Security technologies in home and wireless networking environments

Jan Lucenius, Timo Kyntäjä & Henryka Jormakka

VTT Information Technology

ISBN 951–38–6562–2 (URL: http://www.vtt.fi/inf/pdf/)
ISSN 1459–7683 (URL: http://www.vtt.fi/inf/pdf/)

JULKAISIJA – UTGIVARE – PUBLISHER

VTT, Vuorimiehentie 5, PL 2000, 02044 VTT
puh. vaihde (09) 4561, faksi (09) 456 4374

VTT, Bergsmansvägen 5, PB 2000, 02044 VTT
tel. växel (09) 4561, fax (09) 456 4374

VTT Technical Research Centre of Finland, Vuorimiehentie 5, P.O.Box 2000, FIN–02044 VTT, Finland
phone internat. + 358 9 4561, fax + 358 9 456 4374

VTT Tietotekniikka, Tekniikantie 4 B, PL 1203, 02044 VTT
puh. vaihde (09) 4561, faksi (09) 456 7053,

VTT Informationsteknik, Teknikvägen 4 B, PB 1203, 02044 VTT
tel. växel (09) 4561, fax (09) 456 7053

VTT Information Technology, Tekniikantie 4 B, P.O.Box 1203, FIN–02044 VTT, Finland
phone internat. + 358 9 4561, fax + 358 9 456 7053

Technical editing Marja Kettunen

Published by

Series title, number and report code of publication

VTT Working Papers 10
VTT–WORK–10

Author(s)

Lucenius, Jan, Kyntäjä, Timo & Jormakka, Henryka

Title

Security technologies in home and wireless networking environments

Abstract

The purpose of this document is to describe selected security technologies and applications related especially to home networking environments. The focus is on Wireless Local Area Network technologies. The document defines the home environment and addresses related architectural issues. The security technologies are depicted from the viewpoint of different communication layers. Finally, aspects on authentication, wireless roaming, and remote access to home network are covered. An architecture enabling home access to a roaming user, as well as WLAN market possibilities, is also discussed.

Keywords

home networking, WLAN, security threats, authentication, remote access

Activity unit

VTT Information Technology, Tekniikantie 4 B, P.O.Box 1203, FIN–02044 VTT, Finland

ISBN		Project number
951–38–6562–2 (URL: http://www.vtt.fi/inf/pdf/)		
Date	Language	Pages
October 2004	English	49 p.
Series title and ISSN	Publisher	
VTT Working Papers 1459–7683 (URL: http://www.vtt.fi/inf/pdf/)	VTT Information Service P.O. Box 2000, FIN–02044 VTT, Finland Phone internat. +358 9 456 4404 Fax +358 9 456 4374	

Contents

1. Introduction 9
   - 1.1 Definition of Home Network 9
   - 1.2 Network Architecture and environments 10
   - 1.3 Home Network Technologies 11
     - 1.3.1 IEEE802.3 (Ethernet) 12
     - 1.3.2 IEEE802.11 (WLAN) 13
     - 1.3.3 Other Wireless Technologies 14
   - 1.4 Access 14
     - 1.4.1 Local access 14
     - 1.4.2 Remote access 15
   - 1.5 Security Threats 16
     - 1.5.1 Threat against confidentiality 16
     - 1.5.2 Unauthorized access 17
     - 1.5.3 Man in the middle 17
     - 1.5.4 Availability 18
     - 1.5.5 Trojan Horses, Backdoors, etc 19
     - 1.5.6 Legal aspects 19
   - 1.6 Security Requirements 20

2. Security Solutions on Different Layers 21
   - 2.1 Link Layer 21
     - 2.1.1 IEEE802.11i 21
     - 2.1.2 801.1x 21
   - 2.2 Network Layer 23
     - 2.2.1 IPSec 23
     - 2.2.2 IKEv2 24
     - 2.2.3 Extensible Authentication Protocol (EAP) 24
     - 2.2.4 Host Identity Protocol 25
     - 2.2.5 Other IP security protocols 27
     - 2.2.6 Network Address Translator (NAT) 27
   - 2.3 Transport and Session Layers 28
     - 2.3.1 SSL 28
     - 2.3.2 SSH 29

2.4 Middleware Layer 31
2.4.1 Secure HTTP and HTTP over SSL 31
2.4.2 SIP security 31
2.5 Application Layer 33
2.5.1 Java Security 33
2.5.2 Content Protection 33
3. Authentication, Authorization and Accounting 37
3.1 RADIUS 37
3.2 Diameter 38
4. Wireless access and roaming architecture 39
5. Remote access to home devices scenario 42
Conclusions 46
References 47

Acronyms

3GPP	3rd Generation Partnership Project
A/V	AudioVisual, Audi/Video
AAA	Authentication, Authorization and Accounting
AH	Authentication Header
AKA	Authentication and Key Agreement
ALG	Application Layer Gateway
AP	Access Point
ATM	Asynchronous Transfer Mode
AVP	Attribute Value Pairs
BT	Bluetooth
CA	Certification Authority
CCI	Copy Control Information
CHAP	Challenge Handshake Authentication Protocol
CPRM	Content Protection for Recordable Media
D-H	Diffie-Hellman
DNS	Domain Name System
DoS	Denial of Service
DRM	Digital Rights Management
DTCP	Digital Transmission Content Protection
DVD	Digital Versatile Disc
EAP	Extensible Authentication Protocol
EAPOL	EAP Over LANs
EMI	Encryption Mode Indicator
ESP	Encapsulated Security Payload
FDDI	Fiber Distributed Data Interface
FQDN	Fully Qualified Domain Name
FTP	File Transfer Protocol
GPRS	General Packet Radio Service
GSM	Global System for Mobile Communications
HIP	Host Identity Protocol
HTTP	HyperText Transfer Protocol
HTTPS	Hypertext Transfer Protocol over Secure Socket Layer, HTTP over SSL
HVAC	Heating, Ventilation and Air-Conditioning

IEEE	Institute of Electrical and Electronics Engineers
IETF	Internet Engineering Task Force
IKE (IKEv1, IKEv2)	Internet Key Exchange
IMS	IP Multimedia Core Network Subsystem
IPSec	Internet Protocol Security
ISO	International Standard Organization
ISP	Internet Service Provider
JVM	Java Virtual Machine
LAN	Local Area Network
LEAP	Lightweight Extensible Authentication Protocol
LLC	Limited Liability Company
MAC	Media Access Control
MD (MD5)	Message Digest
MITM	Man-In-The-Middle
MKB	Media Key Block
NAPT	Network Address and Port Translation
NAS	Network Access Server
NAT	Network Address Translator
OSI	Open Systems Interconnection
PAE	Port Access Entity
PAP	Password Authentication Protocol
PDA	Personal Digital Assistant
PDU	Protocol Data Unit
PEAP	Protected Extensible Authentication Protocol
PPP	Point-to-Point Protocol
PSK	Pre-Shared Key Mode
PVR	Personal Video Recorder
RADIUS	Remote Authentication Dial-in User Service
RF	Radio Frequency
RTCP	Real-Time Transport Control Protocol
RTP	Real-Time Protocol
RTSP	Real-Time Streaming Protocol
SA	Secure Association
S-HTTP	Secure Hypertext Transfer Protocol
SIM	Subscriber Identity Module
SIP	Session Initialization Protocol

SLIP	Serial Line Interface Protocol
SPI	Secure Parameter Index
SPS	Secure Packet Shield
SSH	Secure Shell
SSID	Service Set IDentifier
SSL	Secure Socket Layer
STCP	Stream Control Transmission Protocol
TCP	Transmission Control Protocol
TKIP	Temporal Key Integrity Protocol
TLS	Transport Layer Security
TTLS	Tunneled Transport Layer Security
UA	User Agent
UAC	User Agent Client
UAS	User Agent Server
UDP	User Datagram Protocol
UMTS	Universal Mobile Telecommunication System
UPnP	Universal Plug-and-Play
URI	Uniform Resource Identifier
USB	Universal Serial Bus
USIM	UMTS Subscriber Identity Module
WAN	Wide Area Network
VCR	Video Casette Recorder
WEP	Wired Equivalent Privacy
WiFi	Wireless Fidelity
WISP	Wireless Internet Service Provider
WLAN	Wireless Local Area Network
VoIP	Voice Over Internet Protocol
WPA	Wi-Fi Protected Access
VPN	Virtual Private Network
xDSL	(Type x) Digital Subscriber Line

1. Introduction

1.1 Definition of Home Network

In this report, home network means a collection of electronically controlled, interconnected home devices available for the inhabitants of the home and, in some cases, also for other persons. Examples of devices include TV, surveillance camera, and PC with peripherals. More detailed examples are presented in the next section. By means of the home network it should be possible to control, manage and use the devices at home remotely, even over the Internet, or use some services e.g. provided over the Internet. Use scenarios include for example:

- switching on the sauna in advance before arriving home
- controlling a VCR or PVR from an other room or even over the Internet
- sending alarms from home to the inhabitant(s)
- watching a surveillance camera remotely
- remote access to data or media stored in the home PC
- updating software to an intelligent device at home
- chatting between family members and friends independent of their location
- accessing a family calendar (remotely or at home)
- remote monitoring of the temperature etc.

All the services need not necessarily to be implemented in the home computers and devices themselves, but may be distributed or provided by an external service provider, such as an ISP, service company, or community. Because of fast, fixed connections and improving mobile Internet connections, this does not necessarily matter from the usage point of view; internal and external service components integrate seamlessly. For simplicity, however, we usually assume further in this document that the services in the home network are even physically implemented at home, and handle external services like internet services in general (even if they may be of local nature or the access is restricted).

As usual, openness is, also in the case of home networking, a clearly contradictory goal to security. As the networks make life easier in many ways for the inhabitants, this may be at cost of possible misuse. In the worst case an Orwellian scenario where it is possible to track everything from devices existing at home to get more exact knowledge of the family members' activities may be realized. This is why security is so important.

The report presents an overview of existing state-of-the-art security technologies which can be utilized to provide secure access to the home network and secure communication with the home devices. Some of the technologies are included in a scenario where a user accesses his home network devices from a public WLAN provider with whom the user does not have pre-agreement. Additionally, roaming architecture and external WLAN access market possibilities are discussed.

1.2 Network Architecture and environments

In order to understand the security needs of home networking and discuss an adequate framework, we will first present an overview of what a future home may look like.

The network technology itself may be based on Ethernet type LAN (wired or wireless) which connects computers, and thus may be very similar in structure to an office network. Particularly, if some of the family members are private entrepreneurs, the same infrastructure will likely be used both for home and business applications and it may thus not even be possible to distinguish between those two. Even if none of the inhabitants is self-employed, but there are several computers sharing internet access, those are likely interconnected by means of a LAN.

The home network may also use other carriers, with proprietary technology as well, particularly in the case of off-the-shelf solutions targeted for some restricted specific purpose. In this report we focus mainly on the LAN-based approach with Internet protocol suite. Proprietary solutions, and solutions based on other network technology than IP are not studied in detail, and it is further assumed that these are handled through a home server or other appropriate gateway. Connections to dumb devices, sensors, peripherals etc. usually belong to this category.

Even if the network is similar to a small office network, there are some typical home applications, which are slightly different from typical business applications, seen from the users' point of view. Figure 1 shows a view of a possible future home and networks connected to it.

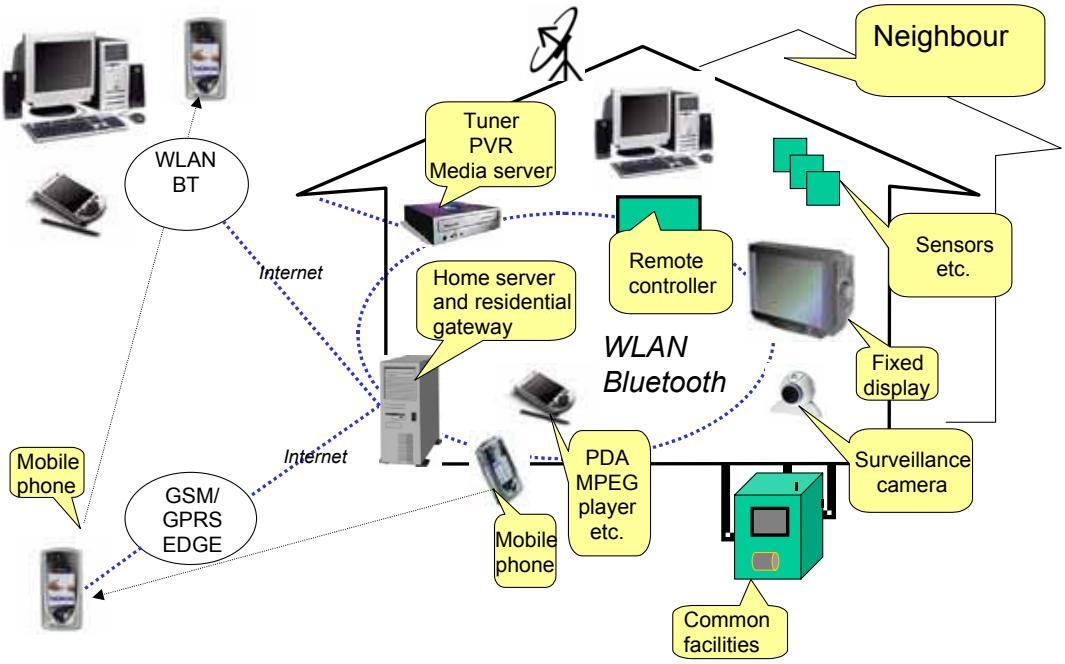

Figure 1. The environment, equipment and networks in connection to a future home. Wireless connections are assumed for the local connections.

