Mobile Systems I

Burkhard STILLER
Oliver BRAUN
Arnd HEURSCH
(Hrsg.)

Institut für Informationstechnische Systeme, IIS

Bericht Nr. 2002-08
Dezember 2002

Universität der Bundeswehr München

Fakultät für

INFORMATIK

Werner-Heisenberg-Weg 39 • D-85577 Neubiberg

B. Stiller, O. Braun, A. Heursch (edts.): Mobile Systems I
Technical Report No. 2002-08, December 2002
Information Systems Laboratory IIS, Departement of Computer Science
University of the Federal Armed Forces Munich
Werner-Heisenberg-Weg 39, D-85577 Neubiberg, Germany
URL: http://www.informatik.unibw-muenchen.de/

Introduction

The Information System Laboratory (Institut für Informationstechnische Systeme, IIS) of the Department of Computer Science, University of the Federal Armed Forces Munich, Germany continues research and teaching during this fall term 2002 (HT02) in the area of communications. An important topics included mobile systems and protocols, where the focus has been laid on wireless technology, their support services, and their applications.

Mobile systems and protocols for data communications are available for some years now. While the Wide Area Network (WAN) and Local Area Network (LAN) show many different technologies, various support services should be based on similar protocols and systems and should be homogenous for both of those areas. Although traditional telecommunications in the wireless domain supported voice communications mainly, advanced technology integrates data communications as well. Important tasks for a mobile communications infrastructure include the support of mobility of users and devices. In particular, ad-hoc networks enable the set-up of a jointly utilized communication "infrastructures" for an arbitrary number of coming in and leaving participants. Furthermore, a secure and accounted for access to a backbone network from a mobile node are of major importance, to ensure a correct accountability as well as authorized access. While Location-based Services (LBS) become an obvious provisioning goal of mobile service operators, Voice-over-IP (VoIP) in the wireless domain may face a number of problems. Finally, the integrated security aspects of wireless technology provide a network level security, which is depending on the particular technology. Therefore, a review of several challenges and weaknesses of this process of migration toward a mobile world is required and the talks in this seminar are providing an approach to judge their dedicated suitability.

Content

This first edition of the seminar entitled "Mobile Systems I" discusses in the first section a larger overview on wireless technologies and beyond. While the major focus is on the description of major characteristics and differences of existing wireless networks, including channel access schemes, IrDA, Bluetooth, Wireless LAN (WLAN), and WAN technologies, a comparison for major characteristics is included as well. Those approaches in support of mobility, in particular the Mobile Internet Protocol (MIP), are presented in detail afterwards. Micro-mobility, cellular IP, and MIPv6 conclude this section. Furthermore, ad-hoc networks are presented in section three. While basic definitions are provided and the principle of operating ad-hoc networks are illustrated, routing strategies for such

networks highlight major problems for a highly flexible and often changing infrastructureless networking approach. Support services for any type of mobile network are presented in the AAA and extensions section. While the tasks of Authentication, Authorization, and Accounting (AAA) are extended by additional charging functionality, their interoperation and efficient provisioning are of major interest to operators due to scalability and efficiency reasons. Different solutions and their functional decomposition are discussed. Having the technology and those support services in place, wireless Location-based Services (LBS) are a logical consequence to be designed and offered in wireless networks. This section introduces LBS and discusses important prerequisites as well as technical challenges to be solved. A set of suitable application scenarios completes this topic. Due to traditional reasons, voice communications have been a major business case for wireless networks. Therefore, even for wireless IP-based networks voice is of interest. While the wired network approach of VoIP (Voice over IP) achieves varying quality levels depending on the topology and technology utilized, a wireless VoIP scenario faces additional challenges. This section presents the H.323 and SIP (Session Initiation Protocol) signaling approaches and highlights major characteristics for the wireless domain. Finally, the security topic summarizes security approaches for WLAN (IEEE 802.11), GSM, and Bluetooth, mainly addressing network-level security concerns. Based on a structure of attack models, suitable and less suitable solutions are discussed.

Seminar Operation

All interested students worked on an initially offered set of papers and book chapters, relating to the topic titles as presented in the Table of Content below. They prepared a written essay as a focussed presentation, an evaluation, and a summary of those topics of interest. Each of these essays is included in this technical report and allows for an overview on important areas of concern, business models in operation, and problems encountered. In addition, every student prepared a slide presentation of approximately 45 minutes to present his findings and summaries to an audience of students attending the seminar and other interested research assistants. Following a general question and answer part, a student-lead discussion debated lively open issues and critical statements.

Local IIS support for preparing talks, reports, and their preparation by students had been granted by Oliver Braun, Arnd Heursch, and Burkhard Stiller. Many pre-presentation discussions have provided valuable insights in the emerging and moving field of Mobile Systems, both for students and supervisors. Many thanks to all people contributing to the success of this event, which has happened in a small group of highly motivated and technically qualified students and people. This seminar has proven that the effort to be put into effect from students, supervisors, and the professor are significant in the sense, that focussing at a good and satisfactory result requires time, reading, and patience.

Neubiberg, December 2002

Inhaltsverzeichnis

Kapitel 1

Wireless Technologies and Beyond

Christian Schachtner

Wireless Technologien sind keine Errungenschaften der letzten Jahre. Schon Jahrzehnte zuvor versuchte man von Kabeln relativ unabhängig zu sein: Hauptsächlich bei der Sprachübertragung, z.B. Funkverbindung zwischen Militäreinheiten bei Einsätzen. Selbstverständlich sind diese Technologien mit den heutigen kaum zu vergleichen: damals analog, heute digital. Jedoch die Ziele waren die gleichen: Unabhängigkeit, Informationsmöglichkeiten unterwegs, usw. Heutzutage werden neben der Sprache jegliche Art von Daten übertragen, die schnell von Punkt A nach Punkt B sollen, egal wo die sich gerade befinden.
Bei Verkehrsverbindungen zum Beispiel ist es entscheidend, wie weit Start- und Zielpunkt auseinanderliegen. Dementsprechend wählt man die Verbindungsart: Das Fahrrad für kurze Strecken, das Flugzeug für sehr weite Strecken. Es ist auch entscheidend, was man transportieren möchte. Für eine Flasche Milch aus der Stadt genügt es mit dem Fahrrad zu fahren und einen Rucksack mitzunehmen. Will man ein Sofa vom Möbelmarkt abholen, benötigt man zumindest ein größeres Auto. In ähnlicher Weise ist dies auch bei Funkverbindungen: Unterschiedliche Technologien für unterschiedliche Entfernungen und unterschiedlichen Datendurchsatz. Oft sind beide Anforderungen nicht in gewünschtem Ausmaß erfüllbar.
Für kurze Entfernungen gibt es Technologien für den sogenannten Personal Area Network (PAN), für mittlere Entfernungen ($\leq$500m) die Local Area Network (LAN) Standards und für alles darüber hinaus gehende die Sparte des Wide Area Network (WAN). Die Preise für diese Techniken staffelt sich genauso: geringe Entfernung: günstig, große Entfernung: teurer, denn jede einzelne Benutzung muss bezahlt werden.
In jeder Kategorie gibt es mehrere Technologien, die jeweils anderen Anforderungen genügen. Neben Reichweite, Datendurchsatz sind Sicherheit, Verfügbarkeit, geringe Sendeleistung (für Mobilgeräte mit Akkubetrieb sehr wichtig), Infrastrukturunabhängigkeit unter anderem mitentscheidende Faktoren. Diese gilt es bei den einzelnen Technologien anhand von unterschiedlichen Szenarien festzustellen. Wichtige technische Daten sind am Ende des Kapitels in drei Tabellen festgehalten.

Inhaltsverzeichnis

1.1 Einleitung und Motivation

Flexibilität ist alles. Musste bislang der Computeranwender flexibel sein und sich sein Arbeitsplatz nach den Anforderungen des PCs richten, so lässt es mittlerweile die Technik zu, dass sich der Benutzer relativ unabhängig mit seinem tragbaren Computer bewegen kann. Sollte jedoch eine Verbindung zu einem Netzwerk oder zum Internet bestehen, war wiederum die Bewegungsfreiheit des Anwenders eingeschränkt. Doch auch dies ist inzwischen kein unlösbares Problem für die Technik. Kaum sind Geräte auf dem Markt, die kabellosen Anschluss an Netze bieten, ist die Nachfrage dafür und das Verlangen nach leistungsfähigeren Technologien groß.
Aber nicht nur die Netzwerk- und Festnetzkabeln können für den Endanwender überflüssig werden, auch die Anschlusskabeln für Peripheriegeräte wie z.B. Drucker, Scanner, Tastatur, Maus usw. werden nicht mehr benötigt. Die Bewegungsfreiheit des Benutzers wird nur noch durch die Akkulaufzeit der mobilen Geräte beschränkt.

Die Möglichkeiten von Wireless Technologien soll folgender Vergleich zeigen: Ein Geschäftsmann, der immer auf dem neuesten Stand der Börsenwerte und seiner für ihn relevanten Daten sein muss. Zuerst zu einer Zeit, in der es noch keine kabellose Verbindung zum Internet gab: Er ist unterwegs bestenfalls über sein analoges Mobiltelefon mit einem Kollegen verbunden, der ihm die Werte durchgeben kann. Sobald er sich in ein Flugzeug begibt, ist der Informationsfluss unterbrochen. Will er auf einer Konferenz mit Kollegen Daten austauschen, ist dies nur via Kabel möglich. Während der Übertragung sind zumindest die Geräte gebunden. Wenn unser Geschäftsmann einerseits die Daten haben, aber andererseits seinen Laptop nicht zurücklassen will, dann muss er die Datenübertragung abwarten.
Im Hotelzimmer hat er mit seinem analogen Mobiltelefon eventuell nur am Fenster einen guten Empfang, was für ihn bedeutet, dass er jedesmal wenn er damit telefoniert, seine Arbeit abbricht und sich nicht mit dem Telefon vom Fenster wegbewegen kann.

Jetzt der gleiche Geschäftsmann mit Wireless Technologie Geräten: Unterwegs ist er über die Börsenwerte und für ihn relevanten Daten gut informiert, sei es über sein digitales Mobiltelefon, das zu seinem Firmennetz bzw. zum Internet verbunden ist, oder aber über WLAN, die ihm an manchen "Points of Interest" wie Bahnhöfen, Flughäfen usw., gegen Bezahlung ebenfalls Anschluss zum Internet bietet. Selbst Flugzeugreisen müssen nicht mehr den Informationsfluss abbrechen. Da die Sendeleistung von WPAN- bzw. WLAN-Geräten maximal 10% von heutigen WWAN-Geräten entspricht, laufen bei ersten Fluggesellschaften Projekte, die eine Internetanbindung für die Passagiere ermöglicht. Die Daten werden zwischen den Flugzeugen und den Backbones über Satelliten übertragen. Dies ist momentan auch noch die größte Problemstelle bei diesem Angebot, da der gebotene Datendurchsatz noch sehr gering ist.
Bei der Konferenz ist unser Geschäftsmann in der Lage, Daten mit Kollegen auszutauschen, ohne seine Bewegungsfreiheit einzuschränken, solange er in der Reichweite der verwendeten Technologie bleibt. Ob er zum Essen geht oder sich einen Vortrag anhört, der Datenabgleich funktioniert trotzdem, ohne dass er sein Gerät zurücklassen muss.
Im Hotelzimmer hat er trotz Digitaltechnik nur am Fenster Empfang, diese Einschränkungen werden trotz modernster Technik noch weiterhin vorkommen. Doch diesmal muss

er sich nicht nach seinem Handy richten, da er mit einem kabellosen Headset die Freiheit besitzt, am Schreibtisch sitzen zu bleiben oder nach Belieben auf und ab zu gehen. Ein Vergessen des Mobiltelefons ist nur beschränkt möglich, da dem Geschäftsmann über ein bei ihm tragendes Gerät, sei es das Headset oder sein PDA, mitgeteilt wird, sobald er sich aus der Reichweite des Telefons bewegt.

Im Folgenden werden Technologien für WPANs, WLANs und WWANs vorgestellt und aufgrund ihrer Fähigkeiten die Einsatzmöglichkeiten dargestellt.
Satellitentechnologien, wie z.B. Iridium, Globalstar oder Teledesic, werden hier nicht behandelt.
Zum besseren Verständnis werden im ersten Teil ein paar wichtige Funktionsgrundlagen für Wireless Technologien erläutert. Am Schluss befinden sich drei Tabellen, die die wichtigsten technischen Merkmale der genannten Standards beinhalten. Die dritte Tabelle "Übersicht WWAN" zeigt Preisabschätzungen für die Benutzung dieser Technologien.
Alle Angaben sind ohne Gewähr und sind auf dem Stand von November 2002.

1.2 Grundlagen

Damit man die einzelnen Wireless-Standards von ihrer Funktionsweise her verstehen kann, ist es notwendig gewisse Grundbegriffe aus den Wireless-Technologien zu kennen. Vorhandene Vor- bzw. Nachteile sind meist ein Resultat aus den verwendeten Technologien, z.B. für den Kanalzugriff. Dieses Kapitel soll die Grundlagen kurz in ihren wesentlichen Merkmalen beschreiben, ohne den Anspruch zu haben vollständig in Hinblick auf Funktionsweise, Spezialfälle und Detailinformationen zu sein.

1.2.1 Kanalzugriffstechnologien

Kanalzugriffstechnologien beschreiben, wie es erreicht wird, dass mehrere Geräte, die auf das gleiche Übertragungsmedium zugreifen, Daten senden bzw. empfangen können, ohne dabei den Datenversand bzw. -empfang von bzw. für andere übermäßig zu stören.

Frequency Division Multiple Access (FDMA)

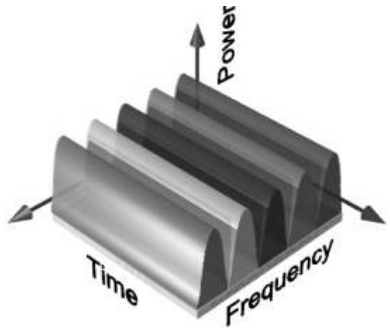

Abbildung 1.1: FDMA: Einteilung nach Frequenzen [29]

Bei FDMA wird die gegebene Bandbreite in mehrere Kanäle unterteilt. Je breiter ein Kanal sein soll, desto weniger Kanäle stehen zur Verfügung. Ein Kanal stellt dann eine direkte Verbindung zwischen zwei Geräten dar, die ohne weitere Einschränkungen ständig für eine unidirektionale Datenübertragung verwendet werden kann. Soll eine bidirektionale Verbindung bestehen, funktioniert dies bei reinem FDMA nur bei Verwendung eines zweiten Kanals. Kanäle können nicht unmittelbar aneinander grenzen, da sonst Störungen auftreten, die

nicht kompensiert werden können. Deswegen muss immer ein gewisser Abstand eingehalten werden: der sogenannte Guard-Abstand. Hierbei können bis zu 50% der gesamten Bandbreite verloren gehen.[29]

Folgende Wireless-Standards verwenden FDMA: u.a. analoge Funkübertragung.

Orthogonal Frequency Division Multiplexing (OFDM)

OFDM ist eine Spezialisierung von FDMA. Auch hier wird die gegebene Bandbreite in mehrere Kanäle unterteilt, mit dem Unterschied, dass der Guard-Abstand wegfällt, da die benachbarten Kanäle orthogonal zueinander gestellt werden. Dadurch können die Kanäle sich teilweise überschneiden. Die Folge ist eine wesentlich bessere Ausnutzung der Bandbreite.
Eine Verbindung zwischen zwei Geräten wird auf mehrere Subkanäle gleichzeitig verteilt, dadurch können auch mehr Daten pro Zeiteinheit übertragen werden. Diese erhöhte Bandbreite in Verbindung mit Forward Error Correction macht OFDM sehr leistungsfähig.

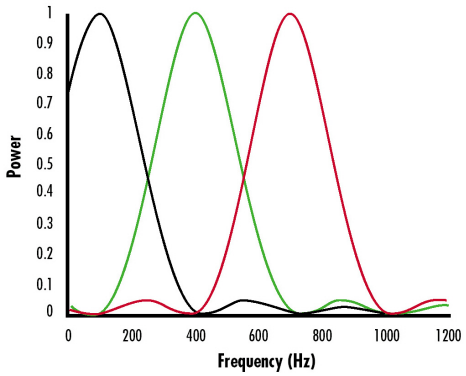

Abbildung 1.2: Überlagerung von Nachbarkanälen möglich [36]

Orthogonal bedeutet hier: Am Spektrumscheitelpunkt jedes einzelnen Kanals sind die Spektren jedes anderen Kanals 0 (siehe Abbildung 1.2). Diese Modulation wird beim Sender mit Hilfe einer Inversen Diskreten Fourier Transformation (IDTF) erreicht. Beim Empfänger werden die überlagerten Kanäle durch eine Fourier Transformation wieder getrennt.[1]

Folgende Wireless-Standards verwenden OFDM: u.a. 802.11a/g/h, HiperLAN2.

Time Division Multiple Access (TDMA)

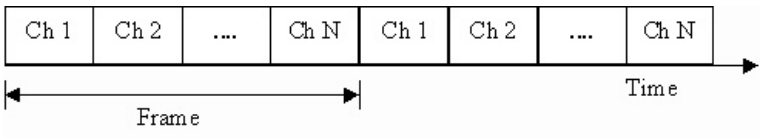

Abbildung 1.3: TDMA: Aufteilung in Zeitschlitze [29]

TDMA unterteilt die Bandbreite in eine gegebene Anzahl von gleichlangen Zeitschlitzen. Jedem zugreifenden Gerät steht (mindestens) ein Zeitschlitz pro Periode zur Verfügung. Ob dieser Zeitschlitz zum Versand oder Empfang von Daten genutzt werden kann, liegt an dem zuteilenden Gerät. Wird pro Periode Versand und Empfang realisiert, wird dies auch als Time Division Duplex (TDD) bezeichnet.

Folgende Wireless-Standards verwenden TDMA: u.a. HiperLAN2.

TDMA/FDMA

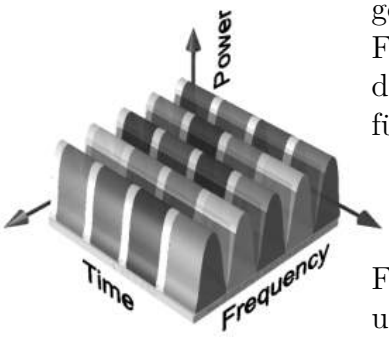

Abbildung 1.4: TDMA/FDMA [29]

TDMA/FDMA ist ein Hybrid von TDMA und FDMA. Die gegebene Bandbreite wird zuerst in mehrere Kanäle nach FDMA unterteilt. Diese entstandenen Kanäle werden nach dem TDMA Prinzip für mehrere zugreifende Geräte zu Verfügung gestellt.

Folgende Wireless-Standards verwenden TDMA/FDMA: u.a. DECT, GSM.[SkyDSP]

Code Division Multiple Access (CDMA)

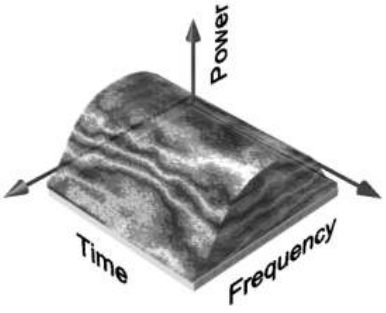

Abbildung 1.5: CDMA: Mehrere Verbindungen pro Kanal möglich [29]

Bei CDMA wird die gesamte Bandbreite für alle zugreifenden Geräte, oder ein Kanal für mehrere Geräte verwendet.

Die zu übertragenden Daten werden auf die gesamte Bandbreite bzw. auf den verfügbaren Kanal gespreizt und mit einem PN-Code (Pseudo random Noise) verschlüsselt. Siehe Abbildung 1.6.

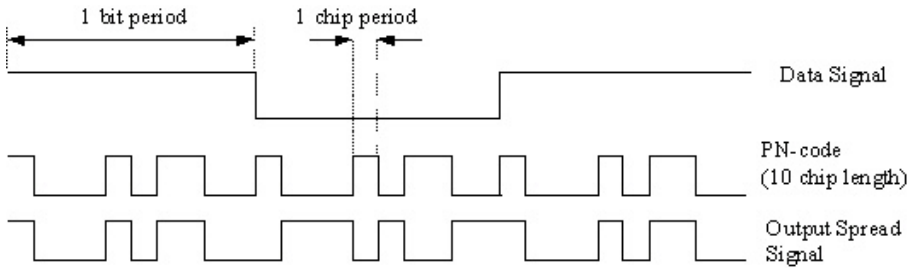

Abbildung 1.6: CDMA: Codierung des Signals [29]

Jede Verbindung hat seinen eigenen PN-Code, der sowohl dem Sender wie auch dem Empfänger bekannt sein muss. Nur damit ist eine Dekodierung auf Empfängerseite möglich. Diese Technik erlaubt das gleichzeitige Senden auf demselben Kanal von mehreren Geräten aus.

