A Reference Model for Context-Aware Mobile Services

Jacqueline Floch, Svein Hallsteinsen, Arne Lie and Hans I. Myrhaug

SINTEF Telecom and Informatics, N-7465 Trondheim, Norway
E-mail: {Jacqueline.Floch, Svein.Hallsteinsen, Arne.Lie, Hans.I.Myrhaug}@informatics.sintef.no

Abstract

The integration of communication and computation technologies, the availability of pocket mobile computers, and the widespread penetration of mobile radio access networks will enable a range of new context-aware mobile services to be offered to the users. Research on methods and technologies for the development of such services requires the involvement of experts in several technological domains such as software engineers and radio experts. In order to enable the development of integrated solutions, the work within these domains should be co-ordinated. Working across technological areas is often difficult. A common model and terminology are required in order to support the unified understanding and formulation of the research issues, and the description of problems and solutions. No overall conceptual framework for mobility exists today. Rather different aspects of mobility are addressed in various frameworks that are developed for different domains, and focus is on different types of system and viewpoints. In a similar way that RM-ODP has been a valuable framework for distributed processing, we believe that a reference model addressing mobility related issues is needed for enabling a unified way of working and integrating results. This paper is intended as a contribution to such a reference model.

Keywords: Mobility, Context, Mobility Reference Model

1 Introduction

With the recent advances in mobile computer technology and the penetration of wireless networks, the nature of the services proposed to the users is moving towards mobile and context-aware services. With the aim to propose methods and technologies for the development of such services and the infrastructure to support them, SINTEF Telecom and Informatics has started LAMA, a strategic research programme on technologies for enabling mobility [LAMA]. The particularity of this programme is that experts from different technological fields are working together in order to enable the development of integrated solutions. Radio experts are working with reliable high-capacity transmission on new wireless mobile radio network infrastructures. Software engineers focus on architectures that support the adaptivity of applications to changes in the means of communication, the available processing power and the user environment. User interface experts work on the adaptation of user interfaces to the changing context of the mobile user.

As the participants in LAMA have different scientific background and are working in distinct technological areas, we believe that it is important to develop a common conceptual framework or reference model for enabling the unified understanding and formulation of mobility-related issues, and describing problems and solutions. Such a model provides a unified way of integrating results. This paper presents parts of the LAMA Reference Model [Flo01]. The model defines central concepts and proposes a common vocabulary of terms for talking about these concepts. The model does not restrict to mobility specific concepts, but also addresses general concepts such as services and networks. As there is no common understanding of these general concepts, we found it necessary to include them in our model. Note that the Reference Model does not aim to specify concrete mechanisms that support the implementation of context-aware mobile services.

We believe that a reference model for context-aware mobile services is valuable not only within LAMA, but also for other system developers, software engineers and telecommunication engineers working with mobility- and context- related issues. Similarly to the widely adopted RM-ODP concepts for distributed processing [ODP], we believe that common concepts for mobility should be of importance. No overall conceptual framework for mobility exists today. Rather different mobility aspects are addressed in various frameworks that are developed for different domains, and focus is on different system viewpoints (e.g. TINA, CORBA, Jini). The LAMA Reference Model is inspired by these existing frameworks, and complements them by seeking to integrate the heterogeneous views of the LAMA participants on mobility.

In this paper, we first introduce informally the concept of mobility. Then we present fundamental concepts related to information and communication systems, and to services offered by these systems. We refine the concept of mobility by classifying mobility according to the types of the entities that move around, and by considering the impact of mobility on service and service availability. The concept of context is then introduced. While context has been discussed within interactive computing [Dey99], context is not yet exploited by mobile communication services. One challenge of mobile services is to cope with the frequent context changes of the mobile users. Finally, we present behavioural aspects and commonly used mobility-support services. All class diagrams that describe entities and relationships are expressed using UML [UML].

2 A glance at mobility[1]

Mobility is about moving around. Different kinds of entities such as persons, computers, information or programs may move around. The mobile entities may communicate and access services ubiquitously or, in the contrary, services may be adapted to various characteristics of the mobility space such as the physical location.