The Home server provides a gateway from the Internet to the home network. It is also the central part of the network, and it may contain the home firewall and the network address translator (NAT). Additionally other functionality supporting the home devices may directly be implemented in it, as most of the devices for a long time may remain quite simple. The home server and residential gateway device may be the home PC, but may possibly be also embedded equipment. Of course there may additionally be other PCs.

Very central role in the home network plays the consumer electronics, consisting of e.g. TV, radio, portable players of different sizes and capabilities, Personal Video Recorders, or media servers (DVD & CD players, VCRs , etc.). The devices may be directly connected to PC, or stand-alone using remote IR controllers; PDAs and mobile phones equipped with Bluetooth or IR may however in some cases replace traditional, dedicated remote controllers in the future.

Not all possible devices are shown in the above figure. For example lamps, white goods etc. may be connected to the home network. The box named "common facilities" stands for equipment which is not specific for only one home, but may be shared, such as central heating and other HVAC (Heating, Ventilating, and Air-Conditioning) systems, yard lights etc. A person living in the home may not have rights to control these: as it may have unwanted consequences for the neighbours, these are possibly controlled by authorized service personnel, for example service company staff or a janitor.

Also access to other equipment in the home, like alarm system, or equipment used for surveillance of sick or elderly people, may be in some cases available to external persons, but it is important here to restrict the users' access rights to prevent unauthorized activities. Additionally, sensors (see Figure 1) may be used to read the state of some material goods, or provide information about different events. Some examples:

- A sensor may alarm if there is a water leakage or a pipe has got blocked.
- A remotely readable label on a food package indicates if the "best before" –date is expired or approaching.
- A remotely readable key card opens the door without need for the family member to use traditional keys.
- A tag on an animal collar unlocks a small door, enabling the animal to walk in.

The sensors do not have any processing power, but the information which they contain may be read, possibly by specific readers connected to the home network. However, secure communication, at a minimum verification of access rights, may be needed also for (remote) access to sensors, card readers, etc. as confidential information otherwise may be disclosed or forged. For example, if the main door is controlled through the home network, a stranger could impersonate the lock mechanism, and open the door.

1.3 Home Network Technologies

This chapter briefly summarizes the common standards that provide the connectivity for LAN-like networks. The focus is in wireless networks. At the end of the chapter there are also shortly mentioned some of the standards that are yet to come.

Protection in the link layer may be necessary for wireless communications, where the link is easily accessible by outsiders. An exception may be a scenario where network layer security is used instead, and unprotected network traffic is either completely disabled or restricted only to isolated applications and/or sites where there are not any considerable threats. In other words, such local networks are very restricted. For example in a home network using unprotected wireless communication, internal addresses may at least be revealed to outsiders. Let us consider for example an intelligent device in a home network which accepts only commands on a protected connection. If this protection is implemented in higher layers and link layer data is transferred in cleartext format, it would still be practically impossible for an outsider to send commands to the device, even if link layer is unprotected. However, it would be possible to cause e.g. DoS attacks by inducing lots of data or invalid data into the link, or by duplicating the MAC address of the device. If IP addresses are revealed, DoS attacks may also be directed towards these specific addresses. Likely, such an attack may cause the device to crash. Encryption may also be solved by brute force, and by guessing the plaintext, if enough data is collected.

On the other hand, link layer protocols apply only to one link at a time; on the next link the traffic is again unprotected. For example encryption on a WLAN or GSM network is decrypted at the base station. In the fixed network, the data is unprotected. The choice, whether to use link layer or higher layer encryption, if any, is a tradeoff between risks and performance, is it likely that someone would attack the network, and what is the worst case scenario if this happens? As long as it is very unlikely that this could cause any harm, one may of course think about using network layer security instead.

In the worst case scenario we still need link layer security protocols in combination with protection mechanisms on higher layers (e.g. IPSec, SSL, etc.)

1.3.1 IEEE802.3 (Ethernet)

Local area network (LAN) is a network where different independent computers communicate with each other using a shared medium. The far most used technology today is Ethernet (others: Token Ring, Fast Ethernet, Fiber Distributed Data Interface (FDDI), Asynchronous Transfer Mode (ATM) and LocalTalk). The size of the LAN (i.e. the number of the computers) can vary between a few and hundreds.

The speed of an Ethernet LAN is 1–10Mbits/s. For faster LANs the Fast Ethernet standard (IEEE 802.3u) has been standardized, which makes it possible to run the LAN with 100Mbits/s speed. Some of the wide area networks (WAN) use even faster technology: Gigabit Ethernet.

Ethernet does not provide mechanisms for security. However, it being a wired media the access is physically restricted. When the home network is (partly) based on Ethernet, the network should be separated from the ones e.g. of the neighboring apartments. The access to this local area network should always be restricted.

In this document the focus is on the wireless standards. As there are a lot of similarities with these wired and wireless standards, one should remember existence of both of them. More detailed information can be acquired from the web; see e.g. [EthernetTutorial].

1.3.2 IEEE802.11 (WLAN)

The IEEE 802.11 standard has been developed by IEEE P802.11 working group that is standardizing Wireless Local Area Networks (WLAN). The standards cover physical and MAC layers of the ISO/OSI reference model. These standards can be compared to IEEE802.3 standard for Ethernet wired LANs. There are several products available on the market, but the standardization work continues, especially with the security issues.

From the user point of view a WLAN consists of Access Points and clients having a wireless access card. The access point(s) needs to be connected to a wired LAN infrastructure at some point of the network, in order to gain access e.g. to the Internet. The clients of the WLAN can get access to the network using either static or dynamic configuration of the networking parameters.

Below the most significant parts of the standards are summarized. More detailed information can be accessed from [IEEE802.11], and the working group has also its own web pages [IEEE802.11wg].

IEEE802.11a

The a-part of the standard is an extension of 802.11 that applies to wireless LANs and provides up to 54 Mbps in the 5GHz band. The first products were introduced late 2002. This standard is the main competitor of the European standard HiperLAN/2. To accommodate this standard for the European markets there exists 802.11h, which focuses on the physical properties of 802.11a.

IEEE802.11b

The b-part of the standard was the first WLAN standard that brought the products available for customers. It is also known as WiFi, Wireless Fidelity. The standard is designed to work using 2.4 GHz frequency and provides up to 11Mbps bandwidth.

While the standard introduced the easiness of establishing WLANs, so far it has not been successful in capturing the markets. The problem has been the lack of security, including authentication of the users of a WLAN. It has also been proved that its encryption mechanism (wired equivalent privacy, WEP) is not strong enough to provide privacy for the users (see also Security: 802.11i, and 802.1x)

IEEE802.11g

The g-part of the standard improved the physical layer of the b-part, so that up to 40Mbps bandwidth could be gained. The products available are also backward compatible with the b-standard products. This means that g-standard access points can support access for b-standard clients. Unfortunately, this will cut down the maximum bandwidth for g-clients to half.

Summary

Local area networking is the main technology used in office environments. Similar networking can be applied for home and factory environment when especially IP-communication is needed. There are further requirements for these standards when considering security issues (e.g. several overlapping WLANs in an apartment house).

The power requirements of the existing products are quite high. The client cards when used e.g. with PDAs use too quickly the capacity of the batteries. The electronics manufactures are producing highly integrated solutions where e.g. the processor and WLAN radio are on one chip. These products are aimed for laptop and PDA devices.

1.3.3 Other Wireless Technologies

Bluetooth was designed to offer a small size and affordable solution for short distance networking connections. The system radio operates on the same spectrum as IEEE802.11b/g. However, it seems that the implementation of the standard for products is not cost effective enough. However, there are products offering Bluetooth connectivity for selected application fields. The most common exploitation of it today is using mobile phones connected with audio or picture reproduction devices. Such solutions are targeted just to replace a wire between two devices; not to connect multiple devices together to form an actual network.

HiperLAN2 has been a European effort to standardize interoperable wireless local area networks. Unfortunately, the standardization work did not advance within the planned schedule. Today IEEE802.11 products are widely available for consumers, when HiperLAN2 first chipsets are produced for other device manufactures. Nevertheless, this standard is offering large guaranteed bandwidth for wireless device communication. Still, the first target of these products is to replace a wire, not to establish a network.

1.4 Access

1.4.1 Local access

By local access, we mean in this document, accessing the home network devices without passing through public networks, like the Internet, public fixed phone network, or GSM network.

Networks in the home environment (inside home and its immediate surroundings) may be wired or wireless. There will not necessarily be available a traditional Ethernet or other dedicated networks, but existing wiring, like power wiring, telephone networks or antenna networks may be utilized to implement a local network (in addition to the use of these lines for public network access, or their original use). For example Homeplug and X-10 use the power wiring.

The most probable is the scenario when the home network contains a mixture of the transmission media. From access point of view, a dedicated wiring is of course the most secure, whereas wireless systems (WLANs, BT and other RF systems) and other wiring may be subject to potential eavesdropping and/or unauthorized access. Especially in blocks of flats and other closely built homes, even unintentional unauthorized access may be possible, especially in auto configuring networks. How can for example new equipment that is brought into the home know whether it is accessing the correct network or the neighbor’s one? This is also the situation where Bluetooth or similar systems are used for interconnecting a mobile phone and a computer. If an intruder gets access to a local wireless LAN (or similar), he may – in principle – see all the traffic. Furthermore, this makes the system vulnerable to all kinds of security attacks.

Local infra red controls are suitable for such handheld devices that are mostly residing in one room and control devices in that surrounding. Nevertheless, IR related issues are deliberately left out of this study.

1.4.2 Remote access

The distinction between local and remote access is not always quite clear, especially as wireless techniques, antenna networks, power lines etc. may be used for implementing both, Internet access and local networks. To make the study simpler, we will assume that remote access is arranged through the Internet or possibly by more conventional means, like phone or SMS, i.e. far access (the latter, however, are of less interest for our study). For example, a person using an internet connection through a GSM phone to switch on his sauna uses remote access, even if he or she is physically located in the garden in front of the home. Reverse, unauthorized access from the neighbor through the power lines, is categorized as local access, if the signals do not pass the operator's facilities.

The security problems may be mirrored, i.e. the same kind of problems may exist at the site from which remote access occurs.

Authorized remote Internet access (in the cases of interest) is made to or from the home through a line or link provided by an internet operator. As home networking also needs inbound access, it is essential that the operator arranges the required channels for accessing the home from the Internet. This may sound self-evident, but for example with current xDSL connections, the home computer(s) are often hidden behind a NAT, and may thus have internal addresses which are not visible in the Internet. Furthermore the addresses may be dynamic, at least the addresses which are used on the Internet side of the NAT. In spite of that, crackers usually find their way to the computers, unless firewalls are used.

According to experience, a computer does not need to be connected to the network more than 15 minutes before someone is trying to attack it. The real home user has to know the name or address of his own home server (or other resources if available) in the way it appears on the Internet side of the NAT (Network Address Translator, see section 2.2.6). Especially if the Internet address is dynamic, the user usually wouldn't know it, unless it is sent to him somehow. Remote access could be provided by means of Session Initiation Protocol (SIP, see section 2.4.2) in a secure and standardized way, but of course other solutions, too, are possible. Some ISPs support static addresses as well. Furthermore, operators may currently explicitly forbid customers to set up some type of servers of their own.

Clearly the access line should be protected by a firewall (or similar) at the home network border. Where the firewall is placed, depends on whether others may access the line or not. As a rule-of-thumb we may state that this should be placed inside the home, but if for example a telephone line or – even better – an own fiber is used, the firewall may be provided by the operator. In some cases it might help if the firewall is integrated with the gateway between internet and the local network, for example in the xDSL modem or cable modem, or in the case of company networks, of course with the router. In some cases, for example if there is only one computer with direct internet access, the firewall may also be implemented by software.

Regardless of how the firewall is implemented, proper configuration of it is essential: how to pass the firewall for authorized access but prevent unauthorized? In ideal cases, only ports for necessary protocols should be open for inbound access. Further, these protocols (enabling access) should either provide protection mechanisms themselves, or otherwise use secure channels.

The networks are also accessed from mobile terminals. The mobile terminal may also have several network interfaces, e.g. WLAN, wired LAN, GSM data, GPRS, Bluetooth. In the ideal case, the terminal should perform a seamless (vertical) handover between these networks, without the user having to reestablish the connection and re-authenticate. The user would then only notice a change in the quality of service.

Some technical arrangements also allow access to secured services from public workstations, for example from libraries and internet cafes. The security requirements for those are clearly stricter and/or the accessible services are more restricted, as public workstations cannot be tailored for access to some specific corporate or home LANs, like personal workstations. In practice, public computers always use some version of Microsoft Windows operating system and the Internet Explorer web browser, which means that the only way to access a service securely from a public computer, is a secure web connection. Access tools and protocols would thus have to be standard on this environment, if this kind of public access is needed. Another question is whether the public computers themself can be considered trusted enough.

1.5 Security Threats

1.5.1 Threat against confidentiality

Confidentiality means that information is not disclosed to unauthorized parties. This means both information in transit and stored information. Although traditionally, confidentiality, meaning that information is encrypted, is not the most important security service, confidentiality threats are at least

- eavesdropping
- reading private information.

As such, these are passive threats, which have no effect as far as the intruder does not forward the information or misuse it. However, in the case of home networking, information may reveal facts about persons and systems in the accessed or accessing networks, which may be misused, and thus, effort has to be put also to ensure confidentiality as far as it's feasible.

Eavesdropping is especially a problem in wireless communication (particularly in wireless LANs), if traffic is not encrypted. It is also possible if protection is too weak, for example using a fixed encryption key which can be found out using brute force, but even this makes eavesdropping much more difficult. Also phone lines, common antenna systems etc. can be tapped, but is more difficult, and thus less probable, since it requires special equipment. Wired LANs are also subject to eavesdropping but those usually span only the local environment.