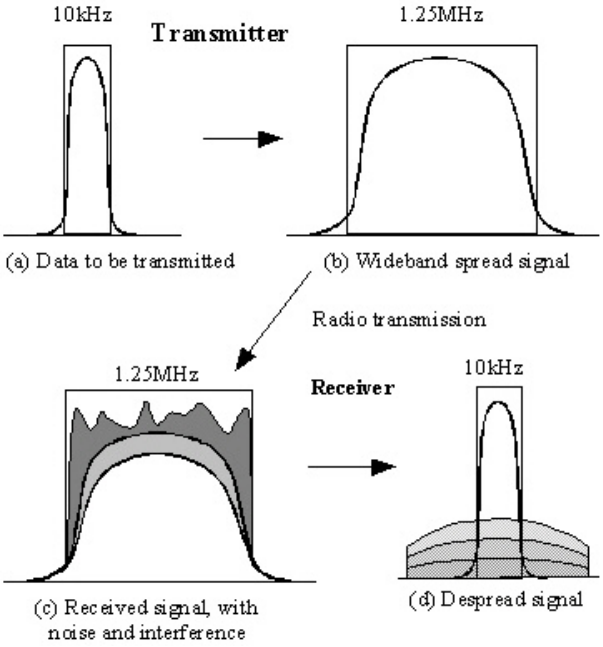

Abbildung 1.7: CDMA: Ablaufdiagramm [29]

Folgende Wireless-Standards verwenden CDMA: u.a. DSSS bei 802.11 (nur zur Verminderung der Störanfälligkeit, kein Multi-Access).

Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA)

Bei CSMA/CA greifen mehrere Geräte auf den gleichen Kanal zu. Allerdings werden im Gegensatz zu CDMA die Übertragungen unbrauchbar, sobald mindestens zwei Geräte auf demselben Kanal gleichzeitig senden. Damit dies nicht passiert, wird eine Vermeidung von Kollisionen soweit wie möglich versucht.
Der Ablauf ist wie folgt: Eine sendewillige Station überprüft den zu verwendenden Kanal auf Verfügbarkeit. Wenn frei, wartet die Station eine DIFS-Zeit (Distributed Inter Frame Space) und überprüft ob der Kanal noch frei ist. Sollte er jetzt belegt sein, wartet die Station bis der Kanal frei ist, hängt wieder eine DIFS-Zeit an und überprüft auf Verfügbarkeit. Dies macht die Station solange, bis dass der Kanal unbelegt ist. Bei einem freien Kanal fängt die Station mit der Datenübertragung an. Der Prozess ist abgeschlossen, wenn die sendende Station nach der Datenübertragung vom Empfänger ein ACK-Paket (acknowledge) erhält, das innerhalb der DIFS-Zeit des eben gesendeten Paketes übertragen wird. Erhält dies der Sender nicht, geht er davon aus, dass eine Kollision beim Übertragen des Pakets stattgefunden hat und beginnt von vorn.

Folgende Wireless-Standards verwenden CSMA/CA: alle 802.11-Standards.

1.2.2 Trägermodulation

Unter Trägermodulation versteht man die Technik, mit der die Phasen eines Frequenzbandes so verändert werden, dass sie Dateninformationen tragen, die am Empfänger wieder entschlüsselt werden können. Die niedrigste Modulationsart ist die, bei der ein Bit pro Phase transferiert wird, z.B. bei BPSK. Es gibt aber auch Möglichkeiten mehrere Bits pro Phase zu übertragen, bei QAM64 sind es z.B. 6 Bits. Je mehr Bits pro Phase übertragen werden, desto anfälliger sind diese Technologien gegenüber Störungen. Passiert dies, wird nach Möglichkeit auf eine niedrigere Modulation gewechselt. Sender und Empfänger müssen nach der gleichen Trägermodulation arbeiten und sich bei jedem Verbindungsaufbau synchronisieren.

Phase-Shift-Keying (PSK)

PSK bewirkt eine Phasenverschiebung bei der zu versendenden Phase. Es gibt unterschiedliche Stufen von PSK, u.a.:

- BPSK (Binary PSK) überträgt ein Bit pro Phase.
  Keine Phasenverschiebung steht für 0.
  Eine Phasenverschiebung von 180° steht für 1
- QPSK (Quadrature PSK) überträgt 2 Bits pro Phase. Siehe Abbildung 1.8.
  Um z.B. die Bitfolge 11 zu modulieren, ist somit eine Phasenverschiebung von 225° notwendig.
- 8-PSK überträgt 3 Bits pro Phase
- 256-PSK überträgt 8 Bits pro Phase

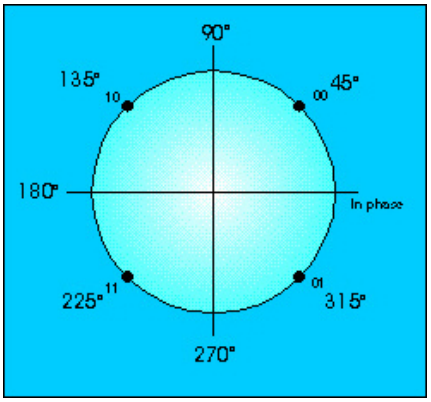

Abbildung 1.8: QPSK [32]

Mit der Zahl bzw. dem Ausdruck vor PSK läßt sich der Winkel zwischen zwei benachbarten Bitfolgen ausrechnen: $\alpha = \frac{360}{PSK-Zahl}°$
Je höher die Stufe des PSK, desto störanfälliger wird die Übertragung, da der tolerierbare Fehlerwinkel, $\frac{\alpha}{2}$, sehr klein wird: bei BPSK beträgt er 90°, bei 256-PSK 0,703°.[SkyDSP]

Folgende Wireless-Standards verwenden PSK: u.a. alle 802.11-Standards, HiperLAN2.

Quadrature Amplitude Modulation (QAM)

Im Gegensatz zu PSK wird bei QAM neben dem Phasenwinkel auch die Amplitude verändert.

- QAM16 überträgt 4 Bits pro Phase
- QAM64 überträgt 6 Bits pro Phase
  Siehe Abbildung 1.9:
  Um z.B. die Bitfolge 010101 zu modulieren, ist eine Phasenverschiebung von 135° und eine Amplitudenerhöhung um das 4,24-fache notwendig

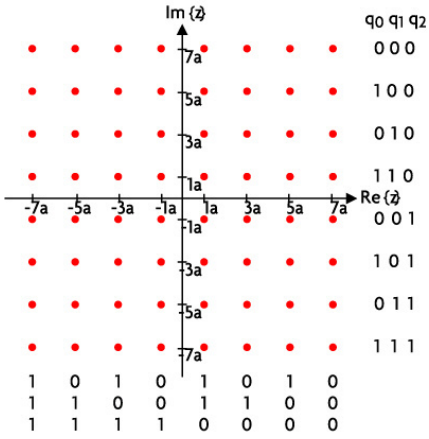

Abbildung 1.9: QAM64 [8]

Folgende Wireless-Standards verwenden QAM: u.a. 802.11a, HiperLAN2.

Gaussian Frequency Shift Keying (GFSK)

Bei GFSK wird das Frequenzspektrum der Trägerfrequenz moduliert. Das heißt, dass die Abweichung zur Trägerfrequenz die kodierte Bitfolge darstellt. Dafür werden zuerst die

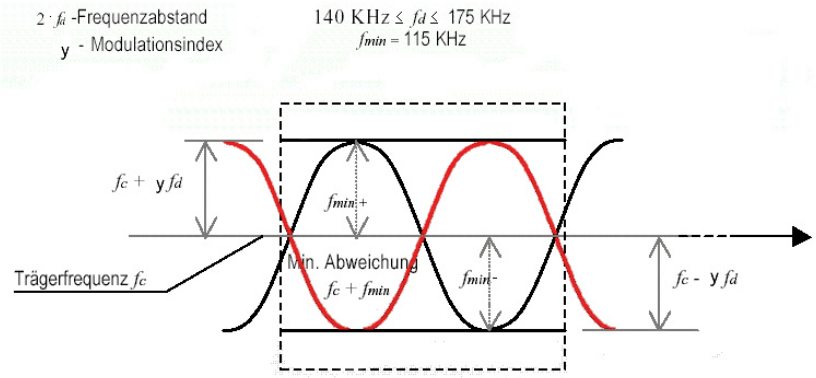

Abbildung 1.10: Frequenzmodulation [20]

binären Rechtecksignale mit Hilfe eines Gauß-Tiefpasses in Signale mit sinusähnlichem Anstieg gewandelt. Grund hierfür ist, dass bei einer Frequenzmodulation von Rechtecksignalen kurzzeitige Störsignale auf benachbarten Frequenzen auftreten. Diese Störsignale treten bei weichen Übergängen nicht auf.
Anschließend wird die Trägerfrequenz mit den entstandenen Signalen moduliert.
Bei 2-Level GFSK: Eine positive Abweichung von der Trägerfrequenz steht für die binäre 1, eine negative für die binäre Null.

Folgende Wireless-Standards verwenden GFSK: u.a. Bluetooth, DECT.

1.3 Beschreibung der Technologien

Die Wireless-Standards werden hier in folgende drei Kategorien eingeteilt: Wireless Personal Area Network (WPAN), Wireless Local Area Network (WLAN) und Wireless Wide Area Network (WWAN). Das Einteilungskriterium ist in erster Linie die Einsatzreichweite. Während WPAN die direkte Umgebung des Benutzers umfasst, ist es bei WLAN ein lokal beschränktes Netz und bei WWAN alles was darüber hinaus geht. Überschneidungen sind möglich und teilweise erwünscht. Im folgenden werden in jeder Kategorie die meist verbreitetsten, innovativsten und voraussichtlich zukunftsträchtigsten Standards in Hinblick auf ihre Funktionsweise, ihren Leistungsmerkmalen und ihrem Einsatzspektrum vorgestellt. Da Daten über z.B. Bandbreite, Reichweite und Sendeleistung allein nicht zwangsläufig entscheidende Faktoren für die Einstufung in Brauchbarkeit und Alltagseinsatz sind, werden alle vorgestellten Technologien an fünf verschiedenen Szenarien geprüft.

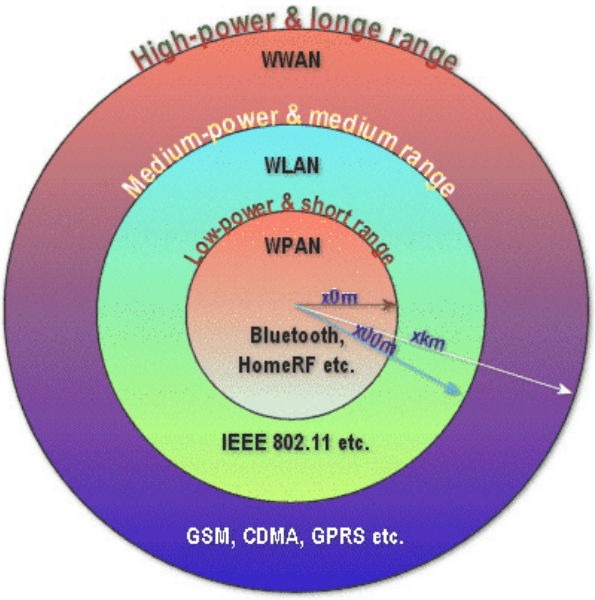

Abbildung 1.11: Die drei Wireless Gebiete

Szenario 1: Ein Benutzer will mit seinem PDA ein Dokument ausdrucken, ohne den Drucker mit einem Kabel an sein PDA anschliessen zu müssen.
Szenario 2: Studenten wollen in der U-Bahn Daten über ihre PDAs austauschen.
Szenario 3: Bei einem Kongress wollen mehrere Teilnehmer Daten untereinander austauschen.
Szenario 4: Ein Benutzer will zuhause mit seinem Laptop immer eine Zugangsmöglichkeit zum Internet haben und dabei ungebunden sein.
Szenario 5: Eine Firma will seinen Mitarbeitern für flexiblere Arbeitsmöglichkeiten einen kabellosen Zugang zum Firmennetz stellen, das sehr vertrauliche Daten enthält.

1.3.1 Wireless Personal Area Network (WPAN)

WPANs decken die direkte Umgebung des Benutzers ab. Wireless-Technik in dieser Kategorie wird hauptsächlich als Kabelersatz für Peripheriegeräte eingesetzt, um die Bewegungsfreiheit am Arbeitsplatz nicht einzuschränken. Ein weiteres wichtiges Kriterium in dieser Kategorie ist die Ad-hoc-Fähigkeit: Ohne vorhandene Infrastruktur eine Netzverbindung zu anderen Benutzern aufbauen zu können. Mehr zu diesem Thema findet man unter Kapitel 3: Ad-hoc Netzwerke.
Alle WPAN-Techniken verwenden unlizensierte Frequenzbänder, somit ist der Betrieb dieser Geräte kostenlos.

IrDA 1.1 - Infrared Data Association

Mit 4 MBit/s, ungefähr 1 m Reichweite und einer Sendeleistung von 40 mW stellt der IrDA 1.1-Standard eine recht schnelle, stromsparende und weit verbreitete, aber auch eine sehr unflexible Technik dar. IrDa wird hauptsächlich zum Datenaustausch und Steuern anderer Geräte verwendet, was allerdings immer einen direkten Sichtkontakt zwischen den Geräten voraussetzt. Dies ist der größte Nachteil von IrDA. Im Alltag wird gerade diese Anforderung als äußerst benutzerunfreundlich angesehen. Der Ad-hoc Betrieb ist standardmäßig nur Point-to-Point möglich. Sollen mehrere Geräte im Ad-hoc-Netz betrieben werden, ist ein Ir-Hub notwendig. Dies ist aber widersprüchlich zu der Anforderung, dass keine Infrastruktur benötigt wird.
IrDA-Anschlüsse findet man nahezu an allen portablen Kommunikations- und Datenverarbeitungsgeräten wie Mobiltelefone, PDAs oder Laptops. Bei Peripheriegeräten gibt es entweder standardmäßig einen IrDA-Port, oder es läßt sich meist mit einem sogenannten IrDA-Dongle nachrüsten wie z.B. bei Druckern.

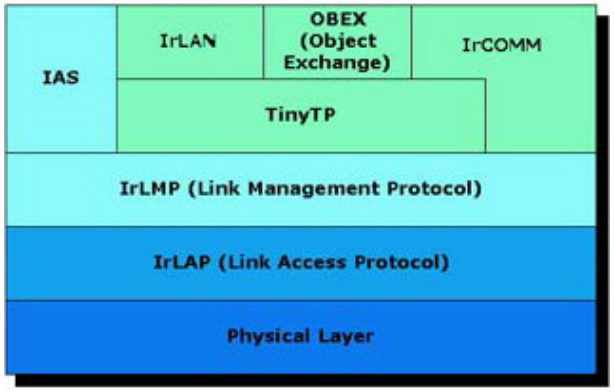

Abbildung 1.12: IrDA-Protokollschichten [39]

- **Physikal Layer:** Infrarotlicht zwischen 850 und 900 nm
- **IrLAP (Infrared Link Access Protocol):**
  - definiert IrDA Frame (8 Bit Adressfeld, 8 Bit Control-Feld, 2045 Byte Payload)
  - verbindungsorientierte und -lose Modi zwischen Kommunikationspartnern
- **IrLMP (Link Management Protocol):**
  - LM-IAS (Link Management Information Access): für Ad-hoc Verbindungen
  - LM-MUX (Link Management Multiplexer): regelt Zugriff auf Ir-Schnittstelle
- **TinyTP (Tiny Transport Protocol):** Einbettung von Kontrollinformationen in die Datenströme (fehlt bei IrLMP)

- **IrLAN:**
  - Anbindung an bestehendes Kabelnetzwerk
  - Access-Point-Mode für Ir-Hubs
  - P2P
- **IrCOMM (Communication Protocol):** Emulation von seriellen und parallelen Ports

Zu Szenario 1: Geeignet, vorausgesetzt der Drucker besitzt einen Ir-Port. Einschränkung: Geringe Reichweite und durch Notwendigkeit des Sichtkontakts zwischen den beiden Ir-Ports unflexibel.
Zu Szenario 2: Gut geeignet. Geringe Reichweite und Notwendigkeit eines Sichtkontakts der Geräte ist hier nicht störend.
Zu Szenario 3: Eingeschränkt geeignet. Ohne Ir-Hub sehr umständlich, da jeweils nur 2 Benutzer Daten austauschen können.
Zu Szenario 4: Nicht geeignet. Mobilität ist hier nicht möglich.
Zu Szenario 5: Nicht geeignet. Anspruch von flexibleren Arbeitsmöglichkeiten kann nicht erfüllt werden.

Eine Weiterentwicklung von IrDA ist das **Fast IrDA** mit einem höheren Datendurchsatz von 16MBit/s. Alle anderen Daten ändern sich nicht.

Bluetooth

Bluetooth/A hat, mit 1 MBit/s Durchsatz, maximal 15 m Reichweite, einer Sendeleistung von 1 mW, volle Ad-hoc-Netzwerk-Fähigkeit und keine Anforderung auf Sichtkontakt zwischen verbundenen Geräten, gute Möglichkeiten sich als der Kabelersatz-Standard für Peripheriegeräte zu etablieren.
Bluetooth/A ist zur Zeit noch in der Anlaufphase, dementsprechend hoch sind die Preise für die wenigen Geräte. Unterstützung ist zur Zeit serienmäßig oder als optionales Zubehör bei Mobiltelefonen, PDAs und Laptops vorhanden. Geplant ist eine große Produktpalette mit hohen Stückzahlen, die eine deutliche Preissenkung zur Folge haben soll.
Neben paketbasierten Datenverbindungen wird auch ein extra Kanal für Sprachübertragung angeboten, der immer zur Verfügung steht und eine Bandbreite von 64 kbit/s in beide Richtungen hat. (Siehe Abbildung 1.13 Audio) Dieser synchrone Kanal wird zum Beispiel von Bluetooth Headsets in Verbindung mit speziellen Mobiltelefonen zur Audioverbindung benutzt.
Der 1 MBit/s Datendurchsatz ist inklusive dem 2x 64 kBit/s breiten Sprachkanal. Diese Ressource kann nicht für paketbasierte Datenübertragung freigegeben werden. Somit steht für den asynchronen Datenkanal ein maximaler Datendurchsatz von jeweils 433,9 kBit/s für den Up- und Downstream zur Verfügung.

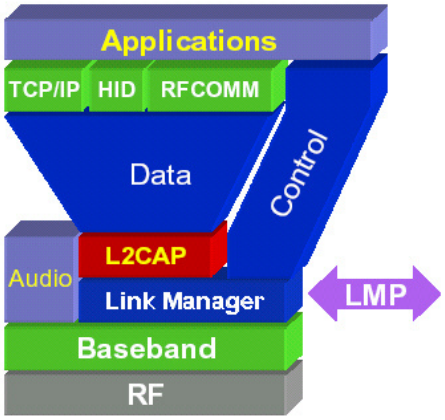

Abbildung 1.13: Bluetooth-Protokollschichten [16]

- **Radio Frequency (RF) und Baseband (PHY):**
  - 2,4 GHz ISM-Band (Industrial, Scientific and Medical)
  - Verwendung von TDD
  - Frequency Hopping (FH): alle 625 μs wird der Kanal gewechselt, 79 Kanäle verfügbar (23 in Frankreich)
  - Trägermodulation mit GFSK
- **Audio:**
  - Direktmodus für PCM-Rohdatenformat
  - diese Audiodaten werden parallel zu anderen Daten übertragen
- **LMP (Link Manager Protocol):**
  - Verbindungsaufbau
  - Steuerung
  - Verschlüsselung
- **L2CAP (Logical Link Control & Adaption Protocol):**
  - Asynchroner verbindungsloser Modus (ACL): Datenübertragung
  - Synchroner verbindungsorienter Modus (SCO): Sprachverbindung
- **RFCOMM:** Emulation von seriellen Ports

Bluetooth unterstützt Point-to-Point (a), Point-to-Multipoint (b), auch Piconet genannt, und Scatternets (c). Scatternets sind mehrere miteinander verbundene Piconets.

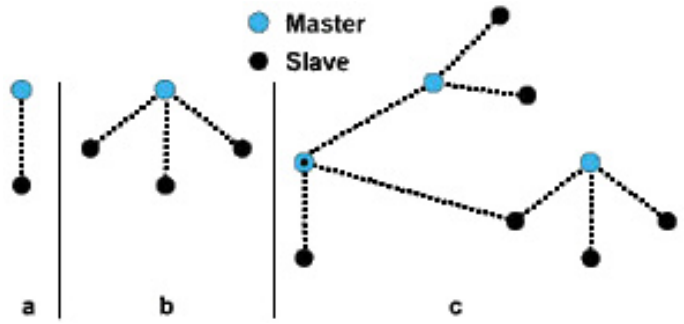

Abbildung 1.14: Bluetooth Ad-hoc Modi [39]

Zu Szenario 1: Sehr gut geeignet. Voraussetzung: Der Drucker hat einen Bluetooth-Port.
Zu Szenario 2: Eingeschränkt geeignet. Einschränkung: Datendurchsatz nicht zu gering.
Zu Szenario 3: Eingeschränkt geeignet. Einschränkung: Datendurchsatz nicht zu gering.
Zu Szenario 4: Gut geeignet. Voraussetzung: Durch die geringe Reichweite müßten ausreichend Geräte mit Bluetooth räumlich gut verteilt sein, damit eine nahtlose Verbindung bestehen kann. Einschränkung: Relativ geringer Datendurchsatz.
Zu Szenario 5: Beschränkt geeignet. Einschränkung: Zu geringer Datendurchsatz um alle Mitarbeiter genügend Bandbreite zu Verfügung zu stellen.

Bluetooth/B hat eine höhere Reichweite von 150m bei gleicher maximalen Bandbreite. Damit würde sich das Einsatzspektrum deutlich erhöhen. Bislang sind noch keine Geräte mit Bluetooth/B auf dem Markt.

