Anonymous Mobile Internet Access

Teemu Rinta-aho
<teemu.rinta-aho@ericsson.fi>

Helsinki 25th September 2001

Master's Thesis

UNIVERSITY OF HELSINKI
Department of Computer Science

Anonymous Mobile Internet Access

Teemu Rinta-aho
<teemu.rinta-aho@ericsson.fi>

Master's Thesis
Department of Computer Science
University of Helsinki
25th September 2001, 53 pages

The operation of current mobile phone networks is based on the availability of massive amounts of subscriber information. Currently there is no way a mobile user can access the Internet without having made a contract with an access network operator. Many business transactions – for example, buying a newspaper from a newspaper stand – are anonymous, if cash is used for payment. Buying with cash, a credit card or a cheque does not require a subscription between the client and the service provider. The payment instrument itself is all that is needed for authorizing the transaction.

This kind of contractless and/or anonymous payment methods could also be available in future public access networks. Payments for mobile network access should also be possible using a credit card or electronic money. In this study, anonymity should be understood to be flexible. *Anonymous mobile internet access* means that the user is anonymous, or the identity of the mobile user is not *necessary* for the operator of the *access* network to know. An access network is a part of the Internet – the part that the mobile user directly connects to at the link layer. The user may be anonymous towards the access network and well known to some other network at the same time.

When the mobile user does not need a subscription with any operator, mobility between different operators and different kinds of access networks (LAN, UMTS, Wireless LAN, etc.) become easier to build. Besides a new business model, this type of universal mobility requires new methods for mobility management which must be in a layer above the access network specific protocols used today, to enable mobility between any network technologies.

This thesis first looks into the current technologies in wireless and mobile Internet computing. Then, different types of electronic payment methods usable in an online environment are studied. Third, a solution to combine these into a new network architecture is presented at a higher level. After that, the study is summarized and concluded.

Classification (Computing Reviews 1998): C.2.1, C.2.2, C.2.3

Keywords: Anonymous access, 4G, e-commerce, multiaccess, multioperator

Contents

1 Introduction

The increasing use of the Internet has grown interest in the possibility of accessing the Internet using mobile nodes which are able to move between access networks. These access networks are networks to which the mobile nodes are connected via fixed cables or via different wireless links. Access networks offer mobile nodes an access to the Internet. Examples of access networks are a cellular mobile data network or a local area network with wireless access points. A mobile user may want to move between these access networks without breaking the ongoing communications. Access networks may be built using different technologies.

Currently, wide-area mobile communication is possible only with a subscription to an access network (usually the mobile phone network's data services). The mobile phone network operator needs its users personal data to be stored in a database to make authentication, authorization and accounting (AAA) possible. Roaming is possible only to those mobile phone networks, for which the operator has made a roaming contract with the operator of the subscriber's "home" network.

The owners of the access networks are usually companies aiming to make a profit, so they want to be paid for the services they provide. Currently charging is based on contracts between the end-users and ISPs (Internet Service Providers), and is usually based on monthly billing. Accessing Internet via a mobile phone network is usually included in the customer's monthly phone bill.

Today the different network operators and network technologies have disparate methods for charging the network users. It is impossible for a mobile user to visit a network without having made a contract with the administrator of the network. It is usually also required that the identity of the mobile user is known to the network. The identity of the user is only needed to guarantee the user's ability to pay for the services. If the mobile user could guarantee that ability by some other means, there wouldn't be any security reasons to make long-term contracts. There is currently no Internet mobility between different public access networks, although the technology to implement it exists [Per00, JP00]. Current solutions which allow the users to move between different types of access networks are based on proprietary solutions. These solutions usually take two existing access network technologies and bind them together in a non-standard and non-open way. The drawback in this kind of solution is that it requires research and design of new protocols every time a new access technology is wanted to combine with the existing technologies. The lack of a common method for payments restricts the user's mobility to the networks he or she has a subscription to – even if it was otherwise technically possible to have mobility between the networks.

Using Internet mobility enables mobility between different access networks, but we also have to establish a trust relationship with the access network so that the operator of the network can trust us. This trust relationship does not necessarily require that we reveal our identity to the network if we can use some other method to guarantee our ability to pay for the service. One way to do that is to present a relationship with a third trusted party, for example with a bank or a credit card company. Another way is to use electronic money or some other prepaid system to pay on-line for the access. Therefore, a protocol for negotiating the payment method is needed. This new protocol, together with the internet mobility protocols, enables a true global mobile Internet (see Figure 1).

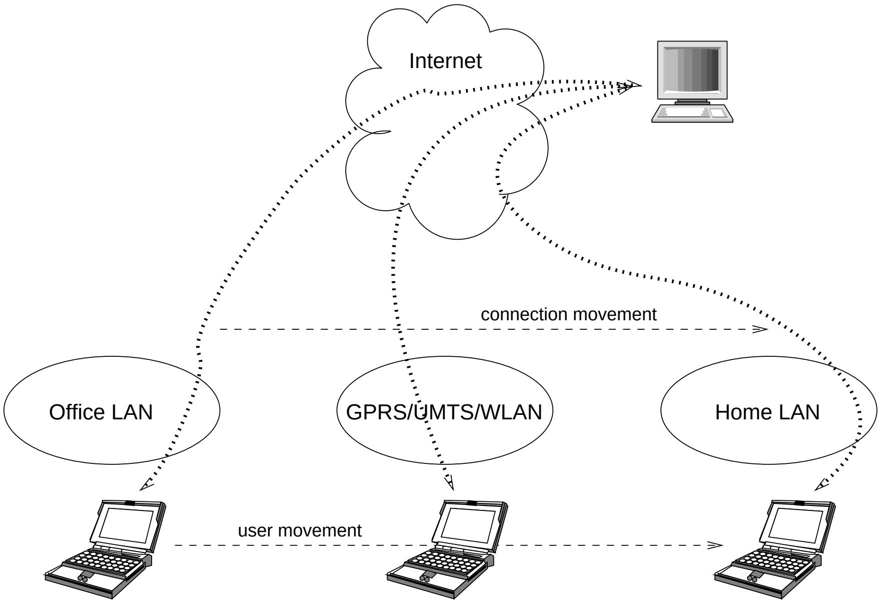

Figure 1: An example use case of desired mobility

Paying for the access on-line enables mobility between networks without any contracts between the user and the operators. A user will be able to use any network as long as he or she can present a valid payment method, or some other means of presenting that he or she can be trusted. This new architecture also enables anonymous access, which means that the access network does not need to know the identity of the user. A benefit of paying on-line is also the reduced billing costs for operators.

If mobile networks had the option to buy services by paying on-line, like when using a phone booth, the operators wouldn't have to make roaming contracts with all other operators and exchange accounting data. This requires that the user of the access network has a method of paying for the services that itself guarantees that the operator will be paid for the services it offers.

In this thesis, current solutions for wireless and mobile networking are studied first. It is important to understand the variety of different access network technologies available today, and that there is no single technology that serves every need. While wireless local area networks are best suited for buildings and small areas, public cellular networks can offer Internet access in larger areas and in the countryside. Wireless access is not neccessarily mobile access, and an access network is not necessarily wireless. Most mobile access networks are wireless, and because the problem is the same for all technologies, only the wireless networks are studied here.

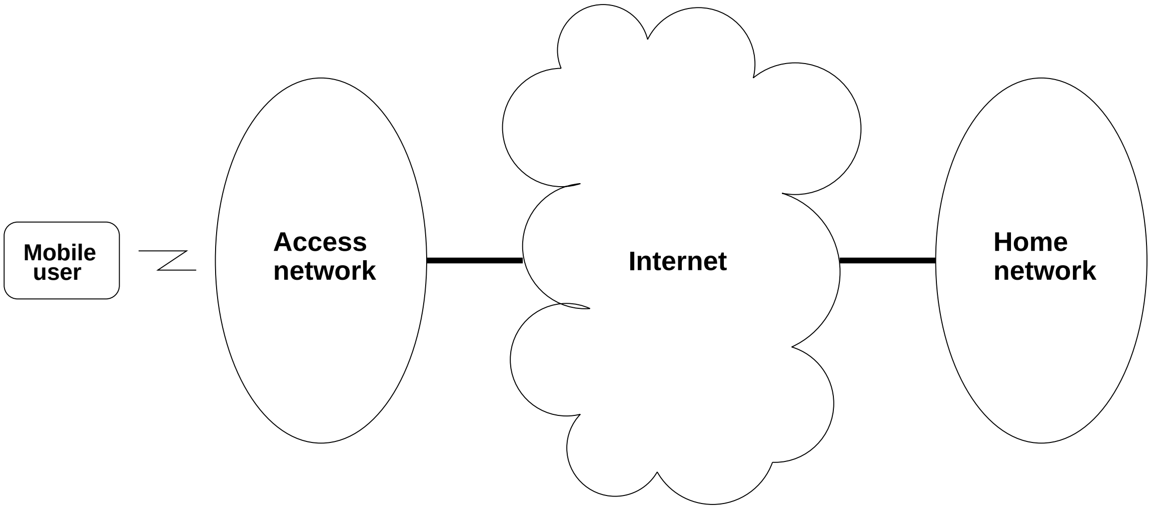

Figure 2: Access and home networks

An access network is a network that is used to access the Internet. It can be, for example, a mobile phone network with data services or a Wireless LAN network. A home network can be for example the mobile phone network that the user has a subscription to, or an Internet Access Provider's (ISP) network, or a corporate intranet (see Figure 2). All mobile nodes don't necessarily have a home network.

Access, transport, and services or "content" can be separate services, and the mobile user could buy all these from separate operators or entities. In this thesis, only the buying of access is considered as it directly affects mobility. Usually the pricing of access and transport are bundled together by the operator. Paying for services is a broad area of research and beyond the scope of this thesis.

Networks are usually built of different layers of protocols, as in the ISO OSI network reference model [Tan96]. All different access networks use different techniques for mobility. Mobility is usually restricted to mobility in one access network, not between access networks. This mobility can be thought of as layer two mobility. Internet Protocol (IP) is a layer three protocol, which can support mobility at a higher level than the layer two protocols, and therefore supports mobility between different types of layer two networks (see Figure 3), assuming that these layer two networks are capable of carrying IP packets. IP itself is not enough for full mobility, but the mobility extensions are required (see section 3.1).

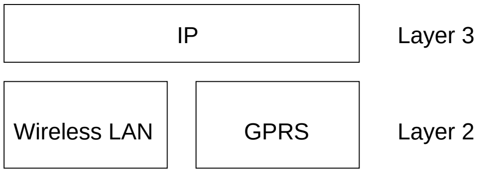

Figure 3: Network layers 2 (link) and 3 (network)

It is very important to first understand how these Internet protocols work, to be able to understand that the mobility and payment problem should be solved with them, and not with proprietary access network specific protocols. That is, if we want to have a true universal mobile Internet. A technically detailed study of IP mobility and security is given in the third section as a basis for understanding the protocol extensions presented later. Public-key cryptography is also the basis for all that is in the fourth section.

In the fourth section, different electronic payment methods are reviewed. While the first is based on credit cards, the subsequent are based on cash. While reading this section, the reader is encouraged to consider how much overhead different on-line payment methods add to handovers between access networks. A handover means the event of changing from an access network to another. The time used for a handover affects the operation of upper layer protocols and possibly applications. Therefore it is important that the handovers don't take too much time.

In the fifth section, a possible solution to combine all different access network technologies together by using internet protocols and on-line payment methods is suggested. The reader should keep in mind that the presentations of different technologies in earlier sections might feel a little disjoint, but they will come together by the end of this study.

Most of the ideas presented in this study have been implemented in a corporate research project which the author has been involved with during the writing. The research project is presented in the appendix.

2 Wireless Computing

Mobile computing is emerging due to the high number of mobile phones, PDAs (Personal Digital Assistants), and hybrids of them. To better understand some of the problems when enabling mobile computing, some degree of knowledge of different wireless access technologies is required.

Wireless Local Area Networks (Wireless LANs or WLANs) are getting more and more popular as the costs decrease and the speeds increase. Most common wireless networks use radio waves as the communication medium. Different WLANs support mobility between access points in the same subnetwork. However, this link-layer mobility doesn't solve the macro-mobility problem which comes when the user wants to move between subnetworks or even different access technologies.

As Internet applications (e-mail, WWW) are becoming the most important computing applications to most people, mobile computing can be thought as mobile *Internet* computing. In the future, most public access networks will offer Internet connectivity. This can be seen in the evolution of mobile phone networks, where for example the GSM networks are upgraded with GPRS service. These mobile phone networks with data services also offer mobility inside the network. This mobility can be very large scale, as operators typically cover a whole country with their network. But this mobility, just as WLAN link-layer mobility, is restricted to a certain network, and is not a solution for the Internet mobility.

Mobile networking differs both from portable networking and from wireless networking. Wireless networking is not neccessarily mobile, if the user can not change to another subnetwork or access network without breaking the ongoing communication. Some applications, for example WWW browsing, work well without mobility, if there are no stateful sessions between the client and the server which store the network address of the mobile client. Most other applications (for example ftp, ssh or telnet) will fail if the mobile user changes the connection from a network to another without network layer mobility. More of this network layer mobility will be studied in the next chapter, and enhancements presented in chapters after that.

In this chapter, some of the most popular wireless network technologies will be studied. When reading this chapter, the reader is advised to keep in mind how these different access networks based on different technologies could be used together in the most efficient way to give the mobile user the best possible connectivity and network coverage. Solutions to some of the problems are presented in the latter chapters of this study.

2.1 Wireless LAN

The IEEE 802 standard family specifies the currently used Local Area Networks (LAN). IEEE 802.11 [IEE99] is the IEEE's standard for wireless LANs. A wireless 802.11 WLAN can be configured in two ways, either in ad-hoc mode or in infrastructure mode. In an ad-hoc network, mobile nodes communicate directly with each other, and the network can change its shape at any time, when mobile nodes come and go (see Figure 4).