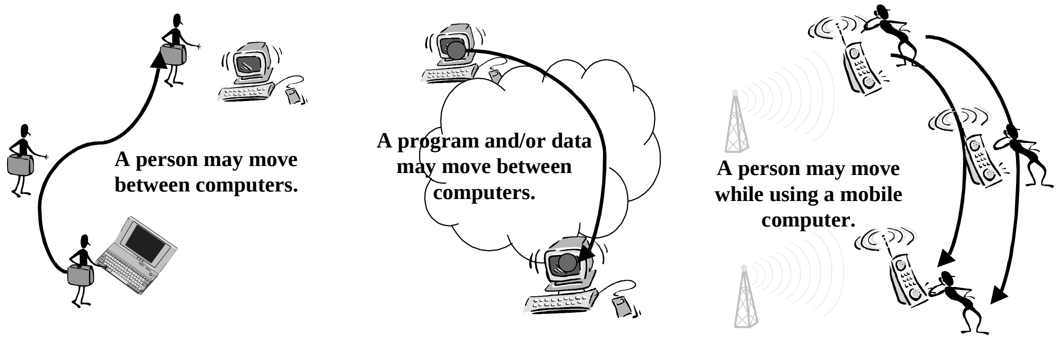

Figure 1 - About moving around.

In the past, user movement has often implied interruption of service. With the advent of pocket size computers and wireless communication, services can be accessed without interruption while the entity using the services is moving. In a mobile radio network, the communication services are provided without interruption under normal conditions.

[1] Note that we do not use a precise terminology in this section. Terms will be precisely defined later in this paper.

In our work, we focus on persons that are moving around carrying small personal mobile computers. These computers may be connected dynamically to various devices such as sensors, printers and information servers forming ad-hoc networks. They are usually also connected to other networks such as local area networks or mobile radio networks that support access to a large global set of services.

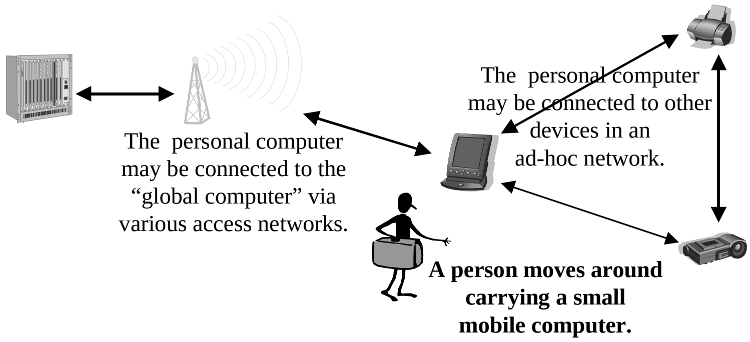

Figure 2 - Personal mobile computers support access to a wide range of services.

3 Fundamental concepts

This section presents fundamental concepts related to information and communication systems, and to services offered by these systems. Their understanding is necessary before further defining specific mobility concepts. These fundamental concepts may seem well-known. However, we observe that they are usually used with different meanings. For example, the term service has several definitions in the literature: services have been presented in various models developed for different domains and viewpoints. The term service is also often used imprecisely or without being defined. For example, the Internet documents use the concept of end-to-end quality of service without defining service. In [Jini], service is first defined in a rather general manner, but it turns out that Jini services are restricted to the computational viewpoint.

We believe that it is important to define concepts with reference to a well-defined viewpoint. As [TINA-O], we distinguish between the enterprise and computational viewpoints. In the enterprise viewpoint, concepts are related to the purpose, scope and policies for the system. In the computational viewpoint, concepts are related to functional system decomposition and to system distribution.

3.1 Networks and services - enterprise viewpoint

A *network* is a computation and communication system. Networks offer *services* to human beings that we call *users*[2]. A service is defined as some set of capabilities provided to the users. *Server* and *client* are defined as roles played in an association to a service. An entity playing the server role provides services, while an entity playing the client role consumes services.