1.3.2 Wireless Local Area Network (WLAN)

WLANs beschreiben Netzwerke, die auf ein Gebäude, Gebäudekomplex oder Gelände beschränkt sind und mehreren Benutzern Zugriff auf Netzdienste gewähren.
Wireless-Technik in dieser Kategorie, soll nicht zwangsläufig das bestehende Netz ersetzen sondern erweitern und den Benutzern mehr Bewegungsfreiraum bieten. Allerdings ist dieser Bewegungsradius auf maximal wenige hundert Meter begrenzt. Da lizenzfreie Frequenzbänder verwendet werden, ist die Nutzung an sich kostenfrei. Ausnahmen können bestehen, da schon Bestrebungen in Gange sind, in Ballungszentren WLAN-Zugänge kommerziell anzubieten, z.B. Englischer Garten, München.
Die Hauptzielgruppen dieser Technologien sind u.a. Firmen und Universitäten. Da sich in diesen Netzen eine große Menge an sensiblen Daten befindet, ist der Sicherheitsaspekt mit die wichtigste Forderung, die an die Wireless-Standards gestellt wird. Besonders die Abhörsicherheit soll hierbei gewährleistet sein, da man davon ausgehen kann, dass Administratoren ihre Netzwerke, die bevorzugt Ziel mehrerer Angriffe sind, durch geeignete Authentisierungs- und Authorisierungsmechanismen schützen. Siehe hierzu Kapitel 4: AAA and Extensions for Wireless Services.

Digital Enhanced Cordless Telecommunications (DECT)

DECT ist mit bis zu 2 MBit/s Datendurchsatz (bei Nutzung aller Kanäle), maximal 300 m Reichweite und einer Sendeleistung von 250 mW ein Standard, der besonders für den privaten Nutzer und für kleine bis mittelgroße Firmen mit geringem Datenaufkommen eine kostengünstige Möglichkeit bietet, einen mobilen Zugang zum Heim- bzw. Firmennetz zu haben.
DECT bietet neben der Möglichkeit für Datendienste auch die Unterstützung von Telefondiensten an. Somit handelt es sich hierbei nicht um eine reine WLAN-Technik.
In größeren Firmennetzen konnte sich DECT nicht durchsetzen, da sich die Bandbreite als zu gering erwiesen hat.
Ohne Access Point ist eine P2P Verbindung zwischen zwei DECT-Geräten möglich.

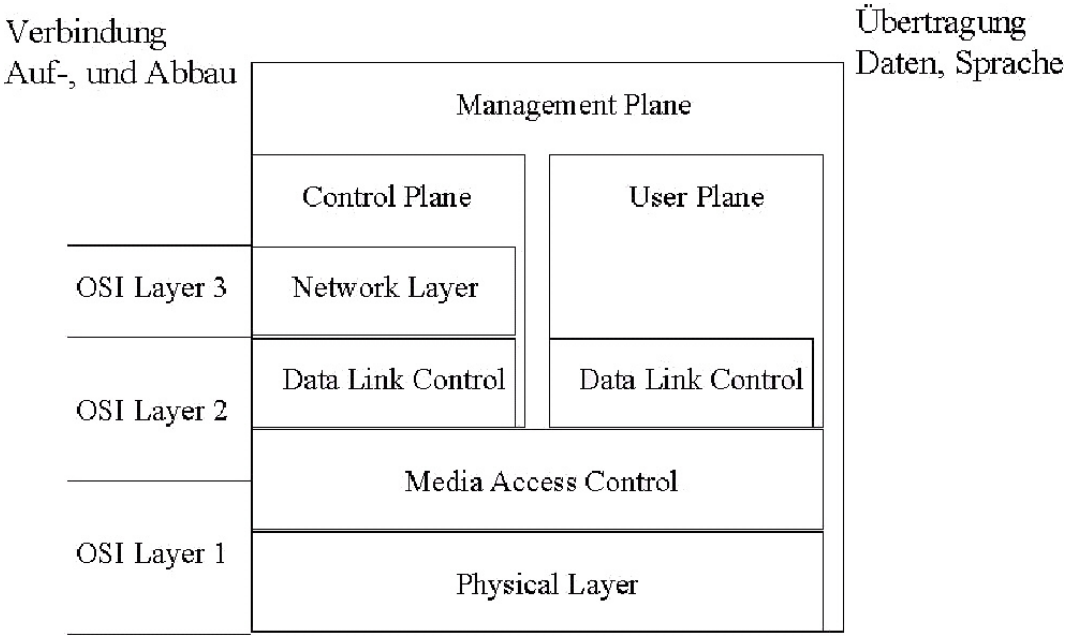

Abbildung 1.15: DECT-Protokollschichten [10]

- **Physical Layer:**
  - 1,88 -1,9 GHz
  - Kanalzugriff nach TDMA/FDMA (10 Kanäle, je 24 Slots)
  - Verwendung von TDD (je 12 Slots)
  - Trägermodulation nach GFSK
- **Media Access Control:**
  - selektiert Funkkanal
  - erstellt Datenpakete (32 Bit Synchronisierung, 64 Bit Signalisierung, 324 Bit Nutzdaten)

- **Data Link control:**
  - Verschlüsselung
  - Frame Switching, Routing, Forward Error Correction
- **Network Layer:** Informationsaustausch für Verbindungsaufbau und -abbau

Zu Szenario 1: Eingeschränkt geeignet. Einschränkung: DECT-fähiges Zubehör für PDAs ist selten. Drucker muss über Drucker-Server angesprochen werden, da es keine DECT-Drucker-Ports gibt.
Zu Szenario 2: Geeignet. DECT-Geräte sind aber größer als bei vergleichbaren Standards.
Zu Szenario 3: Schlecht geeignet. Ohne Infrastruktur ist nur Point-to-Point möglich.
Zu Szenario 4: Gut geeignet. Ausnahme: Datendurchsatz reicht dem einzelnen Nutzer nicht aus.
Zu Szenario 5: Eingeschränkt geeignet. Einschränkung: Datendurchsatz ist für ein Firmennetz meist zu gering.

Die IEEE 802.11 Familie

Die Aufgabe der IEEE (Institute of Electrical and Electronics Engineers) 802.11 Gruppe ist es, Standards für Wireless LANs zu entwickeln, die nur die untersten zwei Schichten der OSI-Struktur definieren. Somit sollen 802.11-WLAN-Standards zu allen anderen IEEE 802 Netzwerkstandards kompatibel sein. Andere wie z.B. ATM werden nicht unterstützt.

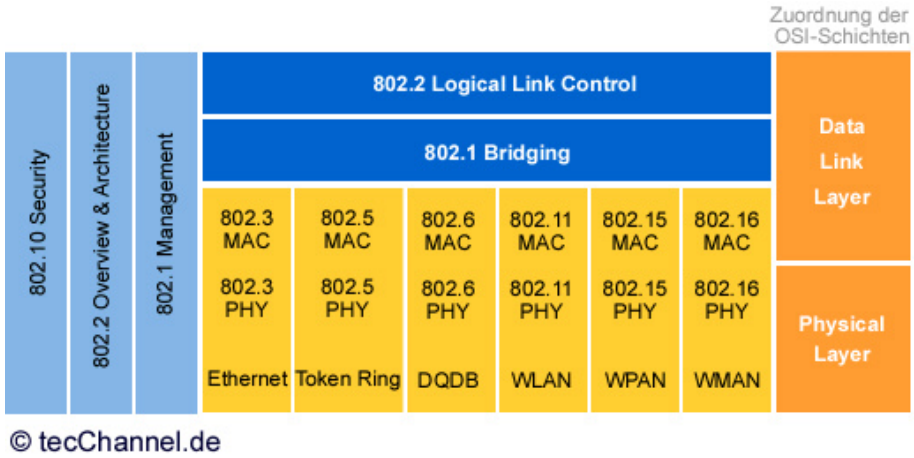

Abbildung 1.16: IEEE 802-Familie [30]

In der IEEE 802.11 Gruppe gibt es verschiedene Task-Groups, die Spezifikationen unterschiedlicher Bereiche bearbeiten.[25] 802.11a/b/g/h sind WLAN-Standards, die in diesem Kapitel noch genauer betrachtet werden.
802.11d Portierung von 802.11b in andere Frequenzbänder.

802.11e Verbesserung der QoS Unterstützung der 802.11 WLAN-Standards
802.11f Verbesserung der Handovermechanismen der 802.11 WLAN-Standards
802.11i Verbesserung der zur Zeit noch unzureichenden Sicherheitseigenschaften. Mehr dazu im Kapitel 7: Security in Wireless Networks.

802.11 WLAN-Standards können ein Ad-hoc Netz bilden, das hier auch als **Independent Basic Service Set (IBSS)** bezeichnet wird. Wird ein Access Point als Repeater zur Verdopplung der Reichweite eingesetzt, wird dies **Basic Service Set (BSS)** genannt. Üblicherweise wird jedoch ein **Extended Service Set (ESS)** gebildet, wobei mehrere Access Points eine große Fläche abdecken, und diese Zugang zu einem bestehenden Netzwerk bieten.
Bei allen 802.11 WLAN-Standards erfolgt der Kanalzugriff über CSMA/CA.
Optional gibt es zur Lösung des sogenannten „hidden node“ Problems das **RTS/CTS (Ready To Send/Clear To Send)** Protokoll. Zu dem „hidden node“ Problem kommt es, wenn zwei Stationen am selben Access Point angemeldet sind, aber zu weit voneinander entfernt sind, um sich gegenseitig zu hören. Hier funktioniert CSMA/CA nicht mehr. RTS/CTS arbeitet wie folgt: Eine Station will Daten senden. Zuerst überprüft sie, dass der Kanal frei ist. Danach schickt sie ein RTS-Signal an den Access Point. Der Access-Point antwortet mit einem CTS-Signal, falls der Kanal frei sein sollte. Darin enthalten sind Informationen über die sendeberechtigte Station und wie lange der Kanal benutzt werden kann. Das CTS-Signal kann von allen Stationen dieses Access Points gehört werden. Somit ist eine Kollision nur noch beim Senden des RTS-Signals möglich. Da aber die sendewilligen Stationen ihre Anfrage wiederholen, wenn nicht nach spätestens einer gewissen Zeit ihr passendes CTS-Signal eingetroffen ist, bereiten die seltenen RTS-Signalkollisionen kaum schwerwiegende Probleme.

Quality of Service (QoS) Unterstützung soll durch die **Point Coordination Function (PCF)** gewährleistet sein. Dabei fragt der Access Point hintereinander alle angemeldeten Stationen ab, ob sie Daten versenden wollen. Der Access Point teilt dann jeder Station mit, wann und wie lange sie den Kanal benutzen kann.
PCF funktioniert nicht im IBSS.
Trotz PCF ist aber noch keine echte QoS Unterstützung gegeben, da PCF versucht die Anforderungen zu erfüllen, aber keine Garantie geben kann. Aber genau das ist ein wichtiger Bestandteil für QoS.

IEEE 802.11 Der erste veröffentlichte Standard hatte einen Datendurchsatz von 2MBit/s, eine Reichweite von maximal 300 m und 100 mW Sendeleistung.
Der Standard nutzt unter anderem das 2,4 GHz ISM-Band. Die Nutzung ist zwar kostenlos, da aber schon sehr viele Geräte in diesem Band betrieben werden, z.B. Bluetooth oder auch die Wellen von Mikrowellenherden, können Interferenzen gehäuft auftreten. Dies hätte höhere Fehlerraten zur Folge. Um eine fehlerfreie Übertragung zu gewährleisten, wird u.a. die Trägermodulation gewechselt, was eine Reduzierung der Bandbreite zur Folge hat.

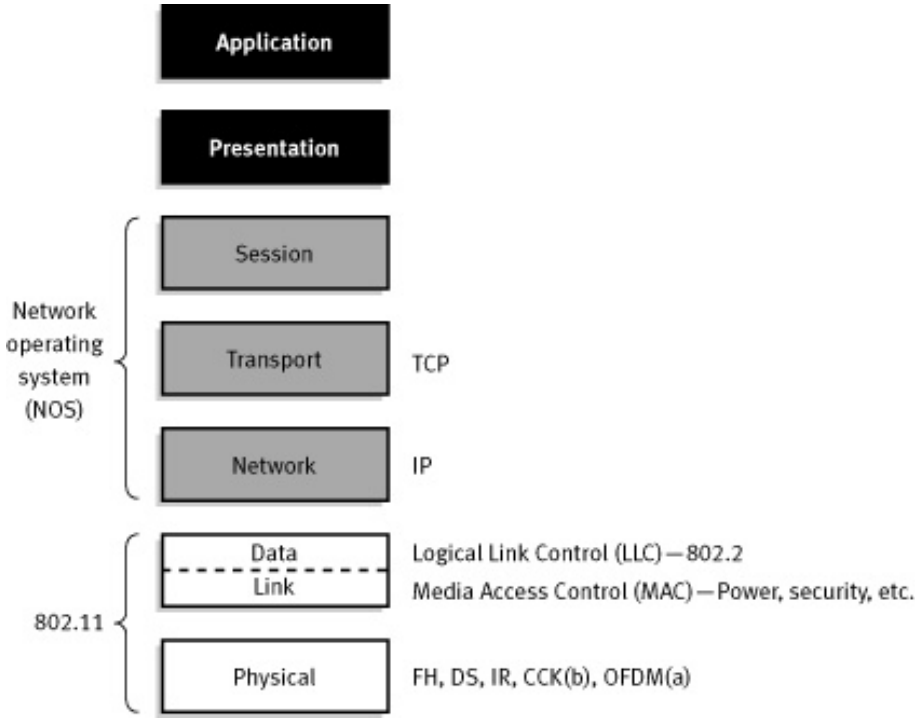

Abbildung 1.17: IEEE 802.11-Protokollschichten [26]

Physical Layer:

- Infrarot oder
- FHSS (Frequence Hopping Service Set) im 2,4 GHz ISM-Band
  - Modulation nach 2-level-GFSK: 1 MBit/s
  - Modulation nach 4-level-GFSK: 2 MBit/s
- DSSS (Direct Sequence Service Set) im 2,4 GHz ISM-Band:
  - Modulation nach BPSK: 1 MBit/s
  - Modulation nach QPSK: 2 MBit/s
  - nutzt CDMA zur Minderung von Störanfälligkeit

Data Link Layer: Media Access Control (MAC):

- Kanalzugriff: CSMA/CA, optional RTS/CTS
- QoS Unterstützung durch PCF
- Verschlüsselung

Zu Szenario 1: Gut geeignet. Voraussetzung: Drucker muss über Drucker-Server angesprochen werden.
Zu Szenario 2: Gut geeignet. Voraussetzung: Datendurchsatz genügt den Ansprüchen.
Zu Szenario 3: Gut geeignet. Voraussetzung: Datendurchsatz genügt den Ansprüchen.
Zu Szenario 4: Gut geeignet. Ausnahme: Datendurchsatz reicht dem einzelnen Nutzer nicht aus.
Zu Szenario 5: Nicht geeignet. Datendurchsatz zu gering. Sicherheitsansprüche werden nicht erfüllt

IEEE 802.11b IEEE 802.11b ist die Weiterentwicklung von 802.11 mit einer höheren Bandbreite von maximal 11 MBit/s. 802.11b ist abwärtskompatibel zu 802.11 DSSS. Infrarot und FHSS werden nicht mehr unterstützt.
Zur Zeit ist dieser Standard die weit verbreitetste WLAN-Technik überhaupt. Dies ist damit zu begründen, dass 802.11b als einzige WLAN-Technik mit mehr als 2 MBit/s Datendurchsatz zur Zeit auf dem Markt verfügbar ist.

Physical Layer:

- Modulation nach BPSK: 1 MBit/s
- Modulation nach QPSK: 2 , 5.5 , 11 MBit/s
- nutzt CDMA zur Minderung von Störanfälligkeit

Data Link Layer: Media Access Control (MAC) wie 802.11

Zu Szenario 1: Gut geeignet. Voraussetzung: Drucker muss über Drucker-Server angesprochen werden.
Zu Szenario 2: Gut geeignet. Voraussetzung: Datendurchsatz genügt den Ansprüchen.
Zu Szenario 3: Gut geeignet. Voraussetzung: Datendurchsatz genügt den Ansprüchen.
Zu Szenario 4: Gut geeignet. Voraussetzung: Datendurchsatz genügt den Ansprüchen
Zu Szenario 5: Eingeschränkt geeignet. Einschränkung: Abhörsicherheit entspricht nicht den Forderungen. Spezielle Maßnahmen durch die Firma kann Abhilfe schaffen.

IEEE 802.11g Um den stetig wachsenden Anspruch an höherer Bandbreite gerecht zu werden, wird an diesem Standard gearbeitet, der einen Datendurchsatz von maximal 54 MBit/s hat und zu den Geräten des 802.11b Standards kompatibel ist. Damit müssen Firmen, die mit 802.11b ausgestattet sind und mit der Zeit den Datendurchsatz erhöhen wollen, nicht ihre WLAN-Infrastruktur auf einmal komplett erneuern.
802.11g nutzt OFDM Technik im 2,4 GHz ISM-Band. Ob die gewonnene Bandbreite nicht durch Reduzierung des Datendurchsatzes wegen gehäuft auftretenden Interferenzen wieder verloren geht, wird die Praxis zeigen. Bislang sind noch keine Geräte auf dem Markt.

Physical Layer:

- OFDM im 2,4 GHz ISM-Band: 6 , 9 , 12 , 18 , 24 , 36 , 48 , 54 MBit/s
- DSSS (Direct Sequence Service Set) im 2,4 GHz ISM-Band:
  - Modulation nach BPSK: 1 MBit/s
  - Modulation nach QPSK: 2 , 5.5 , 11 , 22 , 33 MBit/s
  - nutzt CDMA zur Minderung von Störanfälligkeit

Data Link Layer: Media Access Control (MAC) wie 802.11
[7]

Zu Szenario 1: Gut geeignet. Voraussetzung: Drucker muss über Drucker-Server angesprochen werden.
Zu Szenario 2: Sehr gut geeignet.
Zu Szenario 3: Sehr gut geeignet.
Zu Szenario 4: Sehr gut geeignet.
Zu Szenario 5: Eingeschränkt geeignet. Einschränkung: Abhörsicherheit entspricht nicht den Forderungen. Spezielle Maßnahmen durch die Firma kann Abhilfe schaffen.

IEEE 802.11a Der 802.11a Standard unterscheidet sich von 802.11g insoweit, dass es anstatt des 2,4 GHz ISM-Bands das weniger genutzte 5 GHZ Band verwendet. Dadurch ergeben sich weniger Probleme durch Interferenzen von anderen Geräte. Allerdings ist die Reichweite aufgrund der höheren Frequenz niedriger. Ein weiteres Problem ist, dass die meisten Kanäle, die 802.11a belegt, in Europa für u.a. Rettungsleitfunk oder für die Satellitensteuerung reserviert sind. Damit ist der Betrieb in Europa verboten. Zur Zeit wird geprüft, ob eine Zulassung der unteren 8 der insgesamt 12 Kanäle möglich ist. Dies würde bedeuten, dass 802.11a nur für Indoor-Anwendungen zugelassen wird.[19] Für Europa soll der Substandard 802.11h mit weiteren Funktionen u.a. dieses Problem lösen. In den USA ist der Betrieb von 802.11a erlaubt und erste Produkte sind verfügbar.

Physical Layer: OFDM im 5 GHz Band:

- Modulation nach BPSK: 6 , 9 MBit/s
- Modulation nach QPSK: 12 , 18 MBit/s
- Modulation nach QAM16: 24 , 36 MBit/s
- Modulation nach QAM64: 48 , 54 MBit/s

Data Link Layer: Media Access Control (MAC) wie 802.11

Zu Szenario 1: Gut geeignet. Voraussetzung: Drucker muss über Drucker-Server angesprochen werden.
Zu Szenario 2: Nicht geeignet. Outdoor-Betrieb von 802.11a in Deutschland verboten.
Zu Szenario 3: Sehr gut geeignet.
Zu Szenario 4: Eingeschränkt geeignet. Einschränkung: Durch geringere Reichweite werden evtl. mehrere APs zur vollen Abdeckung benötigt. Keine Nutzung außerhalb von Gebäuden erlaubt.
Zu Szenario 5: Eingeschränkt geeignet. Einschränkung: Abhörsicherheit entspricht nicht den Forderungen. Spezielle Maßnahmen durch die Firma kann Abhilfe schaffen.

IEEE 802.11h Der Substandard 802.11h ist die Erweiterung von 802.11a um zwei weitere Funktionen, die dann den Betrieb eines WLAN nach 802.11 im 5 GHz Band in Europa weitestgehend uneingeschränkt zulässt.
Die zwei Funktionen sind zum einen die dynamische Frequenzwahl (**Dynamic Frequency Selection - DFS**) und die automatische Anpassung der Übertragungsleistung (**Transmission Power Control - TPC**). DFS bietet die Möglichkeit schon durch andere Technik belegte Kanäle nicht zu benutzen. TPC passt die Übertragungsleistung den Bedingungen an. Das heißt z.B. bei geringer Entfernung zum AP eine automatische Reduzierung der Sendeleistung.

Physical Layer: OFDM im 5 GHz Band:

- Modulation nach BPSK: 6 , 9 MBit/s
- Modulation nach QPSK: 12 , 18 MBit/s
- Modulation nach QAM16: 24 , 36 MBit/s
- Modulation nach QAM64: 48 , 54 MBit/s

Data Link Layer: Media Access Control (MAC):

- wie 802.11 zusätzlich Dynamic Frequency Selection (DFS) und Transmission Power Control (TPC))

Zu Szenario 1: Gut geeignet. Voraussetzung: Drucker muss über Drucker-Server angesprochen werden.
Zu Szenario 2: Sehr gut geeignet.
Zu Szenario 3: Sehr gut geeignet.
Zu Szenario 4: Sehr gut geeignet.
Zu Szenario 5: Eingeschränkt geeignet. Einschränkung: Abhörsicherheit entspricht nicht den Forderungen. Spezielle Maßnahmen durch die Firma kann Abhilfe schaffen.