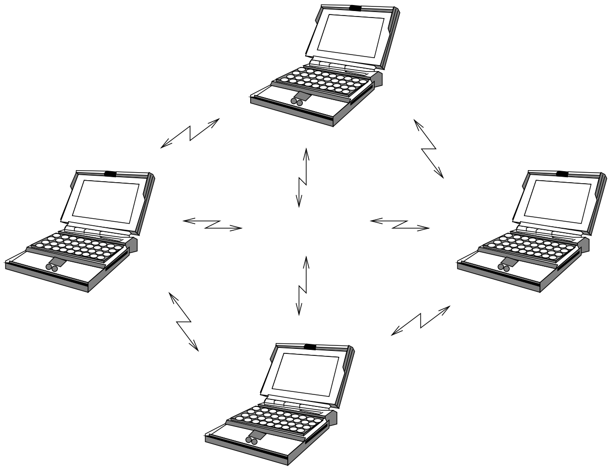

Figure 4: Wireless ad-hoc network

The second type of network structure is to form an infrastructure with fixed network access points with which mobile nodes can communicate. These access points can then be interconnected with landlines to form a larger area of coverage for the wireless LAN and to make it possible for mobile nodes connected to different access points to communicate with others (see Figure 5). When a mobile node moves from one access point to another, a handover occurs, like in today's cellular networks.

The ad-hoc mode is useful if there is no installed infrastructure. Mobile nodes can communicate directly with each other depending only on the physical distance between them, but not on the location of the formed ad-hoc network. A good example could be an ad-hoc network between the nodes carried by a rescue team on the top of a mountain. Managing network level connectivity between the nodes and possible Internet connectivity in an ad-hoc network is an area of active research. In IETF, there is a working group which is standardizing the protocols needed for ad-hoc networking.

The IEEE 802.11 standard specifies both the physical (PHY) and the medium

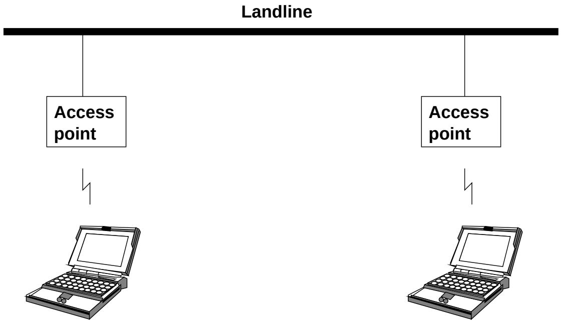

Figure 5: Wireless infrastructure network

access control (MAC) layers of the network. The physical layer can use either direct sequence spread spectrum, frequency-hopping spread spectrum, or infrared pulse position modulation. Data rates can be from 1 Mbps to 11 Mbps. For radio transmission, the frequency band of 2.4 GHz is used, which is an unlicensed band for industrial, scientific and medical (ISM) applications. The same band is also used by Bluetooth, microwave owens, DECT phones and remote controls, so these technologies can interfere with each other. Infrared is better protected against eavesdropping, as it cannot be received from behind a corner or a wall. However, it requires a line-of-sight and can be affected by sunlight.

The medium access control (MAC) layer is a set of protocols which maintain the order in the use of a shared physical medium. The 802.11 standard specifies a carrier sense multiple access with collision avoidance (CSMA/CA) protocol. This protocol specifies that when a node is about to transmit a packet, it must first listen to make sure no other node is transmitting. If it doesn't hear any ongoing transmissions, it then transmits the packet. If it hears an ongoing transmission, it chooses a random backoff factor which gives the amount of time the node must wait until it can try to retransmit the packet. When the channel is clear, the backoff counter is decremented, and when it reaches zero, the node tries to retransmit the packet. Collision detection, which is used for example in the IEEE standard 802.3 (LAN), cannot be used for radio transmissions, as the transmitted signal will hide all arriving signals at the node.

When a packet is to be transmitted, the node first sends out a short ready-to-send (RTS) packet containing the length of the packet to be sent. If the receiving node receives the RTS, it responds with a clear-to-send (CTS) packet. After this, the transmitting node sends its packet. If the packet was received correctly (determined by the CRC check), the receiving node sends an acknowledgement (ACK) packet. This kind of back-and-forth exchange is necessary to avoid the "hidden node" problem (see Figure 6).

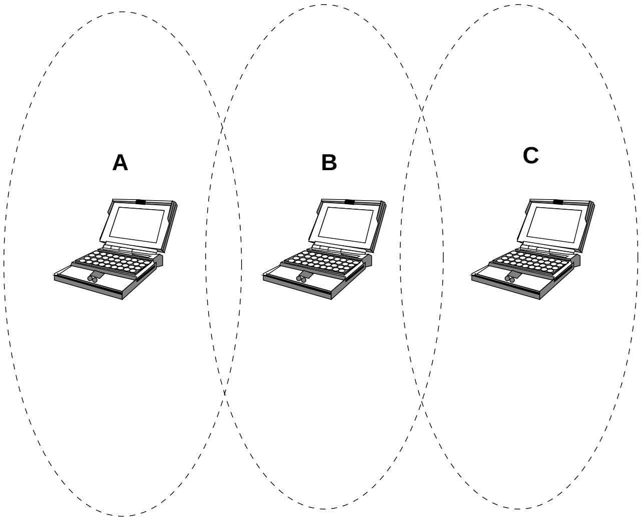

Figure 6: Hidden node problem

In above situation, node A can communicate with node B, and node B can communicate with node C. However, node A cannot communicate with node C, as they are out of others radio/IR range. So, although node A may sense that the channel is clear, node C may in fact be transmitting to node B. The protocol described above alerts node A that node B is busy and that it must wait before transmitting its packet.

Wireless LANs are becoming popular in different hotspot areas, where there are many potential network users. This kind of areas include airports, conferences and campus areas. Recently, many ISPs have started to offer WLAN as a substitute for leased line in highly populated areas. These WLAN access networks are built as infrastructure networks, each access point connected to the ISPs network with a leased line or a microwave link. Access control is usually based on higher level protocols (i.e. IPsec), as it is impossible to prevent a malicious user from forging the hardware address of a network interface card and using the network unauthorized. Although WLAN works fine when the area covered by an access point is relatively small, it suffers from the limited bandwidth in larger areas with many simultaneous users.

2.2 Bluetooth

Bluetooth is a protocol for wireless connectivity for all kinds of devices. It was designed to be cheap to manufacture, so that it would be economically possible to embed bluetooth connectivity in any kind of device. The main application for bluetooth is to replace point-to-point links with wireless connections. This kind of links can be for example the link between a headset and a mobile phone, or a mobile phone and a laptop computer. Bluetooth can be also used for small ad-hoc networks in a way similar to that of WLAN [con99]. Bluetooth has been designed by the Bluetooth SIG (Special Interest Group) that has members such as Ericsson, Nokia, IBM, Toshiba etc.

Bluetooth can have point-to-point and point-to-multipoint connections. Two or more nodes sharing the same channel form a *piconet*. There is one master node and up to seven slave nodes in one piconet. Multiple piconets which overlap form a *scatternet*.

The bluetooth system consists of a radio unit, a link control unit and a support unit. The support unit is responsible for link management and host terminal interface functions. The radio interface operates at the same unlicensed 2.4 GHz frequency band as the WLAN and many other radio applications. The maximum range can be anything between 10 and 100 meters, depending on the transmitting power.

The Bluetooth protocol uses a combination of circuit and packet switching. It supports both connectionless asynchronous links for data transmission and connection-oriented synchronous links for voice services. The asynchronous channel supports the maximum data transfer rate of almost one megabit per second.

Compared to WLAN, bluetooth is cheaper and smaller, and it consumes less power. However, as the operating range is much smaller, it is better suited for homes and personal area networks (PANs), which can be for example the network between a mobile phone and a laptop computer.

2.3 Public Land Mobile Networks

Public Land Mobile Networks (PLMNs) offer larger area mobility but lower bandwidth than the Wireless LANs. Many PLMNs today are cellular mobile phone networks with data services. One of the biggest differences with the current mobile phone networks and the Wireless LANs, for example, is that the phone networks are circuit-switched. Usually people have to pay for the time they allocate the channels, independently of the amount of data they are transferring. However, Quality of Service can be quaranteed, unlike in the Wireless LAN, which offers only best-effort service. For some time, there has been a trend to make the future telephone networks packet-switched. GPRS (General Packet Radio Service) system brings packet-switched connections to mobile phone users, and is simply an add-on to the current GSM mobile phone network architecture.

2.3.1 Mobile phone networks

Current mobile phone networks are organized as cellular networks. This means that the radio network is divided into cells to enable efficient frequency re-use (see Figure 7). Mobile networks are hierarchical: a couple of cells are grouped together to form a larger administrative zone, and then again, these zones are usually grouped together once more. The benefit for this hierarchical model comes in support of mobility - there is less signalling on the top of the hierarchy because there are intermediate nodes that control mobility in smaller areas. The same idea has been applied to the Internet mobility in Hierarchical Mobile IPv6 internet draft [SCEMB01]. Mobility within a network includes handovers between cells. Currently, there is no mobility between different operators in GSM (Global System for Mobile communications) nor in GPRS – connections will be terminated when the network coverage the operator is providing ends. Also, a terminal that could use any operator in a certain area, is manually configured to use just one of them.

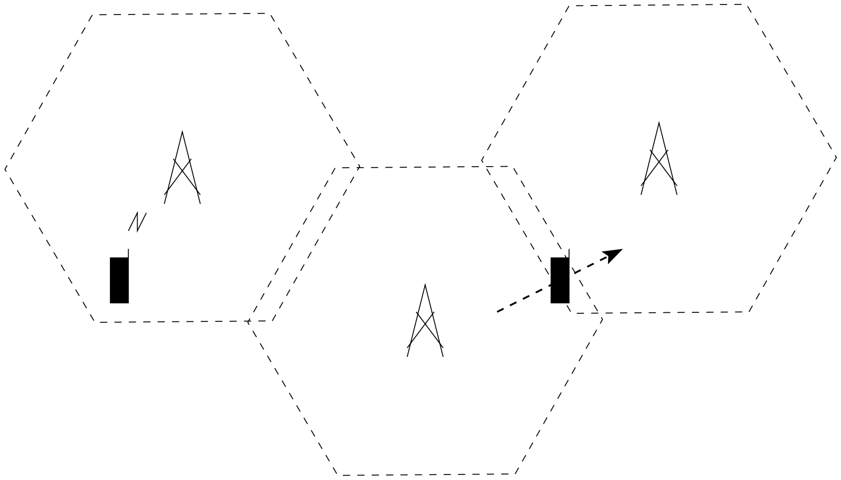

Figure 7: Cellular radio network

2.3.2 3rd Generation Mobile Networks

The so called “3rd generation mobile phone network” standardisation work has been going on for some time in 3GPP (3rd Generation Partnership Project).

3GPP is formed by telecommunications companies: operators and equipment manufacturers.

The goal of the 3GPP is to have a mobile network with larger bandwidth and better services than in current networks. It is also supposed to be a globally accepted standard. The radio technique used in 3G is totally new 2 GHz WCDMA (Wideband Code Division Multiple Access) that theoretically supports bandwidths up to 2 megabits per second.

The 3G mobile system is logically implemented on the GSM/GPRS structure through the addition of a new air interface supported by two network nodes, the RNC (Radio Network Controller) and the Node B (see Figure 8) [3GP99].

The RNCs control radio base stations that are used in UMTS, while BSCs (Base Station Controllers) control radio base stations that are used in GSM and GPRS. The rest of the network is the same for both. Circuit switched phone calls are routed through the MSC (Mobile Switching Centre and possibly also through GMSC (Gateway MSC). Packet switched data is routed through SGSN (Serving GPRS Support Node) and from the GGSN (Gateway GPRS Support Node) to the Internet.

There are several databases in the GSM network. MSCs store a VLR (Visitor Location Register). All nodes in CN (Core Network) share an HLR (Home Location Register), an AuC (Authentication Center) and an EIR (Equipment Identification Register). A small database is also the SIM (Subscriber Identity Module) that is located in the mobile terminal.

The 3G charging architecture is subdivided by the two transmission planes, the CS (Circuit Switched) and the PS (Packet Switched) domain. The call detail records generated by the servicing nodes for the appropriate domain are forwarded to the billing system for processing (see Figure 9) [3GP99].

Current mobile phone networks (GSM and its enhancements), and the 3rd generation mobile networks, are still built on the paradigm that the network operator has long-term relationships with its customers, and the access is tightly coupled with other services. Also, the authentication to the network is done before charging, so it is not possible to just spontaneously buy network access without subscribing to it first.