In the Internet environment, eavesdropping of traffic may be performed especially by espionage organizations. Stored information, in mailboxes, proxies etc. on servers operated by external parties, may be read and misused. Even if we are used to trust the traditional network operators and ISPs, we may in future have untrusted, unreliable hot spot operators etc. Currently, most traffic is unencrypted just because we trust that in any case nobody is interested in what we are doing, as long as it isn't in any way criminal or in the interest of super powers.

When accessing services from public workstations, there is a risk that cached data and even personal information, like passwords, may be disclosed to outsiders. Also, social engineering, like watching when the user enters information is more likely to occur in a public place, either in real time or from video. Users are taught to clear the cache, history etc. after use, but they may forget, or not have time to do so. The best arrangements for this is made in the Opera browser (Windows and Linux versions), where confidential data can be cleared with one command. Unfortunately this is rarely used in public places. The default settings of the browsers may not be ideal from security point of view, for example Netscape 7 remembers passwords by default.

1.5.2 Unauthorized access

Access can be divided into several layers or categories:

- access to the local network, remotely or locally;
- Internet access (outbound);
- access to a specific computer or equipment;
- access to some resource, like file, program, database, shared disk etc. on the computer or equipment.

Protection against unauthorized access is different for each of the different categories above. Internet access, as well as inbound access from Internet to local network, is both a matter of accessing one network from the other, and protection should thus be implemented on the border between these networks. Usually this means passing a firewall and also possible checking of the user's or the computer's rights for the network access.

Local access to a LAN (especially wireless) is subject to the same risks as explained in the previous section.

If the intruder gains access to the local network, he/she may find out what computers and (native IP) equipment are connected to the network, and once the computer is found, the intruder may try to use it. Protection against this may be more difficult as each of the internet applications and daemons running on the computer may form a possible entry point to the computer, and all these possible channels are not even known. Depending on the internal configuration of the computer, permitted channels may be misused to enter programs (like agents, applets, Active-X controls, macros, etc.) which in turn open forbidden channels. Now, the fourth category of access threat, access of some resource internal to or coupled to the computer, is possible.

Another way to get unauthorized access to the local network is through computers and other resources which either are not protected by passwords (or stronger methods) or use easy-to-guess passwords, like manufacturer default passwords. Suggestions for choosing good passwords are not in the scope of this report, but presenting existing methods of authentication, which eliminate the need for separate password entering on every machine, is.

Traditionally, access rights are verified against the identity of the user, using a password or some other key. This is however, not the only way, the access rights may as well be based on authorizations or roles. This can be achieved using attribute certificates. The user could then be anonymous, but must instead prove the ownership of the right to perform the intended task, almost analog to having money loaded on a card.

As mentioned in the previous section, cached data, and especially passwords stored by password managers, may be disclosed to outsiders. These are of course not only a threat against confidentiality, but may enable unauthorized access as well.

1.5.3 Man in the middle

The man-in-the-middle (MITM) threat means that an intruder actively takes part in the communication and modifies the data exchanged between two parties (A and B), identifying

itself as B to A and A to B. Assume that B is a server. Stealing a session and possibly excluding user A from it, the intruder may continue to misuse services, with B still believing that it is communicating with A. Depending on the protocols and algorithms used, secrets like for example passwords may be revealed to the intruder.

If the MITM intercept clear text protocols, especially if these are not protected, it might be relatively easy for him to cause denial of service (see below), or otherwise harm the partners. Especially where proxies and other intermediate servers are used, like in SIP and HTTP, the MITM may act without the users noticing it. For example a false BYE or CANCEL message may end a SIP session or prevent its establishment [Kullenwall]. A false 401 (Unauthorized) response may cause the user to reveal his username and password.

Strictly speaking, we cannot trust system managers not to read workers email messages, monitor traffic in a way which breaks data confidentiality, or at least message flow confidentiality. Some amount of network monitoring may even be made just for network maintenance purposes. Generally, we just have to live with those threats, and trust that the system managers don't read, or at least don't reveal unauthorized information, perhaps avoid for example to use unprotected email for very confidential information. In cases where there are very strict confidentiality requirements, however, additional protection is necessary.

1.5.4 Availability

Availability means that the system works correctly and thus is able to provide the service as intended. The system may fail because of accidents, and mostly because of "natural" reasons like power supply breaks, broken cables or loose sockets, and of course hardware and software crashes. Hardware and software crashes may be more common in a home or small office environment than in average. This is due to the fact that the system may not be managed as well as in larger networks, usage might be less restricted, users are less educated (for example children) etc.

From IT security point-of-view the most common threat against availability is called Denial of Service (DoS) attack. This means as the name says, preventing a system from providing its expected service. If not just shutting down the server or some network device through which the service is accessible, a typical attack may consist of overloading the system or send incorrect packets, which may cause the server (or some network device) to crash or slow down. The most typical examples are perhaps filling up a mailbox with spam, and different kinds of packet flooding (meaning generating more traffic than the network can handle). For example in ping flooding, the device receiving the ICMP packet answers them, doubling the amount of traffic. Intruders may also induce traffic at the MAC layer in LANs, especially into WLANs, which consume bandwidth even if they do not contain valid higher layer data.

DoS attacks may be performed by a MITM, for example by session stealing or sending false messages which may break a specific session. The attacks may be performed using multiple compromised systems, which all attack the same target. This is called Distribute Denial of Service (DDoS).

In networks with fairly slow connections to internet (community, home and company LANs), a user may cause denial of service for other users by consuming most or all of the connection by his own ones, by using streaming media or transferring very large files. A solution may be to

restrict the bandwidth available for each user. This is usually implemented in commercial networks, where each subscriber is charged based on the subscribed bandwidth.

1.5.5 Trojan Horses, Backdoors, etc.

Even if we put a lot of effort on the basic security services (authentication, integrity and confidentiality of data in transit), the data stored in the computers may still be unprotected because of Spyware, Trojan horses or backdoor programs, which may enter the home through email or as add-ons on downloaded or installed files. These are clearly a threat against confidentiality and integrity of the system. It is also possible that software, which is assigned for smart updates, installations, license handling and of course software to share music files, etc. may be misused, or even intentionally equipped with Trojan horses. Even virus protectors like F-secure Antivirus may not be immune against these, as they are not of commercial policy reasons classified viruses. For protection, other tools, like AdAware should be installed [Tech-spot]. Additionally, browsers and other communication software may launch applications automatically. Also email may be used (indirectly) to open a channel for the intruder. It is well-known that email attachments may include viruses. If emails are expressed in html format, unexpected things may happen when the email is read. Communicating program may also notice that for handling a file, it needs additional plug-ins or other components. Particularly this is true for A/V applications, but also e.g. UPnP control points may need to retrieve a description of some device before they are able to process further actions, or a software agent for doing automatic business exchange may need to download protocol profiles and other schemes in order to perform its task. These downloaded components are not only passive data, but parts, or at least instructions for execution of the software.

These techniques certainly make systems more usable and make it possible to expand the capabilities of programs, avoiding the need for a totally new version or a lot of small separate programs. Again, this openness may be a risk from security point of view. There is a risk that the plug-in in fact does something else than the user expects, i.e. it may include a Trojan horse or a virus. Often plug-ins are signed, but these methods are either too expensive, too heavy to use, or badly managed or established. These causes that users probably ignore them, either because they don't care or they don't know what else to do.

Generally, it is likely that the more popular the platform is, the more security attacks it will experience. If the intruder program is capable of spying passwords, it is evident that protecting the network itself is not sufficient. Even if these security risks might present a worse problem than eavesdropping etc., they are left out of the scope of this study.

1.5.6 Legal aspects

Another kind of threat is misuse of copyrighted information. Partly this is related to confidentiality, but material, which is copyrighted, may also be illegally distributed by the home user himself, unintentionally or intentionally. In Finland, it is, as far as we know, legal to download any material and to make personal copies of it (unless accessed in an unauthorized way), but further use of it is restricted by copyright laws. Unintentional distribution may for example happen via Trojan horses or peer-to-peer software, without the user even noticing it. Authorities or copyright owners may hardly assume that an average PC user may prevent further distribution of the files. However, depending on the interests, the country, etc. there exist of course different opinions of what is fair and or legal.

Illegal, intentional distribution is left out of this study, as it is a legal aspect more than a technical one. Some actions could though be taken to minimize the risk for unintentional distribution of as well copyrighted material as own data. Mainly methods for preventing this are the usual, access control and confidentiality assurance. Also techniques like Digital Rights Management (DRM)/content protection are evolving. This is shortly discussed in chapter 2.5.2.

1.6 Security Requirements

The main security requirements that are specific for home networking are the following:

- Users should be authenticated for some devices and services (by home server or independent intelligent devices);
- Unauthorized access to all the home should be prevented, as well remote as through local networks;
- (Authorized) access should be possible through NAT or firewall ;
- Privacy of the communication should be guaranteed;
- The solution should not be restricted to specific operating systems and platforms;
- It should be possible to securely access the home network locally, for example using wireless network technology;
- Remote users should be able to securely access home network devices from different locations using different terminals (company LANs, hotspots, mobile computers, public computers), in other words, security should be provided for both personal and terminal mobility.
- Security breakage attempts should be logged if detected, as well as other detected faults (connection failures etc.), especially those which are (defined as) indications of misconfiguration or technical faults. An average user in a home or small office environment will probably not understand most of the logs, and probably not even see them. Critical events, for example detected breakage of the security policy or hardware failures should be displayed on the user's terminal, if feasible, in words understandable by an average user. Preferably the indications should be combined with further instructions on how to correct the error or fault or advice to consult the seller or the service provider.

2. Security Solutions on Different Layers

2.1 Link Layer

2.1.1 IEEE802.11i

Wired Equivalent Privacy (WEP) suffered from vulnerability: the same encryption key was used repeatedly, and intrusion was thus possible using brute force. Especially if the network uses publicly announced identifier (SSID), and the encryption uses a weak Initialization Vector, it is told to be easy for an intruder to get access [Kalm, Lee]. However, even if SSID is not public, it can be sniffed in plain text by existing LAN detection/traffic analyzer software, of which at least Airsnort does not only detect a WLAN, but also tries to access it. The Wi-Fi Alliance has proposed improvements. These improvements are briefly discussed in the following sections. More profound description can be found in [IEEE802.11iTutorial].

In Wi-Fi Protected Access (WPA), there are two modes for key distribution, an enterprise mode and a home mode. In enterprise mode, master keys are distributed automatically using servers. The home mode, or Pre-Shared Key mode (PSK), uses passwords as master keys. A password should be entered by the user in all nodes attached to the WLAN. The password triggers an encryption process called Temporal Key Integrity Protocol (TKIP). Starting from the master key, TKIP generates new keys periodically, so that the same key is never used twice.

The third key component of 802.11i is 802.1x standard, which provides mechanisms for authorized network access and dynamic key management. Furthermore, it can be applied for both the wired and wireless LANs.

One should note that 802.11i is targeted solely for physical to link layer security. The upper layer mechanisms are needed but not covered by this standard. The upper layer security is needed for creating authorized and encrypted communication between the client and the actual server to be used. For this different variations of Extensible Authentication Protocol (EAP) are proposed.

2.1.2 801.1x

The IEEE standard 801.1x defines authentication, access control and network management for LANs according to the IEEE standards. The standard identifies two roles within the access control interaction, the authenticator, providing services for which authentication is required, and the supplicant, which accesses services. This procedure is thus one-directional, and if mutual authentication is required, the two systems reverse their roles.

The principle of operation is based on port control. Since the port does not support any protocol functionalities, a port access entity (PAE) is defined. The supplicant PAE is responsible for responding to the authenticator's requests for credentials. The authenticator PAE acts as an intermediary between the supplicant's PAE and the authentication server. Each authentication point of attachment to the LAN is divided into two separate ports: an uncontrolled and a controlled one. The uncontrolled port allows exchange of protocol data units (PDUs) only during the authentication stage, while the controlled port is accessible only after the supplicant's successful authentication. That way the services provided by the authenticator are accessible through the controlled port, after this is set to an authorized state. When the port is in an

unauthorized state, DHCP and other IP initialization traffic cannot take place, and authentication is thus performed usually at system initialization time, prior to IP initialization.

The protocol used for authentication between the user's terminal and the access point (AP) is the Extensible Authentication Protocol (EAP) defined in [RFC 2284], carried directly by the LAN MAC services. This is named EAP over LANs (EAPOL). The Authentication is described in [RFC 2869]. The authentication mechanism may be for example MD5-challenge, One-Time Passwords, and Generic Token Card. The MD5-challenge type corresponds closely to the Challenge Handshake Authentication Protocol (CHAP).

An example of 802.1x setup is depicted in Figure 2.

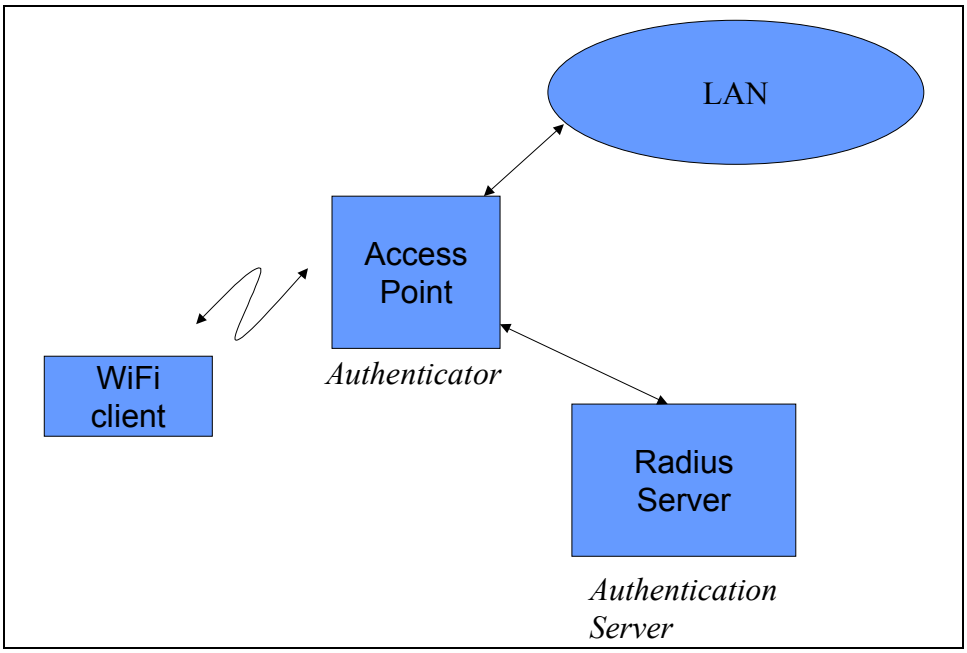

Figure 2. Proposed setup for network access authorization using IEEE802.1x.