HiperLAN2 (High Performance Radio LAN)

HiperLAN2 ist eine WLAN-Technik, die hauptsächlich von europäischen Firmen entwickelt wurde. Sie baut auf das kommerziell erfolglose HiperLAN1 auf. HiperLAN2 nutzt, wie IEEE 802.11a/h das 5 Ghz Band und verfügt wie IEEE 802.11h über Dynamic Frequency Selection (DFS) und Transmission Power Control (TPC) und ist deswegen in Europa, wo das 5 GHz nicht ohne Einschränkungen genutzt werden kann, zugelassen.
Im Physical Layer sind somit kaum Unterschiede zu IEEE 802.11a/h zu erkennen. Dass HiperLAN2 aber doch eine nennenswerte WLAN-Technik ist, liegt somit eine Schicht höher. Während bei den 802.11 WLAN-Standards noch an einer echten QoS Unterstützung gearbeitet wird, soll dies bei HiperLAN2 schon gegeben sein. Auch den hohen Sicherheitsansprüchen, den u.a. Firmen, Behörden und Universitäten an ihre Netzwerktechnik stellen, soll HiperLAN2 gerecht werden. Da es noch keine Geräte für HiperLAN2 auf dem Markt gibt, und es fraglich ist, ob es jemals kommerziell verfügbar sein wird, ist nicht viel über die Qualität der genutzten Verschlüsselung bekannt.[24]

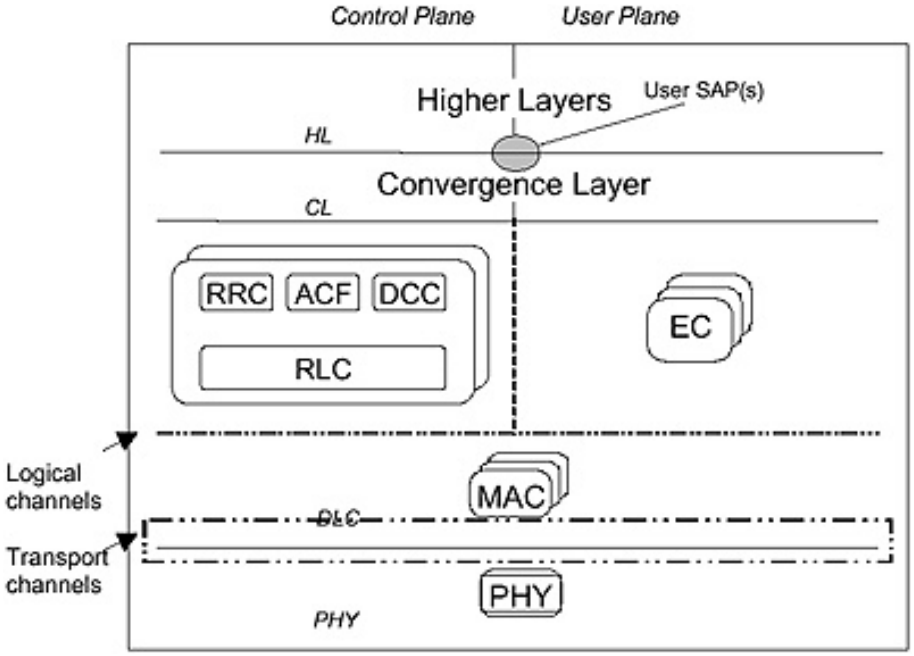

Abbildung 1.18: HiperLAN2-Protokollschichten [12]

Physical Layer (PHY): OFDM im 5 GHz Band:

- Modulation nach BPSK: 6 , 9 MBit/s
- Modulation nach QPSK: 12 , 18 MBit/s
- Modulation nach QAM16: 27 , 36 MBit/s
- Modulation nach QAM64 (optional): 54 MBit/s

Data Link Control Layer (DLC):

- Media Access Control (MAC)
  - nutzt TDMA innerhalb jedes MAC Frames
  - nutzt TDD, dynamische Anpassung der Upload- und Downloadzeiten
- Error Control (EC) Protocol
  - nach Verwendung angepasstes Fehlerprotokoll
- Radio Link Control (RLC)
  - DLC Connection Control (DCC)
  - Radio Resource Control (RRC): u.a. DFS und Handover
  - Association Control Function (ACF): Anmeldung an AP, Absprache nach Verschlüsselung, Modulation

Convergence Layer (CL): Der Convergence Layer verbindet nahezu jede Art von Netzwerkarchitekturen mit HiperLAN2. Damit kann u.a. IP, ATM auf HiperLAN2 betrieben werden. Erreicht wird dies, dass der CL wenn erforderlich, die Pakete der zu verbindenden Netzwerkarchitektur zu den HiperLAN2 Paketen umwandelt oder umgekehrt. HiperLAN2 ist zu UMTS, ATM und Firewire-Netzen kompatibel.[9]

Zu Szenario 1: Gut geeignet. Voraussetzung: Drucker muss über Drucker-Server angesprochen werden.
Zu Szenario 2: Sehr gut geeignet.
Zu Szenario 3: Sehr gut geeignet.
Zu Szenario 4: Sehr gut geeignet.
Zu Szenario 5: Voraussichtlich gut geeignet. Sicherheitsforderungen sollen erfüllt werden. Allerdings gibt es nie einen 100%igen Schutz.

HiperAccess und **HiperLink** sind Weiterentwicklungen von HiperLAN2.
HiperAccess wird im 40 GHz Band betrieben und erreicht bei einer Reichweite von maximal 5 km einen Datendurchsatz von bis zu 25 MBit/s.
HiperLink soll hauptsächlich als Breitband WLAN-Technik Netzwerke verbinden. Bei einer maximalen Reichweite von 150 m wird ein Datendurchsatz von bis zu 155 MBit/s erreicht.

1.3.3 Wireless Wide Area Network (WWAN)

WWANs sind Netze, die die Möglichkeit bieten, mit einer festen, eindeutigen Telefonnummer theoretisch weltweit erreichbar zu sein. Außerdem soll innerhalb eines großen Einsatzradius eine Verbindung u.a. zum Internet oder Firmennetz aufgebaut werden können. Da diese Netze fast ausschließlich kommerziell angeboten werden, ist die Reichweite

nicht durch die technischen Vorgaben beschränkt. Ziel von WWAN-Betreiber ist es, eine flächendeckende Bereitstellung ihrer Netze zu bieten. Im Ausland erhält der Kunde meist mit Roaming die gewünschte Versorgung. An die Wireless-Technik in dieser Kategorie werden mit die höchsten Ansprüche in der ganzen Wireless-Sparte gestellt, da ein relativ hoher Datendurchsatz, eine gute Abhörsicherheit, Handover-Funktionalität und Roaming bei voller Mobilität und bei nahezu jeder Geschwindigkeit geboten werden soll. Diesen Erwartungen wird zur Zeit kein Standard gerecht.
Hohe Anforderungen an die Infrastruktur und die Tatsache, dass die verwendeten Frequenzbänder lizensiert sind, lassen zur Zeit nur eine kostenpflichtige Nutzung zu.
Erwähnenswert ist, besonders in Hinblick auf die Mobilität des Nutzers, dass die Technologien in dieser Kategorie die höchste Sendeleistung haben (GSM basierte Technologien und UMTS mit ca. 2W), und somit der Stromverbrauch deutlich höher als bei WPAN- und WLAN-Standards ist (1mW - 200mW), was wiederum auf die Akkulaufzeit einen enormen Einfluß hat.

WWANs werden in Generationen eingeteilt:

- 1G für die analoge Mobiltelefontechnik
- 2G für die erste digitale Mobiltelefontechnik, wie z.B. GSM
- 2,5G für die Mobildatentechnik, die einen höheren Datendurchsatz als 2G bietet, aber noch deutlich unter dem Maximaldurchsatz von 3G liegt
- 3G ist die kommende Technik mit deutlich höherem Maximaldatendurchsatz, z.B. UMTS
- 4G für den geplanten Nachfolger von UMTS: Flash-OFDM

Global System of Mobile Communication (GSM)

GSM, das eigentlich Groupe Speciale Mobile hieß, und später umbenannt wurde, ist eine 2G-Technologie, die zur Zeit der weltweit weitest verbreitete WWAN-Standard ist. Bis auf Japan und manche Teile Amerikas, wird es in 160 Ländern der Erde zum Teil flächendeckend eingesetzt.
Mit 9,6 kBit/s bietet GSM einen für Datendienste sehr geringen Durchsatz. Seinen großen Erfolg verdankt es allerdings den in erster Linie unterstützten Telefondiensten, die durch Roaming für jeden GSM-Nutzer nahezu weltweit verfügbar sind. Die Datendienste erreichten eine nicht annähernd so große Akzeptanz wie die Telefondienste, da neben der geringen Bandbreite GSM auch verbindungsorientiert arbeitet, also nicht nach Datenvolumen sondern nach Onlinezeit abgerechnet wird. Dies hat dann bei dem geringen Durchsatz relativ hohe Benutzungskosten zur Folge.
Erwähnenswert ist, dass GSM nicht 100%ig abhörsicher ist.[37]

Zu Szenario 1: Eingeschränkt geeignet. Einschränkung: Drucker muss über Drucker-Server angesprochen werden, der zumindest an einem Festnetz angeschlossen ist und angewählt werden kann. Hieraus ergeben sich im Vergleich zu WLAN und WPAN Technologien ganz

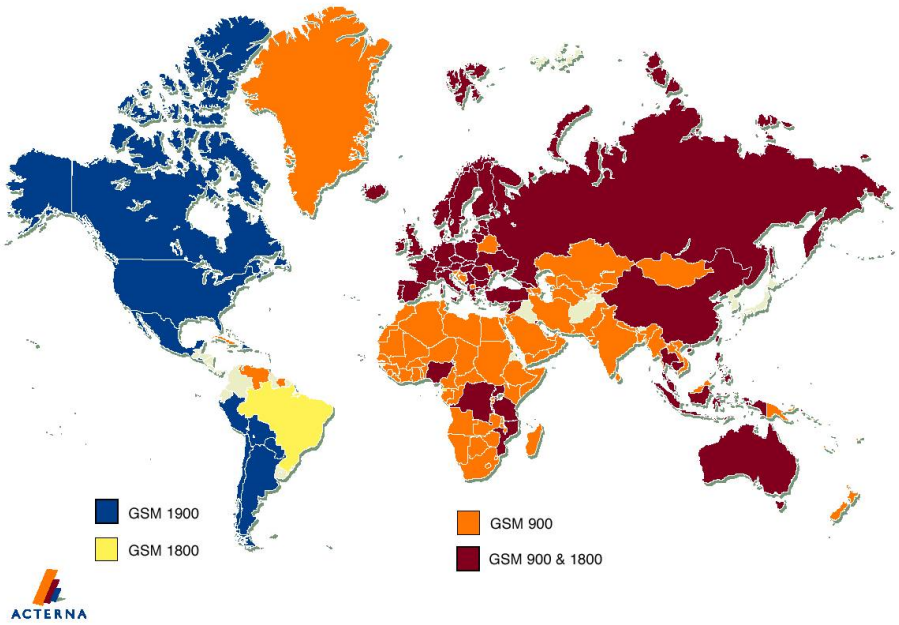

Abbildung 1.19: GSM-Abdeckung weltweit [35]

neue Einsatzmöglichkeiten. Es kann von jedem beliebigen Ort mit Netzzugang z.B. in der Firma ein Dokument gedruckt werden.
Zu Szenario 2: Nicht geeignet. Die meisten U-Bahnen bieten keinen Mobilfunkempfang.
Zu Szenario 3: Eingeschränkt geeignet. Einschränkung: Sehr geringer Datendurchsatz. Direkt können nur jeweils zwei Benutzer Daten austauschen. Wird das Internet genutzt ist dies mit einer theoretisch unbegrenzten Anzahl von Teilnehmern möglich.
Zu Szenario 4: Schlecht geeignet. Datendurchsatz zu gering. Verbindungskosten sehr hoch.
Zu Szenario 5: Nicht geeignet. Datendurchsatz zu gering. Nicht abhörsicher.

High Speed Circuit Switch Data (HSCSD)

HSCSD gehört zu 2,5G. Die maximal erreichbaren 43,2 kBit/s Datendurchsatz werden durch Bündelung von bis zu 4 GSM-Kanälen erreicht. Das ist auch der einzige Unterschied zu GSM. Somit arbeitet auch HSCSD verbindungsorientiert, was in diesem Fall bedeutet, dass für jeden benutzten Kanal bezahlt werden muss. Die Kanalbündelung erfolgt statisch, das heißt, dass bei Verbindungsaufbau eine gewisse Anzahl an Kanälen reserviert wird, die auch bis zum Verbindungsabbruch gehalten werden, unabhängig davon, ob sie genutzt werden oder nicht. HSCSD eignet sich besonders für Datenübertragung, die so schnell wie möglich beendet sein soll, oder für Anwendungen, bei denen ständig die gleiche Bandbreite benötigt wird.

Zu Szenario 1: Eingeschränkt geeignet: Wie GSM
Zu Szenario 2: Nicht geeignet: Wie GSM
Zu Szenario 3: Eingeschränkt geeignet: Wie GSM
Zu Szenario 4: Schlecht geeignet: Wie GSM

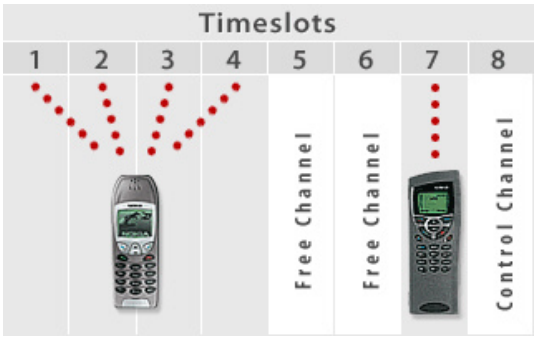

Abbildung 1.20: HSCSD Kanalbündelung [21]

Zu Szenario 5: Eingeschränkt geeignet. Einschränkung: Datendurchsatz sehr gering. Für Außenmitarbeiter für seltene Benutzung eventuell ausreichend. Abhörsicherheit entspricht nicht den Forderungen. Spezielle Codierung durch die Firma kann Abhilfe schaffen.

General Packet Radio Service (GPRS)

GPRS setzt auf GSM auf, gehört mit einer maximalen Datenrate von 170 kBit/s zu 2,5G und unterstützt ausschließlich Datendienste. Der entscheidende Unterschied zu GSM-Datendiensten ist, dass GPRS paketbasiert ist. Das bedeutet für den Anwender, dass er im allgemeinen nicht für die Verbindungszeit, sondern für das Datenvolumen, das er nutzt, zahlt. Somit kann man mit dieser Technologie theoretisch immer online sein. Die 170

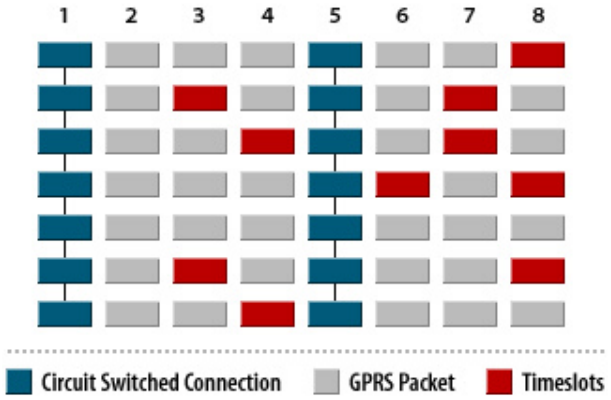

Abbildung 1.21: Diagramm für paketbasierte Übertragung

kBit/s Datendurchsatz werden durch Kanalbündelung erreicht, für die man aber direkt nicht zu zahlen hat. Dass 170 kBit/s ein theoretische Wert ist, liegt daran, dass sich mehrere Benutzer pro Sendezelle die Kanäle teilen. Ist man der einzige Anwender, steht einem die ganze Bandbreite zur Verfügung. Sobald noch andere GPRS-Benutzer in der Sendezelle sind, wird die Bandbreite dynamisch aufgeteilt.

Zu Szenario 1: Eingeschränkt geeignet: Wie GSM
Zu Szenario 2: Nicht geeignet: Wie GSM

Zu Szenario 3: Eingeschränkt geeignet: Wie GSM
Zu Szenario 4: Eingeschränkt geeignet. Einschränkung: Datendurchsatz für viele Anwendungen zu gering. Verbindungskosten abhängig vom Datenvolumen.
Zu Szenario 5: Eingeschränkt geeignet. Einschränkung: Datendurchsatz sehr gering. Für Außenmitarbeiter für seltene Benutzung eventuell ausreichend. Abhörsicherheit entspricht nicht den Forderungen. Spezielle Codierung durch die Firma kann Abhilfe schaffen.

Enhanced Data Rates for GSM Evolution (EDGE)

Die 2,5G Technologie EDGE ist ein modifiziertes GPRS. Der auffälligste Unterschied ist die verwendete Trägermodulation 8-PSK. Damit lassen sich pro Phase 3 Bit übertragen und es ist damit ein theoretischer Maximaldatendurchsatz von 348 kBit/s möglich.
Wie bei GPRS wird nach Datenvolumen abgerechnet. EDGE sollte eine Alternative für alle WWAN-Betreiber sein, die keine UMTS Lizenz erhielten. In Deutschland wird voraussichtlich kein Netzbetreiber EDGE anbieten, da man sich hier auf die Einführung von 3G konzentriert. Allerdings wird in den kommenden Monaten in den USA, wo UMTS wegen belegter Frequenzen nicht zugelassen ist, das erste EDGE-fähige Netz bereitgestellt.[38]

Zu Szenario 1: Eingeschränkt geeignet: Wie GSM
Zu Szenario 2: Nicht geeignet: Wie GSM
Zu Szenario 3: Eingeschränkt geeignet: Wie GSM
Zu Szenario 4: Eingeschränkt geeignet. Einschränkung: Datendurchsatz für manche Anwendungen zu gering. Verbindungskosten abhängig vom Datenvolumen.
Zu Szenario 5: Eingeschränkt geeignet. Einschränkung: Datendurchsatz sehr gering. Für Außenmitarbeiter für seltene Benutzung eventuell ausreichend. Abhörsicherheit entspricht nicht den Forderungen. Spezielle Codierung durch die Firma kann Abhilfe schaffen.

Universal Mobile Telecommunications System (UMTS)

Die 3. Generation soll der Masse der Mobilkunden das mobile Internet durch hohe Datendurchsätze bis maximal 2 MBit/s, neue Anwendungen und akzeptable Preise schmackhaft gemacht werden. Angekündigt war die Einführung dieser Technik auf dem deutschen Markt bis Ende 2002, doch bislang ist nicht absehbar, wann dies in die Tat umgesetzt werden kann. UMTS ist wie GPRS und EDGE paketbasiert. Die maximale Bandbreite ist von der Anzahl der Benutzer pro Sendezelle, der Geschwindigkeit des Anwenders und dem Ort wo er sich aufhält abhängig.[29]

Umgebung	Maximaldaten-durchsatz
städt. Vororte	144kBit/s
Stadt, Fußgänger	144kBit/s
Stadt, Gebäude	384 kBit/s
Innenstädte	2 MBit/s

Ob im durchschnittlichen Alltag die Bandbreite ausreicht, um wie angepriesen sich Videofilme aus dem Internet runterladen zu können bleibt fraglich. Genauso fraglich bleibt allerdings auch, wie sinnvoll und realitätsbezogen solche Vorstellungen zur Zeit sind.

Zu Szenario 1: Eingeschränkt geeignet: Wie GSM
Zu Szenario 2: Nicht geeignet: Wie GSM
Zu Szenario 3: Geeignet: Datendurchsatz hierfür meist ausreichend. Sonst wie GSM.
Zu Szenario 4: Eingeschränkt geeignet. Einschränkung: Datendurchsatz für manche Anwendungen zu gering. Verbindungskosten abhängig vom Datenvolumen.
Zu Szenario 5: Eingeschränkt geeignet. Einschränkung: Datendurchsatz gering. Für Außenmitarbeiter für seltene Benutzung eventuell ausreichend. Keine Infromationen über Abhörsicherheit verfügbar.

Flash Orthogonal Frequency Division Multiplexing (Flash-OFDM)

In die 4. Generation der Mobilkommunikationstechnik werden sehr große Erwartungen gesetzt. Nicht nur dass ein Maximaldatendurchsatz von 10 MBit/s erreicht und QoS Unterstützung geboten werden soll, es soll auch der erste WWAN-Standard sein, der weltweit unterstützt wird, wenn die Pläne des OFDM-Forums in die Realität umgesetzt werden.[23] Flash-OFDM baut auf UMTS auf, wobei OFDM-Technik und andere Frequenzbänder verwendet werden. Durch OFDM, mit eventuell einer höheren Trägermodulation, wie z.B: QAM16, sind die geforderten Datendurchsätze realisierbar. Damit Flash-OFDM auch in den USA zugelassen wird, sind andere Frequenzbänder nötig als bei UMTS.
Es ist noch lange nicht absehbar, wann der Nachfolger von UMTS, das selbst noch große Anlaufschwierigkeiten hat, auf dem Markt erscheinen wird.

Eine Bewertung mit Hilfe der Szenarien ist zur Zeit nicht möglich, da viel zu wenig über den künftigen Standard bekannt ist.