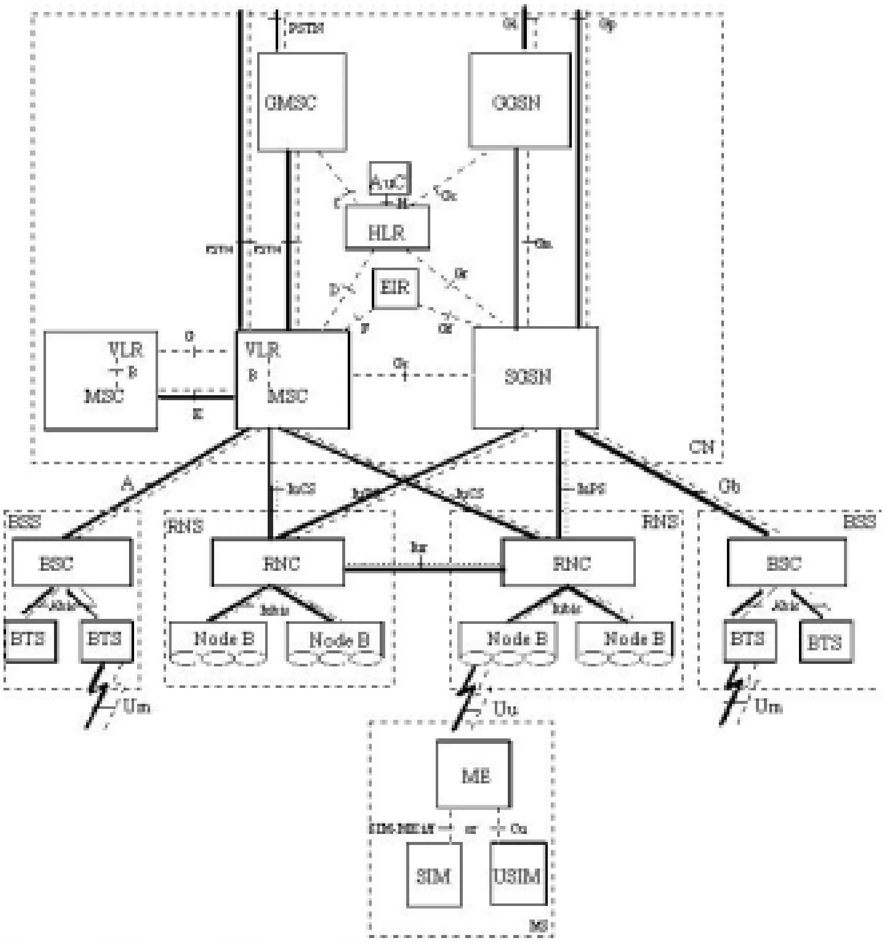

Figure 8: Overview of the 3G logical architecture [3GP99]

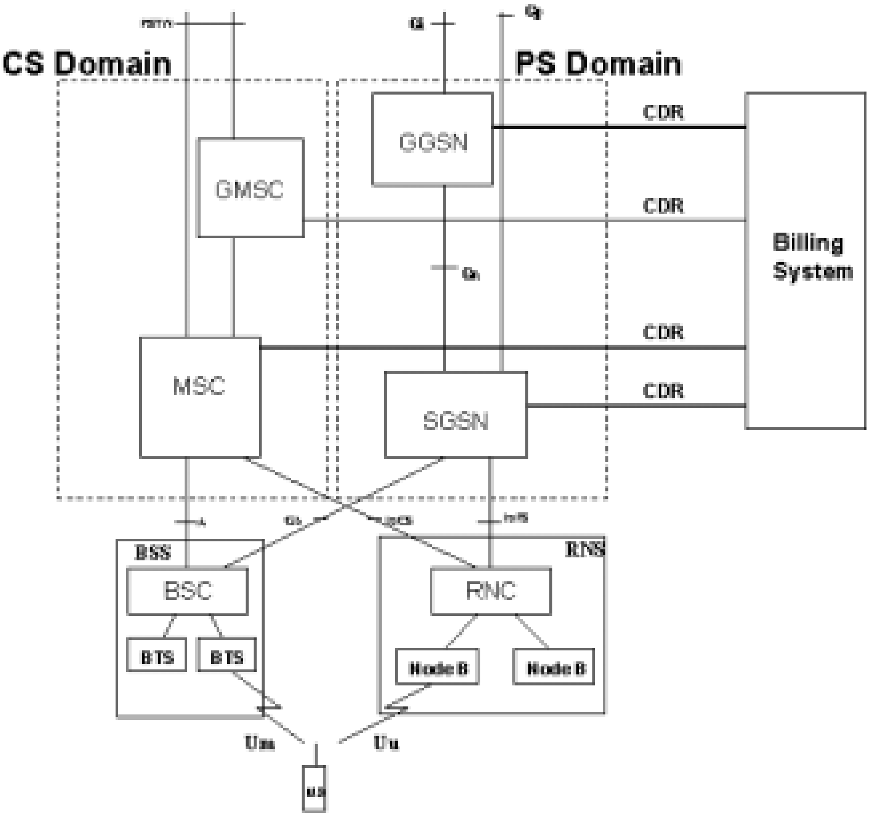

Figure 9: 3G charging logical architecture [3GP99]

3 Mobile Computing

To have mobility in the Internet, the mobility management must be built into the IP (Internet Protocol), which is a network layer protocol [Pos80]. Network layer is the layer above the link layer which hides the differences of different link layers from the transport layer (which in turn is used by applications), and provides global routing of datagrams. IETF (Internet Engineering Task Force) has developed mobility support into both IPv4 [Pos80] and its successor IPv6 [DH98].

3.1 Mobile IP

The Internet Protocol was designed so that the IP address was not only an address, but also the identifier of the node. A node was assumed to be static and to not change its attachment point in the network. If the node changes its attachment point, it can no longer be contacted by other nodes as the regular Internet routing protocols route the packets into the attachment point that the IP address indicates. The node can change its location but then it has to change its IP address too, but then it can no longer be contacted with its old address, and all the ongoing connections break when the address changes. To overcome these problems, an IETF working group has designed extensions to the IP, known as the Mobile IP [Per00].

3.1.1 Terminology

A mobile node is a host or a router that can move from one network to another without changing its IP address [Per00]. A home network is the network that the home IP address belongs to. A home agent is a router on a mobile node's home network. It keeps a database of the locations of mobile nodes that have registered, and captures and encapsulates the IP packets sent to the mobile node's home IP address, and sends them to the current locations of the mobile nodes. A care-of address is an IP address obtained by a mobile node from a foreign network. A care-of address can be a “foreign agent care-of address”, which is the foreign agent's IP address, or a “co-located care-of address” which is externally obtained from the foreign network. A foreign agent is a router on the foreign network which services the visiting mobile nodes. If packets are sent to a mobile node using a foreign agent care-of address, the foreign agent forwards those packets to the mobile node. It can also be the default gateway for a mobile node. Default gateway is used for sending packets outside the local network.

3.1.2 Protocol Overview

The Mobile IP protocol has certain requirements [Per00]. As the connections using transport layer protocols are bound to the IP address of the node, it must not be changed when the node moves. A mobile node must also be able to communicate with nodes that don't implement the Mobile IP protocol. Messages used for updating the mobility information must be authenticated to prevent illegal redirection of data traffic. An important goal in designing the Mobile IP protocol has been to minimize the amount of messages, as mobile nodes may have an error-prone wireless link and a battery as the power source.

The Mobile IP protocol has also been designed with the assumption that the mobile node does not change network more often than once per second [Per00]. The Mobile IP is designed for solving the macromobility, and the handovers are usually much longer than the handovers at the link layer. Macromobility means mobility between different IP subnetworks, while micromobility means mobility between access points where the IP address of the mobile node does not change.

3.1.3 Operation Overview

Mobile IP defines two support services: agent discovery and registration [Per00]. Mobility agents (home and foreign agents) advertise their services by sending Agent Advertisement messages. A mobile node may send an Agent Solicitation message to solicit an Agent Advertisement message. When a mobile node receives an Agent Advertisement, it determines whether it is on its home network or on a foreign network. If the mobile node is on its home network, it operates without mobility services, just like any other Internet host. If the mobile node then has a registration at its home agent, it de-registers by sending a Registration Request message to the home agent. If the mobile node detects that it is at a foreign network, it obtains a care-of address either from a foreign agent or externally, for example by DHCP (Dynamic Host Configuration Protocol [Dro97]). After that the mobile node registers its new care-of address with its home agent by sending a Registration Request, through a foreign agent, or directly. After that, datagrams sent to the mobile node's home address are captured by its home agent and tunneled to its care-of address (either to the foreign agent or to the mobile node, depending on the setup), and finally to the mobile node (see Figure 10). Datagrams sent by the mobile node are delivered to their destination using standard IP routing. These packets do not necessarily go through the home agent, but because of the ingress filtering used in almost all routers today, in practice all the traffic has to go through the home agent. Ingress filtering drops datagrams

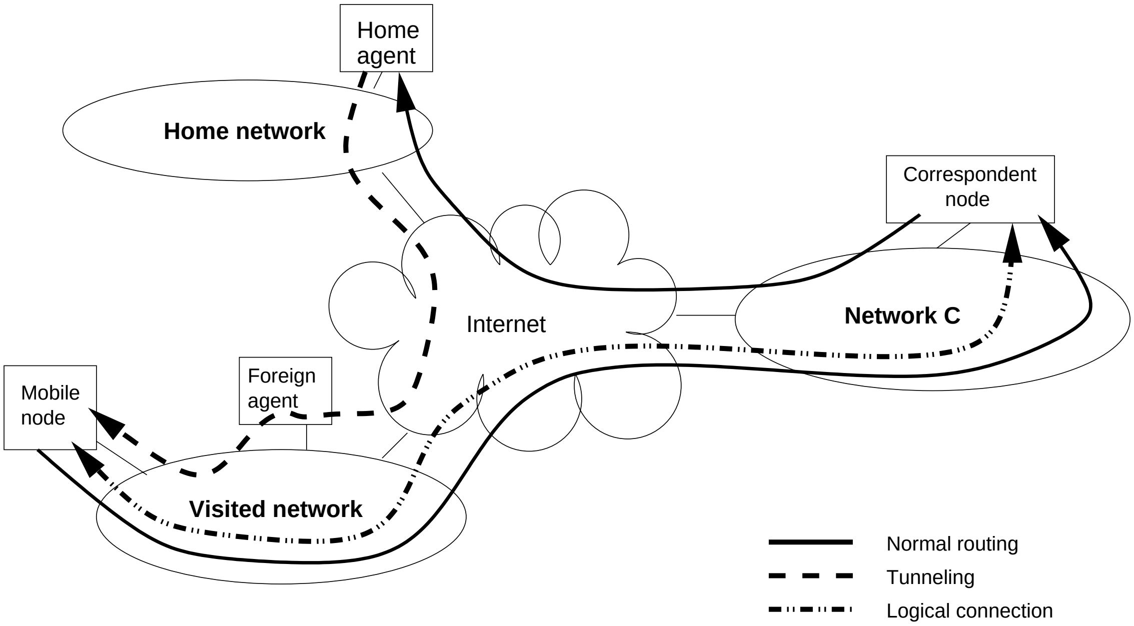

Figure 10: Mobile IP architecture

that have a source address which is not topographically correct, and no visiting node has such a home address by the definition of IP mobility.

3.1.4 Agent Discovery

Agent discovery is used by the mobile nodes to detect whether they are located at the home network or at a foreign network, or if they have moved from a foreign network to another [Per00]. A foreign agent care-of address can also be obtained through the agent discovery mechanism.

In Mobile IP, the ICMP (Internet Control Message Protocol [Pos81]) Router Discovery has been extended to support Agent Discovery. An Agent Advertisement is an ICMP Router Advertisement message extended with the Mobility Agent Advertisement Extension. The Agent Solicitation message is almost identical to the ICMP Router Solicitation [Per00].

Authentication is not required for the advertisement or solicitation messages, although they may be authenticated by using the IP Authentication Header [KA98a].

3.1.5 Registration

Mobile nodes use the registration mechanism to request forwarding services while being at a foreign network, to inform their home agent of the current location, to renew a registration which is close to expiration, and to de-register after returning to the home network [Per00]. A mobile node can also have multiple registrations so that each datagram is copied to each of its care-of addresses. The registration mechanism can also be used to determine the IP address of the home agent.

Registration can be done in two different ways, either via a foreign agent that forwards the registration to the home agent or directly with the home agent. If the mobile node is registering with a foreign agent care-of address, or it has a co-located care-of address, but the foreign agent requests to register via itself, the mobile node must register via that foreign agent. Otherwise, it can register directly with the home agent.

Each mobile node and mobility agent must support mobility security associations between the mobile entities. Registration messages between a mobile node and its home agent are authenticated with the Mobile-Home Authentication Extension [Per00]. The Mobile-Foreign Authentication Extension is not necessarily required in Requests and Replies between the mobile node and the foreign agent.

3.1.6 Routing Considerations

All mobility agents and mobile nodes using co-located care-of addresses must support IP in IP encapsulation [Sim95] (see Figure 11). Minimal encapsulation and GRE encapsulation [HLFT94] may optionally be used [Per00].

A router can also be mobile. This gives a possibility for mobile networks, which in turn may also have visiting nodes. A practical implementation might be a mobile network on an airplane, a ship or a train. Mobile routers themselves are no special cases in the Mobile IP protocol, they operate the same way as any mobile node except that they forward datagrams and may also act as foreign agents [Per00].

A mobile node operates as any fixed node when it is at its home network – all datagrams are routed normally. When the mobile node is at a foreign network it chooses the default router either by using its foreign agent as a first-hop router or some other router learned from a received ICMP Router Advertisement message. Mobile nodes can not broadcast ARP [Plu82] packets into a foreign network. The Mobile IP protocol does not specify how the mobile node can get the MAC address of a router. Similarly, a foreign agent can't use ARP for a mobile node's

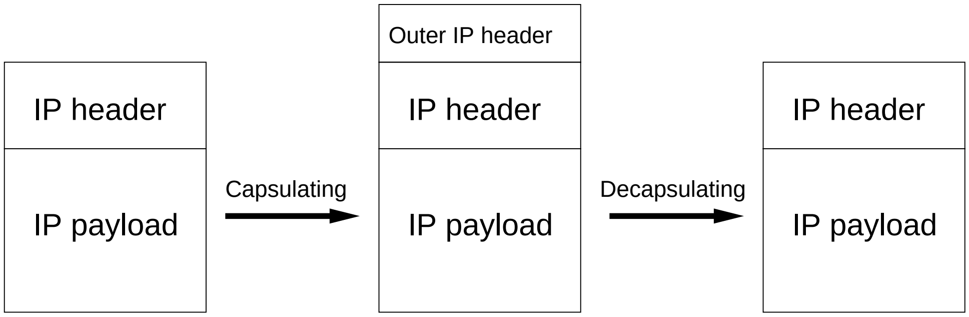

Figure 11: IP tunneling

MAC address on a foreign network. It can obtain the MAC address for example from an Agent Solicitation or a Registration Request.

