ON FUNDAMENTAL CONCEPT OF MOBILITY FOR MOBILE COMMUNICATIONS

Jun-Zhao Sun and Jaakko Sauvola

MediaTeam, Machine Vision and Media Processing Unit, Infotech Oulu
P.O.Box 4500, 4SOINFO, FIN-90014 University of Oulu, Finland
Email: junzhao.sun / jaakko.sauvola @ee.oulu.fi

Abstract – The tremendous demands from social market are pushing the booming development of mobile communi-cations faster than ever before, leading to plenty of new advanced techniques emerging. Mobile communi-cations are changing people's life style in many ways. Behind the scenario, the fantastic characteristic that makes this reality is mobility. This paper studies the basic concepts of mobility for mobile communications, with the attempts of trying to answer the fundamental questions on mobility. A conceptual discussion is made on mobility in the contexts of both computation and communication leading to the illustration of mobile computing. The effects of mobility on both architectures and protocols of networks and communications are analysed. The concept and operations of mobility management for mobile communications are also introduced. New challenges arise in future mobile communi-cations systems with the diversity as the key feature, which lead to the definitions of classified mobility according to different granularities.

Keywords – Mobility, Mobile Communications, Wireless Networks, Mobility Management.

I. INTRODUCTION

Mobility is human's nature. In the field of computing and communication technologies, to be able to communicate with other persons and access and process information simultaneously while moving has been as a long expectation that causes great deal of efforts having been made to turn the fancy into fact. The following advances in different technical areas provide the possibility of realizing the imagination, including:

1) The advances in VLSI, antenna, and battery technologies, which make small and light portable devices like laptop, personal digital assistants (PDAs), and cellular phone becoming more and more popular.

2) The advances in wireless communications theory, which make miscellaneous wireless networks with different air interfaces (e.g. TDMA, CDMA, FDMA, etc.) and wired infrastructures (e.g. Internet, PLMN, ATM, etc.) available.

3) The advances in software technology, e.g. software engineering, language technology, distributed computing, modern database, etc., which make various mobile services with effective supports facilitate human's work and life.

Mobile communications will continue to greatly change the way of people's life. The tremendous demands from social market are pushing the booming development of mobile communications faster than ever before, leading to plenty of new advanced techniques emerging. Various mobile devices, wider transmission bandwidth, manifold wireless and wired networks, and more powerful appliances' processing capability, together with advances in computing technology have brought more and more miscellaneous services to be delivered with more excellent quality.

Next generation mobile systems need the support of all the advances on new theories, algorithms, architectures, standards, and protocols. In the near future, more and more Internet based services like web service can be smoothly accessed with various mobile devices through the wide deployed wireless networks. At present, 3G mobile communications systems are just at the beginning to be deployed for multimedia data applications, while research on the fourth-generation (4G) mobile communications has begun to pave the way for the future [1, 2]. The mobile personal telecommunications and wireless computer networks are converging in the coming new generation of mobile communications. Future mobile communications systems evolve with the trend of global connectivity through the internetworking and interoperability of heterogeneous wireless networks. Roaming in such network architectures is a very complex situation and it causes many new problems. The requirement of smooth and adaptive delivery of real-time and multimedia applications makes the design of mobility management scheme more severe a challenge.

The goal of this paper is to study the basic concepts of mobility, with the endeavour of trying to find the answers to the fundamental questions on mobility for mobile communications. A conceptual discussion on mobility is presented in terms of both computation and communication and then the mobile computing is illustrated accordingly. the effects of mobility on the architectures and protocols of networks and communications are detailed analysed. The main features of the mobility management for mobile communications are also introduced. Diversity is the key feature of the future mobile communications systems, which lead to the definitions of mobility of various granularities.

This paper is organized as follows. Section 2 provides a conceptual discussion on mobility in computing technology. In Section 3 the impacts of mobility to both the architectures and the protocols of mobile communications are analysed in

detail, and then the concept and operations of mobility management mechanism are introduced. In Section 4, we present the diversity of mobile systems as the main feature of the future mobile communications, and then classified mobility according to different granularities. Finally, Section 5 concludes the paper.

II. MOBILITY CONCEPT

Mobility is the characteristic of an object that can be mobile. In the field of computing technology the mobile object can be in both computations and communications, according to which two new paradigms are incurred as mobile computations and mobile communications by extending the features of the objects in these two areas with mobility. The two paradigms then act as the basic components to construct the new research field — mobile computing. This extension is illustrated in Fig. 1. It should be mentioned that computation and communication are always interdependent instead of independent. Mobile computation must base on the support of wireless or wired networks at the same time itself forms the basic techniques for mobile communications.

