The authors present a new networking concept referred to as ambient networks, which aims at enabling the cooperation of heterogeneous networks belonging to different operator or technology domains. We aim to provide a domain-structured edge-to-edge view for the network control to embrace the heterogeneity arising from the different network control technologies. In this way, it appears as homogeneous to the users of the network services. We aim for an instant network composition to allow rapid adaptation of the network domain topology as required for moving networks. This new view of network composition allows us to treat the communication endpoints as a special case of network domains as well. We introduce the ambient control space, which enables the ambient networks concept and introduce its main features. Two ambient control space functions, media delivery and generic link layer, are presented in more detail.

INTRODUCTION

In today’s wired and wireless networks, the trends in networking technology very much point to a dominance of Internet technology with all its flavors. IP is the common lingua franca to enable the exchange of data across various networks. There is, however, an increasing divergence in the network control layer: different control environments are established to facilitate services like virtual private networks (VPNs), security, mobility, quality of service (QoS), network address translation (NAT), multicast, and so on. For a multitude of services, data might still be handled by uniform Internet networking, but the control of such services is becoming increasingly fragmented. More and more, the network as a whole therefore diverts from the pure end-to-end view of the Internet. Furthermore, applications would like to rely on enhanced and consistent support from the network for the complete delivery chain.

This lack of a common control layer for joining the services—in a wide sense of the word—of multiple networks represents a crucial challenge both technically and from a user perspective. Usage scenarios that should be realizable in the mid-term future include utilization of multiple devices, multiple networks, and multiple access technologies in an integrated fashion. This is not easily controllable or manageable with today’s technologies. These usage scenarios partly evolve from the vision of “ambient intelligence,” developed and popularized by various European research efforts. The core notion of this vision is to make technology invisible and usable, whenever necessary. This notion holds in particular for networking technologies: networking should be available whenever required in the most appropriate and affordable form.

Our approach, which we call ambient networking, aims to provide a domain-structured edge-to-edge view for the network control. In this way, an ambient network is expected to embrace the heterogeneity arising from the different network control technologies such that it appears homogeneous to the potential users of network services. The vision is to allow agreement for cooperation between networks on demand, transparent, and without the need of preconfiguration or offline negotiation between network operators. End users are increasingly not just owners of a terminal or PC; they own and effectively operate a network of devices in their homes and offices, and around the body. Consequently, they are included in this network of cooperation, treated as operators of special low-complexity networks. This approach generalizes to many different kinds of networks that are currently appearing, such as intervehicle networks, body area networks, and sensor networks. By making every device a network, the network is the primitive building block of our architecture, allowing all types of networks to be integrated into a larger system.

This concept of ambient networking is guided
by market and business analyses carried out in European research projects as well as by the conclusions and technology concepts discussed within the Wireless World Research Forum (WWRF) [1]. WWRF gathers research efforts targeting mobile networks beyond 3G. The forum provides an excellent opportunity to understand the issue currently discussed in the context of networks beyond third-generation (3G) and the relations between these issues (see Books of Vision [2, 3]).

In the following sections, we briefly describe existing approaches for next-generation wireless networks. Based on this description, we discuss how we imagine the progression from simple all-IP networks to ambient networks. We investigate how ambient networks can be built, with some details highlighted. We then conclude the article.

**PREVIOUS EFFORTS**

To put these goals and requirements into perspective, let us have a brief look at two related research efforts.

**ARCHITECTURE DISCUSSION IN THE INTERNET COMMUNITY**

Research in the Internet community on future network architectures is mainly influenced by the discovered deficiencies of the current Internet, where mobile networks and mobility aspects are treated with comparatively low priority. The motivation for reconsidering the design of the Internet is its current widespread presence and its impact on social, economic, and political aspects [4, 5]. Efforts are made to define a common set of architectural principles and tenets to guide the development of a new Internet architecture. In parallel to these efforts, which adopt a top-down analysis, several technology-focused activities exist to work on solutions for new naming and addressing schemes, QoS support, and routing (e.g., [6]).