The home agent intercepts datagrams destined to the mobile node's home address while the mobile node is at a foreign network. The home agent may use proxy and gratuitous ARP for the interception of packets. Proxy ARP means an ARP reply sent by one node on behalf of another, while gratuitous ARP is a packet that updates other nodes' ARP caches. By changing the mobile node's MAC address in ARP caches to its own, the home agent will receive IP packets destined to the mobile node. This is a non Mobile IP specific security threat in all Local Area Networks. IPv6 has a better solution that replaces ARP.

The home agent only forwards broadcast datagrams if the mobile node has requested such service. If the mobile node is using a co-located care-of address, the home agent tunnels the datagram directly to the mobile node. If the mobile node is using a foreign agent care-of address, the home agent first encapsulates the broadcast datagram into a unicast datagram destined to the mobile node's home address, then it tunnels this datagram to the foreign agent. This mechanism is necessary for the foreign agent to know which mobile node it should forward the datagram. Then the mobile node must decapsulate the original broadcast datagram.

A mobile node may join a multicast group via a local multicast router on the foreign network, if there is a multicast router available. If the mobile node is using a co-located care-of address, it should use it as the source of its IGMP messages [Fen97], otherwise it may use its home address [Per00].

If there is no multicast router on the visited network, the mobile node may join a multicast group with its home agent, if the home agent is a multicast router. The multicast datagram forwarding works identically with the broadcast datagram forwarding described earlier. All IGMP messages go through the home agent, and the multicast datagrams come through the same tunnel. If the mobile node wants to send a multicast datagram directly on a foreign network, it must use a co-located care-of address, and if it sends it through the home agent, it must use its home address as the source address [Per00].

3.1.7 Route optimization

Having a home agent routing datagrams destined to a mobile node presents an un-optimal routing graph between a correspondent node and a mobile node, especially if the home agent is far away from the shortest route between the two communicating nodes. Consider, if a mobile node wants to communicate with a node that is in the same visited network. Every packet between the correspondent node and mobile node will go out of the network all the way to the home agent, and then back to the network to the mobile node. If the home agent is far away in the network topology, this can present a serious performance bottleneck. However, if the correspondent node supports Mobile IP, the mobile node can send it a binding update and it can start sending packets directly to the mobile node's care-of address instead of its home address. When a mobile node moves it must send binding update messages to all nodes it is communicating with, to keep their bindings up-to-date. Route optimization is not yet an official part of the Mobile IP standard, but is widely used in many implementations [PJ00].

3.1.8 Security Considerations

Mobile nodes are often connected to the Internet via wireless links, therefore the possibility of a security threat is even greater than on conventional, static, wired networks. Especially if unauthenticated, the Mobile IP registration protocol is vulnerable to unauthorized traffic redirection. Therefore all control messages between mobile nodes and home agents are authenticated [Per00].

Messages with the foreign agent are not required to be authenticated, mostly because of the lack of a common network key management protocol. In commercial networks, authenticating visiting users is important if anonymous access is not what is desired. Mobile IP does not itself offer a solution to this problem.

To achieve privacy, encryption has to be used. Using IPsec (ESP header) is one way to implement encryption. Mobile IP does not offer special support for encryption. Location privacy can be achieved by tunneling all mobile-node-originating datagrams through the home agent. That way it is difficult to observe where the mobile node is located, as all datagrams seem to originate from the home network.

Timestamp and optionally nonces are used for replay protection in Registration Requests. A nonce is a 32-bit pseudo-random number. The value of a timestamp or a nonce is copied from a Request to the Reply to avoid replay attacks.

3.2 Mobility in IPv6

The next version of the Internet Protocol, IPv6 [DH98], offers not only a larger, 128-bit address space, but also other improvements which help in supporting mobility. All IPv6 nodes have a built-in support for host bindings, so route optimization will work with all connections. In IPv6, source routing will be used instead of tunneling. IPv6 also has a better support for security with the AH and ESP modes.

3.2.1 Comparison with Mobile IPv4

The design of the Mobile IPv6 is based on experiences received from designing Mobile IPv4. The Mobile IPv6 design work has also affected the on-going development of IPv6 itself, making them work more seamlessly [JP00].

Mobile IPv6 is in many ways similar to the Mobile IPv4, but it is fully integrated into IPv6 and has many improvements [JP00]. MIPv6 support for route optimization is built into IPv6. This feature disables what was known in IPv4 as “triangle routing”. Mobile nodes use the care-of address as a source address instead of the home address. This allows the operation in networks behind routers performing ingress filtering. It also simplifies the routing of multicast datagrams. Foreign agents are no longer needed - mobile nodes can operate in any network without any special support. The IPv6 address is usually obtained through Address Autoconfiguration [TN98] for both static and mobile nodes. IPsec is currently the only security mechanism, although it might change after the IESG (Internet Engineering Steering Group) decision where it stated that IPsec AH is a too heavy mechanism to protect just the BUs [Nar01]. The new movement detection mechanism prevents a common situation in wireless networks where the router can reach the mobile but the mobile can’t reach the router and the mobile node starts searching a new router without removing the registration from the old one. This is also known as the “black hole” situation. MIPv6 uses the IPv6 Routing Header instead of the IP encapsulation used in MIPv4. The home agent uses IPv6 Neighbor Discovery [NNS98] instead of ARP to capture packets destined to the mobile node. Because of using the Routing Header and the ICMPv6, the “tunnel soft state” management is no longer needed. IPv6 anycast is used instead of the broadcast to find a home agent. There is a new Advertisement Interval option for the mobile nodes to decide themselves how many Router Advertisements they can miss before searching for a new router. Using IPv6 Destination Options allows all MIPv6 control traffic to be sent along (piggybacked) with any IPv6 packets, and no separate UDP packets are needed as in MIPv4.

3.2.2 Protocol Overview

The Mobile IPv6 protocol operation is closely related to that of Mobile IPv4 [JP00]. While the mobile node is at its home network, it operates as any IPv6 node. While the mobile node is away from home, it registers one of its care-of addresses with a router on its home network, and requests this router to be its home agent. A home agent keeps a list of Mobile Bindings, which relate the mobile node’s home address with a care-of address. A binding registration is done by sending a packet with a Binding Update destination option, and the home agent replies by sending a packet with a Binding Acknowledgement destination option. After that, the home agent uses proxy Neighbor Discovery to capture the datagrams destined to the mobile node’s home address and tunnels them to the registered care-of address using IPv6 encapsulation.

If the mobile node does not know the address of its home agent, or the home agent is replaced while the mobile node is at a foreign network, it uses Dynamic Home Agent Address Discovery by sending an ICMP Home Agent Address Discovery Request to the “Mobile IPv6 Home-Agents” anycast address with its own home network subnet prefix. That causes one of the home agents to reply to the mobile node with a list of home agents.

Mobile IPv6 nodes also send Binding Update destination options to the correspondent nodes to tell their new care-of address. Therefore, mobile nodes keep a list of ongoing connections to other nodes with information of when the last Binding Update for a certain node was sent, and other type of management data.

3.2.3 Correspondent Node

A correspondent node is any node communicating with a mobile node. It can be itself fixed or mobile, or a home agent [JP00].

Packets sent by a mobile node almost always contain a home address option. When any node receives such packet, it must first copy the address from the home address option to the IPv6 header, before passing the packet to the regular IPv6 processing. If a packet contains a Binding Update option, it must be validated, and the Binding Cache updated by deleting any existing bindings for this mobile node and storing a new entry with the given care-of address. If the Binding Update option has the Acknowledge flag set, the correspondent node must send a Binding Acknowledgement option back to mobile node.

If an entry in the Binding Cache expires, it is removed, and further packets destined to the mobile node will go through the mobile node's home network until a new Binding Update is received. A correspondent node can send a Binding Request to a mobile node for which mobility binding is about to expire in the near future.

If the correspondent node has a mobile binding in its cache for a mobile node, any packet sent to the mobile node that causes an ICMP error message will be sent directly to the correspondent node. If the packet is sent via the home agent, and the ICMP error message is caused when the packet is in the tunnel from home agent to mobile node, the ICMP error messages are sent to the home agent which then relays certain ICMP messages to the correspondent node.

3.2.4 Home Agent

Each home agent on a network receive information on other home agents by receiving periodically sent Router Advertisement messages and keeping their Home Agents Lists up to date. When a home agent receives an ICMP Home Agent Address Discovery anycast message, it replies with a reply message containing the list of the home agents [JP00].

When receiving a Binding Update option requesting the router to be a home agent for the mobile node who sent the message, the message must be validated. If the Binding Update is accepted, the home agent creates an entry in its Binding Cache, setting the mobile node's home address from the Home Address option, and the care-of address from the Alternate Care-of Address sub-option, if present, otherwise the care-of address will be taken from the IPv6 header's source address field. Then the home agent will send a Binding Acknowledgement, if requested. Also, when receiving a Binding Update requesting the removal of binding, the home agent must remove the entry for that mobile node in its binding cache.

When a router is serving as a home agent for a mobile node, it must intercept the packets addressed to the mobile node's home address and tunnel them to the registered care-of address. The home agent is using both proxy and gratuitous Neighbor Discovery for packet interception. The packets will be tunneled using IPv6 encapsulation.

3.2.5 Mobile Node

When the mobile node is away from the home network and sending IP packets, the Mobile IP modifies the packets from applications by adding a Home Address option with the home address, and putting the care-of address to the IP header's source address field. For short-term communication, the mobile node can use its care-of address without the Home Address option. This gives less overhead when using for example DNS queries [JP00].

When outbound packets are processed, they are compared to the IPsec Security Policy Database to determine what processing is required for the packet. Binding Updates and Binding Acknowledgements are always protected with Authentication Header. IKE is used as the default key management protocol.

When the mobile node is at a foreign network, it can receive packets either from a correspondent node via a home agent, directly from a correspondent node with a Binding Cache entry for the mobile node, or from a correspondent node which has an outdated care-of address in its Binding Cache. If the packet goes to an old care-of address, and the old default router has the mobile node's new care-of address, it tunnels the packet to the new care-of address.

When the mobile node is moving from a network to another, it detects the movement primarily from the received Router Advertisement messages. It must sent Binding Updates to its home agent and all correspondent nodes after getting a new care-of address. It can also send a Binding Update to a home agent in its previous visited network to allow packet forwarding from old to new care-of address.

When the mobile node returns to its home network, it must send a Binding Update to its home agent requesting a deregistration, and multicast a Neighbor Advertisement message to advertise its own link-layer address as the destination of its home IP address.

3.3 Cryptography in Networks

Most current implementations and suggestions for authentication and privacy support in networks, as well as in electronical commerce, are based on the use of cryptographically strong algorithms.

Cryptography is used for encrypting data so that only the intended receiver can decrypt the data. This can be done in two ways, either by using one symmetric key, or public and private key pairs. The former is called private-key cryptography and the latter is public-key cryptography [Sch96].

Cryptography can also be used for authentication and authorization, with digital signatures and digitally signed certificates, respectively.

Modern cryptography does not rely on the cryptographic algorithms, which are well known, but on large numbers called keys which must be used together with the correct algorithm to be able to encrypt or decrypt the data correctly. Algorithms are usually designed so that decryption with the correct key is easy, but very difficult or impossible for all practical purposes with guessed keys. By effectively using temporary keys, the communication between network entities can be secured.

An infrastructure for creating and distributing keys must be present in the network to allow the use of the abovementioned cryptographic methods.

3.3.1 Symmetric-Key Cryptography

With symmetric-key encryption, the encryption key is the same as the decryption key (see Figure 12).

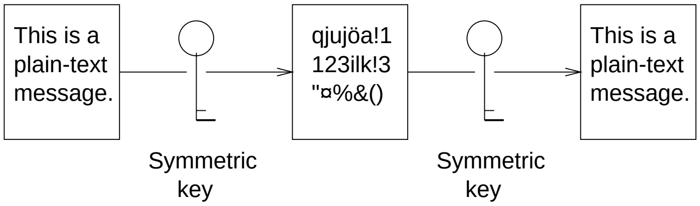

Figure 12: Symmetric-key encryption

Implementations of symmetric-key encryption are usually very efficient, and users do not experience any significant time delay as a result of the encryption or decryption. Symmetric-key encryption can also be used as authentication, since information encrypted with a symmetric key cannot be decrypted with any other symmetric key. So, as long as the symmetric key is kept secret, both sides can be sure that they are communicating with their counterpart as long as the decrypted data continues to make sense.

Symmetric-key encryption is effective only if the symmetric key is kept secret by the two parties involved. If someone else has the key, it affects both privacy and authentication. Anyone with the symmetric key can decrypt messages sent with that key and encrypt new messages which look like they originated from one of the two original nodes.

3.3.2 Public-Key Cryptography

The RSA approach to public-key encryption (also called asymmetric encryption) has two keys: a public key and a private key, which are associated with an entity that needs to authenticate its identity or to sign or encrypt data. The public keys are published, and the corresponding private key is kept secret. Data encrypted with the public key can only be decrypted with the corresponding private key [KS98] (see Figure 13).

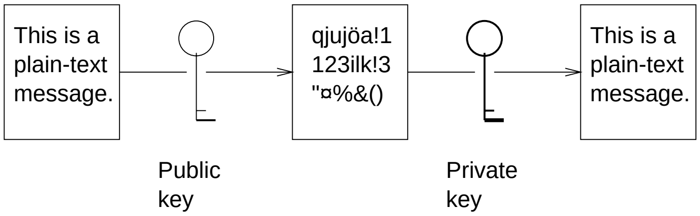

Figure 13: Public-key encryption

Public-key encryption requires more computation compared to symmetric-key encryption, and therefore it is not always appropriate for large amounts of data. In many current implementations public-key encryption is used to send a symmetric key securely, which can then be used to encrypt additional data.

The reverse of the scheme shown above also works: data encrypted with the private key can be decrypted only with the corresponding public key. This would not be a good way to encrypt sensitive data, because it means that anyone with the public key, which is by definition published, could decrypt that data. However, private-key encryption is useful, because it means that you can use your private key to sign data with your digital signature – an important requirement for electronic commerce and other applications of cryptography, including the concept of certificates.