In more detail, many mobile objects can be distinguished in the field of mobile computing. For mobile computation, objects that can be of mobility are usually some logical computing entities (code, data, or state), including [3]:

1) Mobile process, also know as process migration [4], is a concept in the area of OS. A process is the abstraction of a running application that consists of the code, data and OS state, which can be transferred between systems. Load balancing, fault resilience, eased system administration, and data access locality are the main goals of mobile process.

2) Mobile agent is one of the most popular types of distributed and mobile computing environments [5]. It extends the concept of software object with the attributes and capabilities of mobility, reactivity, autonomy, and collaboration, which can carry both code and data and the thread of control. The main goal of mobile agent is to improve performance and reliability.

As to mobile communications, mobile objects are mostly physical components and can span all the path of service delivery. Mobility scenarios include:

1) Service mobility, means that a personalized service available to the user with one mobile device in one network can still be accessible by another mobile device and/or in another network of different region or operator and operate in the new context.

2) Network mobility, refers to the wireless networks that support the connection to mobile devices. Some wireless connections may be based on an infrastructureless architecture—a collection of wireless nodes can dynamically form a network without using any pre-existing fixed network infrastructure, also know as mobile ad hoc networks, in which the networks are physically “mobile”.

3) Terminal mobility, is the ability of a user device that can roam within a network or between networks with on-going or following communications still reachable. Many devices now are portable, e.g. laptop and cellular phone.

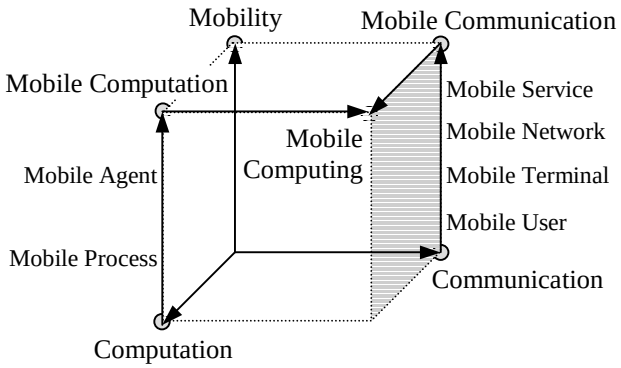


Fig. 1. Mobility and mobile computing

4) User mobility, means that end-users can access personal services regardless of moving to any network or using any terminal, through unique user identification like a universal personal telecommunication (UPT) [6] number.

This paper focuses on the concept of mobility for mobile communications. In more particular, we mainly study such a scenario in which end users are moving with a mobile device, since this is the most common use case and encompass other scenarios.

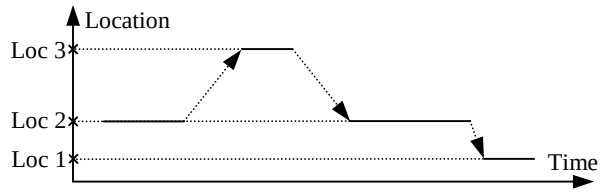
III. MOBILITY FOR MOBILE COMMUNICATIONS

Mobility affects mobile communications on all the components, including devices, networks, and services. To a mobile device, besides the physical requirements like weight, size, power, display, and shape, there still exist other functional requirements e.g. different user interfaces suitable to mobility scenario and the computing and communication capabilities distribution. To a service for mobile case, the most important effect is the requirement on adaptation in which a mobile service should be adaptive to different transmission links, different user mobile devices, and different using contexts. In particular here, we focus on the impacts of mobility on both the architectures and the protocols of networks.

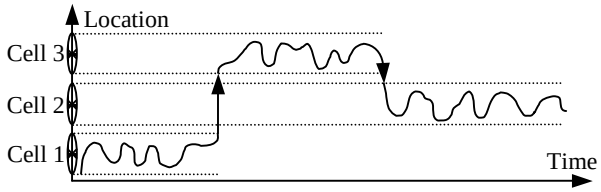
A. Mobility effects to network architectures

For network architectures different mobility modes can be distinguished resulting in different types of network architectures and communication usages. The mobility modes can be divided into three main classes according to the different spatial-temporal relations, as curves illustrated in Fig. 2, including:

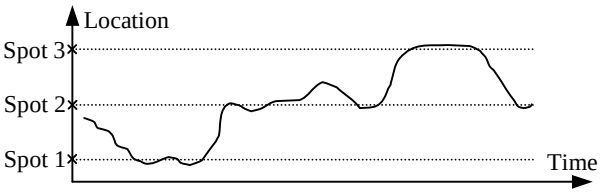
1) Nomadic or portable communications, as shown in Fig. 2(a), in which no network connection is needed during the movement and a new connection will be re-established only after the mobile node has arrived at its new location. Mobility can be either transparent or aware to other nodes. In this scenario only location management is significant and no handoff management is needed. Moreover, portable



(a) Nomadic communications



(b) Cellular communications



(c) Pervasive communications

Fig. 2. Communications of three mobility modes

communications are not necessarily based on wireless networks. This is a quasi-mobile communication mode.