1.4 Zusammenfassung und Ausblick

Während Bluetooth bei WPAN die Nase vorn hat und in ein paar Jahren eventuell ebenso allgegenwärtig wie IrDA heutzutage ist, gibt es bei WLAN keinen Standard, der bei allen wichtigen Faktoren die nötige Unterstützung bietet. Während HiperLAN2 in Sachen Sicherheit, Einsatzvarianz und Robustheit gegen die meisten IEEE 802.11-Standards einen Vorsprung hat, ist diese Technologie wegen mangelnder Produkte keine echte Alternative. Voraussichtlich wird HiperLAN2 nie kommerziell vermarktet werden.[19] Bei allen IEEE 802.11-Standards ist die QoS-Unterstützung nicht gut genug und die Abhörsicherheit praktisch kaum vorhanden. 802.11 und DECT sind veraltet. 802.11b hat einen für viele Firmen zu geringen Datendurchsatz. 802.11g muss sich erst etablieren und zeigen, dass die hohen Datendurchsatzraten selbst im überfüllten 2,4GHz ISM-Band zu erreichen sind. 802.11a ist in Europa noch nicht erlaubt und wird, wenn überhaupt, nur für den Indoor-Bereich zugelassen. Wichtige Anwendungsfelder fallen hierfür dann von vornherein weg. 802.11h hat, mit den genannten Einschränkungen für alle 802.11-Standards, gute

Aussichten erfolgreich zu sein, auch wenn es noch keine Produkte zu erwerben gibt. Diese sollen Mitte 2003 eingeführt werden.
In der WWAN-Sparte ist momentan GPRS bei den Anwendern führend, da die Mehrzahl der Kunden eine Abrechnung nach Datenvolumen bevorzugen und eher Einbußen bei dem Datendurchsatz hinnehmen. Wird eine mobile Verbindung mit einer zugesicherten Bandbreite gefordert, ist allerdings HSCSD besser geeignet. EDGE wird wohl gar nicht erst in Europa eingeführt, und UMTS lässt noch auf sich warten. Wenn es dann soweit sein wird, sind am Anfang die Benutzungsgebühren und Preise wohl weniger etwas für den Durchschnittsbürger. Doch das gilt generell bei allen Geräten dieser recht jungen Sparte. Die relativ hohen Kosten binden noch die meisten Computerbenutzer an ihre Kabel. Allerdings sollten sich die Anschaffungskosten mit der Zeit selbst für viele Heimanwender in erschwingliche Bereiche bewegen. Zu groß ist der Nutzen der neuen Technologien. Dementsprechend wird die Nachfrage steigen und neue Firmen die Konkurrenz beleben. Neue Produkte und Weiterentwicklungen werden entstehen, die am Schluss so sehr in unserem Alltag eingebunden sind, dass es dann nur schwer vorstellbar ist, wie man jemals ohne diese hat auskommen können. Dann ist es möglich, dass Studenten für die Literaturrecherche ihrer Diplomarbeit übers Internet via WLAN im Garten sitzen, von dort aus mit WPAN-Geräten die Stereoanlage im Zimmer steuern oder den Videorekorder programmieren, damit sie anstatt eine wichtige Dokumentation zu sehen eher grillen gehen können. Sollte der Arbeitsplatz zum Beispiel zu einem Baggersee verlegt werden, kann die Literaturrecherche, dank WWAN, trotzdem fortgeführt werden und die Programmierung des Videorekorders um ein paar Filme erweitert werden.
Was die vorgestellten Standards für technische Eigenschaften haben, wird auf den folgenden zwei Tabellen angegeben. Eine dritte Tablle zeigt die zur Zeit marktüblichen Nutzungsgebühren der verfügbaren WWAN-Standards. Die UMTS-Tarife sind nicht für Deutschland gültig, da es bis Dezember 2002 keine Preisankündigungen gab. Die angegebenen Preise sind vom UMTS-Netz von Isle-of-Man, und sollen als grobe Orientierung im Vergleich zu den anderen Netzen gelten.

1.4.1 Überblickstabelle 1

[1]

Technologie	Daten-übertragungs-rate	Reichweite	Sende leistung	Frequenz-band	Kanal-zugriff
IrDA	4 MBit/s	..1 m	40 mW	Licht	NA
Fast IrDA	16 MBit/s	..1 m	40 mW..	Licht	NA
Bluetooth/A	1 MBit/s	..15 m	1 mW	2,4 GHz	TDD/FH
Bluetooth/B	1 MBit/s	..150 m	100 mW	2,4 GHz	TDD/FH
DECT	2 MBit/s	..300 m	250 mW	1880-1900 MHz	TDMA/ FDMA
802.11	2 MBit/s	..300 m	100 mW	2,4 GHz	CSMA/CA
802.11b	11 MBit/s	..100 m	100 mW	2,4 GHz	CSMA/CA
802.11g	54 MBit/s	..100 m	200 mW	2,4 GHz	OFDM mit CSMA/CA
802.11a	54 MBit/s	..100 m	800 mW	5 GHz	OFDM mit CSMA/CA
802.11h	54 MBit/s	..100 m	200 mW	5 GHz	OFDM mit CSMA/CA
HiperLAN2	54 MBit/s	..100 m	200 mW	5 GHz	OFDM mit TDMA/TDD
HiperAccess	25 MBit/s	..5 km	NA	40 GHz	OFDM
HiperLink	155 MBit/s	..150 m	NA	17 GHz	OFDM
GSM	9,6 kBit/s	..35 km	2 W	900, 1800, 1900 MHz	FDMA/TDMA
HSCSD	43,2 kBit/s	..35 km	2 W	900, 1800, 1900 MHz	FDMA/TDMA
GPRS	170 kBit/s	..35 km	2 W	900, 1800, 1900 MHz	FDMA/TDMA
EDGE	384 kBit/s	..35 km	2 W	900, 1800, 1900 MHz	FDMA/TDMA
UMTS	2 MBit/s	..20 km	2 W	2 GHZ	W-CDMA
Flash-OFDM	10 MBit/s	NA	NA	NA	OFDM

[1] Quellen:[2], [3], [4], [10], [11], [12], [27]

1.4.2 Überblickstabelle 2

[2]

Technologie	Träger-modulation	ad-hoc fähig	kompatibel zu	Verfügbar (Dez 2002)	Unterstützung in Europa	USA	Japan
IrDA	NA	ja (P2P)	-	ja	ja	ja	ja
Fast IrDA	NA	ja (P2P)	IrDA	ja	ja	ja	ja
Bluetooth/A	GFSK	ja	-	ja	ja	ja	ja
Bluetooth/B	GFSK	ja	Bluetooth/A	nein	ja	ja	ja
DECT	GFSK	ja	-	ja	ja	ja	ja
802.11	GFSK/ B/Q-PSK	ja	-	ja	ja	ja	ja
802.11b	B/Q-PSK	ja	802.11	ja	ja	ja	ja
802.11g	B/Q-PSK QAM16/64	ja	802.11b	nein	ja	ja	ja
802.11a	B/Q-PSK QAM16/64	ja	-	ja	nein	ja	NA
802.11h	B/Q-PSK QAM16/64	ja	802.11a	nein	ja	ja	ja
HiperLAN2	B/Q-PSK QAM16/64	ja	-	nein	ja	ja	ja
HiperAccess	NA	NA	-	nein	ja	ja	ja
HiperLink	NA	ja (P2P)	-	nein	ja	ja	ja
GSM	GFSK	nein	-	ja	ja	tlw[3]	nein
HSCSD	GFSK	nein	GSM	ja	ja	tlw[3]	nein
GPRS	GFSK	nein	GSM	ja	ja	tlw[3]	nein
EDGE	8-PSK	nein	GSM	nein	nein	ja	nein
UMTS	8-PSK	nein	-	Eu: nein Jp: ja	ja	nein	ja
Flash-OFDM	NA	nein	-	nein	ge-plant	ge-plant	ge-plant

[2]Quellen:[2], [3], [4],[10], [11], [12], [27]
[3]teilweise: Regional abhängig [35]

1.4.3 Übersicht WWAN

Technologie	Datendurchsatz • maximal • minimal	Durchsatz v. Anzahl Nutzer pro Zelle unabhängig?	Kosten: • (Euro pro MB) • (Euro pro h bzw. 1MB pro h in Euro)	Abrechnung nach
GSM	• 9,6 kBit/s • 9,6 kBit/s	ja	• 0,98 - 6,86[4] • 4,20 - 29,40[3]	Zeit
HSCSD	• 43,2 kBit/s • 9,6 kBit/s	nein (statisch)	• 0,70 - 4,90[3] • 4,20 - 58,80[3]	Zeit
GPRS	• 170 kBit/s • 9,6 kBit/s	nein (dynamisch)	• 2,00 - 10,00[3] • 2,00 - 10,00[3]	Daten-volumen
EDGE	• 384 kBit/s • 9,6 kBit/s	nein (dynamisch)	• NA • NA	Daten-volumen
UMTS	• 2 MBit/s • 9,6 kBit/s	nein (dynamisch)	• 1,24 - 7,80[5] • 1,24 - 7,80[4]	Daten-volumen
Flash-OFDM	• 10 MBit/s (geplant) • NA	NA	• NA • NA	vorr. Daten-volumen

[3]T-D1- und Vodafone-Tarife ohne Grundgebühr[31][34]
[4]O2-Tarif inkl. Grundgebühr, Isle-of-Man[33]

Literaturverzeichnis

[1] RWTH Aaachen; Speth, Michael; “OFDM Receivers for Broadband-Transmissions“; www.ert.rwth-aachen.de/Projekte/Theo/OFDM/www_ofdm.html

[2] Bundesamt für Kommunikation; “Faktenblatt Radio Local Area Networks (RLAN)“; www.bakom.ch/imperia/md/md/content/deutsch/telecomdienste/factsheets/10.pdf

[3] Bluetooth Designer; “A Comparison of Bluetooth and IEEE 802.11“; www.btdesigner.com/pdfs/KenNoblittComparison.pdf

[4] DECT Forum; “The standard explained“; www.dect.ch/publicdocs/TechnicalDocument.pdf

[5] Ericsson; “Hintergrundinformation: Immer optimal verbunden“; www.ericsson.de/downloads/pressenews/Hintergrundpapier.pdf

[6] Ericsson; “Infrastruktur Mobilfunknetze“; www.ericcson.de/broschueren/infrastrukur_mobilfunknetze.pdf

[7] elektroniknet; “Der neue Standard IEEE 802.11g“; www.elektroniknet.de/topics/kommunikation/fachthemen/print/02020.htm

[8] elektroniknet; “Digitaler Aufschwung für Kurz-, Mittel- und Langwelle“; www.elektroniknet.de/topics/kommunikation/fachthemen/print/02021.htm

[9] ETSI; “Hiperlan2 - Technical Overview“; www.etsi.org/technicalactiv/Hiperlan/hiperlan2tech.htm

[10] Fernuniversität Hagen; Speth, Michael; “DECT-Standard Digital Enhanced Cordless Telekommunication“; www.fernuni-hagen.de/NT/kurse/sem_1999/dect.pdf

[11] GSM World; www.gsmworld.de

[12] HiperLan2 Global Forum; “HiperLAN/2 - The Broadband Radio Transmission Technology Operating in the 5GHz Frequency Band“; www.hiperlan2.com/presdocs/site/whitepaper.pdf

[13] IEEE Network, Volume 15, Issue 5, Sept.-Oct. 2001, Pages: 28-37; Johansson, P.; Kazantzidis, M.; Kapoor, R.; Gerla, M.; “Bluetooth: an enabler for personal area networking“

[14] IEEE Personal Communication, Volume 8, Issue 6, Dec. 2001, Pages: 58-64; Katsianis, D.; Welling, I.; Ylonen, M.; Varoutas, D.; Sphicopoulos, T.; Elnegaard, N.K.; Olsen, B.T.; Budry, L.; "The financial perspective of the mobile networks in Europe"

[15] IEEE Personal Communication, Volume 8, Issue 5, Oct. 2001, Pages: 10-17; Frodigh, M.; Parkvall, S.; Roobol, C.; Johansson, P.; Larsson, P.; "Future-generation wireless networks"

[16] Intel; "Bluetooth* Architecture Overview"; www.intel.com/technology/itj/q22000/pdf/art_1.pdf

[17] IrDA Specifications; www.irda.org/standards/specifications.asp

[18] ISP Planet; www.isp-planet.com/technology/3g_mobile_wireless.html

[19] LANline 10/2002; "802.11a WLAN in Deutschland"; www.lanline.de/aktuelle-ausgabe/10_2002/lan_1002_008.html

[20] Mulalic, Amir; TU Berlin; Nov. 2001; Diplomarbeit "Erweiterung des Bluetooth Protokolls in einem Embedded Bluetooth System"

[21] MTN; "MTNdataFAST"; www.mtn.co.za/services/df/detail.asp

[22] Universität Oldenburg; "IEEE 802.11 Physical Layers"; http://einstein.offis.uni-oldenburg.de/rechnernetze/physical_layer.htm

[23] O'Reilly Network; "Wireless Carrier Technology Roadmap"; www.oreillynet.com/pub/a/wireless/2000/07/28/magazine/roadmap.html

[24] PCTechGuide; "Communications/Mobile Comms"; www.pctechguide.com/30mobcomms.htm

[25] PersonalTelco; "DifferentStandards"; www.personaltelco.net/index.cgi/DifferentStandards

[26] Pulse; "What is 802.11 & 802.11B?"; www.pulsewan.com/data101/802_11_b_basics.htm

[27] Siemens; Mohr, Werner; Konhäuser, Walter; "Beyond Third Generation Systems"; www.ece.utexas.edu/ yeong/2.pdf

[28] Schiller, Jochen; Mobilkommunikation; 2000

[29] Sky DSP; "The suitability of OFDM as a modulation techique for wireless telecommunications, with a CDMA comparison"; www.skydsp.com/publications/4thyrthesis/index.htm

[30] tecChannel; "802.11: Standard für drahtlose Netze"; www.tecchannel.de/hardware/680/index.html

[31] T-Mobile; www.t-mobile.de

[32] University of Glasgow; “QPSK“;
http://students.dcs.gla.ac.uk/students/hinsheld/TheLastMile/qpsk.html

[33] UMTS-Report; O2 UMTS-Tarife;
www.umts-report.com/index.php4?seite=thema&thema=60

[34] Vodafone; www.vodafone.de

[35] World Cellular Coverage Maps; “World GSM Coverage“;
www.gsmcoverage.co.uk/maps/europe/world.jpg

[36] WiLAN; “Wide-band Orthogonal Frequency Multiplexing (W-OFDM)“;
www.wi-lan.com/library/whitepaper_wofdm_technical.pdf

[37] Wireless KnowHow; www.m-indya.com

[38] Xonio; “Comdex-Star“; www.xonio.com/news/news_8920032.htm

[39] ZDNet TechReport; “Mobile Business“;
www.zdnet.de/mobile/artikel/techreport/mobile-business/mobile-business01-wc.html

Kapitel 2

Mobility and Mobile IP

Stefan Wagenbrenner

Mobile Kommunikation spielt im heutigen Leben eine immer größere Rolle, was zur Entwicklung vieler neuer Techniken geführt hat und noch führen wird. Dies betrifft nicht nur die mobile Telekommunikation, sondern im steigenden Maße auch drahtlose Netzwerke, denn die Grenze zwischen mobiler Telekommunikation und drahtlosen Netzen verschwindet immer mehr. Ein Beispiel ist die Möglichkeit, spezielle Internetseiten auf seinem Handy oder auf einem anderen mobilen Endgerät abzurufen. Dies ist nur einer der Dienste, die von einer steigenden Anzahl von Personen genutzt werden.
Ein besonderer Augenmerk richtet sich daher bei der Entwicklung neuer Systeme und Protokolle auf die Unterstützung der Mobilität der Nutzer und ihrer Geräte.
In dieser Arbeit werden zunächst grundlegende Begriffe der Mobilität, wie z.B. Mikro- und Makro-Mobilität, eingeführt. Im Anschluß wird das Protokoll Mobile IP und seine Funktionsweise erläutert. Mobile IP ist ein Protokoll, welches speziell dazu entwickelt wurde, die Mobilität von Nutzern zu ermöglichen. Zuerst wird Mobile IP in der Version vorgestellt, welche auf dem Internet-Protokoll IPv4 basiert. Ausgehend von den Nachteilen von Mobile IP im Bereich der Mikro-Mobilität, werden die Ziele der Verwirklichung von Protokollen vorgestellt, welche für die Unterstützung von Mikro-Mobilität entwickelt wurden und entwickelt werden. Die Ziele werden abschließend am Beispiel des Protokolls Cellular IP vorgestellt und erläutert. Abschließend werden in einem Ausblick die Möglichkeiten dargestellt, welche bei Einführung von Mobile IPv6 erhalten werden. Mobile IPv6 ist eine neue Version von Mobile IP, welche auf dem Protokoll IPv6 aufbaut.

Inhaltsverzeichnis

2.1 Einleitung

Als das Internet entstand, dachte noch niemand daran, daß es jemals nötig wäre, mobile Systeme zu unterstützen. Es wurden folglich Protokolle für die Kommunikation im Internet entwickelt, welche auf ein statisches Netz ausgelegt waren. Hierbei seien nur die wichtigsten, IP und TCP, genannt.
Heute haben sich viele Nutzer so stark an die Nutzung des Internets gewöhnt, daß sie nirgends mehr darauf verzichten möchten. Als dann auch noch Laptops immer kleiner und leichter wurden, PDAs und weitere transportable Systeme entwickelt wurden, Mobilität also in greifbare Nähe rückte, mußte man erkennen, daß sich TCP/IP bereits als de facto Standard für die Kommunikation in Netzwerken etabliert hat. Da IP aber für statische Netzwerke entwickelt wurde, stand man zunächst vor einem Problem. Aufgrund der starken Verbreitung von IP hat man sich dennoch entschlossen, IP als Grundlage für die Verwirklichung der Mobilität, von Nutzern und deren Systeme, zu nutzen. Auf jeden Fall sollte vermieden werden, ein völlig neues Protokoll zu entwickeln. Ein entscheidendes Ziel war außerdem, keine Änderungen an den nicht mobilen Systemen vornehmen zu müssen. Dies führte zur Entwicklung des Protokolls **Mobile IP**. Mobile IP ist eine Erweiterung des Protokolls IP und wurde 1996 von der Internet Engineering Task Force (IETF), einer Gruppe von Forschern und Netzwerkspezialisten, standardisiert. Durch Mobile IP wird es möglich, ein mobiles System an jedem möglichen Ort, z.B. auf dem Campusgelände oder auf einer Konferenz, an das dort vorhanden Netz anzuschließen und sofort wieder erreichbar zu sein. Dies wäre durch einfaches Ändern der IP-Adresse nicht möglich.
Im folgenden Kapitel werden wir uns zunächst mit einigen Grundbegriffen der Mobilität beschäftigen, bevor wir uns dann, beginnend mit Kapitel 2.3, ausführlich mit dem Protokoll Mobile IP beschäftigen.

2.2 Grundlegende Begriffe der Mobilität

In diesem Teil der Arbeit werden die Begriffe, welche in den weiteren Kapiteln benutzt werden und für das Verständnis von Mobile IP wichtig sind, eingeführt.
Hierbei handelt es sich um die Begriffe Nutzer- und Gerätemobilität und Mikro- und Makro-Mobilität. Desweiteren werden wir IP- und TCP-Mobilität vorstellen.

2.2.1 Device Mobility und User Mobility

Unter dem Begriff **Device Mobility** (Mobilität des Geräts) versteht man die Tatsache, daß ein Benutzer sein Endsystem (wie z.B. einen Laptop oder einen PDA) jederzeit an einen anderen Netzzugangspunkt (Access Point) anschließen kann. Nachdem er sein Gerät an das Teilnetz angeschlossen hat, muß dem Gerät nur noch eine IP-Adresse zugewiesen werden und schon kann der Nutzer die im Teilnetz angebotenen Dienste nutzen. Eine IP-Adresse kann entweder durch den Nutzer manuell eingetragen werden oder durch DHCP zugewiesen werden. Während der Nutzer sich von einen Netzzugangspunkt zu einem anderen bewegt, besteht keine Verbindung.

User Mobility (Mobilität des Nutzers) bezeichnet die Tatsache, daß sich der Benutzer mit seinem Endsystem in einem Netzwerk bewegen kann bzw. zwischen mehreren Netzwerken hin und her wechseln kann, ohne das die Verbindung beendet wird.
Andererseits kann unter dem Begriff der User Mobility auch der Sachverhalt verstanden werden, daß ein Nutzer aktiv von einem Arbeitsplatz zu einem anderen wechselt. Dies betrifft aber nicht die Unterstützung mobiler Systeme und wird deswegen in dieser Arbeit nicht weiter behandelt.
Ein Nutzer kann zum Beispiel über eine WLAN-Schnittstelle (Erklärung siehe Seite 63) auch im Gehen Informationen auf seinen PDA oder seinen Laptop holen, dies geschieht derzeit mit einer Bandbreite von 11 MBit/s (IEEE 802.11b Standard).
Device Mobility und User Mobility werden in der folgenden Abbildung dargestellt.