Interestingly, one of the findings made within these research activities is the concept of separating the Internet into domains or realms that are, on a higher logical level, interconnected to build the Internet. The separate domains are defined to encompass networks deploying different technologies than those used in the existing Internet. This approach is supposed to ease technological advances, as the localized introduction would be made possible. Still, the definition of the required border functions, translating between different domains, has not been provided yet, except for some conceptual approaches such as that described in [7].

A second approach of the Internet research community is the work on overlay networks, introduced to improve network reliability or derive a network structure implementing different routing paradigms. In [8] such an overlay network is described to implement an architecture tailored to the demands of content delivery, which imposes modifications on the addressing and routing schemes implemented.

While all of these research efforts are essential to build on, we believe that the emerging needs of future wireless and mobile networks are not fully satisfied. Issues concerning the (potentially automated) control of larger networks are not conclusively addressed; how to combine smaller networks into larger ones, and support mobility or heterogeneity are other exemplary shortcomings. The following sections will analyze these needs and the solutions proposed by ambient networking.

**IP² NETWORK ARCHITECTURE INITIATIVE**

The migration of mobile networks to an IP-based network architecture is strongly favored by major players in the market [9]. NTT DoCoMo had, for example, already described an IP-based international mobile telecommunications (IMT) network platform (IP²) as the successor to the existing 3G mobile networks in 2001. The concept was introduced to meet the requirements expected to arise when aiming at mass market support for mobile multimedia services. The key requirements identified by IP² are:

- Support of multimedia traffic that equals 80 percent of the overall load in the cellular system
- Efficient support for mobility
- Support for diversified radio accesses
- Support for seamless service provisioning
- Open interfaces to allow external application providers to use the IP² communication infrastructure

The approach taken within IP² is mainly characterized by two design choices. First, it is envisaged to provide an IP-based transport network that provides access to different radio access networks. These networks can deploy various link level technologies, but are expected to provide an IP-based interface to allow them to be connected to the IP² transport network. Second, connectivity and control functions are separated, and all control functions are logically grouped in two domains referred to as the network control platform (NCPF) and service support platform (SSPF).

**FROM ALL-IP NETWORKS TO AMBIENT NETWORKS**

**THE CURRENT VIEW OF ALL-IP**

Ambient networks are based on all-IP-based mobile networks and can be regarded as the outcome of the continued adoption of Internet design principles. All-IP mobile networks can be characterized by, among other aspects, clear separation between transport and control-related tasks. The functions concerned with either of these tasks are grouped in two distinct layers to ease independent development in both areas. Additionally, the transport layer is supposed to be IP-based; thus, IP packets are the smallest denominator of all transport layers.

The ambient network concept adapts these tenets and assumes the presence of a layer to ensure basic connectivity between different networks, constituting diverse addressing domains, participating in the creation of the ambient networks federation. This layer will remain based on IP. The control functions already identified and isolated by all-IP networks are extended to
Ambient Networks extend All-IP networks by several innovations. The extensions build on the demand for enabling communication between different social and economical realms as identified by the Internet research community.

the ambient control space, which embraces a well-defined set of control functions required to guarantee cooperation between networks. The control space is thus no longer concerned only with the management of the transport functions within its own network domain, but is also responsible for the establishment of agreements with neighboring networks to provide end-to-end services through the setup of a service chain involving several networks.

**HOW AMBIENT NETWORKS EXTEND ALL-IP NETWORKS**

Ambient networks extend all-IP networks by several innovations. The extensions build on the demand to enable communication between different social and economic realms identified by the Internet research community. The three main innovations are network composition (beyond simple internetworking), enhanced mobility, and effective support for heterogeneity in networks.

*Network Composition* — One basic mechanism of our approach is the dynamic, instant composition of networks (Fig. 1 shows an example). Composability as required for ambient networking goes beyond what the Internet and mobile networks provide today. Internetworking shall happen not only at the level of basic addressing and routing; additional functions for incorporating higher layer support (e.g., content distribution or service control functions) are required. Ambient networks will deploy a universal framework for network composition, as an approach for building unified communication support out of the resources of individual networks, which can be specialized for particular types of access technologies or business models. Network composition in ambient networks has to function across operator and technology boundaries, provide a security framework, and be executable without user involvement. In addition, the execution of the composition process has to be rapid in order to follow fast topology changes as expected, for example, for mobile personal area networks (PANs).