Instead of EAP, it is also possible to use a web page, where the user has to enter a username and password. Once authentication is succeeded, access is granted to common resources in the LAN and to the Internet. This solution is used by Tampere University of Technology [Keski-Kasari1]. Preferably, this web server should use TSL or SSL (see chapter 4.) The drawback with web based authentication is that it requires an access to the part of the network, from which the authentication page is downloaded, already before authentication can take place.

If the authentication server required is outside of the LAN, it is accessed using higher layer protocols. RADIUS [RFC 2865 – RFC 2869] may be used for this. The WLAN AP, acting as an authenticator (RADIUS client) forwards the user credentials to the RADIUS server. This may also handle accounting information. In case of visiting users, the RADIUS server may forward the request to other RADIUS servers. This is described in more detail later in this report.

2.2 Network Layer

In this chapter, security techniques at the network layer, mainly different variations of IPSec are described. Whereas link layer security protects the local link, network (and higher) layer security protects the connection over the untrusted network (usually internet) between the source and target domains. This may be end-to-end or between gateways at the border of these domains. Security protocols at the network layer still have the advantage that they protect any traffic, independently of the applications. However, this secure channel should be set up first. If communication occurs rarely and/or lasts only momentarily, for example a few request – response pairs related to appliance control, this may, however, be an overkill, especially if the same secure channel cannot be reused next time because either the user or the mobile terminal has changed its location.

2.2.1 IPSec

Internet Protocol Security (IPSec) is specified in RFCs 2401–2411. IPSec provides mutual authentication of the communicating hosts, integrity, confidentiality and key exchange. Once established, IPSec protects all IP-based communication between the communicating hosts or gateways. This indicates, that IPSec should be a good solution for continuous sessions, for example remote working, or when larger amount of data is exchanged, but is clearly unnecessarily heavy for casual connections, such as www browsing or sending a few SIP messages (unless long media sessions are established).

The basic elements of IPSec are:

- Authentication Header (AH)
- Encapsulated Security Payload (ESP)
- Key Management Protocol.

Two modes of operation are defined: the **transport mode** and **tunnel mode**. The tunnel mode is usually used between two gateways, for example in a VPN, whereas the transport mode is used between hosts communicating directly. In transport mode, the ESP header is placed between the original IP header and the encrypted data. The original IP header is thus sent in clear text. In tunnel mode, a new IP header the ESP header and the Authentication Header are placed in front of the original IP packet, which is encrypted as whole (also the original IP header).

As there are two different modes which result in differently structured IP packets, it means that if a host using IPSec wants to communicate both with hosts behind an IPSec gateways, and with other IPSec hosts directly, it should support both modes.

The problem with IPSec is that in the gateway approach (tunnel mode), the traffic between the host and the gateway is unprotected. Additionally, IPSec is claimed to be too heavy for mobile phones and handheld computers. This is partly because the number crunching associated with encryption consumes a lot of power. However, it is included in Symbian 7.0 and at least one product is also available for Palm and Windows CE, the movianVPN [Movian]. As the name indicates, this product uses tunnel mode connections. As this supports various wireless connections, it should, at least in principle, be possible to establish a secure connection between

the handheld device and the home network. In [Aventail] IPSec tunnels are claimed to be difficult to establish through NATs and firewalls, especially if the user changes location.

For the IPSec key exchange, a protocol named IKEv2 (Internet Key Exchange version 2) is specified. This is an improvement of IKEv1, which was specified in RFCs 2407, 2408 and 2409.

2.2.2 IKEv2

The Internet Key Exchange includes two phases:

In the initial handshake an IKE security association (SA - a connection that provides security services to the packets carried by the SA) and a first IPSec SA is established. This means negotiation of cryptographic algorithms, mutual authentication and establishment of a session key. This phase usually contains two message pairs. Version 1 of IKE used three message pairs.

After the first phase, the second phase can be used to establish additional child-SAs. Also informational messages may be sent. This phase can be used for example if there is a need to establish several IPSec SAs between the same hosts. This second phase is not as resource consuming as the first one.

The first message pairs of the initial phase negotiate cryptographic algorithms and establish keys using the Diffie-Hellman algorithm. The second message pair is encrypted and integrity protected using the keys established in the Diffie-Hellman exchange. In this message pair, the communicating hosts are authenticated and the first IPSec SA is established.

2.2.3 Extensible Authentication Protocol (EAP)

Extensible Authentication Protocol is specified by IETF in RFC 2284 as a part of an 802.1x standard. The 802.1x standard provides an architectural framework on top of which one can use various authentication methods such as token cards, one-time passwords, public key authentication using smart cards, or certificates. EAP, which is an extension to the Point-to-Point Protocol (PPP), allows developers to pass security authentication data between RADIUS and the access point (AP) and wireless client.

Originally, 802.1x was conceived as an asymmetric protocol, allowing the network to authenticate the user, but not vice versa, what made possible "man-in-the-middle" attacks. A number of variants of EAP, including: EAP-Tunneled TLS (EAP-TTLS), Lightweight EAP (LEAP), and Protected EAP (PEAP) fix the problem by providing strong mutual authentication between mobile station and access point.

EAP-TLS is a certificate-based authentication protocol and is supported natively in Windows XP. It requires initial configuration by a network administrator to establish the certificate(s) on the user's machine and the authentication server, but no user intervention is required thereafter. The certificates are digital signatures that are used in conjunction with public key encryption techniques to verify the identity of the client.

During an EAP-TLS exchange, the client and authentication server exchange credentials and random data in order to simultaneously synthesize the encryption keys at both ends of the link. Once this has been completed, the server sends the encryption keys to the AP through a secure

Radius channel and the AP exchanges messages with the client to plumb the encryption keys down to the MAC encryption layer.

PEAP is an IETF draft standard and can be used to provide a secure password based authentication mechanism. Although it has not been implemented in any products to date, this is likely to change in the near future.

In a PEAP exchange, only the authentication server (AS) is required to have a certificate. After the initial communication with the authentication server, the public key from the AS certificate is sent to the client computer. The client computer then generates a master encryption key and encrypts this key using the AS's public key and sends the encrypted key to the AS.

Now that the master key is on both ends of the channel, this key can be used as source material to establish a secure tunnel between the AS and the client over which any subsequent authentication method can be used to authenticate the client computer to the AS. In many cases is it expected that this will be some form of a password based authentication protocol.

EAP-TTLS is an IETF draft extending EAP-TLS. It allows legacy password-based authentication protocols to be used against existing authentication databases, while protecting the security of these legacy protocols against eavesdropping, man-in-the-middle and other cryptographic attacks. While in EAP-TLS, a TLS handshake is used to mutually authenticate a client and server, in EAP-TTLS the TLS handshake can be mutual, but it may be one way only, in which case the server is authenticated to the client. EAP-TTLS extends the EAP-TLS authentication by using the secure connection established by the TLS handshake to exchange additional information between the client and server. The connection can be used to allow the server to authenticate the user using existing authentication infrastructures such as e.g. RADIUS. EAP-TTLS has been implemented in some Radius servers and supplicant software designed for use in 802.11 WLAN networks.

LEAP is a proprietary standard developed by Cisco Systems and was designed to be portable across a variety of wireless platforms. It has gained popularity due to the fact that it was the first, and for a long time, the only password based authentication scheme and it also provided this support across several different client operating system platforms.

LEAP is based on a straightforward challenge-password hash exchange where the authentication server issues a challenge to the client and then the client returns the password to the authentication server after first hashing it with the challenge text sent by the AS.

2.2.4 Host Identity Protocol

Traditionally, the IP address has two functions or roles; it is used for routing the packets to the right host, in other words for describing the topological location of the host, and secondly for identifying the host. Even if these two functions are closely related, they are semantically different. The idea to have one identity used for routing does not perform correctly when hosts moving from one network to another, may get a completely different address assigned dynamically. The identity of the host should logically remain the same even when it moves. Host Identity Protocol (HIP) is designed to be used by hosts in end-to-end fashion, it is not for gateways.

HIP [Moskowitz] [Jokela] proposes a solution, where the two mentioned above functions are separated. In HIP a different new name space is introduced for host identities (HI). Basically, the Host Identity is a public cryptographic key of a key pair. The node identified by the Host Identity is the only one that owns the corresponding private key, which makes it possible to prove its identity. Additionally, a host can generate short-term keys to be used for anonymous purposes, i.e. when revealing its actual identity is not necessary. The HI itself is too long to be used as such, and thus it is represented by a 128 bit Host Identity Tag (HIT), which contains a hash of the HI. For more details on how the HIT is calculated, see [Moskowitz]. In order to maintain compatibility with IPv4, a 32-bit Local Scope Identity (LSI) is also defined, but not the node itself but by its peer. The LSIs are allocated from the 1.0.0.0/8 address space, and the least significant 24 bits should be the same as of the corresponding HIT (unless these 24 bits are equal for the hosts; in that case another LSI is selected).

The Host identity layer is placed between the network (IP) and the transport layers (TCP, UDP). Thus, the upper layers see HI, interpreting it as an IPv6 address, but the IP address corresponding to the node's location, is hidden. The mapping between the peer HIT and its IP address can be retrieved for example from a DNS.

One of the goals of HIP is protection against Denial of Service (DoS) and Man in the middle (MITM) attacks. HIP creates an IPSec ESP association between the communicating hosts using four messages, as shown in Figure 3. HIP provides mutual authentication and Diffie-Hellman key establishment. The calculation of the parameters is explained in [Moskowitz].

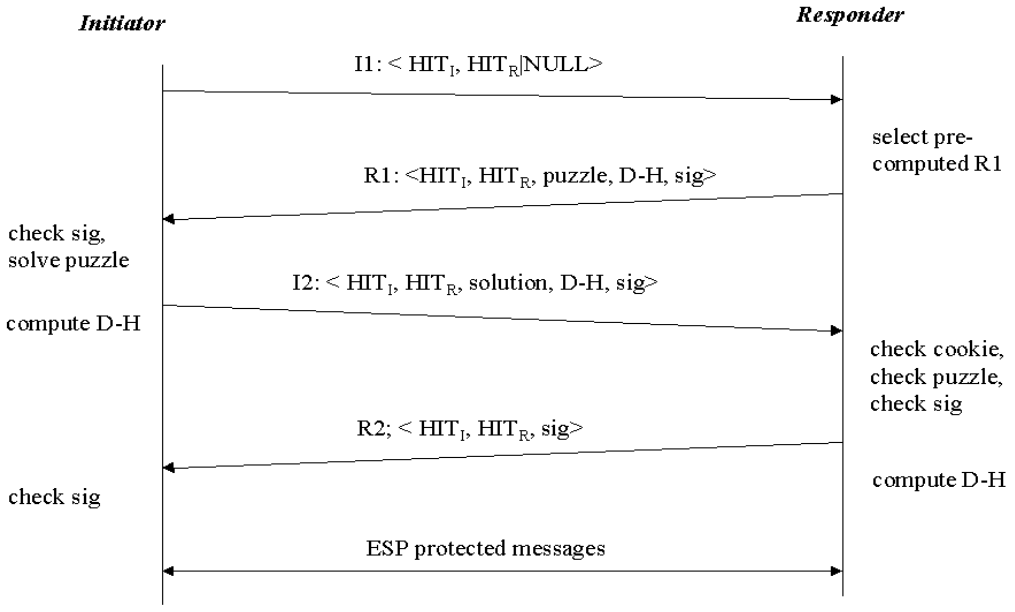

Figure 3. Secure HIP session establishment. Packet I1 triggers the authentication. Packet R1 contains a puzzle, a cryptographic challenge that the Initiator must solve. In its reply, I2, the solution to the puzzle is sent. If I2 message does not contain the solution, it is simply discarded. A standard authenticated Diffie-Hellman key exchange is performed.

The packets are identified using the Secure Parameter Index (SPI) value in IPSec ESP protocol. The SA is controlled by the HIT or the LSI and is not associated with the IP address.

HIP enables roaming in both IPv4 and IPv6 networks, as well as multi-homing, meaning that the node can have several network interfaces.

2.2.5 Other IP security protocols

Examples of proprietary IP security protocols are Secure Packet Shield (SPS) from Fortress Technologies and the Simple Key Management Internet Protocol (SKIP) by Sun Microsystems [Bozoki]. The choice of algorithms is more limited in these protocols than in IPSec.

SPS is a patented protocol, which provides authentication, confidentiality and data integrity and also key establishment using the Diffie-Hellman scheme. According to [Bozoki] SPS scales well and is a "very fast" protocol. In total, SPS seems to be very similar to IPSec tunnel mode.

SKIP provides authorization through access control lists (ACL) in addition to the three basic security functions. SKIP is no longer supported by the newer products.