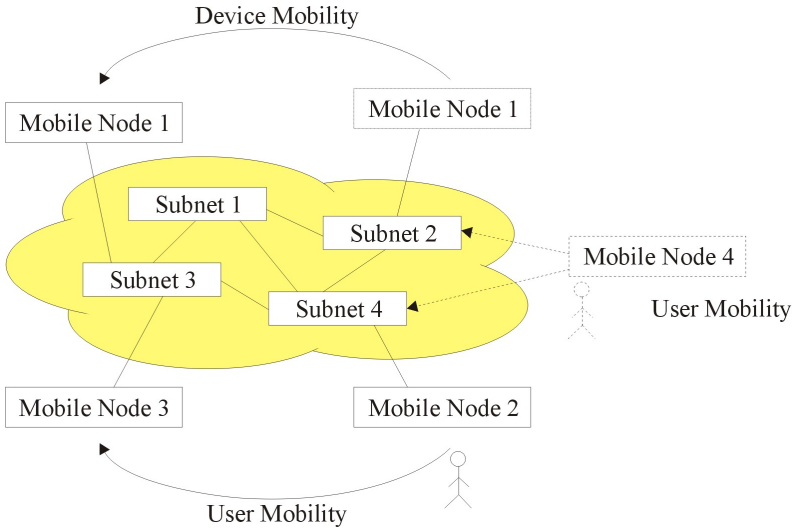

Abbildung 2.1: Benutzer- und Gerätemobilität

2.2.2 Mikro-Mobilität und Makro-Mobilität

So wie man die Mobilität in Nutzer- und Gerätemobilität unterteilen kann, so kann man auch eine Unterscheidung der Mobilität im Bezug auf den Umfang der Bewegung, d.h. wie weit sich der Nutzer von seinem Zugansgpunkt entfernt, vornehmen. Eine Unterscheidung ist nötig, um den Voraussetzungen, wie Flexibilität und Stabilität der Verbindung, gerecht zu werden. Diese Voraussetzungen sind v.a. für Benutzer, welche sich (häufig) in einem kleinen Bereich bewegen, von Bedeutung.
Die erste Art, **Mikro-Mobilität** (micro-mobility), liegt dann vor, wenn sich der mobile Knoten (Mobile Node, MN) innerhalb einer Domäne (domain) bewegt. Hierbei bezeichnet eine Domäne einen logischen Zusammenschluß von Netzwerken unter einheitlicher Administration (nach [11]). Bei dieser Art der Mobilität muß Wert auf schnellen und nahtlosen Übergang zwischen einzelnen Zugangspunkten (Access Points, APs) gelegt werden.
Ziel ist es eine geringe Verzögerung, wenig Paketverluste und einen geringen Signalisierungsaufwand zu erreichen.

Hierfür wurden verschiedene Protokolle entwickelt, welche Mobile IP erweitern:

1. Cellular IP
2. HAWAII
3. Hierarchical Mobile IP
4. Proactive Handoff
5. Fast Handoff
6. TeleMIP
7. EMA

Von diese Protokollen wird in dieser Arbeit nur das Protokoll Cellular IP detailliert vorgestellt. Eine Beschreibung der anderen Protokolle kann z.B. in [11] und [12] gefunden werden.
Im Gegensatz zu Mikro-Mobilität spricht man von **Makro-Mobilität** (macro-mobility), wenn sich das mobile System von einer Domäne zu einer anderen bewegt, wobei die Domänen durch das Internet verbunden sind. Das Protokoll, welches hierfür entwickelt wurde, heißt **Mobile IP**.
Wichtig ist hierbei, daß bei der obigen Definition des Begriff Domäne, keine scharfe Trennung zwischen Mikro- und Makro-Mobilität möglich wäre, denn es können zwei Netzwerke, zwischen welchen sich ein mobiles System hin- und herbewegt (Makro-Mobilität), zu einem logischen Netzwerk zusammengefaßt werden, so daß Mikro-Mobilität vorliegen würde.
Die verschiedenen Arten der Mobilität werden in Abbildung 2.2 dargestellt.

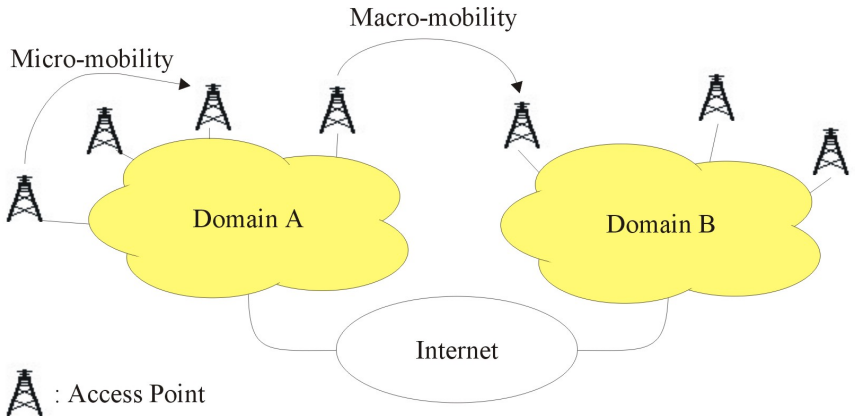

Abbildung 2.2: Mikro- und Makro-Mobilität

2.2.3 IP- und TCP-Mobilität

Wie in der Einleitung bereits erwähnt, wurde das Internet als ein statisches Netz aufgebaut. Um mobile Nutzer zu unterstützen, ist es zwangsläufig nötig, neue Protokolle

einzuführen bzw. Änderungen an bestehenden Protokollen vorzunehmen. Da man keine Änderungen an Systemen vornehmen möchte, die bereits vorhanden sind, hat man sich entschieden, das Protokoll IP durch ein weiteres Protokoll, Mobile IP, zu ergänzen.
Leider reicht die Unterstützung von Mobilität allein auf den unteren Schichten des TCP/IP Referenzmodells (Abbildung 2.3) nicht aus, da es auch möglich sein soll mobil zu sein, ohne eine bestehende TCP-Verbindung zu unterbrechen oder gar abzubrechen.

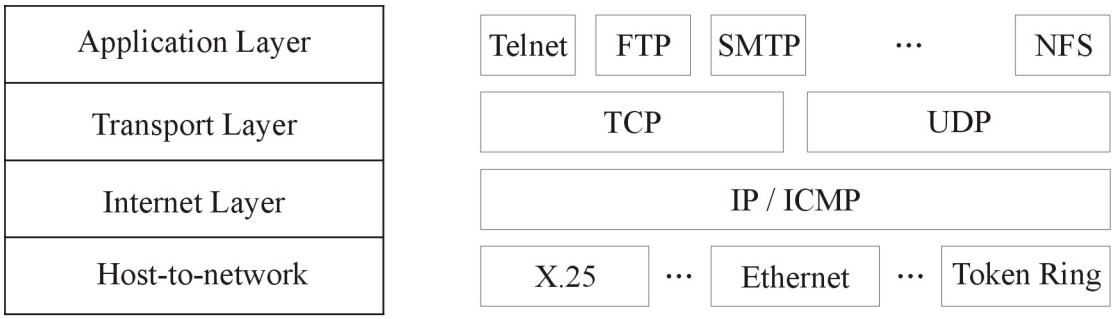

Abbildung 2.3: TCP/IP - Referenzmodell und Beispiele für Protokolle der verschiedenen Ebenen

Viele Anwendungen verwenden TCP (Transmission Control Protocol) zur Übertragung von Daten im Internet. Würde bei der Unterstützung von mobilen Systemen die Anpassung von TCP (und der anderen Protokollen auch) vernachlässigt werden, so könnte dies fatale Folgen für die Übertragung bedeuten. Die Folgen reichen vom Sinken der Übertragungsrate bis hin zum Verbindungsabbruch. Ein Grund hierfür ist, daß ein „sanfter“ Übergang von einem Zugangspunkt zu einem anderen in manchen Situationen nicht erreicht werden kann (z.B. bei sich schnell bewegenden Systemen; siehe auch 2.4.1), was zwangsläufig zu Paketverlusten führt. Paketverluste, bzw. fehlende Empfangsbestätigungen, interpretiert TCP als Stausituation, was dazu führt, daß die Übertragungsrate durch TCP gesenkt wird.
Dies führte zur Entwicklung verschiedener Protokolle, die sowohl zur Erweiterung von IP (eine Liste von Protokollen siehe Abschnitt 2.2.2), als auch von TCP dienen.
Einige Protokolle, welche zur Erweiterung von TCP für mobile Nutzer entwickelt wurden, sind:

1. Indirektes TCP,
2. Snooping TCP,
3. Mobile TCP,
4. Fast Retransmit - Fast Recovery.

2.3 Mobile IP

In diesem Kapitel wird das Protokoll Mobile IP (RFC 2002, [5]) vorgestellt. Mit Hilfe von Mobile IP kann Makro- bzw. Mikro-Mobilität von Systemen, wie z.B. Laptops und PDAs,

gewährleistet werden.
Hierbei sei angemerkt, daß Mobile IP eine Erweiterung von IP ist, wobei bei der Einführung von Mobile IP, wie bereits erwähnt, keine Änderungen an den Systemen vorausgesetzt werden dürfen, welche Mobile IP nicht unterstützen (man spricht hier von der Transparenz von Mobile IP).

2.3.1 Einfache Lösungen des Mobilitäts-Problems?

An dieser Stelle werden Lösungsmöglichkeiten angeführt, welche es scheinbar ermöglichen, mobile Systeme zu unterstützen ohne (möglicherweise umfangreiche) Änderungen am IP-Protokoll vornehmen zu müssen.
Eine erste Möglichkeit sieht vor, mobilen Systemen jedesmal, wenn sie an einen neuen Ort gebracht werden, eine neue IP-Adresse zuzuordnen. Sobald das System eine IP-Adresse erhalten hat, kann es als Client alle angebotenen Dienste nutzen. Ein Problem tritt aber dann auf, wenn man das mobile System erreichen will, um beispielsweise Daten auszutauschen. Da das System eine neue IP-Adresse bekommen hat und alle anderen Teilnehmer nichts davon wissen, ist es unmöglich das mobile System zu finden.
Ein weiteres Problem, das auftritt, wenn man bei jedem Ortswechsel einem mobilen System eine neue IP-Adresse zuordnet, ist, daß Protokolle und Anwendungen auf höheren Schichten einen dynamischen Wechsel der IP-Adresse nicht verkraften, da diese für eine statische IP-Adresse ausgelegt wurden. Als Beispiel betrachten wir eine TCP-Verbindung, welche durch das Tupel

(Quell IP-Adresse, Quell Port-Nummer, Ziel IP-Adresse, Ziel Port-Nummer)

festgelegt ist. Ändert sich während der Verbindung die Ziel IP-Adresse, so wird die Verbindung abgebrochen.
Man muß also erkennen, daß durch einfache Änderung der IP-Adresse eines mobilen Systems, die Realisierung der Mobilität nicht zufriedenstellend verwirklicht werden kann.
Ein weiterer Lösungsvorschlag wäre, spezielle Wege zu einem mobilen System festzulegen. Wechselt ein mobiles System seinen Zugangspunkt, müssen nur die Wegewahltabellen in allen Routern aktualisiert werden und das mobile System wäre wieder erreichbar.
Bei einer Vielzahl von mobilen Systemen ist der hohe Verwaltungsaufwand ein eigentlich unlösbares Problem. Der Aufwand hierfür wäre im Vergleich zum erzielten Nutzen viel zu groß.

2.3.2 Forderungen an Mobile IP

An eine Implementierung von Mobile IP werden zahlreiche Forderungen gestellt:
Als erste Forderung kann angeführt werden, daß Mobile IP keine Änderungen an Systemen erzwingen darf, die bereits verwendet werden, dies gilt sowohl für Hosts als auch für Router. Begründet werden kann dies mit der Vielzahl von Rechnern, die TCP/IP zur Kommunikation einsetzen. Es wäre utopisch zu verlangen, daß Nutzer, die keine mobilen Systeme nutzen, Änderungen an ihren Systemen vornehmen müssen, nur um die Mobilität

anderer zu unterstützen. Deswegen muß bei der Entwicklung von Mobile IP auch darauf geachtet werden, daß eine Kommunikation mit Systemen, welche normales TCP/IP benutzen, weiterhin möglich ist.
Weiterhin muß Mobile IP effizient sein. Das bedeutet, daß nicht zuviele Daten zur Verwaltung des Netzes gesendet werden dürfen. Dies spielt v.a. bei drahtlosen Netzen eine Rolle, da diese eine geringere Bandbreite bieten. Würden diese Netze auch noch mit einer Unzahl von Verwaltungsdaten belastet werden, würde man nicht mehr effizient über das Netz arbeiten können. Auch den Gesichtspunkt der Skalierbarkeit darf nicht vergessen werden, denn es kann davon ausgegangen werden, daß immer mehr mobile Systeme an das Internet angeschlossen werden. Deswegen ist es wichtig, daß Mobile IP auf eine wachsende Anzahl von Nutzern ausgelegt wird.
Ein sehr wichtiger Aspekt in Mobile IP ist Sicherheit. Es muß darauf geachtet werden, daß alle Daten, welche von Mobile IP zur Verwaltung gesendet werden, authentifiziert und authorisiert werden. Man spricht hier von AAA (Authentication, Authorization und Accounting), was in der Seminararbeit „AAA and Extensions for Wireless Services“ behandelt wird. Ein Problem würde z.B. dann entstehen, wenn Pakete zu einem System gelangen würden, das diese Daten nicht erhalten darf. Mit normalem IP kann nur festgestellt werden, ob eine IP-Adresse korrekt ist, nicht aber, ob Daten an einen rechtmäßigen Empfänger gesendet werden.
Würden die genannten Anforderungen nicht erfüllt werden, so würde man riskieren, daß Mobile IP nicht akzeptiert werden würde und folglich zum Scheitern verurteilt wäre.

2.3.3 Begriffsdefinitionen

In diesem Abschnitt werden die Begriffe eingeführt, welche nötig sind, um die Mechanismen von Mobile IP zu verstehen. Hierbei beziehen wir uns auf Mobile IP, wie es in RFC 2002 [5] festgelegt ist.
Abbildung 2.4 zeigt den Aufbau eines Beispielnetzes, welches alle Komponenten enthält, die durch Mobile IP eingeführt wurden.

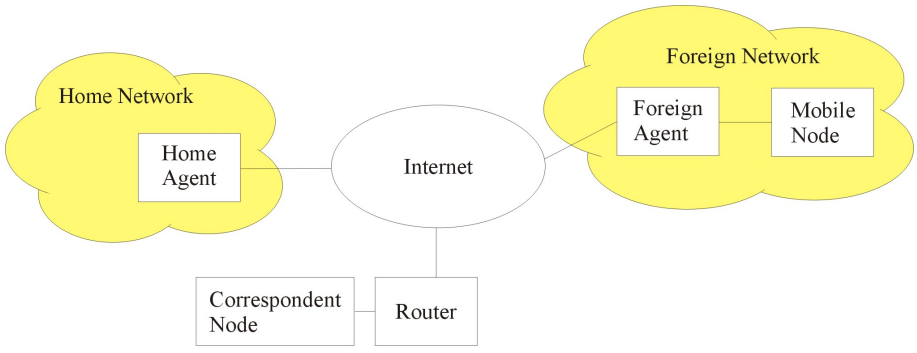

Abbildung 2.4: Aufbau des Beispielnetzes

Hierbei ist ein **Mobile Node** (MN, mobiler Knoten) ein Endsystem (Laptop, PDA, ... oder ein Router), welches seine Zugehörigkeit zu einem Netz von einem Netzwerk zu einem anderen verlagern kann. Ein MN besitzt eine statische IP-Adresse und kann mit anderen Systemen im Internet eine Verbindung aufrechterhalten, egal in welchem Netz der MN

sich befindet, solange eine Verbindung auf Ebene der Vermittlungsschicht vorhanden ist.
Das **Home Network** (HN, Heimatnetzwerk) ist das Netz, welchem der MN aufgrund seiner statischen IP-Adresse angehört. Befindet sich der MN in seinem HN, so ist keine Unterstützung durch Mobile IP notwendig.
Das **Foreign Network** (FN, Fremdnetzwerk) ist das Netzwerk, in welchem sich der MN gegenwärtig aufhält und nicht sein HN ist.
Ein **Correspondent Node** (CN, Kommunikationspartner) ist ein System, mit welchem der MN eine Verbindung unterhält. Dieser kann sowohl ein stationäres als auch ein mobiles System sein.
Der **Home Agent** (HA, Heimatagent) bezeichnet einen Router, welcher sich im Heimatnetz des MN befindet. Seine Aufgabe ist es, alle Pakete, welche für den MN bestimmt sind, falls sich dieser nicht in seinem Heimatnetz aufhält, abzufangen und an den MN weiterzuleiten. Um die Weiterleitung von Datenpaketen zu ermöglichen, unterhält der HA eine Liste mit den Aufenthaltsorten aller MNs.
Ein **Foreign Agent** (FA, Fremdagent) ist ein Router in dem Netz, in welchem sich der MN zur Zeit befindet. Dieser hat die Aufgabe Datenpakete, welche für den MN bestimmt sind zu entpacken und an diesen weiterzuleiten. Für Pakete, die der MN aus dem Fremdnetz heraussendet, fungiert der FA als Standard-Router.

Desweiteren müssen zwei Adressen definiert werden, welche den MN eindeutig identifizieren:

1. **Home Address**
   Dies ist die statische IP-Adresse, welche dem MN zugeordnet ist. Diese Adresse ändert sich nie, unabhängig davon wo sich der MN befindet.

2. **Care-of Address** (COA)
   Die COA gibt den aktuellen Aufenthaltsort des MN an. Alle Pakete, welche an den MN gesendet werden, werden an die COA geliefert und nicht direkt an die Home Address des MN. Die Zustellung geschieht durch sogenanntes tunneln, welches in Abschnitt 2.3.5 erläutert wird. Es werden zwei Arten von COA unterschieden:

   (a) Foreign Agent COA:
   Hierbei ist die COA die Adresse eines FA. Der FA ist das Ende des Tunnels und für die weitere Zustellung der Pakete an den MN verantwortlich.

   (b) Co-located COA:
   In diesem Fall wird dem MN eine IP-Adresse zugeordnet, welche dem Adress-Pool des FN entstammt. Diese Adresse kann der MN beispielsweise durch DHCP (Dynamic Host Configuration Protocol, RFC 1541, [13]) erhalten.
   Wenn dem MN eine co-located COA zugeordnet wird, so ist der MN selbst der Tunnelendpunkt und für das Entpacken der Pakete verantwortlich.
   Ein Nachteil ist bei dieser Art der COA, daß ein Adreß-Pool vorhanden sein muß, aus welchem der MN eine Adresse erhält. Dies führt, falls IPv4 verwendet wird, zu einer Verknappung der IP-Adressen.

2.3.4 Ablauf der Paketzustellung

Nachdem die grundlegenden Begriffe eingeführt wurden, wird an dieser Stelle die Weiterleitung von Paketen von einem CN zu einem MN erläutert (siehe Abbildung 2.5). Es

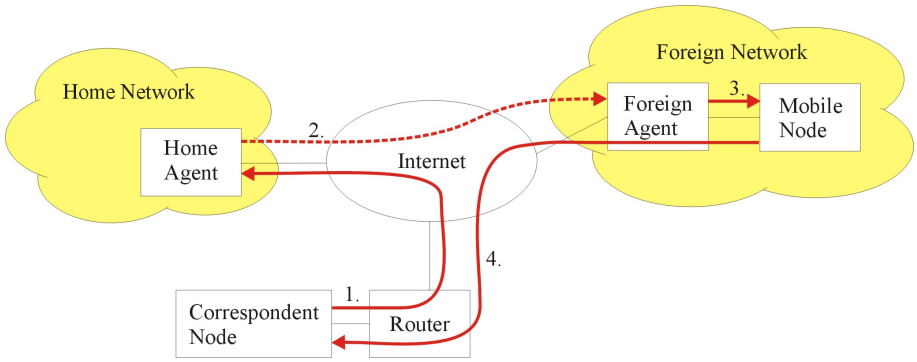

Abbildung 2.5: Ablauf der Paketzustellung (nach [10])

ist anzumerken, daß dem CN nicht bekannt sein muß, ob der MN mobil ist oder nicht. Desweiteren muß der CN nicht wissen, wo sich der MN zur Zeit befindet, er muß nur dessen Home Address kennen.
Im ersten Schritt (1. in der Abbildung) sendet der CN ein Datenpaket an die Home Address des MN. Im Internet wird das IP-Paket weitergeleitet. Nachdem das Paket im HN angekommen ist, fängt der HA es ab. Diesem ist bekannt, daß sich der MN nicht im HN aufhält und er sendet deswegen das Datenpaket an die COA des MN (hierbei nehmen wir an, daß dem MN eine Foreign Agent COA zugeordnet wurde). Dies geschieht, indem er das Datenpaket an den FA durch einen Tunnel weiterleitet (2.). Der genaue Ablauf wird in Abschnitt 2.3.5 dargestellt. Nachdem das Datenpaket am FA angekommen ist, leitet der FA das Datenpaket an den MN weiter (3.).
Will der MN, nachdem er das Datenpaket erhalten hat, Daten an den CN senden, so sendet der MN Pakete mit seiner Home Address als Quelladresse an die Adresse des CN (4.). Hierbei könnten Probleme auftreten, welche hier kurz vorgestellt werden.

1. Das FN, in welchem sich der MN befindet, wird durch eine Firewall vom Internet abgeschirmt. Diese Firewall könnte Pakete abblocken, welche eine Sendeadresse besitzen, welche nicht zum Adreßbereich des Netzwerks gehört (d.h. die Adressen stellen für das Netzwerk eine topologisch nicht korrekte Adresse dar). Da der MN aber mit seiner Home Address als Sende-Adresse Pakete verschickt, werden seine Pakete nicht durchgelassen.

2. Weiterhin könnte die TTL (Time to Live) so eingestellt sein, daß z.B. ein Manager im Netzwerk des Verwaltungsgebäude (hier das HN) alle Rechner erreicht. Befindet sich der Manager aber auf einer Konferenz, so kann es passieren, daß er auf Grund der zu geringen TTL nicht mehr alle Rechner im HN erreicht. In diesem Fall müßte man die TTL manuell, je nach Ort, einstellen. Dies verstößt aber gegen die Transparenz, welche man von Mobile IP verlangt.

Die aufgeführten Probleme führten zur Entwicklung von Reverse Tunneling (RFC 2344, Reverse Tunneling for Mobile IP), welches die obengenannten Problem löst, aber neue

(Sicherheits-) Probleme einführt, die teilweise bis jetzt noch nicht gelöst wurden. Für eine vertiefte Darstellung sei auf den RFC 2344 [9] verwiesen.