Another example where instant network composition is relevant is the joining of operator networks, which today are based on explicit human-negotiated and -executed agreements and are therefore too slow and cumbersome to set up for, say, rapid service roaming. Rather, access, interconnectivity, and service level will be associated on the fly between two networks.

*Mobility* — Existing IP-based mobility solutions target either intradomain mobility within a static network architecture or roaming solutions across domain boundaries. Ambient networks focus on integrated mobility concepts applicable not only to both of the above scenarios, but also integrating localized communications (e.g., in PANs) and device-device interactions. In dynamically composed network architectures, mobility of user group clusters can support effective local communication. Furthermore, mobility must interact efficiently with the control interfaces needed to enable QoS and optimal routing and rerouting of individual multimedia flows. An ambient networks mobility solution will have to work well across business and administrative boundaries, which requires solutions for the security issues in interdomain operation.

*Heterogeneity* — Ambient networks will be based on a federation of multiple networks of different operators and technologies. On one hand, this leads to increased affordability of ubiquitous communication, as the user has full freedom to select technology and service offerings, and the investment needs for new networks are reduced. On the other hand, networks will have to integrate the capabilities of different technologies end to end for a seamless and secure solution for the user.

Ambient networks take a new approach to
embrace heterogeneity visible on different levels, such as link technologies, IP versions, media formats, and user contexts. Diversity of access links, especially of links provided by mobile networks, is supported by a generic link layer concept, which will efficiently enable the use of multiple existing and new air interfaces. Ambient networks also consider the implications of heterogeneous wireless systems on the overall network, especially the impact on end-to-end QoS and multimedia delivery. In particular, the novel concept of network composition will include negotiation between different networks regarding their capabilities (e.g., regarding QoS). Ambient networks provide an integrated framework for enhanced support of multimedia delivery in heterogeneous environments by embedding novel media flow routing and transport functionalities into the overall ambient network architecture.

**HOW TO BUILD AMBIENT NETWORKS**

The vision of ambient networks is that users and operators can jointly exploit available radio and network resources for a broad range of services. To enable this vision, the concept builds on a common set of control functions, dynamically (re-)configurable and universally available.

**DESIGN PRINCIPLES FOR AMBIENT NETWORKS**

To evolve today’s mobile networking into this vision, current systems need to be analyzed, especially from a user’s perspective as outlined above. Currently, the lack of commonly available and configurable control functions and dependence on network technologies and ownership are major obstacles to further rapid growth. Ambient networks take up the challenge of defining an essential set of universally available and usable control functions. To achieve this, a conceptual framework is defined, including the control functions necessary to achieve the required network capabilities. This framework is based on the following three principles.

**Principle #1: Ambient Networks Build upon Open Connectivity and Open Networking Functions** — One basically new way of defining networking in ambient networks is to remove architectural restrictions on who or what can connect to what. Compared to existing internetworking, the goal is to enable all networking services for connected networks instead of connected nodes (e.g., QoS or media delivery support). This is motivated by the observation that current node-centric designs fail for many scenarios including PANs, moving networks, or sensor networks, when connecting such networks to other networks. In general, we will always assume a network at the end of the communication flow. Hence, we talk about end environments rather than end nodes.

The ambient networks concept defines a set of support functions required to satisfy the business needs of the operators of such end environments, where operators can be commercial entities as well as end users. For such an end environment, capability offerings are broken down to their nucleus, which is the end-to-end relationship between control functions. The challenge is to provide suitable mechanisms to enable any such relationship regardless of whether the partners reside in an operator’s network or an end user’s terminal.

**Principle #2: Ambient Networks Are Based on Self-Composition and Self-Management** — While composing networks is an easy task when only packet forwarding is concerned, advanced end-to-end functions like QoS and security as well as mobility are currently very difficult to establish across network boundaries. Ambient networks treat network composition and reconfiguration in a self-managed way as a guiding design principle for the research work. The goal is to use network composition and self-management as basic, locally founded building blocks of a networking architecture. This includes self-composition across different business domains. These features will broaden the business case for the operator and enable fast introduction of new services in all connected networks.