2.2.6 Network Address Translator (NAT)

Network address Translator (NAT) means in short, that the IP address and possibly ports are changed, or translated, at the border between two networks, usually the public internet and a private LAN or for example a home PC. There are two reasons for using NAT. One is because the address space of IPv4 is insufficient. Private networks may internally need more addresses than are available from the network operator, but they do not need be all coupled to the Internet continuously. The operator provides perhaps a few, either dynamically or statically assigned internet addresses for those nodes, which are accessing the net and/or need to be visible from the Internet. The private network may internally use either dynamic or static addresses, depending on the needs. The other reason is that it hides the network elements from the Internet, which may provide additional security. Although NAT is not really a security protocol in itself, we include it in this report because of its impact on the traffic between the private and the public network. NATs are nowadays typically used between the home, or small office LAN and the Internet. Also the GPRS networks use private addresses. On the long run, IPv6 may of course eliminate the need for NAT, but a secure network infrastructure today would not be widely usable, unless it also works over NAT.

There are different variations of NAT, Basic NAT and NAPT [RFC 3022]. The Basic NAT translates only the IP address, whereas NAPT translates both the IP address and the port.

NAT is clearly designed for outbound access, i.e. from the private network to the Internet (or more public network). The addresses assigned to the communicating nodes on the Internet side may be dynamic, or even if they are static, they are not necessarily publicly known. However, this does not mean that the computers in the private network would be safe from attacks, unless a firewall is combined or installed in series with the NAT. Experience with fast internet connections in homes, using e.g. ADSL have clearly shown that there are frequent attacks towards the home network, particularly NetBIOS attacks. In order to provide useful inbound access, special arrangements have to be provided. For example, if external access to an http server located inside the NAT is required, a fixed address and port might to be assigned for this. The same applies to inbound peer-to-peer connections. If the peer to be contacted is behind a NAT, either its external address IP address or its FQDN (Fully Qualified Domain Name) has to

be known. The latter is possible only if the NAT supports DNS-ALG (Domain Name System – Application Layer Gateway). The functionality of a DNS-ALG is described in [RFC 2694].

Not all higher layer protocols pass the NAT without problems. In more detail complications are handled in [RFC 3027]. Current IPSec implementations may not work, but efforts to resolve this are under work. New IPSec specifications of January 2004 should support also NAT traversal. All protocols, where address and port information is exchanged on higher layers, may break without application layer gateway (ALG) functionality, as NAT would handle each of the connections independently and cannot know their inter-dependency. Examples of these are applications, where a control connection is used to establish the data flow, like FTP, H.323, SIP and RTP/RTSP. The Session Initiation Protocol (SIP) described in section 2.4.2, an ALG or, preferably, a SIP proxy should enable communication over NAT.

2.3 Transport and Session Layers

Logically, transport layer is a good choice to put security mechanisms, as the communication is end-to-end, and the mechanisms, at least in theory, may be used by several end-user applications. One problem here is though, that all applications may not use TCP, some may use UDP, which may not be reliable enough to provide any security. Also, because IP packets are not encrypted, more about the communication is revealed to a potential eavesdropper.

2.3.1 SSL

The Secure Socket Layer (SSL) protocol [Reyes et al.] was designed by Netscape Communications for use with Netscape Navigator www browser. Version 2.0 was used with early browsers. Some weaknesses of SSL 2.0 were fixed by Microsoft and these are incorporated in the current version of SSL, 3.0. IETF published in year 1999 a protocol called TLS 1.0 (Transport Layer Security) [RFC 2246] which is based on SSL, and in fact equivalent to SSL 3.1. TLS is backward compatible with previous versions of SSL. Most browsers support both SSL and TLS.

SSL is placed between the TCP layer and the application, which usually is a www browser. However, SSL is not restricted to this application but is assigned as a generic security protocol to be used with almost any application. SSL requires though a reliable transport connection and does not work on UDP. Because of this, SSL cannot be used for SNMP, NFS, DNS and VoIP according to ITU-T recommendation 323.

As SSL is placed above the transport layer, one may think of it as a session layer protocol. Functionally this is also true in the sense that a session connection is established, starting with the login and ending with the logout. In other words, the authentication and some other parameters are valid for the period of a session. Further, the session may last over several transport connections. This is somewhat contradictory to the ISO/OSI reference model, which states security techniques shall not be placed in the session layer.

SSL has four sub-protocols: SSL Record, SSL Handshake, ChangeCipherSpec and Alert. SSL Record takes care of the transmission of the information blocks between the communicating computers. The Record protocol is also responsible for fragmentation, compression, encryption and MAC (Message Authentication Code) generation. The Handshake protocol handles the authentication and negotiation of encryption and hash algorithms and key generation. The SSL

ChangeCipherSpec protocol is just a simple indication to the recorded protocol to start encryption with the negotiated parameters, in other words to make a pending Cipher Spec state active. Alert protocol is used to indicate all kinds of failures, which can be either warnings or fatal errors. A fatal causes the connection to be terminated immediately, whereas warnings don't require any specific action.

According to [Reyes et al.] SSL is very well designed, recognized weaknesses depend more on security policies, and on how secure the implementation of the protocol is, and of course on how secure the environments are. For example if a browser complains that a certificate has expired on some host, an average user would probably not understand, neither bother to read the requester, and just click OK to continue as before. Some possible vulnerability exists though, SSL is for example vulnerable to certificate injection and MITM attacks.

The most common factors that can make a web server insecure are:

- small keys (512 bit public key, should be 1000 bit)
- SSL version 2.0
- weak cipher suites (40 bit, should be 64 bit)
- self-signed or expired certificates.

As in case of IPSec, it is also possible to set up virtual private networks (VPNs) using SSL [Aventail]. At least in the simplest fashion, this means using the web browser interface for the most important applications. This has the following advantages compared to IPSec [Reyes et al.]:

- The resources in the home or company network can be accessed from almost anywhere, even from internet cafes or libraries
- SSL connections are considered to be easier to maintain than IPSec
- IPSec implemented on the network layer grants access to the whole network (in tunnel mode) during the whole user's session. With SSL each application is secured separately.

There are also disadvantages:

- SSL provides access only for applications that supports it, mainly web applications
- Suspectible to MITM attacks
- Authentication in SSL is inferior to that of IPSec.

2.3.2 SSH

Secure Shell (SSH) has become popular as a secure remote login protocol to offices. In addition to login and telnet type text based terminal emulation, SSH also supports secure file transfer and secure TCP/IP and X11 forwarding. With some SSH implementations it is also possible to only execute one command on the remote host without opening an interactive terminal window. SSH authenticates the parties, and encrypts, compresses, and ensures the integrity of transferred data. [SSH-IETF]. SSH consists of a transport protocol, an authentication protocol and a connection protocol.

The transport protocol performs server host authentication, key exchange, encryption, and integrity protection. A session key for encryption is negotiated using Diffie-Hellman algorithm.

The authentication protocol authenticates the client user. The authentication protocol makes use of the security provided by the transport protocol.

The connection protocol enables multiplexing of multiple streams (or channels) over the secure transport protocol. A user may thus have an interactive shell simultaneously active as 'proxy-forwarded' external protocols and X11 sessions. In practice, ssh provides similar kind of functionality as IPSec, but implemented on a higher layer. As the IP packet itself is not protected, NATs should not cause as much problems as for IPSec.

The algorithms used for authentication, integrity and confidentiality are negotiated between the user's workstation and the host, new algorithms may be applied without changing the basic protocol and it is even possible to leave some security features out, although this is of course not recommended for other than debugging purposes. Strong algorithms and long enough key lengths are recommended.

Especially, SSH has been designed to prevent MITM attacks, eavesdropping, replay, etc., and enabling confidential communication over untrusted networks.

Some weaknesses found in older versions of SSH are fixed in SSH2 [SSH vulnerabilities].

Both, password and public key authentication may be used; usually these two are used in combination. A problem in SSH may be the initial authentication, performed the first time a user establishes an SSH connection to a specific host [SSH tutor]. As the user's workstation does not know the server's public key, a fingerprint is displayed, and the user is giving the options to quit or continue. The same kind of situation occurs if the server's key has changed. In principle, these fingerprints may be checked, but usually, the user would not know, or may not be able to verify, that this actually is the case and the server is the intended one. There is a risk, that the server is fake, as well. A user can generate the keys with the keygen program that comes with the SSH. The user's public key may be uploaded to the server. As the user may initialize sessions from several different workstations as well, the server does not necessarily know the user's key in advance. Another possibility is to copy the private and public keys between the different terminals. At this stage the authentication protocol usually trusts on user – password authentication here. Once the public key is installed on the server, public key authentication may be used instead of password authentication. For the user it looks almost the same, the difference is that now the user enters a pass phrase for his/her private key instead of a server password [Miller]. The keys can also be loaded to an SSH-agent, which in turn makes them available for the SSH client, eliminating the need to enter the pass phrase for each SSH connection. It is also possible to set up restricted keys on server, which forces specific commands, for example in software development projects, cvs access only may be allowed, but no port forwarding.

Servers requiring higher degree of security may also use one-time passwords, for example generated by a SecurID card.

Port forwarding is one of the main purposes with the SSH protocol. SSH uses local port forwarding to enable applications to communicate over the secure channel. Improperly used, this is also a potential weakness. The channel is secure, but the port in the client may be either opened for access in the client's LAN environment (if specified as gateway port), or restricted to the same host. If the access to the port is not restricted to local host only, there is a risk that an

intruder may access the client's LAN environment or even access the client host from outside, and get access to the secure channel through this forwarded port.

Remote forwarding is possible at the server's side also; this enables a port on the server to be forwarded to specific host on the local network.

2.4 Middleware Layer

2.4.1 Secure HTTP and HTTP over SSL

There are two separate approaches to protecting information transmitted between a web client and a server. Secure HTTP (S-HTTP) [RFC 2660] is an extension to the standard HTTP protocol. The other common approach, HTTPS, uses HTTP on top of the Secure Socket Layer (SSL).

S-HTTP is a message-oriented communications protocol designed to transfer encrypted information over the world-wide web. It supports confidentiality and integrity services, sender authentication and can also support non-repudiation of origin. S-HTTP is very flexible in choice of key management mechanisms, cryptographic algorithms, modes and parameters. Security services, algorithms and options to be used may be negotiated between the client and the server on per-message basis. S-HTTP may take advantage of certification infrastructures, but does not require the use of client public key certificates. As S-HTTP is message-oriented, it may be used for example for securing web-forms.

In the other approach, HTTPS (HTTP over SSL), a secure channel is established between the server and the client, and authentication will be required, when a secured portion on the server is accessed. In this case all the data (higher than the Secure Socket Layer) will be protected inside the SSL tunnel. SSL was described in section 2.3.1. HTTPS uses port 443.

2.4.2 SIP security

Session Initiation Protocol (SIP) is described in [RFC3261]. Network elements, which may be involved in the SIP signalling are the User Agents (UA), proxies, registrars and redirect servers. A user agent must be able to act both as a user agent client (UAC) and user agent server (UAS), depending on whether it initiates or responses to a call. Although the primary task of SIP is signalling in IP network, it may be used for sending short control messages to devices (using extensions), retrieve status information, instant messaging and setting up sessions, both over Internet and telecomm networks, or even combining these. SIP supports user mobility by means of registrars. This means that users are accessible by the same name but may change terminals and locations. This may also be applied to services. When a connection is set up, usually the proxy retrieves the contact from the registrar. This approach makes SIP very scalable compared with multicast based systems, like UPnP.

A proxy and a registrar in combination with a firewall enable secure external access to the home. Authentication and/or authorization should in this case be performed at the latest on the servers, but at least partly already at the proxy in order to prevent DoS attacks.

As with bundled protocols in general, NAT may cause problems also for SIP. One problem is that the port specified in the SIP request may not be valid for the response. The same applies to

the IP address and port specified in the SDP payload for the multimedia connection. Thus, the SIP proxy or ALG should be able to decipher this information, or the communication should use pre-defined, globally significant ports. The proxy should co-reside with the NAT.

SIP may use a "Digest" authentication mechanism, which is nearly identical to HTTP authentication [RFC 2617], [RFC 3261], [Kullenwall]. This provides only message authentication and replay protection, but not message integrity and confidentiality. The protection domain differs slightly from HTTP. In SIP it is specified by the realm string, user info, host and port of the request-URI. There is also a simpler authentication scheme based on username and password called basic authentication. In the basic authentication scheme, username and password is sent in clear text, and thus it cannot be recommended if security requirements are high.

If an UAS or a proxy server requires authentication it responds to an unauthenticated request with a 401 (Unauthorized) (or 407) response, including a WWW-authenticate or Proxy-Authenticate header, which contains challenge information. The UAC sends then a new request containing a response to the challenge. The parameters in these messages are described in more detail in e.g. [Kullenwall] and will not be repeated here. It may be worth mentioning that the default message authentication algorithm used here is MD5.

Another possibility than the HTTP authentication is to used EAP in SIP. EAP may further be based on IMS AKA or UMTS AKA (Authentication and Key Agreement). This scheme is proposed by 3GPP (3rd Generation Partnership Project) for mutual authentication between the servicing network (SN) and the user equipment (UE). This is based on HTTP Digest authentication but the algorithm is different. It is based on a 128 bit shared secret between the Home Environment (HE) and the UE. The algorithm requires that HE must trust SN to handle authentication correctly. The authentication is based on authentication vectors, which the SN receives from the HE.

As proxies and redirect servers must be able to read and modify SIP messages, they cannot be encrypted or checked for integrity as such. The payload of the SIP messages may, however, be protected using for example S/MIME (Secure Multipurpose Internet Mail Extension). This is a quite natural choice, because of the similarity between SIP and email messages. It is also possible to protect the whole message by tunneling it in the SIP message itself, but result looks quite complicated. Like HTTP, also SIP may rely on lower layer protocols, mainly TLS or IPSec to provide confidentiality and integrity. The latter solutions may, however, not be used on end-to-end or peer-to-peer basis for the reasons mentioned above, but only on hop-to-hop basis. The secure lower layer connections must also be setup in advance of the SIP signalling in order to provide protection.