2) Cellular communications, see Fig. 2(b), in which the wireless network is organized as a cellular structure in order to enable frequency reuse. This is also traditionally known as mobile communications. Each cell encompasses a certain distance and coverage. Continuous connectivity should be provided when an on-served mobile node is moving from one cell into another (maybe either neighbouring or overlapping cell). Both location management and handoff management are necessary in the scenario, while handoff management is only invoked when on-serving movement out of a certain area.

3) Pervasive communications, see Fig. 2(c), in which the communications between mobile nodes are ubiquitous and even invisible. The scenario is mostly based on a dynamic on-the-fly set-up without using any pre-existing network infrastructure, known as mobile ad hoc networking. It is an autonomous system in which mobile hosts connected by wireless links are free to move randomly and often act as routers at the same time. Location management scheme in ad hoc networks is involved into the routing strategies and mainly used in the cluster-based hierarchical multi-hop ad hoc networks, while handoff management is only treated with fast location management.

B. Mobility effects to protocol stack

The feature of mobility also affects the whole protocol stack, from the physical, data link, and network layers up to the transport and application layers.

1) At the physical layer, mobility influences are remarkable since most mobile communications are based on wireless media like radio. A wireless channel varies with most mobility factors e.g. velocity, direction, place (outdoor or indoor), etc. Resource reuse and avoiding interference are two important problems at the physical layer.

2) At the data link layer, mobility based on wireless networks brings problems of bandwidth, reliability, and security, for which compression, encryption, and error correction techniques are needed. Other problems include fixed or dynamic channel allocation algorithms, collision detection and avoidance measures, QoS resource management, etc.

3) At the network layer, mobility of mobile nodes means that new routing algorithms are needed in order to change the routing of packets destined for a moving node to its new point of attachment in networks. How to track a node's movement and how to keep the moving node's connectivity are two basic issues at the network layer. This in turn forms the two main operations of mobility management.

4) At the transport layer, a end-to-end connection may mix wired and wireless links. This makes congestion control a complex task due to the different characteristics of wired and wireless networks, since packet loss is caused mainly by high error rates and handoff in wireless networks instead of because of congestion—the situation on wired links. Retransmission mechanism based on increasing interval may lead to an unnecessary drop in the date rate. Function distribution between the transport and the data link layer is a new problem caused by mobility.

5) At the middleware and application layer, mobility brings new requirements on middleware supports. Examples include service discovery schemes, QoS management, and environment autoconfiguration. Device-aware applications are important to adapt to different types of user devices, while connection-aware applications are needed to adapt to the changing conditions of network connectivity. Besides these challenges, mobility also brings new opportunities to applications. Context-aware applications are possibly based on the measures for sensing miscellaneous context information of mobile end users.

C. Mobility management for mobile communications

Mobility management is the essential technology that supports roaming users with mobile terminals to enjoy their services through wireless networks when they are moving into a new service area. From the viewpoint of functionality, mobility management enables communication networks to track and locate roaming terminals in order to deliver data packets to the new destination and maintain connections with terminals moving into new areas [7]. According to the concept above, mobility management mainly contains two distinct but related components: location management and handoff management.

Location management is to locate roaming terminals in order to deliver data packets despite that the locations of them may change from time to time. Operations of location management include location registration and location paging. Addressing, location database structure, location

update time, and paging scheme are the key research issues for location management. Handoff management is to control the change of a mobile node's attachment point to network in order to maintain connection with the moving node during active data transmission. Operations of handoff management include handoff triggering, connection re-establishing, and packet routing. Fast and seamless handoff schemes and efficient routing algorithms are two important problems for handoff management. Note that many issues in location management are not protocol dependent, while handoff algorithms are much related to the network protocols of e.g. routing and resource management. A survey on different schemes designed or studied for mobility management for PLMN (PCS), wireless ATM, wireless Internet (Mobile IP), and satellite networks can be found in [8].

There are still many other aspects concerning the network management of mobility, e.g. mobile QoS and resource management, mobile security and privacy, billing, device power management, etc.