[TINA-S] considers different classes of services: *information services* that handle information resources, *telecommunication services* that support communication between users (including information transport and establishment of connections), *management services* that support the FCAPS functionalities (fault, configuration, accounting,

---

[2] Our definition differs from the definition usually adopted in the telecommunication domain [TINA-G], where the term "user" may designate both a human user and an application.

performance and security), and the service lifecycle management. This service classification is given as a guideline. Actually many services combine capabilities for information processing, telecommunication and management.

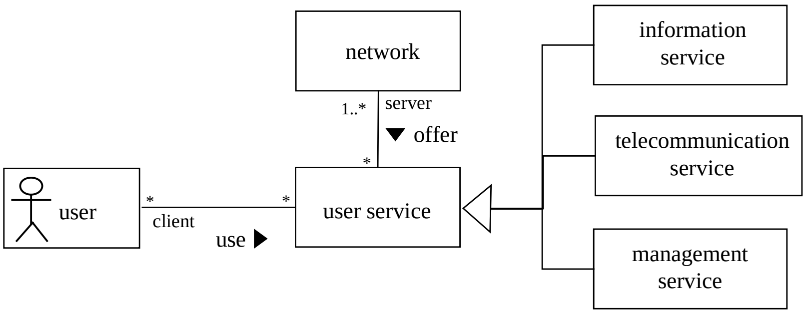

Figure 3 - Networks and services.

Our Reference Model does not currently discuss business issues. The TINA business model would provide a good starting point [TINA-B]. This model defines the interactions between business administrative domains in terms of contracts, business roles (e.g. consumer, retailer, broker, service provider) and relationships.

3.2 Network structure - computational viewpoint

In the computational viewpoint, we decompose networks into *nodes* and *links*. Nodes run *software components*. Links are communication paths that provide connectivity between nodes. Software components interact by offering *computational services* to other components. A computational service is defined as a set of capabilities provided by a component (e.g. a computational object) to other components. We do not define any hardware component. We assume that a software component such a driver or a wrapper is always associated with a hardware component. Services provided by the hardware components are accessible through that software.

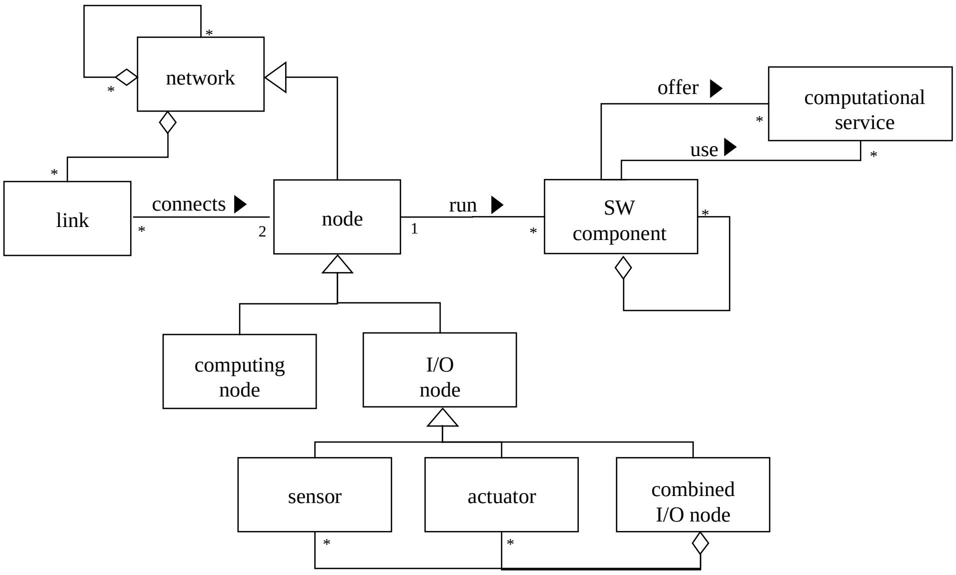

Figure 4 - Network structure.

In the telecommunication domain, a network is usually a part of a telecommunication system that provides capabilities for the transport of information [TINA-G]. Our approach uses a wider definition for network. *We use the term network to designate* *both transport networks or any other set of interconnected devices*. For us, the user equipment consisting of interconnected computers and I/O nodes is also seen as a network.