Also RADIUS (see section 3.1) may be used for authentication [Keski-Kasari2]. However, SIP clients do not, according to current specifications, act as authenticators for RADIUS authentication servers. There are differences between the HTTP Digest and the RADIUS authentication protocols. However, if the RADIUS server supports the HTTP Digest attributes, the SIP server may act as a RADIUS authenticator. If RADIUS is used, the SIP server makes the protocol conversions between the HTTP Digest and RADIUS protocols.

For just sending short messages and responses, like SIP's DO or MESSAGE including short messages e.g. for controlling an appliance in the home network, the additional overhead for authentication and authorization, confidentiality and integrity would cause quite a lot of overhead and several roundtrips, compared with one for a similar control message exchange in a very simple approach, where SIP is used over UDP and only the payload and some critical

header information are secured using a shared secret. It is though questionable if it is worth to support protection schemes which mainly differ from the common praxis, and thus we perhaps just have to accept either some reasonable risks, or the additional overhead. After all, the same increase of overhead is already accepted for other similar protocols.

2.5 Application Layer

The distinction between application and middleware related security is not at this point very clearly defined. One point which causes confusion is that almost any application may use communication, even if it is not really a networked application. Once some media sharing mechanism is set up, any application may usually access files from these shared medium. Security problems related to sharing network drives are solved at lower layers. Solutions, which are handled in this section, are more related to operations and access rights of applications themselves.

2.5.1 Java Security

In the design of Java programming language much care has been taken to ensure security. The code is executed in an isolated "sandbox", which should disable external access to the Java programs from outside during execution, as well as prevent the Java programs to harm the system, in which it is running. In reality, of course, the security depends on how well the Java Virtual Machine (JVM) is implemented and which rights are given to the Java programs. A runtime error in a Java program usually ends the program if the error was not caught, or may end the JVM, but rarely chokes up the whole computer.

Especially, strict security restrictions were assigned for applets, small Java programs which are executed on the browser. For example, applets were not allowed to write files. Later, the security requirements have slightly been loosened, and applets may be given more rights. For example, in Microsoft's Internet Explorer, high safety should be given to Java, if one doesn't want to give more rights to the applets than initially planned.

Because of the "sandbox" design, internal error handling and memory allocation, it is less likely to make bad Java code by mistake, where for example buffer overflows may open possibilities to "run arbitrary code with root privileges". This type of security hole is maybe the most common in protocol implementations made with C/C++, and may not be easy to identify, even by an experienced programmer. Also, the way classes, methods and variables are referenced in Java, is clearer than in C/C++.

Additionally, it is possible to impose internal access restrictions for the building blocks (classes, methods, variables) of Java programs, using packages and attributes: public, private, protected. For classes, methods and variables not declared public, the access is local and restricted.

2.5.2 Content Protection

Because the entertainment related services will be important in future homes, this report reviews shortly existing technologies for protection of content and property rights.

According to the law, the originator of a work such as a computer program, a musical composition, a video, etc. has the immaterial rights for it. It is up to the originator how (and if) the work may be distributed. Especially for audio-visual media this is essential.

Nowadays, the digital revolution is changing the way media works are stored and transferred, and also how it is consumed. In addition to traditional analogue ways, media may be stored and distributed digitally, for example transferred over internet or mobile telephone networks. Computers play an increasing role when producing, storing, transferring, copying, and reproducing media by all the parties involved from the originator to the consumer. For the home network it is essential that the computer can communicate without hinders and loss of quality with the other consumer electronic equipment and the mobile equipment also for the media. Giving new possibilities, it is also likely that an ever increasing part our income will be spent on different media.

This is where the interests of the companies producing and distributing media are in great conflict with the interest of the customers. As digital media can be compressed, transferred over the net, copied and reproduced without any significant loss of quality, also illegal copying is common. Laws in this respect may also differ between countries in the interpretation of what is legal or illegal. A great fraction of the bandwidth in the Internet is nowadays believed to be used for peer-to-peer communication, of which the main part consists of music distribution of questionable legality (Napster, Grokster etc.). However, it is still fair, that a user who has legally bought some A/V material for some device (mobile phone, CD player, etc.) also could copy it to other devices for his/her own use, or make a backup copy. Royalties are already paid for empty media. This also means that protection should not prevent the use in computers and older players.

As the situation is today, computers and consumer electronics do not usually cooperate digitally, because they lack compatible connectors (and cheap devices may lack analog connectors as well). Exceptions, or at least steps in right directions, are video and digital cameras with firewire and/or USB connectors and some audio equipment with USB connectors. Some modern CD players can also, in addition to audio CDs, play MP3 disks, which may be burned on the computer. In addition to the novelty of the digital media technology, an inhibitor for an integrated (digital and analog) media infrastructure may be the unstabilized situation of the content protection issues. Manufacturers of electronics and media companies are naturally dependent of each other as their products have to be compatible to be of any use. The most protective content protection requirements will in practice require licensed hardware and software, since one requirement is that the information should be kept secret also when it traverses busses within the device, and also, that it shall not be possible for the user to disable protection for the content. Clearly, this is a threat against open source efforts.

As neither the technology, nor the business situation has yet stabilized, media companies probably try to find out all the possibilities to make business. Buying music legally through the Internet is still quite expensive because the royalties; customers may not want to pay almost the same price as for the CD, particularly if they have a slow and unreliable connection and cannot ensure the quality. Especially this is true if the users may get the corresponding product for free, in spite of the fact that it may be illegally distributed.

Technically a middle-way may be possible: as long as the originators and artists cannot be sure to get their royalties and in particular ensure somehow that their immaterial rights are not misused, their restraints may be high. A fair digital rights management (DRM) method combined with an infrastructure to share royalties may thus have an encouraging effect in the use of internet as a distribution channel for media. However, a condition for this is that it would

be very cheap for the consumer, as publishing on a web page or a peer-to-peer communication system also is cheap. This scenario may, however, not be the ideal for the traditional media companies, as they may lose business.

Like encryption technology, DRM is also as a challenge for hackers. During playback, the protection has to be stripped off, and raw data fed for the D/A conversion. Unless all the decoding as well as the decompression occurs in a protected piece of hardware, hackers may anyhow access the data at some point. Raw data may be captured before it is fed to the amplifier when it is uncompressed and equivalent to the sound recorded analogically. Protected hardware however, has the risk that it restricts the playback too much, and thus slows down the standardization of DRM technology. The same applies to protected software, as it may lead to monopoly. Preferably, the methods for DRM should be open, like the encryption algorithms are, and the copyright indicated by the keys, not by the method itself, the difficulty to invent some system is, however, clear from the discussion above.

CPRM [Intel 6-4]

Content Protection for Recordable Media (CPRM) is developed by Intel, IBM, Matsushita and Toshiba (formerly 4C). CPRM is a method for protecting content on portable/removable media. The two main technical components of CPRM are the C2 cipher and the Media Key Block (MKB). The C2 cipher algorithm is licensed by 4C. MKBs are tables of cryptographic values. Each licensed product also has a device key, LLC. MKBs and LLC are used to calculate a media key. If a device key is compromised, it will be revoked, as it cannot longer calculate a correct media key when MKBs are updated.

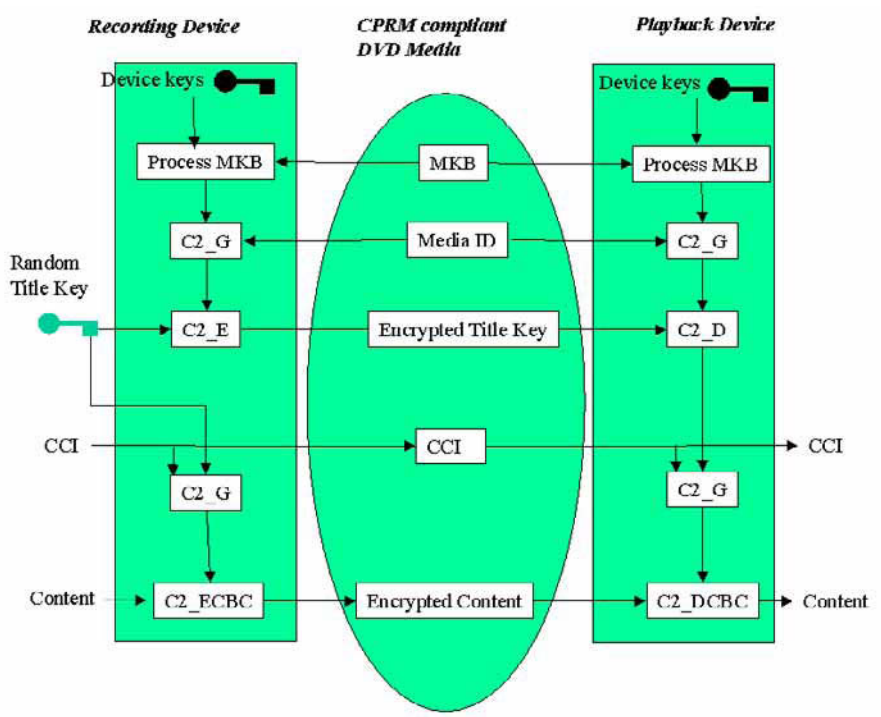

Figure 4. Operation of CPRM [Intel 6-4]. C2_G = C2 cipher one way function, C2_E, C2_D = C2 encrypting / decrypting function, CCI = copy control information, C2_ECBC, C2_DCBC = C2 encrypting / decrypting functions using cipher block chaining.

DCTP [Intel 6-4]

Digital Transmission Content Protection (DCTP) provides encrypted exchange of contents and Copy Control Information (CCI) between authenticated devices. Supported interfaces are IEEE 1394 and USB, as well as MOST (Media oriented system transport) but the technology is not limited to these. A challenge when extending this technology to a WLAN environment is to prevent anonymous sharing of content.

CCI may be carried for example in a MPEG-2 transport descriptor. The technology may restrict copying of material using an Encryption Mode Indicator (EMI), which may take values "copy_once", "copy_never" and "copy_no_more". Authentication is performed according to this value. If the value is "copy_never", full authentication will occur, in the two other cases restricted authentication is sufficient. Full authentication uses Diffie-Hellman key exchange whereas restricted access uses shared secrets.

The content may also contain System Renewability Messages with certificate revocation lists. These prevent distribution of content to compromised devices.

Itunes [MacWorld 6/03]

Also Apple has made up its own system trying to balance between interests of the record companies and the consumers. Apple uses protected AAC files for their iTunes music store. Music may be purchased and downloaded over Internet for 99 c / song. (However, songs are not always available separately and the consumer must then buy the whole album.) Here, the Macs on which the music is to be played, have to register to the music store. Three Macs are permitted to play the tracks. The music can be listened on an unlimited number of iPods (Apple's portable player). It can also be burned to CD, but only 10 times.

DRM by Open Mobile Alliance

Distribution of media and other files has become a popular business in the area of mobile phones. Examples of files available for mobile phones are ringing tones, background images and Java games. Customers are usually charged for these files, and so far they have accepted this more than paying for material distributed on the Internet. Traditionally, ringing tones and background images have been type dependent and fairly simple, say monophonic tones and black and white pictures. As the capabilities and power of mobile phones and communication has grown, also audio and video clips, and colour pictures are distributed. Further, also mobile phones may display/replay media of the same file formats used in computers.

In such an environment there is also a demand for DRM. In this case, these standards may be built and implemented almost from scratch. The scope of OMA’s DRM specification [OMADRM], [OMADRMREL] is to enable content providers to express access and usage rights and for example prevent forwarding of downloaded material.

3. Authentication, Authorization and Accounting

3.1 RADIUS

Remote Authentication Dial In User Service (RADIUS) is described in [RFC 2865] Accounting and extensions are additionally defined in [RFC 2866] – [RFC 2869]. RADIUS is designed to manage authentication, authorization and accounting (AAA), originally for modem pools. It may be used for authentication and configuration information of various services (for example, SLIP, PPP, telnet, rlogin). Also the authenticator in a wireless LAN may act as a RADIUS client. Additionally proxy-RADIUS servers may forward the information to other RADIUS servers.

The RADIUS server can support PPP PAP or CHAP, UNIX login, and other authentication mechanisms.

The protocol is simple. The network access server (NAS) sends an Access request packet to the RADIUS server, containing User-Name, password (User-Password or CHAP-Password), and description of the services which the user wants to access. The RADIUS server replies with Access-Accept, Access-Reject or Access-Challenge. The Access-Challenge asks for further information from the NAS, which is sent in a new Access-Request to the RADIUS server (provided that the NAS supports challenge/response, of course). In this Access-Request message the password attribute is replaced with the response calculated from the challenge. This procedure may be repeated until the RADIUS server finally answers with an Access-Accept or Access-Reject packet.

A typical case for a proxy scenario is roaming. The following scenario [RFC 2865] illustrates a proxy RADIUS communication between a NAS and the forwarding and remote RADIUS servers:

1. A NAS sends its access-request to the forwarding server.

2. The forwarding server forwards the access-request to the remote server.

3. The remote server sends an access-accept, access-reject or access-challenge back to the forwarding server. For this example, an access-accept is sent.

4. The forwarding server sends the access-accept to the NAS.

A RADIUS server can function as both a forwarding server and a remote server, depending on the authentication realm or other criteria.