**Principle #3: Ambient Networks Functions Can Be Added to Existing Networks** — In today’s networks, there is a large degree of homogeneity on the very basic connectivity/packet forwarding functions, but network control is distributed over multiple layers and specific to network technology, operator, and even implementation. As a starting point for ambient networks, connectivity and control level are logically separated. The control level, referred to as the ambient control space (discussed next), can enhance existing technologies with distinct control functions that are compatible across all domains of an ambient network.

An exemplary networking scenario of a composed ambient network is shown in Fig. 2. With the above design principles, ambient networks will provide end-to-end services for such a composition of heterogeneous networks.

This approach yields the following advantages.

**Network operators** can decide what level of support they want to give to users and business partners, based on flexible sets of network control and composition functions. Affordable and simple as well as service-rich networks can be built without restricting integration into global connectivity. Existing networks can be integrated, and furthermore, new technologies can be introduced in a coherent way.

**End users** select between different control functions on their devices or personal networks, which provides the feature-rich service environ-
ment they like to use. Innovation is nurtured by the opportunity to develop any service on top of the basic connectivity with as few barriers as possible because of the simple but extendable connectivity interfaces.

Service providers can easily address large numbers of users and deliver specific services to them without having to worry about their users’ access or network capabilities.

**THE AMBIENT CONTROL SPACE CONCEPT**

Following the three principles discussed above, the notion of the ambient control space is introduced to encompass all control functions in a certain network domain. The ambient control space together with a (possibly legacy) connectivity network is called an ambient network.

The ambient control space (Fig. 3) hosts a set of control functions that might be missing in certain instantiations of the control space. Examples of such functions include support for mobility or multi-access networks, as well as more abstract functions like the provisioning of context information.

The main characteristics of an ambient network are:

- It provides well defined control interfaces to other ambient networks and service platforms or applications.
- It provides (at least a subset of) the ambient control space functions.
- It can be dynamically composed of several other ambient networks to form a new ambient network.

Cooperating ambient networks could potentially belong to separate administrative or economic entities. Hence, ambient networks provide network services in a cooperative as well as competitive way. When ambient networks and their control functions are composed, care must be taken that each individual function controls the same resources as before: by composing two ambient networks, resources shall not become a common good but rather a good that can be traded.

**AMBIENT NETWORK INTERFACES**

Control spaces spanning several domains are required, as users will move among them and want to exploit the capabilities of as many domains as possible; they will assume their personal services to be at their disposal regardless of their location and physical connections utilized. The intrinsic complexity of domains might vary, and economic considerations will decide the feasibility of different approaches. To allow cooperation across different ambient networks, the ambient network interface (ANI) is provided.
It offers standardized means to connect the functions of an ambient control space with functions of another domain (Fig. 3). It also advertises the presence of control space functions in adjacent domains to allow for a dynamic set of control functions in each.

Accessing the services of an ambient network happens via the ambient service interface (ASI). Even in a composed ambient network, only a single homogeneous control space is visible to external entities (Fig. 3). An application or service, making use of the functions in the control space, will always find the same environment regardless of to which ASI it is currently connected. The mechanisms implemented in the ambient control space, and the ANI will ensure this.

A particular issue is to ensure scalability of these interfaces to be applicable to small personal networks up to large-scale networks; a related issue is the universality of the approach to cover all types of networks.

In summary, the ANI and ASI have the following properties:
- Similar functions regardless of the nature of the ambient network
- A simple plug&play connection between ambient networks
- Network reconfigurability
- A single interface to the outside
- Support moving networks

**Ambient Networks in Detail**

In this section two functions envisaged for ambient networks are presented in more detail. They have been selected as they present functions logically located at the upper and lower boundaries of the ambient networks scope. Several other key functions, such as QoS, self-management, context management, and security, are not detailed in this article.