A node is itself a network. This recursive definition enables us to look at a compound network either as a single node/network or as a set of nodes/networks that are interconnected. The interconnection of networks can be derived from the model in Figure 4. Networks are usually partitioned or decomposed into sub-networks. We use the single term "network", also for designating sub-networks. Links, computing nodes and I/O nodes are further described in [Flo01].

3.3 Components and computational services

We distinguish between different types of software components that provide different types of computational services. *Facade components* are the components the users interact with when using a service provided by a network, e.g. a dialog window on a computer. *Application domain components* support some function within a particular domain. A domain may be related to a market segment or to a family of systems that handle the same types of phenomena. For example, we may consider application domain components within the accounting or medical imaging domains. *Computational infrastructure components* provide general purpose computational services that are fundamental for developing distributed applications. Trading, naming, notification and transaction services are such services defined in [CORBA-S].

3.4 Network types

We distinguish between different types of networks. Networks may be classified according to the size of the area they span over (e.g. local vs. global networks) or the type of function they provide (e.g. information transport between networks or interconnection of small I/O nodes). There may be overlap between two or more types (e.g. between a personal network and an ad-hoc network). It is difficult to model such overlap as the overlap may vary according to the way we look at a network.

Personal networks

Services are delivered to the users on nodes (or networks) that we call *personal networks*. A personal network consists in one or several I/O nodes and at least a computing node. These nodes are connected either over a radio link (e.g. using Bluetooth technology) or a fixed link to. An illustration is given in Figure 5.

A personal network supports access to both local services i.e. services provided by the nodes that are part of the personal network, and services provided by other networks that the personal network is connected to.

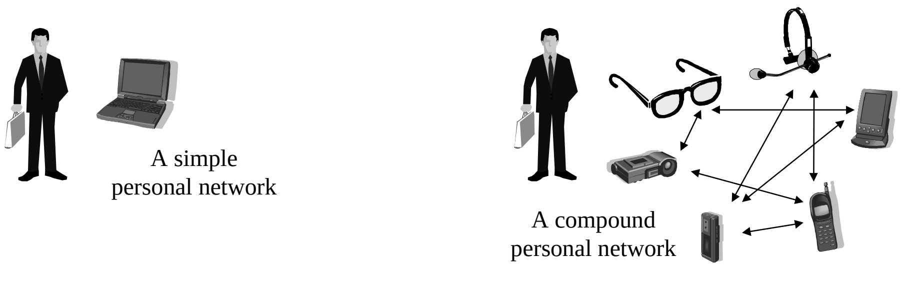

Figure 5 - A personal network may have a more or less complex structure.

Access and core networks

When considering telecommunication transport networks, it is usual to distinguish between *access* and *core* networks. Access networks provide an entry to a transport network and an interface (or access) to a wider set of networks and services. Examples of access networks are GSM, Bluetooth, Cable TV networks. Core networks are larger networks with high capacity that form the backbone infrastructure of a telecommunication network. *Our focus is on mobile radio access networks.*