3.3.3 Certificates

Encryption and decryption address eavesdropping, but do not solve the other two important problems: tampering and impersonation.

Tamper detection and related authentication techniques are based on a mathematical function called a one-way hash. A one-way hash is a number of fixed length with two important characteristics. First is that the value of the hash is unique for the hashed data. Any change in the data, even changing the value of one bit, results in a different hash. Another one is that the content of the hashed data cannot be computed from the hash, which is why the hash is called "one-way".

It is possible to encrypt the data with the private key and decrypt the date with the public key. This method of using keys is not usable for protecting sensitive data, but it enables the digital signatures. Instead of encrypting the data, a one-way hash of it is created and then encrypted with the private key. The encrypted hash and other necessary information, such as the hashing algorithm used, bundled together, form a digital signature.

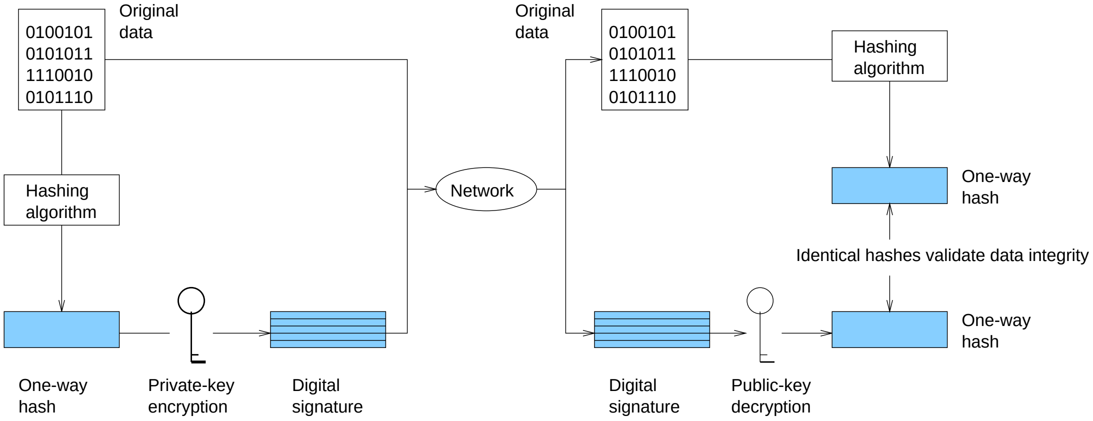

Figure 14: Using a digital signature

Two items are transferred to the recipient: the original data and the digital signature (see Figure 14). The receiver first decrypts the hash with the signer's public key and then uses the same hashing algorithm to generate a new hash of the data. If the received and computed hashes match, the receiver can be certain that the data was sent by the owner of the public key and that it was not tampered with during the transport.

Confirming the identity of the signer also requires some way of confirming that the public key really belongs to the sender. Certificates can be used for the authentication of a key.

A digital signature is comparable to a handwritten signature. Once something is signed, it is difficult to deny doing it so later, assumed that the private key has not been compromised.

3.3.4 Certificates and Authentication

A certificate is an electronic document that is used to identify an entity and to associate that entity with a public key. A certificate is comparable to a driver's license or a passport that can be used to prove one's identity.

A passport, for example, is typically issued at the local police department, which verifies the applicant's identity, address, and other information before issuing the passport. Certificates are somewhat similar to these familiar forms of identification. Certificate Authorities (CAs) are organizations that check the identities and issue certificates.

A certificate binds a public key to the name of the entity the certificate identifies. They prevent the use of fake public keys that could be used for impersonation. Only the public key certified by the certificate will work with the corresponding private key that the identified entity is using.

In addition to a public key, a certificate always includes the name of the entity it identifies, an expiration date, the name of the CA that issued the certificate, a serial number, and other information. A certificate always includes the digital signature of the issuing CA. The CA's digital signature allows the certificate to be trusted by entities that don't trust the identified entity but trust the CA.

Authentication means the process of confirming an identity. In network interactions, authentication also requires the identification of peer. Certificates are one way of solving the authentication problem in networks.

Authentication in networks is typically required for communication between a client and a server. The client needs to have a confident identification of the server and vice versa. Client and server authentication is not the only application of certificates. For example, a digital signature on an e-mail message authenticate the sender of the e-mail as well as the integrity of data contained in the message. Digital money can also be constructed with certificates and digital signatures (see section 4.2).

3.3.5 IPsec

IPsec (IP security protocol) is a set of protocols developed by the Internet Engineering Task Force (IETF) to provide security services for IP traffic at the IP layer [KA98c].

The goal of IPsec is to provide interoperable, high quality, strong cryptography based security for IPv4 and IPv6. To meet this goal, IPsec defines two traffic security protocols: the Authentication Header (AH) [KA98a] and the Encapsulating Security Payload (ESP) [KA98b], cryptographic key management procedures and protocols, Security Policy Database (SPD) that defines the security services that can be used and Security Associations (SA), which are relationships between two or more systems which describe how the systems will use the security services [KA98c].

The IPsec protocols operate in two modes: in transport mode or in tunnel mode. IPsec tunnel mode is designed to protect one or more tunnels between a pair of hosts, between a pair of security devices or security gateways (SGs), or between a security gateway and a host. IPsec transport mode applies to communications between two hosts [KA98c].

IPsec defines two traffic security protocols. The Authentication Header (AH) defines authentication methods for IP payloads. It also provides connectionless integrity, data origin authentication and an optional anti-replay service. The Encapsulating Security Payload (ESP) protocol defines encryption methods for IP payloads and (in tunnel mode) for part of the IP header. It provides encryption, and limited traffic flow confidentiality. It also may provide connectionless integrity, data origin authentication, and an anti-replay service [KA98c].

AH and ESP can be used alone or in combination during an IPsec communication session. Both protocols use encryption keys to protect data. The difference between AH and ESP authentication is that AH authenticates the entire IP packet, including any tunnel header while ESP only authenticates from the payload of the ESP encapsulation. This is acceptable for intranet packets that are encapsulated completely.

IPsec allows both manual and automatic key exchange. However, when IPsec is used on a wide scale, automatic key exchange becomes necessary.

Automatic key exchange is defined by a number of Internet drafts, but the main framework is described by the Internet Security Association Key Management Protocol (ISAKMP) [MSST98]. ISAKMP is designed to be key exchange independent; that is, it is designed to support many different types of key exchange. However, there is an Internet draft defining a specific public key based approach for automatic key management, known as Internet Key Exchange (IKE) [HC98].

For authentication, IKE supports digital signatures based on public key cryptography. IKE provides a facility for identification of different certificate authorities (CAs), several certificate types, and the exchange of the certificates identified.

A Security Association (SA) is the method IPsec uses to track a given communication session. It defines how the communicating systems will use security services, including information about the traffic security protocol, the authentication algorithm, and the encryption algorithm to be used. SAs also contain information on dataflow and lifetime as well as sequence numbering for anti-replay.

SAs are negotiated between two IPsec systems. The two IPsec systems can also negotiate the level of authorization for a range of addresses, protocols and ports that will be protected by the SA.

An SA is unidirectional. For each pair of communicating systems A and B there are at least two security connections: one from A to B and one from B to A. A given SA can use ESP or AH; not both. If a connection needs both protocols, it needs to establish two SAs for each direction; four for a bi-directional connection.

A Security Association is identified by a combination of a Security Parameter Index (SPI), which is a randomly chosen unique number, the destination IP address of the packet and the traffic security protocol to be used (AH or ESP).

Security Associations require two databases. A Security Policy Database (SPD) specifies the security services that will be provided for IP packets. This database contains an ordered list of policy entries. Each entry includes information about the type of packets the policy should apply to, such as source and destination address. An IPsec policy entry also would include an SA specification, listing the IPsec protocols, mode (tunnel or transport) and the security algorithms to be used. In a Security Association Database (SAD) each entry defines the parameters associated with one SA. Each SA has an entry in the SAD. For outbound packets, entries are pointed to by entries in the SPD. For inbound processing, each entry in the SAD is indexed by a destination IP address, protocol type, and SPI. Each SA specification in the SPD points to an SA, or a bundle of SAs.

When an IPsec system sends a packet, it first matches the packet against the entries in the Security Policy Database to see if there is an SA for the packet in the SAD. If a SA does not exist, the IPsec system creates one. It then applies the processing specified, and then inserts the SPI from the Security Association into the IPsec header. When the IPsec peer receives the packet, it looks up the Security Association in its database by source address and SPI (and security protocol) and then processes the packet as required.

Different users may have different security needs, and are likely to be using different security protocols. The Security Association enables users to negotiate a common set of security attributes in an authenticated and protected manner. An SA contains a set of information that must be agreed upon and shared between communicating systems.

The way in which SAs are negotiated is defined by ISAKMP, which provides a framework for negotiating a wide range of encryption algorithms, authentication mechanisms, key management algorithms, and other security features. During negotiation, the communicating systems determine exactly which algorithms to use (for example, DES or IDEA to encrypt data and headers; MD5 or SHA for authentication). After deciding on the algorithms, the two devices negotiate how they will share session keys.

Once the basic set of security attributes has been agreed upon, initial identity authenticated, and required keys generated, the SA can be used for subsequent communications by the system that invoked ISAKMP.

ISAKMP is not bound to any specific encryption algorithm, key generation technique, or security mechanism. However, ISAKMP has basic requirements for its authentication and key exchange components. For example, ISAKMP requires strong authentication mechanisms – specifically, digital signatures and digital certificates created by a trusted third party or certificate authority.

A Security Association can be in one of two modes of use: transport mode or tunnel mode. Both modes are supported by both AH and ESP. Tunnel mode allows a remote client or network device such as a router to encapsulate, encrypt and forward packets through an IPsec "tunnel" to a destination Security Gateway, typically a router or firewall with IPsec. The Security Gateway decrypts the original IP datagrams and forwards them on to the destination host. Tunnel mode is more secure, but it adds overhead to communication, as it essentially is just packaging an IP packet into another IP packet.

Transport mode is only used between hosts. It encapsulates and encrypts only the data portion (payload) of each IP packet, but leaves the header untouched. When a host runs ESP or AH, the payload field is the data that normally follows the IP header (for example, a TCP or UDP header followed by user data). Transport mode is less secure than tunnel mode, since it does not conceal or encapsulate the IP control information, but it is also more lightweight as it doesn't add as much overhead.

4 Methods for On-line Payments

Electronic commerce has emerged during the last years. A lot of shopping is expected to be done on the Internet in the future. Even today, a lot of different goods can be bought from on-line shops: cd-roms, books, software, etc. One of the most interesting aspects in electronic commerce is the market of "digital goods". If someone wants, for example, to buy a movie from the other side of the world, it can be delivered almost immediately, via the network. This kind of commerce has given rise to the development of new electronic payment methods. The current use of cheques and fiat money (bills and coins) is not scalable to the so-called "micro purchases" in networks.

Many organizations are currently developing implementations of electronic money or electronic cash. Electronic cash, or E-cash, would be a very flexible payment instrument in networks. It can be stored in very small spaces, and can be transferred very quickly over the networks. Electronic money can be stored not only in computers, but also in smartcards or mobile phones' SIM cards. The biggest problem in electronic money is that it is even more difficult to trust than regular money. It is more difficult to counterfeit, but it is a lot easier to copy and use simultaneously at many points of the network. If money is checked from the issuer every time it is used, a serious scalability problem arises.

Using an electronic payment method to pay for network access on-line is a new interesting concept. If it was trusted by all parties, anonymous mobile internet access could be provided. Also, one way to pay for on-line access is to use some kind of pre-paid system, where the user buys some kind of tickets which can be used to buy access. In mobile networks, the tickets would be cryptographically protected certificates [BH99].

4.1 SET

Secure Electronic Transactions (SET) is an open, independent standard, which is based on credit card payments. It has been developed by VISA, Mastercard, Microsoft, IBM, Netscape and American Express.

SET's main goals are [VM97]:

- Secure method for transferring payment information
- Guarantee the integrity of payment information
- Authenticate the user as the owner of the card

- Authenticate vendor
- Use the best known security practises
- Create a protocol that does not depend on the transport layer security
- Encourage co-operation between software companies

4.1.1 Implementation

The implementation of SET is based on cryptography. SET uses for example the technique of two digital signatures. The laws for encryption in different countries differ, but most countries allow the use of encryption for commercial transactions, which are well defined and use fixed length messages and can not easily be used for other purposes.

Using SET for on-line payments is based on a client software that stores all the needed user information, cryptographic keys, receipts, etc. This software communicates with the organization that issued the credit card and with the vendor's software. Each user has a digital key for each credit card, which is issued by the issuer of the card.

4.1.2 Operation

When a user wants to start a payment transaction, the client software authenticates the vendor based on its certificate. After this the client software uses the vendor's public key to encrypt all messages it sends to the vendor software. Client sends the order information and the credit card number and the amount to the vendor. The vendor can read the order information, but the billing information is encrypted with the credit card issuer's public key, and the vendor can't read them. Vendor passes the billing information to the issuer, and gets a confirmation of the billing. After that the vendor sends a confirmation of the order to the client. All the messages are digitally signed, and for example the confirmation of the order is stored in the client software as a receipt of the purchase.

The benefits of SET compared to traditional credit card billing are clear: neither credit card number nor any other valuable information is transferred unencrypted at any time. The vendor does not get the client's credit card number, but both sides of the transaction can authenticate each other.

4.2 Electronic Money

Digital, or electronic money is the electronic counterpart of the current bills and coins. In practice, electronic money can be handled as any digital data on computers. For the electronic money to be useful, it should be analogous to current money. The most important attributes of money are anonymity and liquidity [Pan96].

Anonymity means that when a client pays a vendor, no others than the vendor can recognize the client. In certain cases, recognition of the client by the vendor is neither necessary nor possible. The transaction is not necessarily recorded. This is analogous to buying a magazine from a news stand. The money itself is enough to certify the transaction.

Liquidity means the acceptability of money as a payment instrument – every vendor should accept the same money.

