communicate this information to the cluster gateway nodes and this information will trigger the creation of secure tunnels between the clusters, as shown in Fig. 10.5 For simplicity the PN agent is presented as a central entity, but it could be as well distributed. The purpose of the tunnels is twofold. First, they provide secure inter-cluster communication by shielding the intra-PN communication from the outside world. Secondly, these tunnels are established dynamically. When clusters move, the information in the PN agent will be kept up-to-date. As a consequence, the PN agent will function as a secure database that tracks the PN clusters.

The information in the PN agent is tightly coupled to other components of the architecture such as the naming system, resource and service discovery framework and tunnel management mechanism. The PN agent can also be used by foreign nodes that wish to communicate with the PN. The only thing foreign nodes need to know in order to communicate with the PN is the contact information of the PN agent. The PN agent will find the appropriate service or node to connect with.



(a) Direct inter-P-PAN communication between X and Y

(b) Indirect inter-P-PAN communication between X and Z

Fig. 10.6. Inter-P-PAN communication

Multi-PN Communication

A first type of multi-PN communication is inter-P-PAN communication. Inter-P-PAN communication is the communication between personal devices of two persons that are in each other's neighborhood. This communication can be direct or indirect. Direct means that the communication paths do not include devices belonging to a third party. Indirect means that the communication paths include nodes that belong to other persons. Figure 10.6a, b illustrates direct and indirect inter-P-PAN communication.

Inter-P-PAN communication does not rely on the fixed infrastructure, resulting in multi-hop wireless communication paths. This implies that all network mechanisms such as resource and service discovery, naming, addressing, and routing should be able to operate in such an infrastructureless environment. For instance, nodes need to have unique addresses for communication, and intra-cluster routing should be extendable to neighboring P-PANs in a secure way, etc. Of course, access to the fixed infrastructure could be available, but will probably not be the most efficient way of networking.

Inter-PN communication is the communication between personal devices of two persons that are at remote locations. Figure 10.7 illustrates this type of communication.

When one PN wants to communicate with another PN at a remote location, it first has to find out how to reach this PN. To this end, naming and the PN agent concept will play a major role. By identifying the PN by its name and contacting the PN agent, the agent can provide the necessary information to establish communication with the remote PN, as it stores the IP addresses

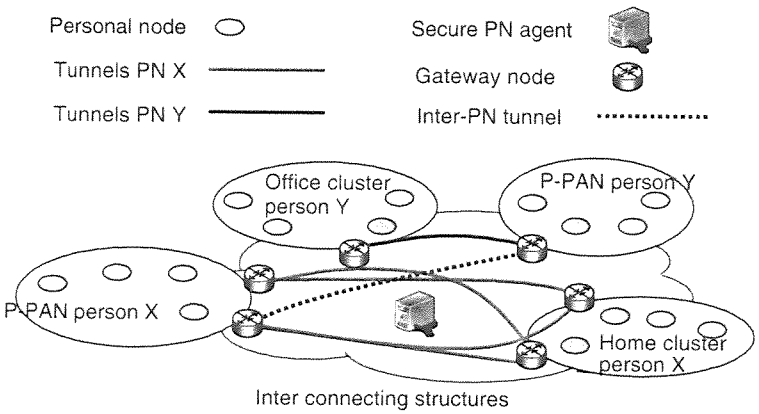


Fig. 10.7. Inter-PN communication

of the gateway nodes. This means that the central agent has to function as a kind of broker for establishing communication between PNs. Based on the information provided by the PN agent, a secure tunnel between the PNs can be established. The addressing solutions developed for intra-PN addressing should be extensible to inter-PN communication. In addition, routing protocols should be able to create connections, over the inter-PN tunnel, to the remote PN based on the assigned PN addresses.

10.5 Conclusions

In this chapter we have shown how the concept of personal networking, as studied within the IST MAGNET project, may bring a solution for the realization of the connectivity ether. The PN concept creates a personal distributed environment where persons can interact with various devices not only in the close vicinity, but potentially anywhere. This general PN architecture bridges heterogeneous (wired and wireless) networks and different wireless technologies, and hence hides the complexity of the connectivity to the end-user. The network layer is the glue that binds all a person's devices together into one PN. The network layer is based on a long term trust relationship that can offer communication between all a person's devices in a secure way. In this chapter, some solutions were presented to some of the most important networking issues needed to realize intra-PN and multi-PN communications that are context aware, and further support the dynamics of the PN and the user mobility. Those solutions based on results from the IST MAGNET project represent one of the possible implementations of AmI networking.