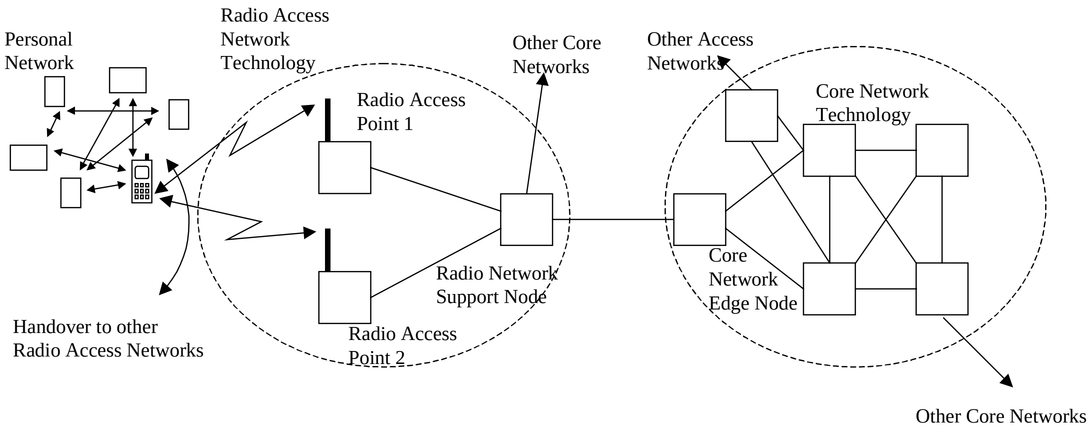

Figure 6 - Personal, access and core networks.

Server networks

Similarly to personal networks, we define *server networks*. A server network is a common resource in a network that provides services to various users. A server node is always part of a network such as a local area network or a telecommunication transport network.

Ad-hoc networks

An ad-hoc network is a network that supports temporary "plug-in" connections of various computing and I/O nodes; the connections are usually wireless. An ad-hoc network offers support that facilitates the connection to the network and the easy access to the services offered by the network. The plugged-in nodes usually connect to the network for a short duration. Ad-hoc networks do not necessarily involve mobile I/O nodes, but become more actual in a mobile context where mobile computing nodes need to connect temporary I/O nodes when entering a new location.

3.5 Sessions

The access and usage to a service involve several components and collaboration between them. A *session* is a temporary relationship among a group of software components that are assigned to collectively deliver some services for a period of time [TINA-S]. A session can be described by the states of the components involved in the session and of the relations between these components. Sessions may be suspended and resumed. The concept of session is useful for dealing with the changing context of mobile entities, and elaborating solutions for fault management.

4 What is mobility?

In this section, we define the concept of mobility and look at the impact of mobility on services. Mobility is classified according to the types of the mobile entities.

4.1 Mobile and stationary entities

Not all entities are able to move. We distinguish between *mobile* entities that can move and *stationary* entities that cannot move. For example, a light mobile phone may move while a heavy PC remains stationary within an office. Usually software components closely related to the hardware infrastructure (such as operating systems and drivers) cannot move. We introduce two UML stereotypes of class: "mobile" and "stationary" in order to represent mobile and stationary entities.

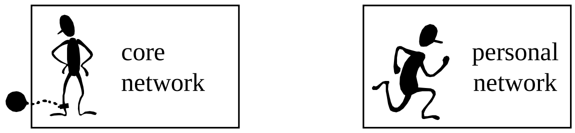

Figure 7 - The class stereotypes "stationary" and "mobile".

We also introduce extensions to UML for showing that association and class can be instantiated at different times. The instantiation order is indicated by a natural number. Examples are given in Figure 8 and Figure 9.

4.2 Types of mobility

Personal mobility enables a person (i.e. a user) to use services that are personalised with their preferences and identity ubiquitously, independently of both physical location and specific user equipment [TINA-S]. This definition encompasses both the mobility of a person between terminals and the mobility of a person that moves around with a personal terminal. *In our work we assume that the access to a service and its usage are dependent on the user context and preferences*. Moving between terminals is illustrated by the object diagram of Figure 8.

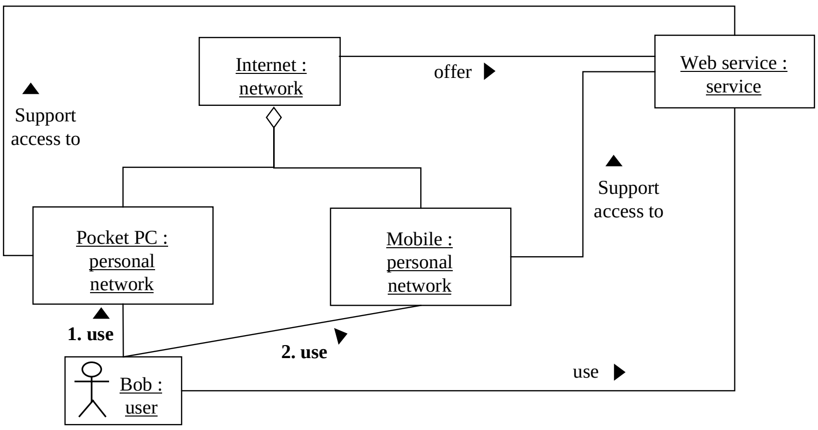

Figure 8 - Using UML object diagram for representing personal mobility.