Electronic money has several benefits over regular money. It is a lot easier to store and transfer. It is more difficult to counterfeit than regular money, but a lot easier to copy and use simultaneously at different locations. A lot of research has been done to provide solutions to these problems [HP98, BH99, AMSW97, ZL98, Sim96].

4.2.1 Implementation

Electronic money can be implemented with pre-bought smartcards (for example phonecards) or with an all-electronic approach. Future smartcards can act as wallets, which store the electronic money, and can be used both in regular shops and attached to a personal computer or a mobile phone.

The major technical problem in implementing electronic money is the requirement of safe and scalable systems. In current implementations, the money must be checked from the issuer when it is used. This means that the vendor must have a connection to the issuer to be able to accept electronic money. In transactions between private consumers, and in small, mobile business like a hot dog stand, a connection to the bank may be difficult to maintain. Currently it is impossible to make sure the money is not used twice if it is not checked from the bank. Every time money is used, a record must be made at the bank.

4.2.2 Electronic Cash

Ecash is an implementation of electronic money by DigiCash. Ecash uses public key cryptography and blind signatures scheme for security and privacy [Cha92].

Ecash is used like normal cash. It can be deposited and drawn, and it can be transferred between accounts. The difference with normal cash is that the bank has to be involved in every transaction.

The money in Ecash is presented as bitstrings. Every money is a different string with a certain value. When client gets the money from bank, it is actually created by the client and only signed by the bank, after taking that amount away from client's account.

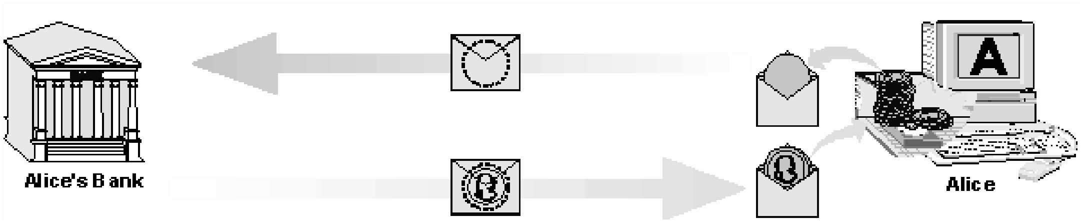

Figure 15: Blind Signature [tec00]

To provide anonymity without decreasing the level of security, a blind signatures scheme is presented [Cha92]. The idea in blind signatures is that the client multiplies the bitstring that represents the money with a random number before sending it to the bank for signing. After receiving the signed money, the client divides the "money" with the same random number. This way the bank's signature remains in the money, but the bank won't be able to recognize the money anymore, except that it is signed by its private key, and the money becomes untraceable.

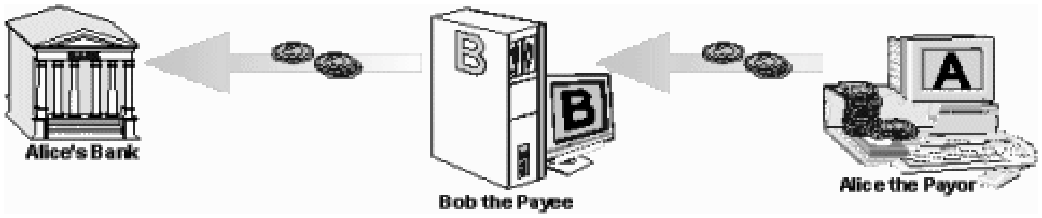

Figure 16: Paying by Ecash [tec00]

A paying transaction in Ecash system is started when the client gets a request to pay from the vendor, and after the client accepts the request, the client software transfers a money bundle from the client's hard disk to vendor. After that the vendor checks the moneys from the bank. If the client and the vendor are using different banks, the client's bank checks the money from the vendor's bank. If the money is accepted, it is marked as used in the bank's database.

4.2.3 Money on Smartcards

Electronic money can be stored on smartcards. One example is Mondex from Mastercard. The Mondex card is the same size as a credit card but it is equipped with a microprocessor. The card conforms to the ISO 7816 standard. The first specification of the card was released in 1994. Today over 450 companies in over 40 countries are working with them.

The software on the card is acting as a electronic wallet. Money can be transferred to and from the card with a card reader equipment. The card is protected with the user's own private PIN.

In Finland it has been possible to have a similar microchip on a bank or a credit card. Money can be loaded on to that card from regular ATMs, and can be spent at any shop having the equipment to communicate with the card. There is also available a reading device for computers which enables payments over the network.

In the future the SIM cards in mobile phones will most likely include a wallet for electronic money. As people carry their mobile phones with them all the time, it is natural that the phones will at some point merge with wallets. That same money could easily be used to pay for Internet access also.

5 Anonymous Mobile Networking

The ideas in this study are largely based on the work done in a corporate research project with which the author was involved when the writing of this study begun. The project and its backgrounds are presented in greater detail as an appendix at the end of this paper. Some of the problems identified in that project still exist, some have been circumvented and new problems have arisen. However, it has been recognized that these issues are important in the near future and the work is currently pursued in many other research projects.

Anonymity in this study should be understood a bit differently than normally in the area of pure security research. Anonymity here means that the trust between two network entities, which in this case are a mobile user and an access network, is based on something else than just the *identification* of the entities. Anonymity is not desired here with the possibly faulty assumption that most people want to hide their identities. It is desirable because it enables the creation of new trust relationships in a way that doesn't restrict the speed of the mobile node movement. It also broadens the variety of access networks the mobile node can choose from. If identification was the only way to build a trust relationship, it wouldn't be possible to build it during a handover without affecting the ongoing real-time data streams a user might have.

However, as a side-effect, anonymity also promotes for privacy. This is currently respected to some degree on networks. For example, in the GSM mobile phone network the actual phone number is never sent over the air, but a temporary identification number is used. That prevents an eavesdropper to follow a certain user. In the Internet, anonymous e-mail, news and chat are very popular. For the same reason some people go to a drug store to buy a magazine with cash, they might want to use the access networks anonymously. Being anonymous gives people the privacy and in the case of mobile networking, location privacy – users don't necessarily want anyone to know where they were located at a certain time.

Anonymity should be flexible in networks. We generally don't care if we are anonymous to the access networks we are using, but would like to identify ourselves to a location server somewhere on the Internet, possibly in our own home network, so that other users could contact us while we are mobile.

The most common method of billing today is subscriptions with monthly billing. Most users have to pay for the use of the telephone network, and possibly also to a separate ISP for the Internet traffic. This comes from the age when all telephones were fixed devices, and could be even owned by the same operator that owned the lines they were attached to.

For example, anyone who would like to offer a fast wireless LAN access to a mobile phone network's users suffers from the inflexibility of the current billing mechanisms. New access providers would benefit from a new, more flexible way of charging for the network services.

5.1 General Access Negotiation Protocol

If we want to have truly universal mobile Internet access, we can't rely on any technology that lies under the network layer to solve the mobility or billing requirements. A general access negotiation protocol is required in the network layer. This protocol should be flexible and not depend on any specific payment method. It also shouldn't change the existing protocols in any other way than possibly extending them. It should be a framework protocol towards different payment protocols just like the ISAKMP (see section 3.3.5 on page 28) is a framework protocol towards different key exchange protocols in the IP security area.

One solution is to extend the IPv6 Router Advertisement (RA) message [NNS98] with a new header option. This option would include information of the available payment methods, pricing information and QoS to mobile nodes. This would not add too much to the current protocols and would not break anything that already exists. The mobile node would then easily see if there is a common payment protocol available that could be used for payments, and also evaluate whether the pricing is acceptable. If the mobile node receives Router Advertisement messages from more than one router, it can use this extra information to select the most suitable one.

If the mobile node accepts this offer received in a Router Advertisement (RA) it then starts standard IPsec negotiations to generate a Security Association (SA) with the access router. After an SA is generated, communication can go on in an authenticated manner using the Authentication Header (AH).

The final stage in the protocol is to use the payment method specific protocol to set up and maintain the billing information and the billing state. The access router, or some other node in the access network, would then control the mobile node's access based on the billing state. For example, electronic money could be used to pay on certain intervals to keep the mobile node on the access control list of the router. This specific payment protocol would be independent of the General Access Negotiation Protocol which is only used to transfer the minimal amount of parameters necessary for starting the specific payment protocol. These payment protocols can be, for example, SET or electronic money transfer protocols. A similar approach is taken in Internet Open Trading Protocol (IOTP) [Bur00], but it is not especially designed to be integrated with the access discovery and is at application layer.

This new protocol adds to handover times, as the negotiation has to be made with every access router separately. The overhead can be reduced if handover takes place between access routers in the same access network (see Figure 17). The technique to reduce handover time is to add co-operation between the access routers. The new access router can get the negotiated parameters from the old access router withouth having to re-negotiate. This requires that the mobile node can be securely associated with the connection parameters. This "context transfer" technique is studied at least by the IETF SeaMoby (Seamless Mobility) working group. The context that is transferred can include for example the QoS, header compression, and, payment method states. Proactive context transfer would also help to decrease the delay as the context could be transferred to the new access router before the mobile node makes the handover. This requires a heuristic algorithm to calculate which router will be the new router in the mobile node's path of movement. It is also under development in the IETF.

It would make the handover even faster if the new access router could inherit the IPsec security association from the old access router without having to re-negotiate with the mobile node, but the way IPsec is designed makes it impossible. That is because the SA is bound to the home IP address of a node. One solution to this would be not to use IPsec protocols, but some other suitable security mechanism to authenticate the received IP packets.

The applicability of context transfer for handovers between access networks depends on the trust relationship between the routers in different access networks. If the networks are operated by the same operator, context transfer may be a good solution. Otherwise, the handover time will be much longer than inside an access network (see Figure 18).

Even if the handover takes some time, this architecture offers a serious advantage over the existing mobile networks. For example in standard GPRS, the mobile user can not usually roam between operators without breaking the ongoing connections, because the operators have not enabled the feature. With this kind of new open architecture the user is freed of the underlying relationships between different operators and different access technologies.

If the legacy billing should be preserved, or wanted to be reused with the new one, a gateway from IP layer general access negotiation protocol to the proprietary

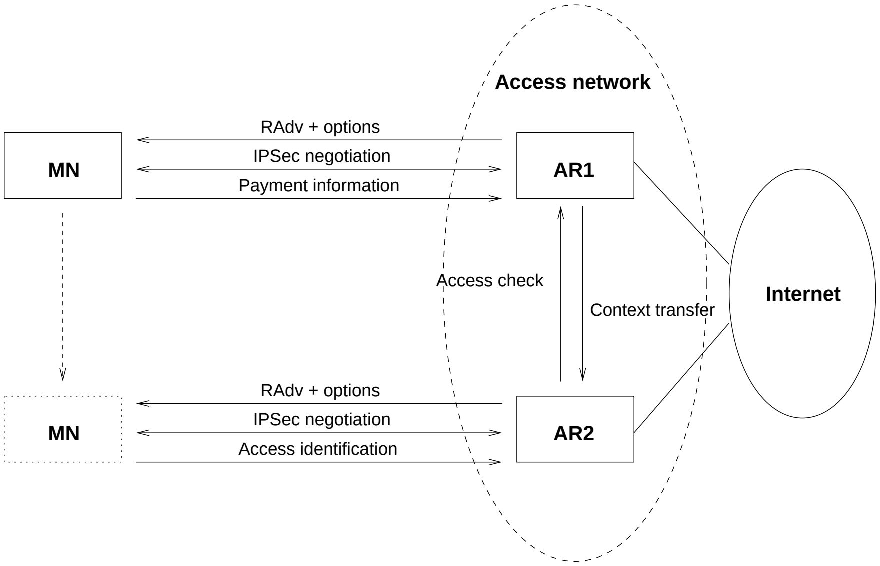

Figure 17: Mobility within an access network

protocol can be designed. For example, it would not be very difficult to design a gateway between the mobile user and the GPRS billing system. This could be even implemented in the GGSN node, which is accessible to the mobile user at IP layer (see chapter 2.3.2 on page 10).

In Figure 19 the communication sequence of access discovery and handover is shown at a higher level. The included parties are Mobile Node (MN), Access Router 1 (AR1) and Access Router 2 (AR2). IPsec and access negotiation phases consist of several message exchanges. The optional context transfer phase between the access routers shortens the latter access negotiation phase significantly, and if a clever solution is found, also the IPsec negotiation phase. But that is beyond the scope of this study.

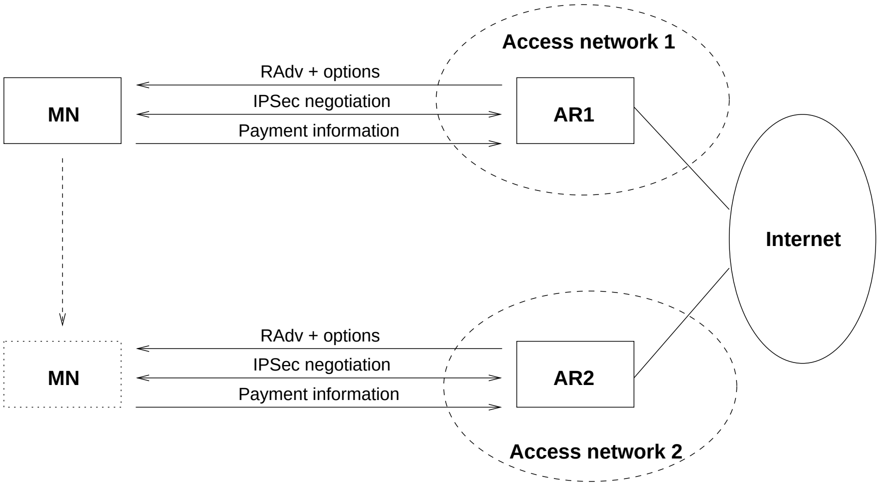

Figure 18: Mobility between access networks

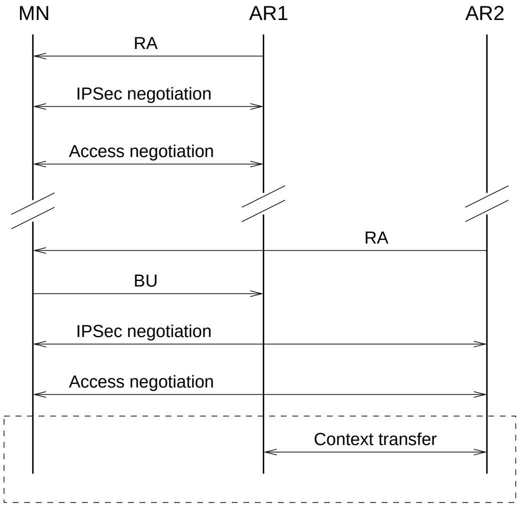

Figure 19: Sequence diagram of a handover

5.2 The Future of Mobile Networking

Here I will present an example which combines the technologies presented earlier. This example, although somewhat idealistic, should give an idea what the future of mobile networking at its best might be.