**Ambient Media Delivery Overlay**

Media distribution and delivery is foreseen to become an important task within ambient networks. Multimedia services have different requirements when it comes to required media delivery support functions (e.g., conversational multimedia, download, messaging, streaming, multicast/broadcast). Synergies can only be expected from providing a common media delivery support layer to all kinds of multimedia services. However, this media delivery layer needs to fulfill all requirements posed by existing multimedia services and be extendable in order to cope with future service requirements. A particular challenge is the support of ambient multimedia services, integrating different end devices and various access technologies in a seamless way for the sake of increased service performance and end-user convenience.

In ambient networks the media delivery support layer described above is realized as an overlay network.

The ambient network logical nodes constitut-
A MediaPort can provide the required adaptation functions by itself or it manages and controls the required adaptation resources. The needs of commercial content providers need to be considered carefully, since they often would like to avoid uncontrolled transformation of their content.

Figure 5. The multiple link layer scenario compared to the generic link layer scenario.

The multiple link layer scenario compared to the generic link layer scenario.

Figure 5. The multiple link layer scenario compared to the generic link layer scenario.

The needs of commercial content providers need to be considered carefully, since they often would like to avoid uncontrolled transformation of their content.

The Generic Link Layer for Multiradio Access

For today’s mobile communication systems, where specific mobile networks use particular radio access technologies, the approach of highly optimized radio protocol layers has proven to be an efficient design choice. However, the heterogeneity of future ambient networks means that connectivity to mobile terminals will be provided via a multitude of radio access technologies. This will enable dynamic selection and configuration of different radio access technologies to allow cost-effective data transmission based on service requirements, link quality, and radio resource consumption. To account for this new networking paradigm, it is necessary to provide methods for cooperation of different radio access technologies and networks. Ambient networks will provide two such approaches. First, multiradio resource management functions allow coordinated allocation of heterogeneous types of radio resources. This will enable load balancing across different radio access technologies and increase the wireless system capacity. Second, a generic link layer (GLL) will perform seamless and efficient radio protocol reconfiguration for dynamic changes of radio access technologies. The GLL will exploit commonalities of radio access technologies and allow cooperation of radio access networks on the link layer. The GLL provides a toolbox of configurable link layer processing functions, in order to achieve universal link layer processing for a multitude of radio access technologies. For example, the generic functions comprise compression of higher layer protocol headers, segmentation of higher layer data, recovery of transmission errors by means of automatic repeat request (ARQ), and data scheduling. A GLL configuration for a particular radio access technology includes the selection of new metrics on which routing decisions should be made. Those metrics could, for instance, take into account characteristics of the media content, link characteristics, and user preferences for delivering media content in the best and most efficient way. Utilizing information from lower transport layers within the ambient media delivery network, the routing and adaptation decisions can be further optimized.
the appropriated link layer functions from the toolbox of functions, as well as determining suitable protocol parameters and timers. The configuration of the GLL toolbox has to be controlled by a configuration management function in case parameters cannot be determined autonomously by, say, measurements of link round-trip times. Still, the GLL will also allow the inclusion of radio specific functions for cases in which a generalization is impossible or inefficient. An example of a radio-specific function is a medium access scheme designed for the particular properties of a certain physical layer technology. A reconfiguration of the GLL due to a change of radio access technology will be seamless, without a loss of link layer state or buffered data. It is therefore required that the GLL reconfiguration be coordinated with mobility management functions and context transfer at intersystem handover be supported. As a result, the GLL provides compatible radio link layers for different radio access technologies, which allows the handover of link state information. This enables seamless data transmission even across access network boundaries (Fig. 5). A more concise overview of the GLL can be found in [11].

CONCLUSIONS

We have presented a new networking concept referred to as ambient networks, which aims at enabling the cooperation of heterogeneous networks belonging to different operator domains. It can be regarded as an extension and continuation of network development starting from the all-IP networking vision and the current trend in the Internet community to structure the Internet into smaller domains. It will enable the introduction of new networking technologies as well as new services in a flexible way. Cooperation between networks plays a key role in our concept, as communication endpoints are treated as networks as well. Additionally, we demand the possibility of instant network composition to allow rapid adaptation of network topology as required for moving networks. The ambient control space turning today’s networks into ambient networks will enable this.