Personal network mobility is the ability of a personal network to change location. Some personal networks continue to support access to services while moving, some do not. A traditional view of personal network mobility encompasses seamless access to services i.e. the ability to access services independently of the location of the personal network [UMTS]. In our work, we distinguish between services provided independently of the personal network location from those that depend of the terminal location and other context parameters.

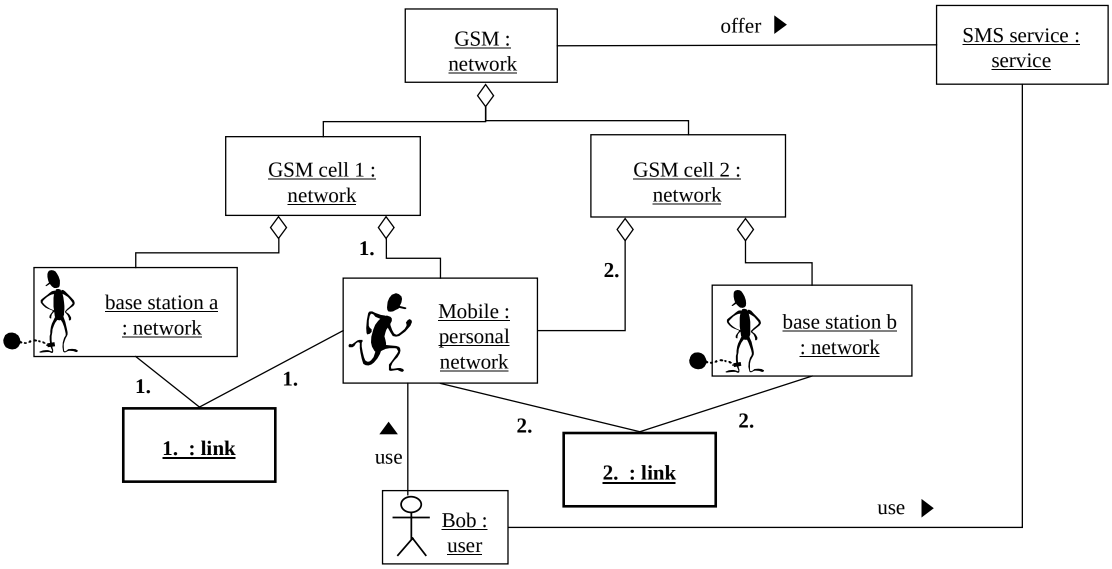

Figure 9 - Using UML object diagram for personal network mobility.

Software component mobility: A software component may be relocated from one computing node to another. A program or some database content may be moved. A hardware component and the software associated to this component may be unplugged and plugged at some new location.

Session mobility: Session mobility is motivated by the mobility types described above. It ensures that active sessions are not undesirably disrupted while terminals, persons or applications are moving or being relocated. Sessions may have to be suspended momentary and then resumed in order to enable mobility. It may be necessary to adapt the services delivered in the session to the new environment. Our definition covers both the mobility of a user session and the mobility of a radio communication session.

4.3 Availability of services

[IP-2] distinguishes between "portable computing and networking" and "mobile computing and networking" where mobile computing activities are or are not interrupted. In a slightly similar way, we distinguish between services that are continuously available and those that are not. The type of availability supported depends on the characteristics of the transport protocols and other context parameters.

Continuous availability of services enables services to be provided without interruption while the entity accessing the services moves and changes context. Continuous availability of voice and data services is supported in cellular radio networks under normal conditions.

Discrete availability of services enables services to be provided within certain areas, contexts and for certain access points. For example, some services may be available at home and in the office, but not while moving from home to the office. Session mobility is required in order to ensure that sessions are suspended in well-defined states and resumed correctly while the personal network or user are being relocated.