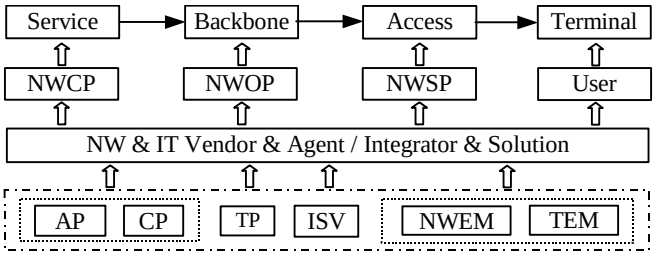
IV. MOBILITY IN FUTURE MOBILE SYSTEMS

A. Features of future mobile systems

The key feature of the future mobile communications systems is the seamless integration of terminals, networks, and applications (together with users), based on the adaptive management of the diversity [9]. Diverse resources come from several aspects. As the value chain of communications industry shown in Fig. 3, multiple operators and providers and producers together with various end users lead at last to the multiple diversity in technologies including:

- Service diversity, voice, data, multimedia, etc.
- Backbone diversity, by different operators, different regions, and/or based on different technologies.
- Access diversity, including both wired and wireless network infrastructures and air interfaces.
- Terminal diversity, different capability combination of computing, processing, storage, and communication, with different user interfaces.

As an important feature, the developing trend of the wireless network architecture is base on an all-IP core with asymmetry accesses, where mobile ad hoc networks can be treated as a special class of access networks since many mobile ad hoc networks have some internetworking with other infrastructures like the Internet. From the cellular structure point of view, future mobile networks can be divided into different sizes of cellular coverage, as shown in Table 1. The basic idea behind this is to seamlessly integrate two categories of wireless network technologies together, i.e. those that can provide low-bandwidth over a wide geographic area and those that can provide a high bandwidth over a narrow geographic area. The various cellular techniques consequently causes the complex intercarrier and/or intersystem roaming of a mobile node in a heterogeneous and overlay wireless cellular environment, as illustrated in Fig. 4, which is a great challenge to the design



NWCP: Network Content Provider, e.g. Yahoo
NWOP: Network Operator, e.g. AT&T
NWSP: Network Service Provider, e.g. AOL
AP: Application Provider, e.g. Netscape
CP: Content Provider, e.g. BBC
TP: Technique Provider, e.g. IETF/ITU/3GPP
ISV: Independent Software Vendor, e.g. Sun
NWEM: Network Equipment Manufacturer , e.g. Lucent
TEM: Terminal Equipment Manufacturer, e.g. IBM

Fig. 3. Value chain of communication industry

of mobility management schemes for the future mobile systems.

B. Mobility granularity

Mobile host should be able to roaming within a whole mobile communications system or between different systems as long as their networks are interconnected. Networks can firstly be classified according to different providers and/or technologies. Then one symmetric network can be further divided into domains, location areas, access point regions, zones of access points, and logical channels within one access point. Different mobility levels/granularities can then be defined accordingly, including

1) Mega-mobility, is the mobility between the networks of different providers or technologies (heterogeneous or asymmetry networks), e.g. satellite to UMTS to WLAN to Bluetooth, etc. One network may includes some domains.

2) Macro-mobility, is the mobility between different "visited domains" but still within one network. One visited domain usually includes several "location areas".

3) Micro-mobility, is the mobility between different "location areas" but still within one visited domain. One location area may includes several "access points regions".

Table 1. Cellular coverage division

Cell name	Place	Coverage	Speed	Techniques
Mega-cell	Global	Global coverage	>200km/h Airplane	satellite
Macro-cell	Suburban, rural	1km-10km	20-200km/h Vehicle/train	2G/3G PCS
Micro-cell	Urban	100m-1km	10-50km/h Vehicle	PCS, WLAN, HiperLAN
Pico-cell	In-building	10m-100m	<10km/h Walk	WLAN, HomeRF, Bluetooth
Nano-cell	Personal area	1m-10m	Nearly stationary	Bluetooth, IrDA

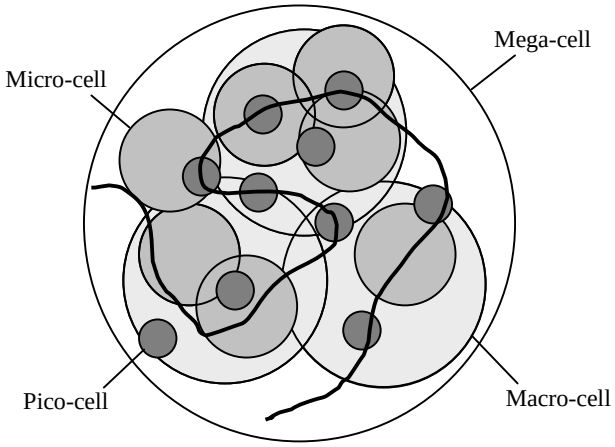


Fig. 4. Scenario of node roaming in future mobile networks

4) Mini-mobility, is the mobility between different "access point regions" but still within one location area. One access point region usually includes several "access points".