Within the European context, the IST project Ambient Networks (IST Ambient Networks, online), also develops complete and coherent solutions for

ambient networking, enabling the easy and dynamic composition of disparate networks amid an ever-increasing heterogeneity of technologies and provider structures. The main objectives of the project are to define a set of adaptive and self-configuring mobile network components, which will reduce planning, deployment, configuration, and network maintenance costs and a comprehensive, integrated security framework, preserving end-to-end network protection and robustness against attacks.

Another related IST project is WINNER (Wireless World Initiative New Radio). This project is rather focusing on providing new radio technologies to make mobile communication systems more adaptable to user needs anytime and anywhere. WINNER is working toward enhancing the performance of mobile communication systems through improvements of radio transmission.

A third related IST project is E2R (End-to-End Reconfigurability) (IST E2R, online). The key objective of the E2R project is to devise, develop, and trial architectural design of reconfigurable devices and supporting system functions to offer an expanded set of operational choices to the users, applications and service providers, operators, regulators in the context of heterogeneous mobile radio systems.

The Ambient Networks, WINNER, and E2R projects are also part of the Wireless World Initiative (WWI) (WWI, online).

In the framework of AmI networking, two more IST projects should be mentioned here – although there may be more IST projects dealing with some aspects of Pervasive Computing and Ambient Intelligence – are IST VESPER (Virtual Home Environment for Service Personalization and Roaming Users) (IST VESPER, online) and IST RUNES (Reconfigurable Ubiquitous Networked Embedded Systems) (IST VESPER, online). The purpose of the VESPER project is to develop a service architecture for provision of a Virtual Home Environment (VHE) across a multi-provider, heterogeneous network, and system infrastructure. The VHE (Bougant et al. 2003) is defined as a concept for personal service environment portability across network boundaries and between terminals. This concept includes that users are consistently presented with the same personalized features, user interface customizations, and services in any network and any terminal, wherever the user may be located.

The vision within the RUNES project is to enable the creation of large-scale, widely distributed, heterogeneous networked embedded systems that interoperate and adapt to their environments. The inherent complexity of such systems must be simplified for programmers if the full potential for networked embedded systems is to be realized. The widespread use of network embedded systems requires a standardized architecture that allows self-organization to suit a changeable environment. RUNES aims to provide an adaptive middleware platform and corresponding application development tools, which provide programmers the flexibility to interact with the environment where necessary, whilst affording a level of abstraction that

facilitates ease of application construction and use. This will allow for a dramatic cut in the cost of new application development and a much faster time to market.

Another European initiative is the Eurescom project Personal Nets (Eurescom, online). Personal Nets are a generic concept for providing an individualized, user-centric solution to the longer term integration of all communication and information services. The Personal Nets project is a feasibility study to test the concept of Personal Nets. The main objectives of this study are to identify and clarify from a customer point of view, the technical and business models, and concepts of Personal Nets bearing the changing paradigm of the digital economy in mind (i.e., different models for cooperation in service provisioning) and to investigate the underlying technologies, services and applications, and the business model of the Personal Nets concept.

In the United States there are also several research initiatives on the connectivity and cooperation between personal devices, such as the MOPED Project at University of Illinois at Urbana-Champaign (Kravets et al. 2001) and the Oxygen project at MIT (Oxygen, online). The goal of the MOPED Project is to enable cooperation between personal devices. The project presents a networking model that treats a user's set of personal devices as a MOPED, an autonomous set of MObile grouPEd Devices, which appears as a single entity to the rest of the Internet. All traffic for a MOPED user is delivered to the MOPED, where the final disposition of traffic is determined. To the outside world, this MOPED appears as a single device with a single interface or identifier. In reality, the group of devices cooperates to provide better services to the user. A MOPED component however needs constant contact with the infrastructure to function properly. A MOPED component cannot facilitate further ad hoc communication with MOPED components from other persons.

The vision of the Oxygen project, which is a project about pervasive human-centered computing, is that in the future computation will be human-centered and will be freely available everywhere, like batteries and power sockets, or oxygen in the air we breathe. Oxygen enables pervasive, human-centered computing through a combination of specific user and system technologies. Oxygen's user technologies directly address human needs. Speech and vision technologies enable humans to communicate with Oxygen as if they are interacting with another person, saving much time and effort. Automation, individualized knowledge access, and collaboration technologies help humans to perform a wide variety of tasks that they want to do in the ways they like to do them. Computational devices embedded in homes, offices, and cars sense and affect the immediate environment. Handheld devices empower humans to communicate and compute no matter where they are. Dynamic, self-configuring networks help machines locate each other as well as the people, services, and resources they want to reach. Software that adapts to changes in

the environment or in user requirements helps humans to do what they want when they want to do it.

It is clear that the "connectivity ether" described in the introduction is no longer a distant dream. Various research projects are rapidly developing architectures that will transform the current wireless and wire-line communication networks into this connectivity ether.

ISBN 3-540-28972-0

springer.com