The motivation for our concept has been discussed and its main features introduced. Two control space functions, media delivery and the generic link layer, have been presented in more detail.

The issues presented in this article are topics for research in the Ambient Networks project, which is part of the EU’s 6th Framework Program. In total 37 organizations from Europe, Canada, Australia, and Japan are involved in this Integrated Project, which will run from 2004 to 2005 in its first phase.

ACKNOWLEDGMENTS

The authors would like to thank Bengt Ahlgren, Robert Hancock, Toon Norp, Frank Reichert, and Jens Zander, whose contributions substantially shaped the vision of ambient networks.

REFERENCES


ADDITIONAL READING


BIOGRAPHIES

Norbert Niebert (Norbert.Niebert@ericsson.com) studied electrical engineering at the Technical University of Aachen, and received his M.Sc. degree in 1986 and his Ph.D. degree in 1991. He joined Ericsson Research in 1991 in the research area of UMTS. He was project manager of the RACE Mobilize project, 1992–1995. He was chair of the Intelligence in Services and Networks IS&N project line during that time. He was a founding member of the UMTS Forum. He has been group leader of the mobile multimedia networks research group at Ericsson Research in Germany since 1997. He was rapporteur of the WSI Think Tank and a founding member of the Wireless World Research Forum (WWRF). His current work interests are in network performance and new mobile networking architecture.

Andreas Schieder (Andreas.Schieder@ericsson.com) received his M.Sc. and Ph.D. degrees in electrical engineering from the Technical University of Aachen in 1995 and 2003, respectively. During his occupational career within Ericsson Research he has been involved in research on UMTS and GPRS performance evaluation. He has concentrated on support of real-time applications and actively contributed to the standardization of the 3G radio access network, GERAN. He is vice working group chair of the WWRF.

Henrik Abramowicz (Henrik.Abramowicz@ericsson.com) graduated with his M.Sc. degree in electrical engineering in 1976 from Lund Institute of Technology. He joined Ericsson that year, working on product management for telephony functions and data communications. 1981 he joined Philips in Sweden and became responsible for Teletex development at Philips, and contributed strongly to the Teletex standards within the International Telecommunications Union (ITU) as well as development of OSI. 1982 he returned to Ericsson as an expert in data communications and became responsible for coordination of the development of data communications within Ericsson Information Systems. During that period he contributed to the development of IEEE 802.3, 4, and 5 standards. In 1987 he became department manager for the systems and architecture group. He has since then worked in a managerial capacity in different positions. He contributed strongly to the establishment of ATM infrastructure for UMTS and was the chair of ATM Forum’s AAL2 group. In 1996 he initiated a business development project for wireless LAN within Ericsson Radio and became responsible for establishing a product unit 1997. He later became general manager for that unit that was also responsible for the development of HIPerLAN 2 standards and products. He is presently overall responsible for Ericsson’s Beyond 3G EU activities for the 6th
Co-operation between networks plays a key role in our concept, as communication end-points are treated as networks as well. Additionally, we demand the possibility for instant network composition to allow rapid adaptation of the network topology as required for moving networks.
In today's wired and wireless networks, the trends in networking technology very much point to a dominance of Internet technology with all its flavors. IP is the common lingua franca to enable the exchange of data across various networks. There is, however, an increasing divergence in the network control layer: different control environments are established to facilitate services like virtual private networks (VPNs), security, mobility, quality of service (QoS), network address translation (NAT), multicast, and so on.

NEC Network Laboratories Heidelberg. the 3G/UMTS packet service as well. This has allowed many European mobile network manufacturers to hold on to their strong market position when the 3G technology entered the market a few years ago. To maintain the competitive advantage of European manufacturers in mobile networks, the European government is strongly supporting collaborative research towards new open, IP-based mobile networks within its 6th Research and Development Framework program (FP6).}


Introduction In today’s wired and wireless networks, the Ambient Networks will deploy a universal framework for network composition, as an approach for building a unified communication support out of the resources of individual networks, which can be specialised for particular types of access technologies or business models.