2.3.5 Tunneln

In Abschnitt 2.3.4 wurde bereits kurz angemerkt, wie der HA Pakete an den FA weiterleitet (siehe 2. in Abbildung 2.5). An dieser Stelle wird die Vorgehensweise der Weiterleitung detaillierter dargestellt. Das Weiterleiten von Datenpaketen bezeichnet man im Allgemeinen als „tunneln“. Der Weg eines Datenpakets hat folgende Gestalt:

$$\text{Quelle} \rightarrow \underbrace{\text{Verpacker (encapsulator)} \longrightarrow \text{Entpacker (decapsulator)}}_{\text{Tunnel}} \rightarrow \text{Ziel.}$$

Der Tunnel beginnt beim Verpacker und endet beim Entpacker. Der Verpacker wird auch als Tunneleingangspunkt, der Entpacker als Tunnelausgangspunkt bezeichnet.
Die Quelle ist im Fall von Mobile IP der CN, welcher Daten an den MN senden möchte. Befindet sich der MN nicht in seinem HN, so werden die Pakete vom HA abgefangen, der die Rolle des Verpackers inne hat. Der HA sendet dann das Datenpaket als ein neues Datenpaket an die COA. Dem Ursprungs-Paket wird hierzu einfach ein neuer IP-Header vorangestellt, der sogenannte äußere IP-Header (Outer IP Header, siehe Abbildung 2.6). Dieser Vorgang wird als IP in IP Encapsulation bezeichnet (vgl. RFC 2003, [4]).

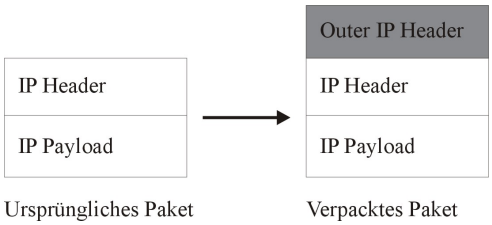

Abbildung 2.6: Verpacken eines Pakets (aus [4])

Die folgende Abbildung zeigt den Aufbau des Pakets, wenn es sich im Tunnel befindet. An dem IP-Header des eingegangenen Pakets wird nur eine Änderung vorgenommen, die TTL wird um 1 vermindert. Dies hat zur Folge, daß der Tunnel, aus der Sicht des Pakets, eine Länge von 1 hat.
Kommt das Paket beim Entpacker, dem Ende des Tunnels, an, so entfernt dieser den äußeren IP-Header und leitet das Paket an den MN weiter.
Wir können also feststellen, daß mit Hilfe von IP in IP Encapsulation die Zustellung von Paketen vom CN zum MN bewerkstelligt werden kann. Beim Tunneln muß an folgende Punkte gedacht werden:

1. Die Größe des Datenpakets nimmt aufgrund des Hinzufügens des äußeren IP-Headers zu.

2. Man muß vor dem Senden wissen, daß das System am Ende des Tunnels das Entpacken durchführen kann.

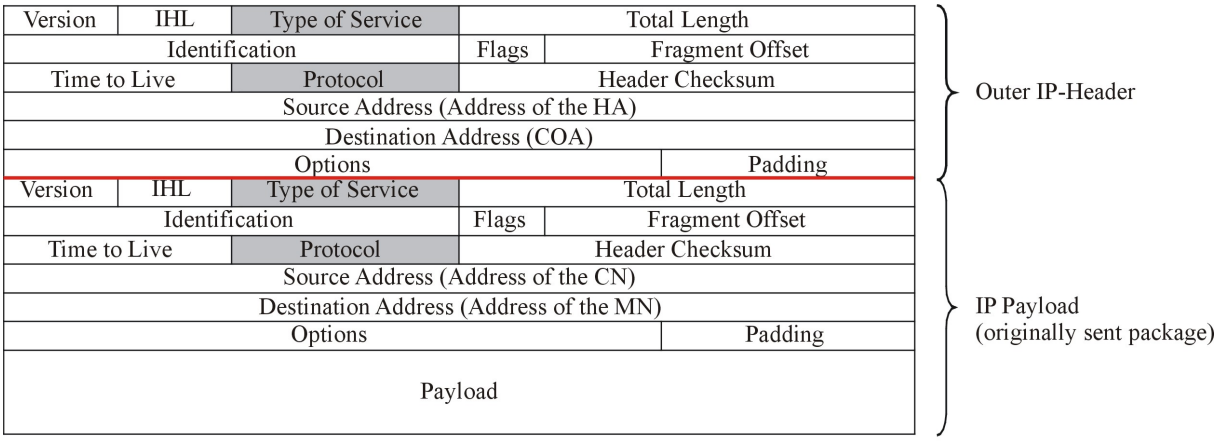

Abbildung 2.7: Aufbau des Pakets im Tunnel (nach [4])

Desweiteren kann festgestellt werden, daß es Informationen gibt, die bei der IP in IP Encapsulation in beiden IP-Headern vorkommen, z.B. wird der Type of Service-Eintrag und der Protocol-Eintrag vom inneren in den äußeren IP-Header kopiert (beide Felder sind in der Abbildung 2.7 grau schattiert dargestellt), also doppelt übertragen werden. Diese Tatsache führte zur Entwicklung des Protokolls Minimal Encapsulation within IP, welches im RFC 2004, [6], festgelegt ist.

2.3.6 Agent Discovering

Eine wichtige Frage ist, wie ein MN überhaupt feststellen kann, wo er sich zur Zeit befindet, d.h. vor allem, ob er sich von einem Netzwerk in ein anderes bewegt hat. Eine Beantwortung der Frage ist vor allem deswegen wichtig, da der MN bei einem Wechsel seines Zugangspunktes eine neue COA erhalten muß, um seine Erreichbarkeit sicherzustellen. Die angesprochene Frage kann der MN mittels Agent Discovery („Entdecken eines Agenten“) beantworten. FAs und HAs senden spezielle ICMP Router Advertisement Messages (hier kurz als Advertisement Message bezeichnet). Wenn diese Nachrichten in gleichen Abständen gesendet werden, so sollte der Abstand zwischen den Nachrichten $^1/_3$ der Lebensdauer/Gültigkeitsdauer der Advertisement Message betragen. Dies erlaubt dem MN 3 Nachrichten zu verpassen, bevor er den Agenten für sich als ungültig erklärt. In diesen Advertisement Messages wird die sogenannte Mobility Agent Advertisement Extension als Nutzlast transportiert (siehe Abbildung 2.8).

Ein wichtiges Feld des ICMP Headers ist der Eintrag Lifetime, welches die bereits erwähnte Gültigkeitsdauer der Advertisement Message angibt.

Die wichtigsten Einträge der Agent Advertisement Extension sind Registration Lifetime, welche die maximale Dauer der Gültigkeit einer Registrierung angibt (gemessen in Sekunden), die der jeweilige Agent akzeptiert, und die Einträge für die COAs, welche die durch den Agenten angebotenen COAs enthalten (in Abbildung 2.8 sind es in diesem Fall 2 COAs, es können aber auch keine, eine oder mehr als zwei angeboten werden).

Empfängt der MN eine Advertisement Message, so kann er daraus seinen Aufenthaltsort entnehmen und dann auf einen möglichen Ortswechsel reagieren.

Andererseits kann der MN auch selbst nach einem Agenten suchen. Hierfür verwendet er Agent Solicitation Messages. Diese Nachrichten sollten aber nur dann von einem MN ge-

Typ	Code	Checksum
Num Addrs	Addr Entry Size	Lifetime
Router Address [1]		
Preference Level [1]		
Router Address [2]		
Preference Level [2]		

Typ	Length	Sequence Number							
Registration Lifetime		R	B	H	F	M	G	V	Reserved
COA 1									
COA 2									

Abbildung 2.8: Aufbau einer Agent Advertisement Message (nach [14] und [5])

sendet werden, wenn er keine Advertisement Messages empfängt und wenn er auch nicht durch andere Maßnahmen, wie DHCP, eine COA erhalten hat. Der MN verwendet für das Suchen die gleichen Nachrichten, wie sie für ICMP Router Solicitation Messages festgelegt sind. Zu beachten ist hier außerdem, daß der MN die Rate, mit welcher er die Solicitation Messages sendet, nicht zu hoch wählen sollte, um das Netzwerk nicht unnötig zu belasten (für weitere Details sei hier auf [5] verwiesen).
Abschließend muß der MN seinem HA seine erhaltene COA mitteilen. Dieser Vorgang wird in dem folgenden Abschnitt erläutert.

2.3.7 Registrierung der Care-of Address

Nachdem der MN eine neue COA erhalten hat, muß er diese dem HA mitteilen, um seine Erreichbarkeit sicherzustellen. Diesen Vorgang bezeichnet man als Registrierung.
Bevor wir den Ablauf der Registrierung vorstellen, muß darauf hingewiesen werden, daß alle Pakete, welche für den Registrierungsvorgang verschickt werden, authorisiert werden müssen. Würde dies nicht geschehen, wäre der Vorgang anfällig für Angriffe.
Der Ablauf der Registrierung findet auf zwei verschiedene Arten statt. Das Unterscheidungsmerkmal ist die Art der COA (co-located COA bzw. Foreign Agent COA).
Wurde einem MN eine co-located COA zugewiesen, so sendet der MN eine Registrierungsanfrage (registration request) direkt an den HA. Der HA sendet dem MN dann eine Registrierungsantwort (registration reply) an die COA. Mit dieser Nachricht teilt der HA dem MN mit, ob er der Anfrage zustimmt oder ob er diese ablehnt. Den Ablauf dieser Art der Registrierung zeigt Abbildung 2.9.

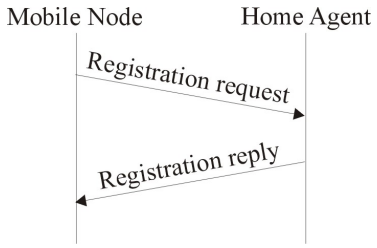

Abbildung 2.9: Ablauf der Registrierung bei Vorliegen einer co-located COA (nach [10])

Besitzt der MN hingegen eine Foreign Agent COA, so besteht der Ablauf der Registrierung

aus folgenden Schritten:

1. Der MN sendet eine Registrierungsanfrage an seinen FA.
2. Der FA verarbeitet die Anfrage und leitet diese dann an den HA des MN weiter.
3. Der HA sendet nach der Bearbeitung der Anfrage eine Registrierungsantwort an den FA. Der HA hat die Möglichkeit die Anfrage abzulehnen oder ihr zuzustimmen.
4. Zum Abschluß der Registrierung sendet der FA die Antwort des HA an den MN.

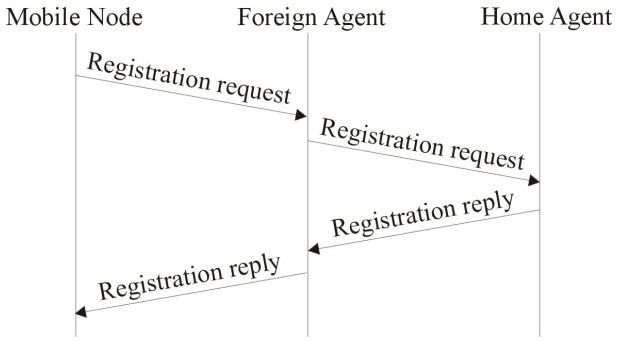

Abbildung 2.10: Ablauf der Registrierung, wenn eine Foreign Agent COA vorliegt (nach [10])

Während des Registrierungsprozeßes wird auch die Dauer der Gültigkeit der Registrierung verhandelt. Der HA speichert die Gültigkeitsdauer zusammen mit dem Aufenthaltsort des MN in einer Liste. Nach Ablauf der Gültigkeitsdauer löscht der HA den Eintrag, deswegen muß sich der MN regelmäßig beim HA registrieren.
Für Registrierungsanfragen und -antworten werden UDP-Pakete eingesetzt. Den Aufbau der jeweiligen Nutzdaten zeigen die folgenden beiden Abbildungen (2.11 und 2.12). An dieser Stelle soll nur die Bedeutung des Eintrags „Lifetime“ erwähnt werden, denn dieser gibt die Gültigkeitsdauer der Registrierung an. Die Bedeutung der anderen Felder ist entweder selbsterklärend oder kann [5] entnommen werden.

Type	S	B	D	M	G	V	Rsv	Lifetime
Home Address								
Home Agent								
COA								
Identification								
Extensions								

Abbildung 2.11: Aufbau der Nutzlast einer Registrierungsanfrage (aus [5])

2.4 Protokolle für die Mikro-Mobilität

In diesem Kapitel werden wir Protokolle vorstellen, welche speziell dafür entwickelt wurden, die Mikro-Mobilität von Systemen zu unterstützen. Bevor wir diese Protokolle näher betrachten, werden wir in dem folgenden Abschnitt die Probleme von Mobile IP erläutern.

Type	Code	Lifetime
Home Address		
Home Agent		
Identification		
Extensions		

Abbildung 2.12: Aufbau der Nutzlast einer Registrierungsantwort (aus [5])

2.4.1 Probleme von Mobile IP

Durch Mobile IP wird zwar das Problem mobiler Systeme gelöst, das Protokoll hat aber dennoch einige Nachteile, gerade im Bereich der Mikro-Mobilität, welche im folgenden vorgestellt werden.
Ein erstes Problem ist, daß eine schnelle Übergabe (**fast handoff**) von einem Zugangspunkt zu einem anderen nicht möglich ist. Jedesmal, wenn der MN seine COA ändert, muß er dies seinem HA mitteilen. Es kann aber lange dauern, bis der HA von der Änderung der COA informiert wird, was durch Verzögerungen bei der Übertragung bedingt ist.
Desweiteren wird der nahtlose Übergang (**seamless handoff**) von einem Zugangspunkt zu einem anderen durch Mobile IP nicht erreicht. Dies liegt an dem umfangreichen Registrierungsprozeß, bei dem es zwangsläufig auch zu Paketverlusten kommt. Dies führt auch zu einem erhöhten Verkehrsaufkommen, da verlorengegangene Pakete erneut gesendet werden müssen.
Ein weiteres Problem ist die große Menge an Signalisierungsnachrichten (**signaling traffic overhead**), welche sich v.a. bei mobilen Systemen ergibt, welche häufig ihren Zugangspunkt wechseln.
Bei Echtzeit- und Multimedianwendungen spielt auch **Quality of Service** (QoS) eine große Rolle. Bewegt sich der MN an einen neuen Ort, so wird zwischen HA und neuem FA eine Verhandlung der QoS-Parameter stattfinden, obwohl möglicherweise ein Großteil der Strecke zwischen einem CN und MN gleich geblieben ist.
Insgesamt kann man also feststellen, daß Mobile IP im Bereich der Makro-Mobilität das Mobilitätsproblem löst. Bewegt sich der MN aber nur in einem kleinem Bereich und wechselt er in diesem auch noch häufig seinen Zugangspunkt, so treten die zuvor erwähnten Probleme auf. Dies führte zur Entwicklung von Protokollen, welche das Problem der Mikro-Mobilität lösen sollen.

2.4.2 Cellular IP

An dieser Stelle wird das Protokoll Cellular IP und seine Arbeitsweise vorgestellt. Cellular IP erweitert Mobile IP im Bereich der Mikro-Mobilität. Abbildung 2.13 zeigt den grundlegenden Aufbau eines Netzwerkes, in welchem Cellular IP eingesetzt wird.
Die wichtigsten Komponenten in einem Cellular IP-Netzwerk sind die Base Stations (Basisstationen, BS), welche einen drahtlosen Zugangspunkt zum Netzwerk bieten.
Das Netzwerk ist durch einen Gateway an das Internet angebunden. Dieser Gateway hat die Rolle eines FA inne, wobei ein MN in diesem Netzwerk die Adresse des Gateway als seine COA nutzt. Bewegt sich ein MN in einem Cellular IP- Netzwerk von einer Basisstation zu einer anderen, so beeinflußt dies nur das Routing in dem Netzwerk selbst, der

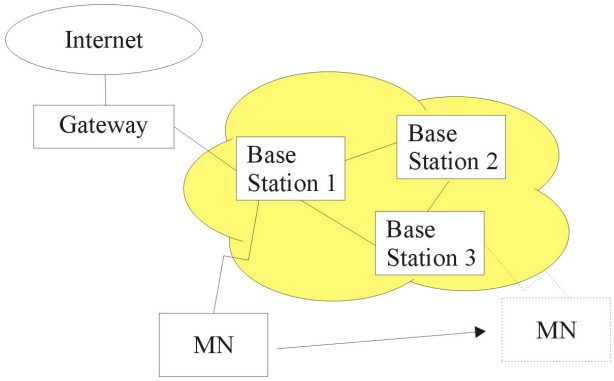

Abbildung 2.13: Aufbau eines Cellular IP-Netzwerks

Sachverhalt wird aber nicht seinem HA mitgeteilt (Mikro-Mobilität). Nur wenn sich die COA ändert wird dies dem HA mitgeteilt, d.h. wenn der MN sich in eine andere Domäne begibt (Makro-Mobilität).

Routing in Cellular IP-Netzwerken

In einem Netzwerk, in welchem Cellular IP eingesetzt wird, sendet der Gateway in regelmäßigen Abständen Pakete in das Netzwerk (Beacon Packets). Die einzelnen BS merken sich daraufhin, von welcher BS sie zuletzt ein solches Paket erhalten haben. Diese Informationen nutzen die BS dann, um Pakete eines MN mittels hop-by-hop Routing zum Gateway zu leiten.
Sendet ein MN Pakete, so werden diese von BS zu BS bis hin zum Gateway weitergeleitet. Hierbei speichert jede BS in einem Routing Cache die Sendeadresse des MN und die BS, von welcher das Paket erhalten wurde. Als Beispiel würde die BS_2 in Abbildung 2.13 den Eintrag (Adr_{MN}, BS_3) speichern, wenn sie ein Paket von einem MN mit der Adresse Adr_{MN} von der BS_3 erhalten hat. Es ist anzumerken, daß die Einträge in einem Routing Cache nicht dauerhaft gültig sind, nach einem Zeitabschnitt, dem sogenannten Route-Timeout (gem. [1]: 9s), werden die Einträge für ungültig erklärt. Solange der MN ständig Pakete sendet, stellt dieser Sachverhalt kein Problem dar. Wenn der MN aber länger als eine Zeiteinheit von 9s nichts sendet, würden die Einträge in den Routing Caches in den einzelnen BS dieses MN ungültig werden. Deswegen sendet ein MN, der eigentlich keine Daten zu senden hat, in periodischen Abständen (gem. [1] alle 3s; route-update time) route-update Pakete an den Gateway, wodurch die Einträge in den Routing Caches erneuert werden.

Handoff in Cellular IP

Durch Cellular IP werden 2 Arten von Übergaben (Handoffs) eines MN von einem Zugangspunkt an einen anderen eingeführt:

1. Hard Handoff und
2. Semisoft Handoff

Der Unterschied zwischen diesen beiden Ansätzen besteht darin, daß man bei Hard Handoff Paketverluste in Kauf nimmt, wohingegen man dies bei Semisoft Handoff zu vermeiden versucht. Im folgenden werden beide Ansätze kurz erläutert.

Hard Handoff: Wenn ein MN eine andere BS als neuen Zugangspunkt wählt, sendet der MN ein Route-Update Paket, wodurch neue Einträge in den Routing Caches entstehen. Einträge in den Routing Caches, welche auf den alten Zugangspunkt des MN verweisen, werden nicht gelöscht, diese werden automatisch nach Ablauf des Route-Timeout Zeitintervalls aus dem Cache entfernt. Während der Übergabe können Pakete verloren gehen, was daran liegt, daß Pakete weiterhin an die alte BS gesendet werden, da noch Einträge in den Routing Caches darauf verweisen. Diese Tatsache akzeptiert man aber, da die Zeit, welche benötigt wird, um Pakete „umzuleiten“ geringer ist als bei Mobile IP.
Semisoft Handoff: Die Tatsache, daß bei Hard Handoff die Einträge in den Routing Caches erhalten bleiben, kann dazu genutzt werden, einen anderen Übergabe-Mechanismus zu implementieren.
Anders als bei Hard Handoff sendet der MN vor der eigentlichen Übergabe ein sogenanntes Semisoft-Paket an die neu BS, wodurch neue Einträge in den Routing Caches der BSs, die sich zwischen der neuen BS und dem Gateway befinden, initiiert werden. Währenddessen bleibt der MN aber mit seiner „alten“ BS verbunden. Erst nach einem Zeitintervall, der semisoft delay-Zeit, führt der MN eine normale Übergabe durch. Dies führt insgesamt dazu, daß der MN während dieser Zeit Pakete von beiden BSs zugestellt bekommt.