Bob has a PDA which is running an operating system with IPv6 support. It also has Wireless LAN and LAN connections through network cards and a UMTS connection via a mobile phone. The PDA is connected to the mobile phone through bluetooth (see Figure 20). The previously sketched GANP (General Access Negotiation Protocol) is also drawn in the picture as it is implemented in both IP layer (the extended RA, see section 5.1 on page 37) and in application layer (decision making and communication with servers).

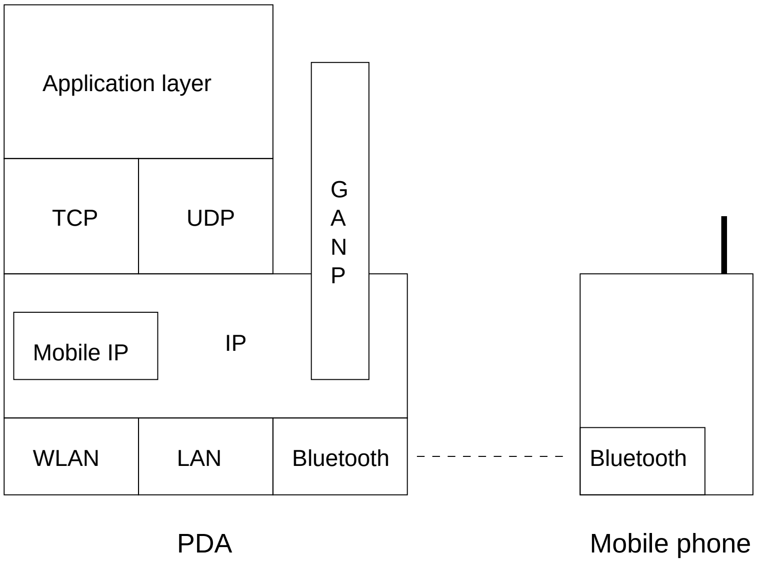

Figure 20: An example setup

When the user is at the office, he is using the office WLAN, without any kind of charging. The only restriction is to use IPsec ESP until the WLAN firewall to prevent unauthorized use (see Figure 21).

After the working day, the user leaves office and takes a train home. After leaving the WLAN coverage at the office, the Mobile IPv6 in the PDA decides to roam to the UMTS as there is no WLAN available.

The UMTS phone has a special SIM which works with the consortium of operators that support IP level mobility and access negotiation. The operating system of

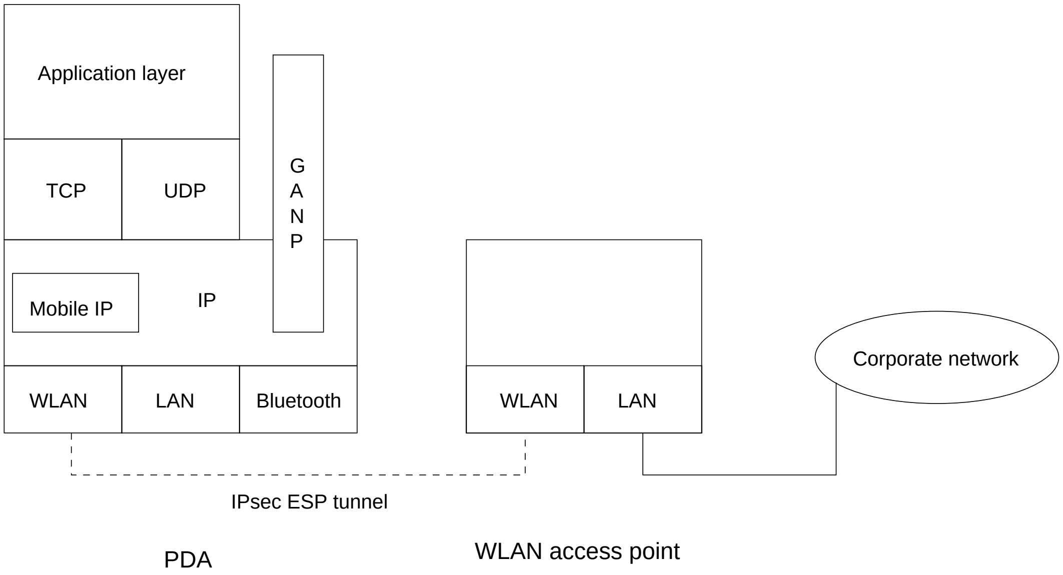

Figure 21: Using office WLAN

the PDA sees the UMTS device as any other network interface. It receives a Router Advertisement from the GGSN which configures the new care-of address and gives the pricing information and other parameters. After negotiating for the access with the PDA, the GGSN authorizes the mobile terminal with this particular SIM card to use IP services. A certificate stored in the SIM is used for authentication at the IP layer. Bob has a wallet of electronic money in his laptop, and it is used to pay for the UMTS access. The mobile node does not have to break any connections when roaming from the WLAN to the UMTS (see Figure 22).

After arriving at the central railway station, Bob decides to go for coffee. He chooses an Internet cafe that offers Wireless LAN access. When the software in the PDA detects that there is WLAN coverage by receiving a Router Advertisement, it decides on the policies set by the user, that it should roam to the WLAN network. Based on the billing information received in the RA, it starts setting up the billing state with the WLAN network while maintaining the UMTS connection until the negotiation with the WLAN is finished. When the WLAN access is ready to be used, the Mobile IPv6 in the PDA sends a binding update to the old router in the UMTS network telling the new location where to forward packets coming from correspondent nodes that have outdated mobility bindings. After that the general access negotiation protocol ends the billing at the UMTS network and Mobile IPv6 sends new binding updates to all correspondent nodes

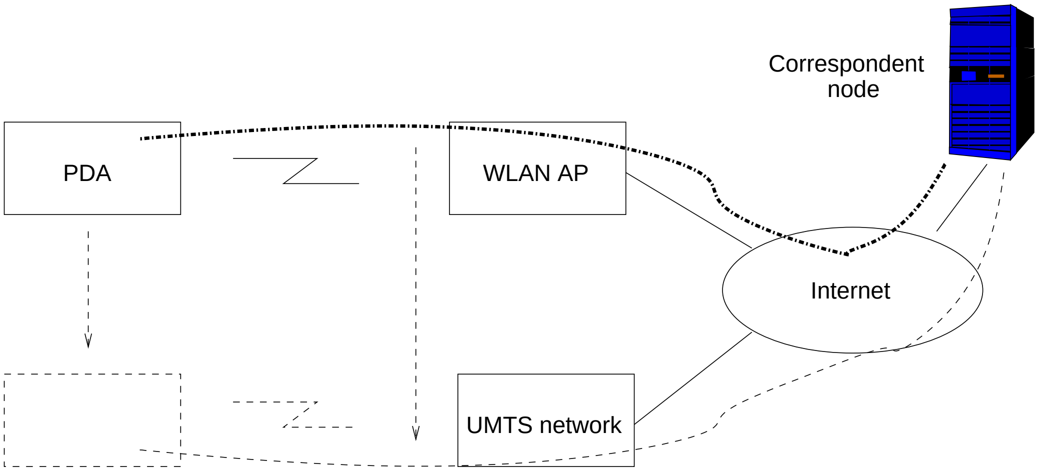

Figure 22: Handover from Office WLAN to UMTS

and the home agent to tell the new location at the Internet cafe's WLAN network. Packets belonging to existing connections start flowing through the WLAN network without interrupting the applications.

After leaving the Internet cafe, Bob roams again to the UMTS network in a similar fashion as described earlier, except that the billing in the WLAN can not be ended through WLAN, because the network coverage just disappears. However, the billing server in the WLAN network can be accessed through the Internet and the PDA closes the billing state by checking with the billing server that all the sent packets are being paid. Bob continues using UMTS until arriving home, where he plugs his PDA into the home LAN. Obviously there is no billing in the home LAN, so only regular Mobile IPv6 functionality is needed to start using the LAN. The billing with the UMTS network can be stopped after roaming to LAN.

6 Conclusions

In enabling mobile Internet, architectures of access networks play a major role. Access networks offer the users an access to the rest of the Internet. Mobile access networks usually offer wireless access and mobility within the network. Most popular wireless technologies were reviewed in the second chapter.

Larger scale Internet mobility requires that the mobility management is done on the IP layer, which is above the access network specific link layers. Using Mobile IP enables mobility between access networks. Mobile IP for both IPv4 and IPv6 were studied in the third chapter.

Almost all network protocols today contain a certain amount of cryptography. Therefore it is important to understand the basic technologies and what they enable. Cryptography was studied in the third chapter.

Methods for on-line payments were studied in the fourth chapter. Most interesting were E-cash and SET. These new payment instruments can be used to pay for the Internet services as well as for Internet access.

Large scale mobility also requires that we have good coverage of network access. This can not be achieved by relying on any single operator with which we have a contract. A moving mobile node goes through several types of access networks, some of which it might visit for the first time. Therefore it is necessary that there is a common framework protocol which can be used to negotiate access quickly, so that the Mobile IP can move the ongoing connections to the network that best suits the situation. The anonymous networking concept and a sketch for a General Access Negotiation Protocol were presented in the fifth chapter.

All the problems with Internet mobility are not solved. One of the biggest is the latency in handovers which affects real-time communication applications. This problem is recognized and worked on in the IETF. There is a suggestion for authentication in IPv6 networks [AFPE01], but it relies on a contactable AAA service in the mobile node's home network, and it does not support anonymous access. Some kind of generally accepted access/payment negotiation protocol that is built *into* the network layer is required before mobile Internet access can become an everyday thing.

References

3GP99 3GPP. 3G TS 32.105 V0.0.1 - 3G charging. Technical report, 3GPP, 12 1999.

AFPE01 N. Asokan, P. Flykt, C. Perkins, and T. Eklund. AAA for IPv6 Network Access. Internet draft, IETF, Jan 2001. draft-perkins-aaav6-02.txt.

AMSW97 B. Askwith, M. Merabti, Q. Shi, and K. Whiteley. Achieving user privacy in mobile networks. In *Proceedings of the 13th Annual Computer Security Applications Conference (ACSAC '97)*, pages 108–116. IEEE, 1997.

BH99 L. Buttyán and J. Hubaux. Accountable Anonymous Access to Services in Mobile Communication Systems. In *Proceedings of the 18th IEEE Symposium on Reliable Distributed Systems*. IEEE, 1999.

Bur00 D. Burdett. Internet Open Trading Protocol - IOTP. RFC 2801, IETF, Apr 2000.

Cha92 D. Chaum. Achieving Electronic Privacy. *Scientific American*, pages 96–101, August 1992.

con99 Bluetooth consortium. Bluetooth Specification Version 1.0 B - Vol. 1. Technical report, Bluetooth consortium, Dec 1999.

DH98 S. Deering and R. Hinden. Internet Protocol, Version 6 (IPv6). RFC 2460, IETF, Dec 1998.

Dro97 R. Droms. Dynamic Host Configuration Protocol. RFC 2131, IETF, Mar 1997.

Fen97 W. Fenner. Internet Group Management Protocol, Version 2. RFC 2236, IETF, Nov 1997.

HC98 D. Harkins and D. Carrel. The Internet Key Exchange (IKE). RFC 2409, IETF, Nov 1998.

HLFT94 S. Hanks, T. Li, D. Farinacci, and P. Traina. Generic Routing Encapsulation (GRE). RFC 1701, IETF, Oct 1994.

HP98 G. Horn and B. Preneel. Authentication and Payment in Future Mobile Systems. In *Proceedings of the 5th European Symposium on Research in Computer Security (ESORICS '98)*, 1998.

IEE99 IEEE. ANSI/IEEE Std 802.11, 1999 Edition. Standard, IEEE, 1999.

Jok99 P. Jokela. Wireless Internet Access Using Anonymous Access Methods. In *6th IEEE International Workshop on Mobile Multimedia Communications (MOMUC'99)*, Dec 1999.

JP00 D. Johnson and C. Perkins. Mobility Support in IPv6. Internet draft, IETF, Nov 2000. draft-ietf-mobileip-ipv6-13.txt.

KA98a S. Kent and R. Atkinson. IP Authentication Header. RFC 2402, IETF, Nov 1998.

KA98b S. Kent and R. Atkinson. IP Encapsulating Security Payload (ESP). RFC 2406, IETF, Nov 1998.

KA98c S. Kent and R. Atkinson. Security Architecture for the Internet Protocol. RFC 2401, IETF, Nov 1998.

KS98 B. Kaliski and J. Staddon. PKCS #1: RSA Cryptography Specifications Version 2.0. RFC 2437, IETF, Oct 1998.

MSST98 D. Maughan, M. Schertler, M. Schneider, and J. Turner. Internet Security Association and Key Management Protocol (ISAKMP). RFC 2408, IETF, Nov 1998.

Nar01 T. Narten. IESG security concerns with MIPv6. e-mail, IETF mobile-ip mailing list, Mar 2001.

NNS98 T. Narten, E. Nordmark, and W. Simpson. Neighbor Discovery for IP Version 6. RFC 2461, IETF, Dec 1998.