The access requestor is sent unencrypted in the access request, access accept and access reject packets. In [Hill] this protocol is claimed to be vulnerable to eavesdropping and MITM attacks. Particularly poor implementations, where authenticators and shared secrets are predictable, may be vulnerable. RADIUS servers contain a database associating the User-Names with authentication information. As different methods might be used in different circumstances, it is recommended, that also different User-Names are used for this, as it otherwise might be possible to attack the server using the least secure method. As several different identities and mechanisms makes the system more complicated and less manageable, we suggest to use both the User-Name and service description attributes, for example port-type and number as keys when authenticating the user.

3.2 Diameter

RADIUS is originally designed for PPP connections over serial lines, and has been criticized as not being enough for VPN connections with roaming handheld devices. The most important drawback of RADIUS is packet confidentiality not defined, meaning that RADIUS may be vulnerable to replays. Some features are left open, like support for IPSec, support for proxy, redirect and relay agents (brokers), server initiated messages, behavior in fail situations and auditability.

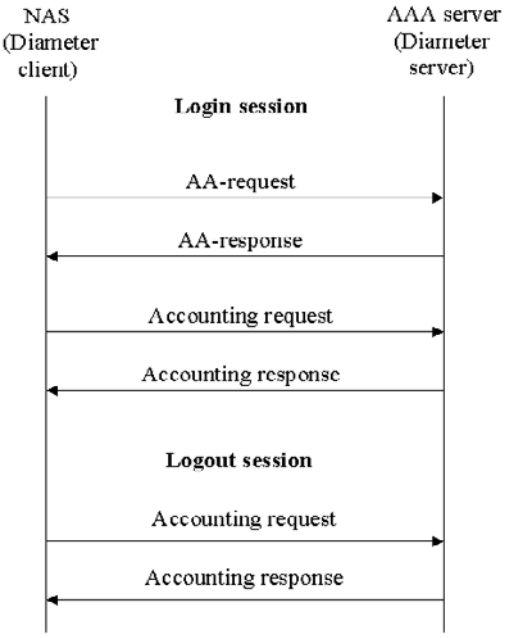

Figure 5. Sequence diagram of the basic diameter protocol [Utwente]. NAS = Network Access Server, AAA = Authentication, Authorization, Accounting.

An improvement based on RADIUS, called Diameter is proposed [RFC 3588]. The basic Diameter protocol shown in Figure 5 is very similar to RADIUS, but Diameter uses reliable transport (TCP, STCP) instead of UDP, as RADIUS does.

Diameter also supports capability negotiation, peer discovery and roaming, which RADIUS lacks. Another difference between RADIUS and Diameter is that in diameter also the server may initiate the authentication procedure [Keski-Kasari2].

The attribute value pairs (AVP) in RADIUS are limited to 255. Diameter extends the capabilities of the RADIUS protocol. Diameter uses a 32 bit address space for AVP [Roy]. The base protocol may be used with Mobile IPv4 or network access, but it may also be extended for new applications, by creating new AVPs and accounting applications.

Diameter defines agents, which may be used between the AP and the Diameter server. The agents take care of forwarding the AAA request to the Diameter server of the user's home network. Further, some agents may convert AAA information between different networks. [Keski-Kasari2].

4. Wireless access and roaming architecture

Typically, wireless public access has been provided by cellular systems owned by cellular operators. The situation is changing with the development of the public WLAN service market. The relatively low investment costs and usage of unlicensed frequency band causes that the potential hotspot providers can belong to one of the following categories:

- Mobile network operator (GSM/GPRS/3G) offering additionally public WLAN access.
- Added value Wireless Internet Service Provider (added value WISP) - a site owner offering public WLAN service as added value to its core services as e.g. an airport host, energy supplier, or a hotel owner,
- Pure WISP - offering public WLAN service as the only product.
- Non-public WLAN access provider - offering WLAN access to limited group of users as e.g. employees, or block of flats inhabitants.

In each of these cases the WLAN providers have different strategies and different target market, but many of the problems stay similar. As WLAN was not intended for public services, standardized authentication technology, billing system, or roaming possibilities are still to be provided. Also service handover to other access networks has to be considered, although it may not be very important as the typical usage of laptop, being be the main WLAN device, does not require frequent seamless handover.

In spite of the fact that the size and value of the target market is uncertain, and many failures were already accounted among the WISPs, public WLAN may have strategic value for the mobile network operators. Their 3G offer will be more attractive with public WLAN access as a value-added service. Additionally, the usage of mobile Internet may raise the usage of the core mobile services, while not having the public WLAN access can lead to customers moving to competitors who offer such possibility. In these circumstances the Third Generation Partnership Project (3GPP) has recently taken the initiative to develop an inter-working cellular-WLAN architecture. The goal is to enable 3G operators to provide public WLAN access as an integral part of their cellular system. This includes the reuse of 3GPP subscription, network selection, 3GPP system-based authentication, authorization and security key agreement using SIM/USIM card, user data routing and service access, as well as end user charging (see Figure 6). This is going to be achieved without any 3GPP specific requirements for the WLAN system, but relaying on the IEEE 802.11 standards.

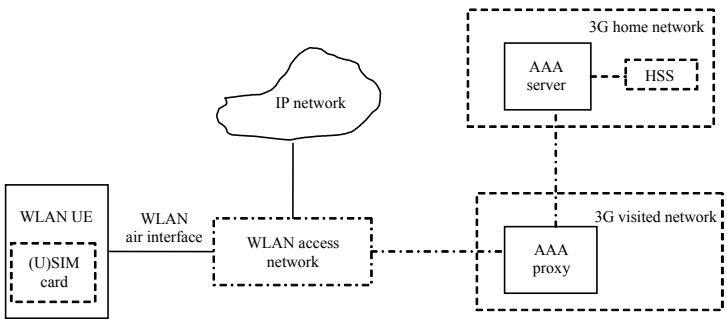

Figure 6. 3GPP-WLAN authentication and roaming architecture.

Facing strong competition with mobile operators and encountering earlier losses, the added-value WISPs are directing their offers to selected group of users. Although the target market for public WLAN services is still rather unknown, the corporate travellers are believed to be the group that will most frequently require access to e-mail, or corporate access while travelling. The mobility pattern of those users is predictable, so hotspots are located in places visited by the group for a longer time. To this belong airports, hotels, or conference centers, but not many other public places. The provided services are considered as added value to the main offer (accommodation, or transport) and are provided by the main service provider, i.e. site owner. The access is typically based on one-time accounts sold to the user at the service usage place, which means that the roaming to other networks is not offered.

The most difficult situation from the business point of view is the situation of pure WLAN providers. They may find it hard to operate while competitors can afford some losses - mobile operators relaying on other service revenues, or added-value WISPs treating WLAN access as a small addition to their main services. So far, to assure constant income, the pure WISPs are following the work pattern of the value-added WISPS and are seeking for cooperation with hotels, or airports, asking for a guaranteed monthly minimum fee, or monthly support fee. Some hotel WISPs have a revenue-sharing agreement on public meeting rooms. Example of DECT/GSM dual mode phones shows that voice over IP technology may not necessary change the situation, especially when the customers are used to high quality cellular connections. Also expanding the network of hotspots and offering roaming possibilities with other WISP networks does not guarantee the success for the pure WISPs, as the operators may rather lose 3G revenues to their own subsidiaries than to the competing WISPs. In any case, the roaming requires an industry agreement on service levels, billing systems, or authentication technology, which do not exist.

Possible authentication and roaming architecture is presented in Figure 7, where the numbers indicate the logical order of requests for user authentication. In this scenario the assumptions are that the user has a pre-signed agreement with a WISP called here Home WISP, while all the WISPs providing roaming possibilities have federation agreements. As the number of WISPs can be considerably big and signing bilateral agreements practically impossible, a trusted entity, called here Clearing House, providing federation possibility may be required. The Clearing House would have agreements with all the WISPs and would act as a mediator between different WISPs, and a Certification Authority (CA) issuing certificates to WISPs and their customers. When a user contacts a visiting domain, the local AAA server would pass the user's credentials to the Clearing House that, in turn, would contact the user's home AAA server asking for authentication. However, as the introduction of the Clearing House concept is against the open, decentralized Internet philosophy, the business role of the entity requires further investigation. Today the few WLAN roaming possibilities are very limited, and from the user point of view the situation is far from a one-click solution, while the users are used to the level that they experienced by usage of cellular networks.

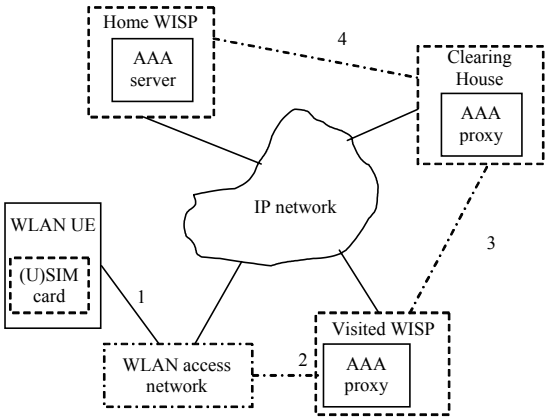

Figure 7. Inter WISP authentication and roaming architecture.

Different from the WISPs case is the situation of the non-public WLAN providers. Offering services to a limited group of people does not require complicated standardized billing, or authentication systems. It is easier to make an agreement between the involved parties, even if personal mobility of the users between different hotspots of the provider is required. As non-public WLAN access will be most often provided by corporative users to its workers, or universities to their students, the revenue will not be of primary concern. In this case, the authentication scenario enabling access to the Internet will remind the one presented in Figure 7, where the Home, Visited, and Clearing Home AAAs are replaced by the company distributed database enabling authentication of the users.

5. Remote access to home devices scenario

In case of home network the most challenging from the security point of view is the situation when the user is accessing home appliances from a remote place. In this section an overall view of home network security with stress on secure access from an external WLAN hot-spot to IP-enabled home appliances is presented.

Let us first assume that the user has an agreement with a WISP, called Home WISP (WISP_h), but is requesting Internet access from external public WLAN hotspot operated by another WISP. The WISP providing the user with WLAN access is called a Visited WISP (WISP_v). Both, Home and Visited WISPs, have federation agreements with Clearing House.

Additionally, it is assumed that the user has a certificate issued by the CA and several keys stored at his personal mobile terminal, or smart card. Among the keys are the Home WISP's and the CA public keys and the user's private key, all certified by the Clearing House. They will provide WLAN roaming and generic Internet access. For accessing the home network the user will have symmetric key which he will generate himself. The protocol used for exchange of messages during the user access phase could be 802.11i when it is standardized, or EAP-TTLS. In Figure 8 presented is the second case, where EAP-TTLS is proposed for mutual authentication between client and server. RADIUS is used as the protocol supporting communication with AAA server. The components presented in the figure are logical entities and may correspond to separate network components, but except for the UserTerminal may be combined into a single device.

The EAP-TTLS protocol takes advantage of the Transport Layer Security (TLS). The TLS handshake is used to provide mutual authentication between the user's terminal and the server, based on the certificates. The conversation begins when the access point requests the identity of the client trying to connect to the network. After sending the EAP_Req/Identity the access point acts as a pass-through device, allowing direct negotiation between the mobile terminal and the TTLS server. The client (UserTerminal) responds by sending its identity, which according the EAP-TTLS specification should not include the user name, but may point to the user's trusted WISP by e.g. including "@WISP_h.com" part. For the privacy of the user, the name can be sent only after an encrypted channel between the user terminal and the TTLS server is established. After receiving the client identity, the server initiates the TTLS authentication by issuing an EAP_Req/TTLS/Start message. The client replies with a TTLS client hello message that is part of the TLS handshake protocol used to negotiate the secure attributes of the session. The server responds with an EAP_Req/TTLS/ message containing among others the server certificate signed by CA, and request for client certificate.

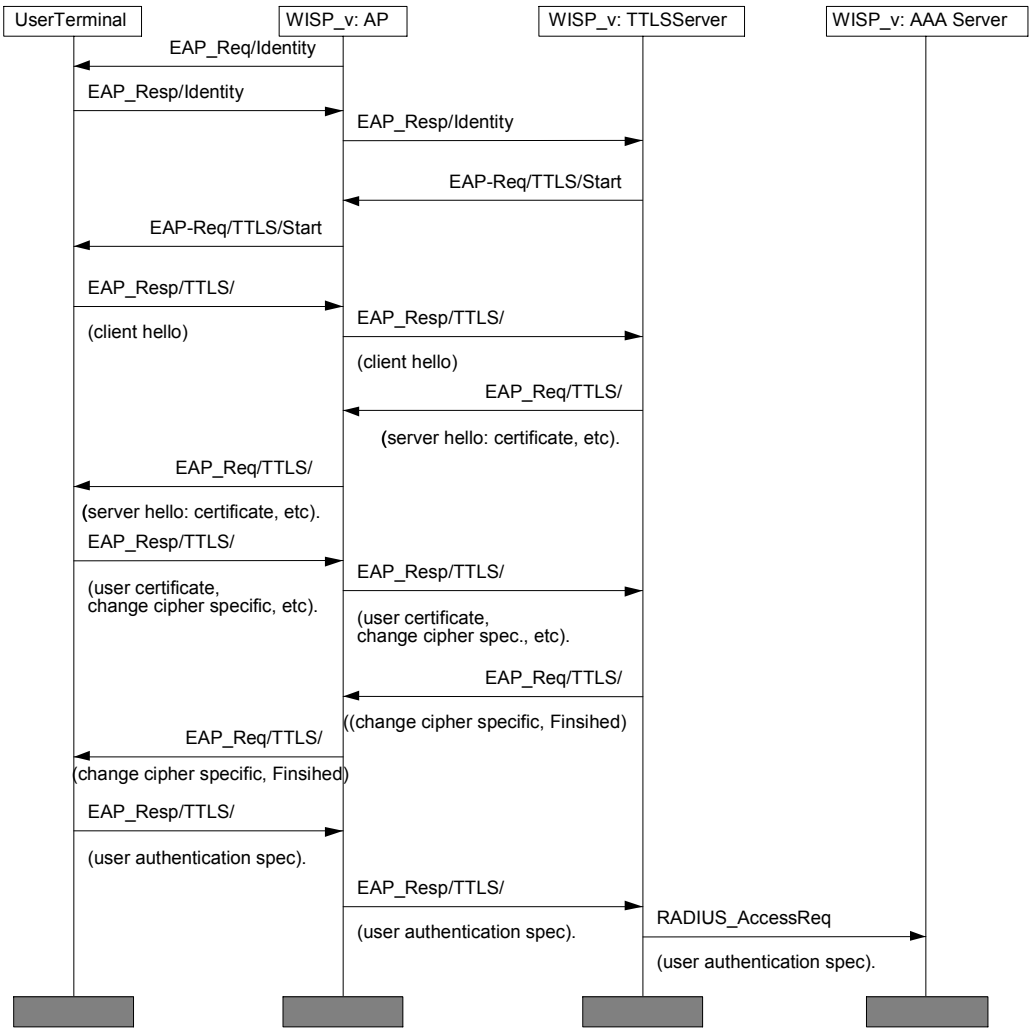

Figure 8. Establishment of a secure wireless communication channel between user terminal and WISP server.