4.4 Mobility and services

Service and service provision may depend on the user location, the service usage context and the user preferences.

Mobility-insensitive services are services that do not depend on the location, the movement of the client user or the context changes following the user mobility.

Mobility-based services are services that make use of some characteristics or context parameters that change with mobility (e.g. information about location).

Mobility-support services are services that enable systems to support mobility. Several mobility-support services are described in Section 6.

5 Context

The use of context is becoming important in interactive computing, but still context is a poorly utilised source of information [Dey99]. Context information can be used to facilitate the communication in human-computer interaction. However the concept of context is not yet well understood or defined, and there exists no commonly accepted architecture that supports the management of context information. Context information often changes for the mobile user. One challenge of mobile services is to cope with the frequent context changes of the mobile users and exploit them.

We differentiate between the *user context*, which covers the context elements that are closely related to the user and the space around the user from the *network context* which covers the context elements related to computing and network equipment. For example, the user mood and role, and the user physical location are elements of the user context. The available network capacity and the accessible shared I/O devices are elements of the network context. While the user context information is usually collected by specialised sensors or with the help of the user, the network context information can be collected by analysing the state of the network and information already available in the network for other purposes, such as the radio signal level.

When entities move around, context changes occur more frequently than if they were motionless. The location of a mobile user changes, the persons and other entities surrounding the mobile user leading to entities being added or removed from the user context. Location is currently exploited in mobile communication and computing. However most mobile applications have hidden location and location changes rather than exploiting it [Schmi98]. Whether mobility and context changes should be hidden or exploiting depend on the type of the services proposed.

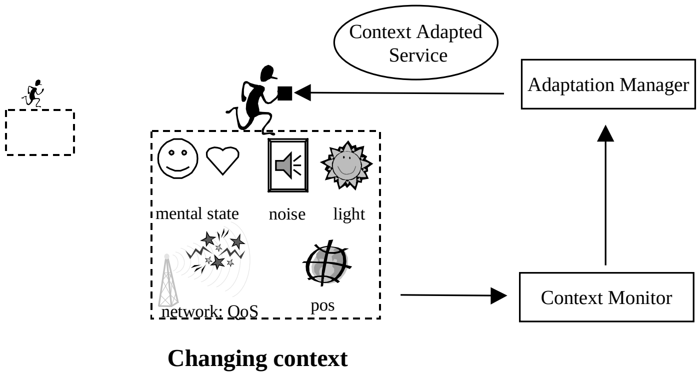

Figure 10 – Service adaptation to context changes.

A classification of the user context provides help to the application developer for identifying relevant context information. An initial classification is identified in [Flo01].

The *context history* contains all the past user contexts. A *context trace* is a subset of user contexts in the context history that may be exploited in order to predict possible future contexts [Rahl01].

6 Mobility-support services

This section briefly presents some general services for service discovery and management. Further details can be found in [Flo01].

6.1 Service trading

Mechanisms are needed in order to provide components with support for retrieving and accessing computational services. Advertising a capability or offering a service are called *export*. Matching against needs or discovering services are called *import* [CORBA-S]. Export and import facilitate dynamic discovery of services. *Trading* is the process of exporting and importing services [TINA-G].

6.2 Service naming

It is usual to give names to services. A name is usually a word (it may be a phrase) by which an entity can be distinguished from other entities. A name is valid within a limited area. Such area is called a naming context in [CORBA-S] and a naming domain in [TINA-G]. Name *binding* is the association of a name with a particular entity. To *resolve* a name is to determine or isolate the entity associated with that name. A *naming service* is a service that enables to locate entities from its name.

6.3 Service subscription

A *subscription* is an agreement between the user of service and the provider of the service. A subscription gives the right of usage. It may include rules for billing, usage restrictions and user preferences. Subscriptions are central in traditional telecommunication networks. The deregulation of the telecommunication network and the Internet with open service availability create needs for more flexible agreement models for service access and usage.