5) Pico-mobility, is the mobility between different "access points" but still within one access point region. One access point usually covers a certain zone.

6) Nano-mobility, is the mobility within the zone covered by one access point, where the cell zone can be various from mega-cell down to nano-cell, see Table 1. One access point may employ several logical channels.

Note that this division does not mean there exists any direct correspondence with either the cellular division above that based on cellular size, or the moving speed or range of a mobile device. Instead, the division is totally based on the mobility granularity from the network point of view. For example for a spot in overlay cells of heterogeneous networks, mega-mobility may happen from cell A (e.g. picocell) to cell B (e.g. macrocell) without necessarily moving out of the coverage area of cell A or it can even be stationary. This is also known as vertical handoff. There are also other examples of mobility while stationary: for pico-mobility handoff may occur in the area overlapped by two contiguous access points, and for nano-mobility a logical handoff may happen between different logical channels in one physical zone.

The significance of this division is that different mobility granularities may have different effects on the related mobility management schemes invoked. For example, nano-mobility, pico-mobility and mini-mobility often do not need the joint of location update; micro-mobility in a certain domain can limit the location updates into the domain and leave other domains uninfluenced and unaware; in case of mega-mobility, vertical handoff and location management strategies must be involved in order to treat with the asymmetry network structures and protocols. Note that these definitions are not exclusive with each other. Some of them can be converged in some systems where the mobility granularities have not been distinguished in such detail.

V. CONCLUSTIONS

This paper studies the basic concepts of mobility for mobile communications. Mobility is defined in the context of computation and communication, which lead to the concept of mobile computing. The impacts of mobility on mobile communications are analysed and the main ideas of mobility management are also introduced. New challenges arise in future mobile communications systems that with the diversity as the key feature. Various mobility granularities are then defined for the description of mobility in the future mobile networks.

The new challenges of the future mobile systems make the management of mobility more complex. Many new problems need to be carefully taken into account during the design of mobility management mechanisms. Moreover, new effective mobility and traffic models, together with simulation and measurement based experiments, for the performance evaluation of new management methods are very important issues under carefully studies.

ACKNOWLEDGMENT

Financial support by the National Technology Agency of Finland is gratefully acknowledged.

REFERENCES

1. J.M. Irvien, r. Tafazolli, and I.S. Groves, "Future mobile networks," *Electronics & Communication Engineering Journal*, Vol. 12, No. 6, pp. 262–270, Dec. 2000.
2. B.G. Evans and K. Baughan, "Visions of 4G," *Electronics & Communication Engineering Journal*, Vol. 12, No. 6, pp. 293–303, Dec. 2000.
3. D.S. Milojicic, F. Douglis, and R. Wheeler, "Mobility, an Edited Collection", *Addison Wesley*, February 1999.
4. D.S. Milojicic, F. Douglis, Y. Paindaveine, R. Wheeler , and S. Zhou, "Process Migration Survey", *ACM Computing Surveys*, Vol. 32 No. 8, pp. 241–299, Sep. 2000.
5. Vu Anh Pham and A. Karmouch, "Mobile software agents: an overview," *IEEE Communications Magazine*, Vol. 36, No. 7, pp. 26–37, July 1998.
6. R. Pandya, D. Grillo, E. Lycksell, P. Mieybegue, H. Okinaka, and M. Yabusaki, "IMT-2000 standards: network aspects," *IEEE Personal Communications*, Vol. 4, No. 4, pp. 20–29, Aug. 1997.
7. J.Z. Sun, D. Howie, and J. Sauvola, "Mobility management techniques for the next generation wireless networks," In *Proc. SPIE Asia-Pacific Optical and Wireless Communications Conference*, Beijing, China, Nov. 2001, Vol. 4586, 155-166.
8. I. F. Akyildiz, J. S. Ho, and W. Wang, "Mobility management in next-generation wireless systems," *Proceedings of the IEEE*, Vol. 87, No. 8, pp. 1347-1384, Aug. 1999
9. J.Z. Sun, D. Howie, and J. Sauvola, "Features in future: 4G visions from a technical perspective," In *Proc. IEEE Global Communications Conference*, San Antonio, TX, Nov. 2001, Vol. 6, pp. 3533 –3537.