As the client has not been granted connection to the network, in order to verify the server's identity, the client has to be pre-configured with the certificate of the trusted CA to perform the authentication. In case of mutual authentication, after validation of the server, the client sends response containing its own certificate and the TLS change cipher message. The TLS change cipher specific data exchange between the terminal and the server is used for transition in ciphering strategies to inform the receiving party that the subsequent records will be protected under the newly negotiated cipher and keys. At this phase the client and TTLS server have shared key and an agreed cipher suite to secure subsequent communication. Now, the secure TLS channel can be used as a tunnel to perform additional functions as user authentication or authorization, for instance for accounting purposes, based on commonly used protocols as e.g. RADIUS. The communication enabling the user's authentication and providing the user with access to his home devices is presented in Figure 9.

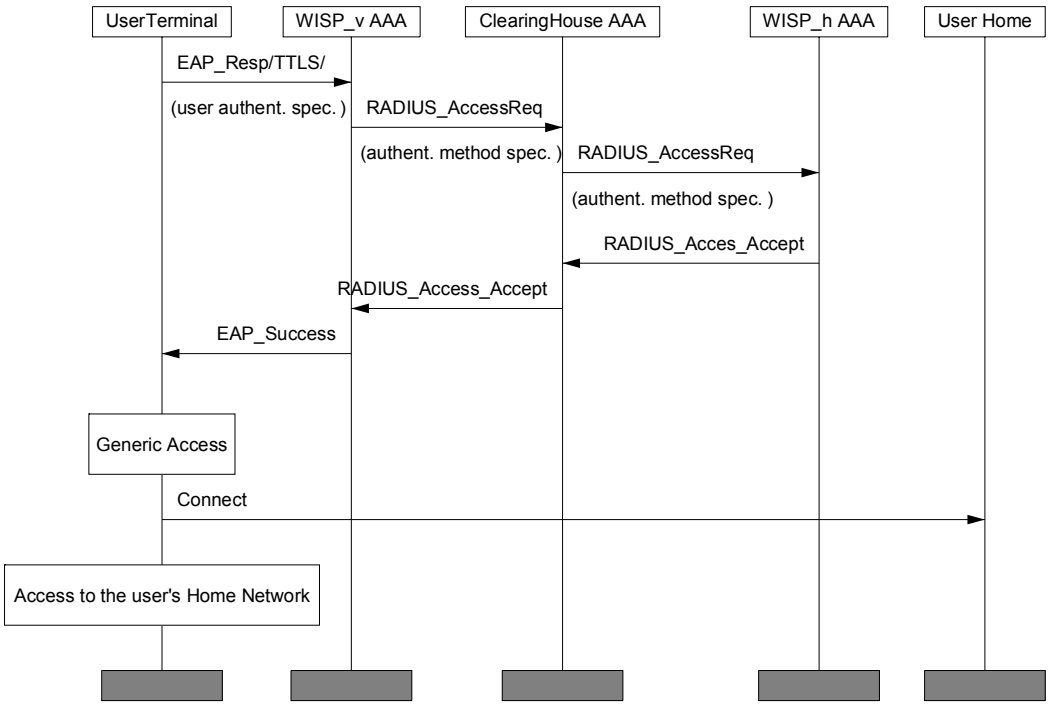

Figure 9. Authentication phases in Internet access and Home Network access from a visited LAN.

The authentication request issued by the client depends on the authentication method. However, as the WISP_v server de-tunnels the data sent over the wireless link, methods like PAP where the user's name and password is sent through the tunnel in clear text should be avoided. The Visited WISP after checking that it does not have an agreement with the user, nor with the user's Home WISP, forwards the user request to the Clearing House with request for authentication. During the communication the two entities are using IPSec. The request is forwarded to the user's Home WISP. IPSec protocol is used between the Clearing House and WISP_h. After WISP_h confirms the authentication of the user, the Clearing House informs WISP_v about the authentication result. The user is allowed to access the Internet, and access his home network. For the whole communication between the user terminal and the home server IPSec with the user's secret key is used. The information exchange is completely transparent to the other parties.

In case of WLAN provided by a 3G operator the access is provided by reusing 3GPP system subscription and authentication, while in case of free Internet access, the authentication phase is obsolete.

After the user gains access to the Internet, for the whole communication between his terminal and the home server, S-HTTP can be used. In case when the user home is at the same an office and applications different than only web applications are required, the user is running IPSec with the user's secret key. The information exchange is completely transparent to the other parties. The user's home network is protected against an unauthorized access by a firewall which at the same time acts as NAT. The authorization for accessing the services is performed by the home server.

WLAN installed at home exposes the network and from a security point of view must be treated as an access network. For that reason to prohibit an unauthorized usage of the home appliances, not only the access to the home server, but also the internal, at home communication has to be secured when WLAN is used. The home server, or appliances must verify the identity of the user to stop an outsider from accessing home WLAN, and privacy against an eavesdropper carrying man-in-the-middle attacks must be provided. Additionally, the unauthorized access to the home server potentially serving as NAT has to be secured, which means that the security should be provided at the link layer. This can be done with 802.11i, but an upgrade of existing hardware may be required to implement Advanced Encryption Standard.

Conclusions

This report addresses security techniques relevant for home networks, and especially remote access to the home from public networks and/or terminals. As it is assumed that the core of the home network will be a standard Ethernet type LAN, much of this is also applicable for example for small office networks. The authentication, authorization and accounting for visited WLAN are also studied. Such public WLAN hotspots differ from a typical home network in the sense that public access will rarely be provided in a home network. Consequently, the number of users having access to the home network is fairly small, and authentication may alternately be based on shared secrets. However, the family members are not the only ones who may have access to the home services, also friends and for example service personnel may be granted rights, although usually limited.

The security measures are studied based on layers (according to the OSI reference model). However, because of different tunneling solutions and a tendency to build "everything over IP and IP over everything", the protocol stacks may no longer appear quite that straightforward, there may be stacks inside stacks and duplicated layers. Also, say when a LAN operates on layer 2, the authentication and access control still need higher layer functionality.

One special point of study in this report was the fact that the same kind of security measures are specified on several layers, resulting in overlapping solutions, less efficiency and more incompatibility. However, recommending only one solution seems not feasible. Depending on the case and the purpose, solutions likely differ from each other. If performance is a bottleneck in handheld computers and mobile phones of today, these problems are likely to be solved in a few years.

References

[Aventail] Comparing Secure Remote Access Options: IPSec VPNs vs. SSL VPNs, Aventail Techincal White Paper. http://www.sphinxcst.co.uk/images/files/IPSecvsSSL.pdf

[Bozoki] Bozoki, E. IP Security Protocols, Dr. Jobb's Journal, December 1999

[EthernetTutorial] Short introduction to wired networking standards and technologies, http://www.lantronix.com/learning/tutorials/

[Hill] Hill, J. An Analysis of the RADIUS Authentication Protocol, InfoGard Laboratories, 2001. http://www.untruth.org/~josh/security/radius/

[IEEE802.11] WLAN standards on-line, http://standards.ieee.org/getieee802/802.11.html

[IEEE802.11iTutorial] Brief technical introduction to the WiFi Protected Access standard, http://www.commsdesign.com/design_center/wireless/design_corner/OEG20021126S0003

[IEEE802.11wg] Working group web pages, http://grouper.ieee.org/groups/802/11/index.html

[Intel 6-4] Intel Technology Journal, Vol. 6, Issue 4, 2002, http://www.intel.com/technology/itj/2002/volume06issue04/preface.htm

[Jokela] Jokela, P. et. al. Host Identity Protocol: Achieving Ipv4 – Ipv6 Handovers without tunneling.

[Kalm] Kalm, J. IEEE 802.11b–verkkojen tietoturva, Elisa Oyj, 2003.

[Keski-Kasari] Keski-Kasari, S., Huhtanen, K. Inventory of web-based solution for inter-NREN roaming, Terena Mobility Taskforce, Deliverable F, revision 2.

[Keski-Kasari2] Keski-Kasari, S. Verkkopalveluiden autentikointi yhteisen käyttäjätietokannan avulla, diplomityö, Tampereen Teknillinen Korkeakoulu, Tietotekniikan osasto, 2002. http://www.wirlab.net/pdf/di_tyo_samikk.pdf

[Kullenwall] Kullenwall, J. Study of security aspects for Session Initiation Protocol, Master thesis, LiTH-ISY-EX-3234-2002, Linköping University, 2002-04-19

[Lee] Lee, B. Wireless security. 802.11 With a focus on security. http://www.cc.gatech.edu/classes/AY2005/cs4803cns_fall/Wireless_Security.ppt

[MacWorld 6/03] Mac World, , July 2003, Behind the Music, p. 56

[Moskowitz] Host Identity Protocol, http://www.tml.hut.fi/~pnr/HIP/draft-moskowitz-hip-08.html

[Movian] http://www.certicom.com/products/movian/movianvpn.html

[OMADRM] DRM Content Format Version 1.0. 08-July-2002, Open Mobile Alliance™. OMA-Download-DRMCF-v1_0-20020708-p http://xml.coverpages.org/OMA-Download-DRMCF-V10-20020708.pdf

[OMADRMREL] "DRM Rights Expression Language". Open Mobile Alliance™. OMA-download-DRMREL-v1_0. http://www.openmobilealliance.org/release_program/enabler_releases.html#Download

[Reyes et al.] Reyes et al. Secure Socket Layer Protocol, project paper, ENTS-650 Network Security, 05/08/2003, http://www.ece.umd.edu/class/ents650/SSL.pdf

[RFC 2617] Franks, J. et al. HTTP Authentication: Basic and Digest Access Authentication. Network Working Group. 1999 http://www.faqs.org/rfcs/rfc2617.html

[RFC 2246] Dierks, T., Allen, C. The TLS Protocol Version 1.0, IETF Network Working Group, RFC 2246, January 1999.

[RFC 2284] Blunk, L. et al. PPP Extensible Authentication Protocol (EAP). Network Working Group http://www.faqs.org/rfcs/rfc2284.html

[RFC 2660] The Secure HyperText Transfer Protocol http://www.faqs.org/rfcs/rfc2660.html

[RFC 2694] Srisuresh, P. et. al. DNS extensions to Network Address Translators (DNS_ALG), RFC 2694.

[RFC 2865] Remote Authentication Dial In User Service (RADIUS)

[RFC 2866] RADIUS Accounting

[RFC 2867] RADIUS Accounting Modifications for Tunnel Protocol Support

[RFC 2868] RADIUS Attributes for Tunnel Protocol Support

[RFC 2869] RADIUS Extensions

[RFC 3022] Srisuresh, P., Egevang, K. Traditional IP Network Address Translator (Traditional NAT). http://www.faqs.org/rfcs/rfc3022.html

[RFC 3027] Holdrege, M., Srisuresh, P. Protocol Complications with the IP Network Address Translator. http://www.faqs.org/rfcs/rfc3027.html

[RFC 3588] Diameter Base Protocol, ftp://ftp.rfc-editor.org/in-notes/rfc3588.txt

[RFC 3261] Rosenberg, J. et. al. SIP: Session Initiation Protocol http://www.faqs.org/rfcs/rfc3261.html

[Roy] Roy, M. Diameter extends remote authentication, Network World, 01/31/00, http://www.nwfusion.com/news/tech/0131tech.html

[SSH Tutor] Simple SSH Tutorial Outline, http://aperiodic.net/phil/ssh/

[SSH vulnerabilities] SSH Vulnerabilities, http://www.physnet.uni-hamburg.de/physnet/security/vulnerability/SSH_vulnerabilities.html

[SSH-IETF] Secure shell (secsh) http://www.ietf.org/html.charters/secsh-charter.html

[Tech-spot] http://www.techspot.com/vb/showthread/t-9646.html

[Utwente] The Internet NG Project, Diameter http://ing.ctit.utwente.nl/WU5/D5.1/Technology/diameter

VTT WORKING PAPERS

VTT TIETOTEKNIIKKA – VTT INFORMATIONSTEKNIK –
VTT INFORMATION TECHNOLOGY

1 Petäkoski-Hult, Tuula, Strömberg, Hanna & Kuukkanen, Hannu. Virike. Ikääntyneet Internet- ja digi-tv-palvelujen käyttäjinä. 2004. 67 s.

10 Lucenius, Jan, Kyntäjä, Timo & Jormakka, Henryka. Security technologies in home and wireless networking environments. 2004. 49 p.

ISBN 951–38–6562–2 (URL: http://www.vtt.fi/inf/pdf/)
ISSN 1459–7683 (URL: http://www.vtt.fi/inf/pdf/)