6.4 Service leasing

Service leasing is introduced in [Jini]. A lease is a grant of guaranteed access over a time period. A lease is negotiated between the user of a service and the service provider. Leasing is a convenient feature in networks where users are frequently appearing and leaving. Leasing also facilitate the de-installation of services. Similarly a registration lifetime is introduced in mobile IP for restricting the availability of a mobile IP node.

6.5 Mobility management

The main purpose of *mobility management* is to keep track of track of the location of the mobile users, enabling the mobile users to be accessed by other users. A usual approach to mobility management is based on the concept of *home* network where the mobile user or terminal can register his/its movements. This approach is both used in radio mobile networks (e.g. GSM) and Mobile IP.

7 Conclusion

In a similar way that RM-ODP has been a valuable framework for distributed processing, we believe that a reference model for context-aware mobile services is needed for enabling the unified understanding of mobility-related issues. This paper has presented a contribution to such a reference model. Our model uses concepts defined in existing frameworks and complements them with the heterogeneous views of experts within various technological domains. Integrating these views was in no way an easy task: we experienced how difficult it is to achieve understanding across technological areas. This illustrates the importance of our work.

Our reference model includes definitions related to fundamental concepts related to systems and services and a description of mobility and context. Most concepts have been modelled using the UML notation. In order to model the dynamics of mobile entities formally, we had to introduce extensions to UML.

We do not currently consider business issues. Also, we do not discuss security, dependability and performance. We have defined - but not in detail - some properties of the entities defined in the model. The properties are represented by attributes. We will refine the definition of these properties during the validation of the reference model by the participants in the LAMA programme.

References

[Bluetooth] The official Bluetooth Website. Information available at http://www.bluetooth.com/

[CORBA] OMG. The Common Object Request Broker: Architecture and Specification. CORBA V2.2.

[CORBA-S] OMG. CORBA Services: Common Object Services Specification. CORBA V2.2.

[Dey99] Dey A. K and al. An Architecture to Support Context-Aware Applications. *GVU Technical Report GIT-GVU-99-23, College of Computing, Georgia Institute of Technology*.

[Flo01] Floch J., Hallsteinsen S., Lie A. and Myrhaug H. The LAMA Reference Model. *SINTEF report*. January 2001.

[ISDN] ITU-T. General Aspects of Services in ISDN, Rec. I.120, 1988.

[Jini] Jini Connection Technology. Information available at http://www.sun.com.jini/

[IP-1] IETF - Mobile IP Working Group. IP Mobility Support. RFC 2002, October 1996.

[IP-2] C. Perkins, "Mobile IP," IEEE Comm., Vol. 35, No. 5, 1997, pp. 84-99.

[LAMA] A strategic research programme on technologies for enabling mobility. Information available at http://www.informatics.sintef.no/projects/LAMA/

[ODP] ITU-T. Information technology - Open distributed processing - Reference Model: Overview, Rec. X.901 (08/97).

[Rahl01] Rahlff, O. W. and al.. Using Personal Traces in Context Space: Towards Context Trace Technology. Springer's Personal and Ubiquitous Computing, Vol. 5, Number 1, p. 50-53, 2001.

[Schilit94] Bill N. Schilit, Norman I. Adams, and Roy Want. Context-Aware Computing Applications. In *Proc. of the Workshop on Mobile Computing Systems and Applications*. December 1994.

[Schmi98] Schmidt A. and al. There is more to Context than Location. In *Proc. of the International Workshop on Interactive Applications of Mobile Computing.* November 1998.

[TINA-B] TINA-C. Business Model and Reference Points. Version 4.0, May 1997.

[TINA-G] TINA-C. Glossary of terms. Version 2.0, January 1997.

[TINA-O] TINA-C. Overall Concepts and Principles of TINA. Version 1.0, February 1995.

[TINA-S] TINA-C. Service Architecture. Version 5.0, June 1997.

[UML] OMG Unified Modeling Language Specification. Available at http://www.omg.org/

[UMTS] ETSI. Framework for services to be supported by the Universal Mobile Telecommunication System (UMTS). Draft UMTS DTR/SMG-05201, July 1995.

[UPT] ITU-T. Universal Personal Telecommunication, Rec. F.850.