Pan96 P. Panurach. Money in Electronic Commerce. *Communications of the ACM*, 39(6):45–50, June 1996.

Per00 C. Perkins. IP Mobility Support for IPv4, revised. Internet draft, IETF, Sep 2000. draft-ietf-mobileip-rfc2002-bis-03.txt.

PJ00 C. Perkins and D. Johnson. Route Optimization in Mobile IP. Internet draft, IETF, Nov 2000. draft-ietf-mobileip-optim-10.txt.

Plu82 D. Plummer. Address Resolution Protocol. RFC 826, IETF, Nov 1982.

Pos80 J. Postell. Internet Protocol. RFC 791, IETF, Sep 1980.

Pos81 J. Postell. Internet Control Message Protocol. RFC 792, IETF, Sep 1981.

SCEMB01 H. Soliman, C. Castelluccia, K. El-Malki, and L. Bellier. Hierarchical MIPv6 mobility management. draft-ietf-mobileip-hmipv6-03.txt, IETF, Feb 2001.

Sch96 B. Schneier. *Applied Cryptography.* Wiley, 1996.

Sim95 W. Simpson. IP in IP Tunneling. RFC 1853, IETF, Oct 1995.

Sim96 D. Simon. Anonymous Communication and Anonymous Cash. In *Proceedings of Crypto '96*, 1996.

Tan96 A. Tanenbaum. *Computer Networks, Third Edition.* Prentice Hall, 1996.

tec00 Ecash technologies. http://www.ecashtechnologies.com. HTML page, Ecash technologies, Mar 2000.

TN98 S. Thomson and T. Narten. IPv6 Stateless Address Autoconfiguration. RFC 2462, IETF, Dec 1998.

VM97 Visa and Mastercard. SET Secure Electronic Transaction Specification Book 1: Business Description. Technical report, Visa and Mastercard, May 1997.

ZL98 J. Zhou and K. Lam. Undeniable billing in mobile communication. In *Proceedings of the Fourth Annual ACM/IEEE International Conference on Mobile Computing and Networking (MOBICOM '98)*, pages 284–290. ACM/IEEE, Oct 1998.

A 4th Generation Mobile Networks

This section describes the thoughts and ideas that were behind the DIA project at Ericsson Research.

Current phone networks based mobile networks are and will be outdated as the requirements for using Internet as the primary means of communication become more and more important. Phone networks are built on communication protocols that are designed for circuit switched connections and equipment that is expensive to build and difficult to administer. To bring broadband wireless access to as many locations as possible, it is necessary that the access part of the network is “lighter” than currently. It must be possible to set up a wireless access network by offering the radio access and all the services within the existing infrastructure somewhere on the Internet. Practically that means that in the future, a wireless access point should be connected directly to the Internet (Figure 23), or to a private IP-based network with an Internet gateway (Figure 24). There should also be no reason to differentiate Internet services from the services in a mobile phone network. All the mobile phone network services can be implemented with Internet techniques using standardized and proven techniques, for example DNS (Domain Name Server).

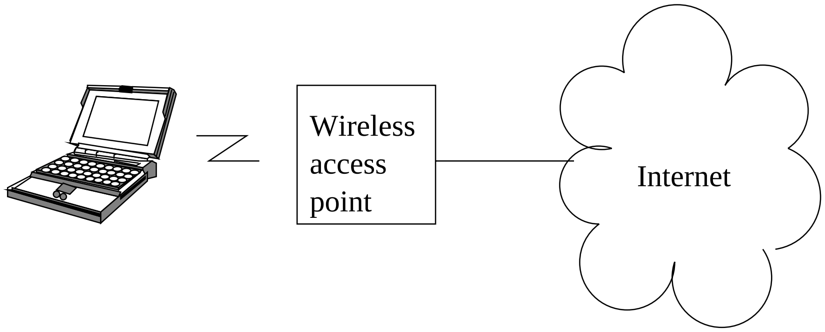

Figure 23: The most simple case of wireless Internet access

It can also be seen that the protocol stack, for example in GPRS (which looks effectively the same as it will in 3G packet service), is huge (see Figure 25). There are many protocol layers, the necessity of which can be questioned if aiming for offering the best possible packet switched connection for mobile nodes. The future (4G) networks will most likely be based on Internet protocols, and all services, like voice calls, will be offered as applications on top of internet protocol layers. Still open questions are how to support mobility, security, charging and quality of

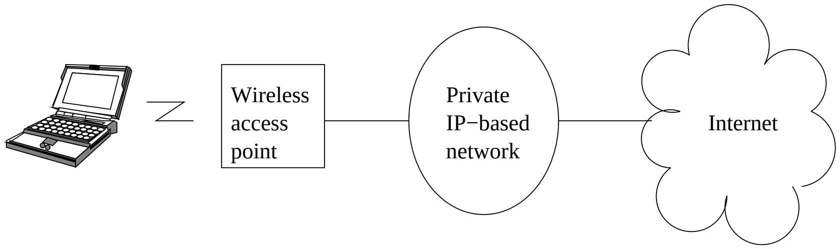

Figure 24: An IP-based core network

service in a multiaccess and multiprovider environment. Multiaccess means that the mobile terminal can have network access through many different link layers. For example, a future laptop computer may use wireless LAN while in the office area, and 3G mobile phone network while out of the office network. Multioperator means that roaming is between networks owned by different operators, but not necessarily between different types of access technologies.

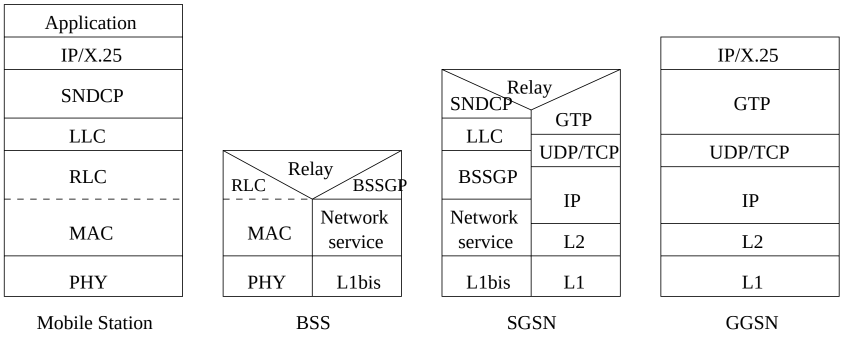

Figure 25: GPRS protocol stack

A.1 DIA project: An All-IP Approach

DIA (Direct Internet Access) was a project of Ericsson Research [Jok99]. The purpose of this project was to study the possibilities in replacing the UMTS core network with IP-based solutions (compare pictures 25 and 26 on pages 49 and 50). Another big difference between a "DIA" and "traditional" 3G networks is that DIA supports anonymous access. The mobile user does not need to be a subscriber to the access network. Following these two principles brings out new problems into the network architecture: what techniques, what protocols and

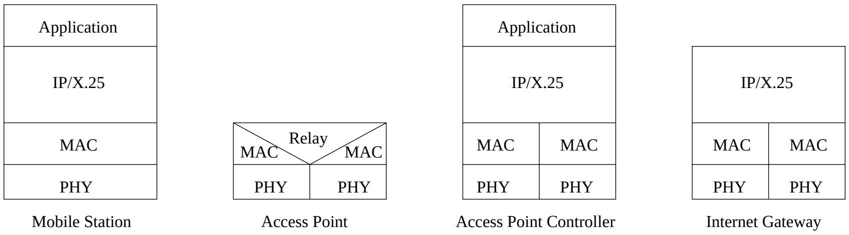

Figure 26: 4G protocol stack

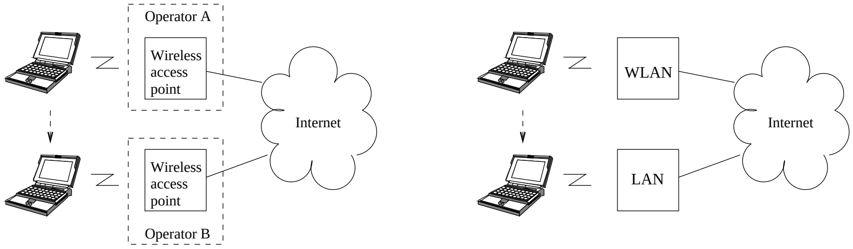

Figure 27: Roaming between multiple providers and multiple accesses

what kind of business models are required to support DIA networking. Having a DIA architecture available makes offering Internet access easier. The access network operator does not have to deal with telephone network agreements and protocols. Internet protocols are open and easier to implement than currently used SS7 (Signalling System 7).

In DIA it was decided to use the Internet Protocol version 6 (IPv6) as the basis for the network as it offers better security, larger address space and better support for mobility than IPv4. In IP networks, mobile nodes need to have a home address and therefore a home network. So the mobile user needs to have an ISP (Internet Service Provider). Mobile node receives a care-of address from the currently used access network. During roaming, new care-of addresses are allocated when necessary.

One of the main goals for the DIA project was to allow anonymous access to the Internet. Two separate methods to provide anonymous access were identified during the project. The first one relies on the existence of Home ISP nodes. The second one does not require any home ISP at all. It also enables a new mobility management method, the method of dynamically allocating a Mobile IP Home Agent for mobile terminals in the access networks.

According to initial decisions, anonymity would be achieved by using certificates, i.e. digitally signed and ciphered messages between mobile terminals and the home ISP (service provider). The user would remain anonymous in the access network if he only reveals his service provider's identity in the initial connection request. The access network is then able to contact the service provider and request for authentication of the user. Together with the service provider's identity, the user also sends a ciphered certificate to the access network in the initial access request message. The access network then passes the certificate to the service provider. The service provider is able to decode the certificate and validate the access request. It is then at the service provider's responsibility to authenticate and authorize the user. The access network and the service provider then negotiate on accounting issues and the access network may e.g. get an accounting permission up to some amount of money. This can be agreed on by using non-repudiable certificates. The user is able to connect to the Internet and to send and receive data after the authorisation phase. However, in the DIA model, the data is routed directly to the Internet. It must be emphasized that the data is not going through any service provider's nodes. So, it is not possible for the service provider to perform accounting by itself.

Another way to provide totally anonymous access is to use electronic money, e-cash. According to this scenario, e-cash is used to pay for the access. The access network does not need to know the identity of a roaming user since the user must pay immediately for the access. In this case, there is no need to contact the home ISP. In fact, there is no need for a home ISP at all.

According to this scenario, mobile users are allocated a temporary IP address that can be used as a home address. Together with the IP address, a home agent is also dynamically allocated in the access network. This home address can be stored for example in a Dynamic DNS (Domain Name System) server in the Internet. This enables other users to reach the mobile node. Other nodes in the Internet may look for a mobile node by its name, e.g. mymobile.ericsson.com. Standard address resolving methods (DNS query) can be used to retrieve a valid IP address for the mobile node.

When a mobile node roams within the access network or between two access networks, a care-of address is allocated to the mobile node. This care-of address is updated only to the dynamically allocated home agent (in the first access network). This means that the DNS server is not aware of the address change but it will always return the dynamically allocated home address as a response for a DNS query. So, when a correspondent node starts to communicate with the mobile node, the communication always starts through the dynamically allocated home agent if the correspondent node does not have an up-to-date mobility binding in its database.

A.2 Anonymous Access

Certificate based anonymous access was planned at the first stage. The solution would have been based on certificates implemented in another research project. However, during the year, a new concept was invented: anonymous access using electronic cash. There exists many different e-cash systems, one of them providing a completely anonymous payment method without a tamper proof device. This method (developed by DigiCash, now owned by Ecash Technologies Inc.) is based on blind signatures, which hide the user identity from the network. Even the bank cannot know to whom it has issued the money. It can verify that a coin is signed by itself, and it is not used before.

The DIA anonymous access method does not prevent using other means (e.g. non-anonymous e-cash methods, credit card numbers, etc.) to pay for the access. Depending on the e-cash system, it can even be used to pay for other services provided in the Internet. The method used here provides that kind of possibilities.

A.3 Mobility

Two aspects of user's mobility management were considered in the DIA project: macro and micromobility. Macromobility provides user location handling when the user changes access network and location. Micromobility handles user roaming within the visited network.

The first aspect, macromobility, is covered using a location server. This server can be for example a SIP (Session Initiation Protocol) server, providing current location information to nodes which are willing to call our mobile user. The used implementation of the location server in second and third prototypes was a dynamic DNS (Domain Name System) server. It was updated by the mobile user after getting the first address under the visited network. At this stage, authentication of the user was also performed between the user and dynamic DNS, but user identity was not revealed to the visited radio operator. During the last stages of the project, a real location server was implemented.

Roaming under the visited network when all existing connections have to be maintained cannot be handled by only using a location server. User location is resolved only once before a connection is established. This second aspect in mobility management, micromobility, is handled using Mobile IPv6. Mobile IPv6 relies on Home Agents (HAs) which are usually located at the home network of the mobile node. In the DIA solution, however, there doesn't exist a home network and the mobile node doesn't have a home IPv6 address. This means that we have to create a Home Agent at the visited network (actually at the first point of connection) and the address given at that point is used as Mobile Node's home IPv6 address.

A.4 AAA

The combination of an anonymous access scheme and mobility handling methods provides a new style of AAA. In traditional telecommunication networks Authentication is handled first. User identity is sent to the visited radio operator, which performs the authentication with help of the user's home operator. After authentication, the user can be authorized by the home network to use the visited network (all bills have been paid, etc.). After this, accounting is started.

In the DIA concept, sending e-cash to the visited radio operator starts the accounting part. After the visited operator has validated the e-cash from the bank, the user is authorized to use the network. At this stage, authentication is not necessarily performed at all. If the user wants to update his current address to a location server, he has to authenticate himself to that server. This final step accomplishes the AAA.

A.5 Security

IPsec was used to provide the needed security in the network. This solution can provide end-to-end security in communications, which cannot be achieved in current telecommunications systems.