Paging in Cellular IP

Unter paging versteht man das schnelle und effiziente Suchen und Finden von MNs, welche sich im Stand-by Modus befinden. Ein MN befindet sich genau dann im Stand-by Modus, wenn er länger als eine gewisse Zeiteinheit (active-state-timeout) keine Pakete gesendet hat.
Das Problem einen MN zu finden, würde nicht auftreten, wenn ein MN ständig Pakete senden würde, da dadurch die Einträge im Routing Cache aktuell bleiben. Dieser Ansatz bringt aber ein Problem mit sich. Durch das (unnötige) Senden von Paketen wird Bandbreite verschwendet.
Deswegen unterteilt man das Netzwerk in Paging-Bereiche (paging areas, PA), welche mehrere BS zusammenfassen (siehe Abbildung 2.14). Ziel dieses Ansatzes ist es, daß der MN nur noch dann seinen Standort im Netzwerk bekanntmachen muß, wenn er sich von einer PA zu einer anderen bewegt hat. Würde der MN im Stand-by Modus keine Pakete senden, würden alle Einträge in den Routing Caches, welche diesen MN betreffen, im Laufe der Zeit ungültig. Um aber weiterhin seine Erreichbarkeit sicherzustellen, sendet der MN periodisch paging-update Pakete (dies sind ICMP-Pakete), welche an den Gateway adressiert sind und zu diesem durch hop-by-hop routing geleitet werden. Manche BS besitzen neben dem Routing Cache zusätzlich noch einen paging cache, welcher durch ein paging-update Paket aktualisiert wird. Dieser paging cache besitzt die folgenden Eigenschaften:

1. Sowohl der Routing Cache als auch der Paging Cache besitzen den gleichen Aufbau für die Einträge.

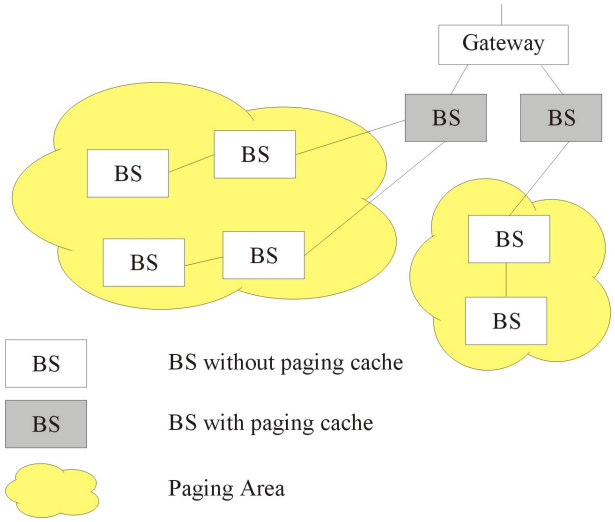

Abbildung 2.14: Unterteilung des Netzwerks in Paging-Areas (nach [3])

2. Einträge in einem Paging Cache bleiben länger gültig, müssen also seltener aufgefrischt werden.

3. Jedes Paket, was von einem MN gesendet wird, frischt den Paging Cache auf.

Der paging-Mechanismus wird nur dann zum Finden eines MN eingesetzt, wenn Einträge für diesen MN in keinem Routing Cache vorhanden sind. Besitzt eine BS keinen Paging Cache, so sendet diese das Paket an alle seine Nachbarn, nur nicht an die BS, von welcher sie das Paket erhalten hat.
Das erste Paket an einen MN wird in einem Netzwerk, in welchem weder Einträge im Routing Cache bzw. Paging Cache für diesen MN vorhanden sind, per Broadcast im Netzwerk verteilt.

2.5 Ausblick - Mobile IPv6

Mobile IP, welches wir in dieser Arbeit vorgestellt haben, basiert auf dem Internet Protokoll in der Version 4 (IPv4). An dieser Stelle werden wir kurz einige wesentliche Vorteile darstellen, die sich beim Umstieg von Mobile IP auf Mobile IPv6, das auf dem Internet Protokoll Version 6 (IPv6) basiert, ergeben. In die Entwicklung von Mobile IPv6 sind die Erfahrungen, welche bei der Entwicklung und dem Einsatz von Mobile IP gemacht wurden, eingeflossen. Weiterhin profitiert Mobile IPv6 auch von den Möglichkeiten, die IPv6 bietet.
Als erstes kann festgestellt werden, daß der Adreßraum von IPv6 größer ist als der von IPv4 (die Adreßlänge von IPv4 beträgt 32 bit, die von IPv6 hingegen 128 bit). Dadurch wird die Adreßknappheit beseitigt, welche bei der Vergabe von co-located COAs auftreten könnte.
Zusätzlich bietet IPv6 Mechanismen, welche die Zuordnung einer COA zu einem MN erleichtern. Es handelt sich hierbei um das Neighbor Discovery-Verfahren, mit welchem jeder

Knoten seine Nachbarn finden kann, und Stateless Address Autoconfiguration, wodurch dem MN eine IP-Adresse zugeordnet wird. Diese beiden Verfahren führen dazu, daß weder DHCP noch FAs benötigt werden, um einem MN in einem FN eine COA zuzuordnen.
Weiterhin implementieren alle IPv6-Knoten Authentifizierungs- und Verschlüsselungsverfahren, um die Sicherheit im Internet zu gewährleisten. Diese Verfahren können von Mobile IPv6 genutzt werden und müssen nicht zusätzlich, wie in Mobile IP, implementiert werden.
Um der Verschwendung von Bandbreite bei der Datenübertragung entgegenzuwirken, wurde zusätzlich das Verfahren Route Optimization (Wege-Optimierung) zum Protokoll von Mobile IP hinzugefügt. Route Optimization ist integraler Bestandteil von Mobile IPv6, in Mobile IP mußte dieses Verfahren zusätzlich implementiert werden. Das Verfahren, welches hinter Route Optimization steht, wird in dem folgenden Beispiel erläutert (siehe Abbildung 2.15): Zwei Manager treffen sich auf einer Konferenz und wollen Daten

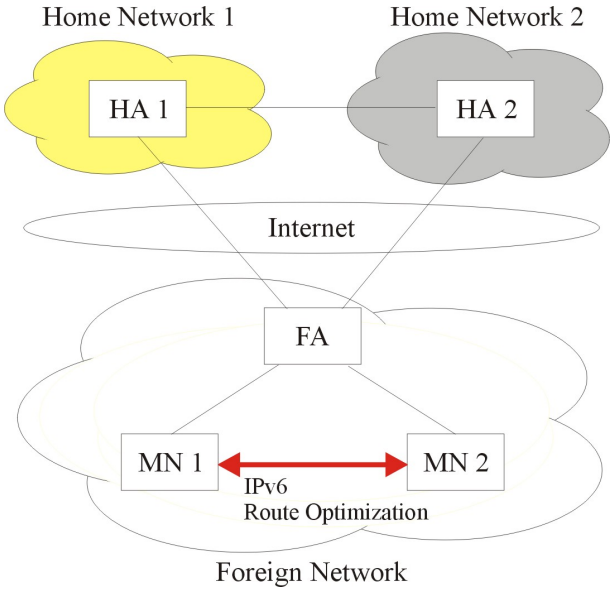

Abbildung 2.15: Aufbau des Beispiel-Netzwerks

zwischen ihren PDAs (MN1 bzw. MN2) austauschen. Im Normalfall läuft die Datenübertragung über die jeweiligen HAs (HA1 bzw. HA2) der Manager. Daß dies nicht sinnvoll sein kann, ist sofort einsichtig, denn die Manager sind vielleicht 1m voneinander entfernt, die jeweiligen HAs können aber beliebig weit entfernt sein. Es wäre also besser, die Daten direkt zwischen den beiden PDAs auszutauschen, als diese über weite Strecken zu verschicken. Deswegen gibt es in Mobile IPv6 auch die Möglichkeit einem CN (der sowohl stationär als auch mobil sein kann), mit welchem der MN verbunden ist, die COA des MN mitzuteilen. Das heißt, daß die Manager ihre Daten direkt austauschen können, was Bandbreite spart.
Ein weiterer Mechanismus, welcher durch Mobile IPv6 angeboten wird, kommt dann zum Einsatz, wenn ein MN keine Adresse eines HAs kennt. Dieser Mechanismus, Dynamic Home Agent Discovery, wird verwendet, um die Adresse eines HA zu ermitteln. Hierzu sendet der MN eine Nachricht an die Home Agents anycast address (RFC 2526, Reserved IPv6 Subnet Anycast Addresses), wodurch der MN einen Router in seinem HN erreicht, der auch als HA fungiert. Dieser HA sendet dem MN eine Liste mit allen HAs im HN.

Daraufhin sendet der MN dem ersten HA auf der Liste eine Nachricht, um sich bei diesem anzumelden. Antwortet dieser nicht oder lehnt der potentielle HA eine Anbindung des MN ab, so geht er die Liste von oben nach unten durch, bis er eine Bestätigung der Anbindung an den HA erhält.
In diesem Abschnitt haben wir nur einige wenige Vorteile von Mobile IPv6 vorgestellt. Aber bereits die genannten Vorteile lassen das Potential erahnen, welches bei der Einführung von Mobile IPv6 zur Verfügung steht. Dennoch darf nicht vergessen werden, daß Mobile IPv6 nicht alle Probleme löst, welche bei der Nutzung mobiler Systeme auftreten. Zu diesen Problemen zählen nach [7] u.a. die folgenden Beispiele:

1. Es ist weiterhin nicht möglich, zwischen Paketverlusten aufgrund von Bitfehlern und Verlusten aufgrund von Stausituationen zu unterscheiden.

2. Es wird auch kein Verfahren genannt, mit welchem es möglich ist, angebotene Dienste zu finden.

Weitere Beispiele für Vorteile von Mobile IPv6 und auch für die Unvollständigkeit von Mobile IPv6 finden sich in [7].

Glossar

P

Protokoll Ein Protokoll ist eine Menge von Regeln, mit welchen der Austausch von Informationen zwischen Systemen geregelt wird. Hierbei dient der Austausch von Daten dazu, einen Dienst anzubieten.
Bekannte Beispiele für Protokolle sind HTTP (HyperText Transfer Protocol), TCP (Transmisson Control Protocol) oder IP (Internet Protocol).

Q

Quality of Service (QoS) QoS umfaßt alle Verfahren, die geeignet sind, das Datenaufkommen in einem Netzwerk so zu steuern, daß ein Dienst mit einer genau spezifizierten Qualität beim Empfänger ankommt.
Es gibt aber keine einheitliche Definition von QoS. So haben z.B. die Organisationen ITU, IETF, das ATM-Forum und auch die OSI jeweils eigene Definitionen.

W

Wireless LAN (WLAN) Ein WLAN ist ein Netzwerk, welches ohne Kabelverbindung arbeitet. WLAN wurde von der IEEE als Standard 802.11 festgelegt. Die Übertragung erfolgt mittels Infrarot oder über Funkfrequenzen.

Z

Zugangspunkt An einem Zugangspunkt (auch: Netzwerkzugangspunkt) hat ein Nutzer die Möglichkeit auf ein Netzwerk zuzugreifen.

Literaturverzeichnis

[1] A. Campbell, J. Gomez, C.-Y. Wan, S. Kim, Z. Turanyi, A. Valko: *„Cellular IP“, draft-ietf-mobileip-cellularip-00.txt.* Draft, Dezember 1999.

[2] A. T. Campbell, J. Gomez, S. Kim, C.-Y. Wan, Z. R. Turanyi, A. G. Valko: *Comparison of IP Micromobility Protocols.* IEEE Wireless Communications, Volume 9(1):Seiten 72–82, Februar 2002.

[3] A. T. Campbell, J. Gomez, S. Kim, G. Valko, C-Y Wan, Z. R. Turanyi: *Design, Implementation, and Evaluation of Cellular IP.* IEEE Personal Communications, Volume 7(4):Seiten 42–49, August 2000.

[4] C. Perkins: *„IP Encapsulation within IP“, RFC 2003*, Oktober 1996.

[5] C. Perkins: *„IP Mobility Support“, RFC 2002*, Oktober 1996.

[6] C. Perkins: *„Minimal Encapsulation within IP“, RFC 2004*, Oktober 1996.

[7] D. B. Johnson, C. E. Perkins, J. Arkko: *„Mobility Support in IPv6 “, draft-ietf-mobileip-ipv6-18.txt.* Draft, Juni 2002.

[8] E. Nett, M. Mock, M. Gergeleit: *Das drahtlose Ethernet - Der IEEE 802.11 Standard: Grundlagen und Anwendung.* Addison-Wesley, 2001.

[9] G. Montenegro (Herausgeber): *„Reverse Tunneling for Mobile IP“, RFC 2344*, Mai 1998.

[10] J. Schiller: *Mobilkommunikation.* Addison-Wesley, München, 2000.

[11] J-Z Sun, D. Howie, J. Sauvola: *Mobility managment techniques for the next generation wireless networks.* Technischer Bericht, Infotech Oulu, University of Oulu, Finland.

[12] P. Reinbold, O. Bonaventure: *A Comparison of IP mobility protocols.* Technischer Bericht infonet-TR-13, Infonet group, University of Namur, Belgium, Dezember 2001.

[13] R. Droms: *„Dynamic Host Configuration Protocol“, RFC 1541*, Oktober 1993.

[14] S. Deering (Herausgeber): *„ICMP Router Discovery Messages“, RFC 1256*, September 1991.

[15] W. Fritsche, F. Heissenhuber: *Mobile IPv6 - Mobility support for the Next Generation Internet.* White Paper.

Kapitel 3

Routing in mobilen Ad-hoc Netzwerken

Klaus Schumacher

Drahtlose Netzwerke sind eine aufkommende neue Technologie die dem Benutzer unabhängig von seinem geographischen Aufenthaltsort auf elektronischem Weg den Zugang zu Informationen und Diensten ermöglicht. Drahtlose Netzwerke können in zwei Kategorien eingeteilt werden: Infrastrukturnetzwerke und infrastrukturlose mobile Ad-hoc Netzwerke (MANETs). Infrastrukturnetzwerken bestehen aus einem Netz verkabelter Gateways, ein mobiles Endgerät kommuniziert mit einer Bridge (auch Basisstation oder Access Point) die in seinem Sendebereich liegt. Das mobile Endgerät kann sich während der Kommunikation geographisch bewegen, sobald es den Sendebereich des Access Points verlässt, sucht und verbindet es sich mit einem anderen Access Point und setzt die Kommunikation durch diesen fort. Der Wechsel von einer Basisstation zur nächsten wird auch Handoff genannt. In Infrastrukturnetzwerken sind die Access Points unbeweglich. Im Gegensatz zu dazu sind in infrastrukturlosen Netzwerken alle Netzknoten mobil und können dynamisch und selbstständig untereinander Verbindungen aufbauen. Alle Partizipanten in einem solchen Netzwerk verhalten sich wie Router und beteiligen sich an der Suche von neuen und der Pflege von bestehenden Pfaden. Ad-hoc Netzwerke sind sehr nützlich bei Militäroperationen, Notfall Such- und Rettungseinsätzen, Meetings oder Konferenzen bei denen beteiligte Personen schnell und einfach Informationen austauschen wollen oder bei der Datengewinnung in schwer zugänglichem Gelände, da sie überall sofort einsatzbereit sind. Diese Arbeit befasst sich schwerpunktmäßig mit dem Routing in MANETs, den beiden Ansätzen auf Routing Protokolle aufbauen und stellt einige Reasilierungen dieser vor.

Inhaltsverzeichnis

3.1 Einleitung

Ein mobiles Ad-hoc Netzwerk besteht aus schnurlosen Endgeräten die über Funk miteinander kommunizieren ohne von infrastrukturellen Gegebenheiten abhängig zu sein. Die Mobilität der Netzknoten in einem MANET kann stetige unvorhersehbare Veränderungen in der Netzwerktopologie verursachen, wodurch besonders das Routing im Gegensatz zu kabelgebundenen Netzwerken sehr kompliziert wird. Umfassende Forschungsprojekte wurden der Entwicklung effektiver und effizienter Routing Protokolle für MANETs gewidmet. Im folgenden werden die Voraussetzungen for effizientes Routing in MANETs und die grundlegenden Routingstrategien erläutert und einige ausgewählte Realisierungen von Routing Protokollen vorgestellt. Mobile Geräte benötigen portable Energiequellen (üblicherweise Batterien oder Akkus), ihnen steht daher nur eine begrenzte Betriebsenergie zur Verfügung. Daraus resultieren die ersten beiden Anforderungen an MANETs: die einzelnen Geräte sollen mit geringer Sendeleistung auskommen und möglichst nur dann senden, wenn Nutzdaten zu einem anderen Gerät übertragen werden sollen. Ein weiterer Punkt ist die sich dynamische ändernde Anordnung der Endgeräte zueinander aufgrund der Mobilität dieser, besonders das Entstehen und Zerbrechen von Links (Zwei Geräte die in direkter Sendereichweite zueinander sind, verbindet ein Link) darf das Netzwerk nicht auseinander brechen lassen. Und letztendlich zeichnen sich MANETs wie alle Funknetzwerke durch start schwankende Durchsatzraten aus, abhängig von der Anzahl der Sender die gleichzeitig Daten übertragen wollen, Funkschatten oder atmosphärischen Störungen. All dies muss ein Routing Protokoll berücksichtigen um zuverlässiges und ökonomisches Routing zu ermöglichen.

3.2 Routing im MANET

Routing in mobilen Netzwerken stellt aufgrund deren charakteristischen Eigenschaften ganz besondere Anforderungen an die Algoritmen und Protokolle bein Routing. Es müssen Pfade vom Sender zum Empfänger gefunden werden ohne die Energiequellen der einzelnen Geräte übermäßig zu belasten. Trotzdem soll die Verbindung robust sein, nach Möglichkeit eine gewisse quality-of-service garantieren (Durchsatzrate, Antwortzeiten, etc.) und letztendlich möglichst schnell verfügbar sein wenn Daten verschickt werden sollen. Je nachdem welcher Anforderung man höchste Priorität gibt stehen verschiedene Ansätze zur Verfügung. Grundsätzlich gibt es zwei Arten von Routing Protokollen in mobilen Ad-hoc Netzwerken, die sich klassifizieren lassen nach: on-demand (reaktives) oder table-driven (proaktives) Routing. Die beiden Verfahren unterscheiden sich in der Art und Weise, wie und wann Routen gefunden werden. Bei table-driven Routing Protokollen wird von jedem Endgerät eine Liste mit aktuellen Informationen zur Topologie im gesamten Netzwerk verwaltet, wogegen beim on-demand Routing die Pfade erst dann erstellt werden, wenn zwei Endgeräte tatsächlich Daten untereinander austauschen wollen. In den nächsten beiden Abschnitten werden verschiedene table-driven und on-demand Protokolle vorgestellt.

3.2.1 Table-Driven (proaktives) Routing

Beim proaktiven Routing verwaltet jedes Endgerät eine oder mehrere Listen mit Routing Informationen zu jedem anderen Netzknoten. Diese Listen werden ständig aktualisiert um ein konsistentes möglichst aktuelles Modell des Netzwerks aufrecht zu erhalten. Wenn sich die Netztopologie ändert, werden Updateinformationen durch das gesamte Netz gesendet, damit die Konsistenz und Qualität der Routing Informationen der einzellnen Endgeräte gewahrt bleibt. Table-driven Routing Protokolle unterscheiden sich in der Art und Weise in der die Topologieänderungen im Netzwerk bekannt gemacht werden und in der Anzahl der Routingtabellen. Später werden einige existierende table-driven Routing Protokolle für Ad-hoc Netzwerke vorgestellt.

3.2.2 On-demand (reaktives) Routing

Diese Protokolle verfolgen den "bequemeren" Ansatz beim Routing. Im Gegensatz zum proaktiven Routing werden keine aktuellen Pfadinformationen von jedem Endgerät bereitgehalten und gepflegt, sondern Routen gerade erst dann gesucht, wenn sie tatsächlich gebraucht werden. Will ein Teilnehmer Daten an einen anderen senden, so löst er einen Routenfindungsmechanismus aus, um einen günstigen Pfad zum Ziel aufzuspüren. Die Route bleibt gültig bis sie nicht mehr benötigt wird, oder das Ziel in direkte Sendereichweite rückt. Das Auffinden von Routen zu einem bestimmten Ziel erfolgt unter Anwendung von Flooding. Eine Routenanfrage, die Absender- und Zieladresse enthält wird durch das gesamte Netzwerk geleitet bis sie ihr Ziel erreicht und dieses eine Routenbestätigung mit dem bevorzugten Pfad zurücksendet.

3.2.3 Welcher Ansatz ist besser?

Beide Routingstrategien haben sowohl Vor- als auch Nachteile. So wird bei einem table-driven Routing Protokoll der Pfad von einem Netzknoten zu einem anderen stets sofort verfügbar sein, da jedes Gerät aus seinen eigenen aktuellen Informationen über die momentane Netztopologie sofort die beste Route berechnen kann, während bei on-demand Routing Protokollen erst ein Pfad gefunden werden muss, was durchaus einiges an Zeit in Anspruch nehmen kann. Auf der anderen Seite erzeugt das ständige Aktualisieren der Routing Listen beim proaktiven Routing Protokollen gegebenenfalls eine sehr hohe Netzlast, durch die der eigentliche Austausch von Nutzdaten eingeschränkt oder sogar blockiert wird (viel Overhead). Reaktive Routing Protokolle dagegen kommen mit sehr wenig Routingdaten aus. Auch können sich proaktive Routing Protokolle nur langsam an Veränderungen in der Topologie anpassen. On-demand Routing Protokolle werden dafür einen einmal gefundenen Pfad nicht mehr aktualisieren, auch wenn sich später bessere Möglichkeiten ergäben. Neue Forschungsprojekte versuchen eine Symbiose aus beiden Strategien in einem Protokoll. So sollen die Vorteile von proaktivem und reaktivem Routing vereint werden. Im Grunde kommt es immer auf die Beschaffenheit und Eigenarten des jeweiligen Netzweks an, auf die Anzahl der Knoten insgesamt, auf den Grad der Dynamik in der Topologie und die Homogenität der Netzlast, welche Routingstrategie die bessere Wahl ist.

3.2.4 Flooding

Flooding ist kein Routing Protokoll im eigentlichen Sinne, jedoch wichter Bestandteil von sowohl table-driven als auch on-demand Routing Protokollen. Der Mechanismus ist simpel und effektiv: Der Sender schickt ein Datenpaket an jedes andere Gerät in Sendereichweite. Diese schicken das Paket ebenfalls an alle Nachbarn und so erreicht das Paket irgendwann seinen eigentlichen Empfänger. Dabei wird das Paket von jedem Knoten nur einmal versendet, mehrfach empfangene Kopien werden verworfen, wodurch sichergestellt ist, dass der Vorgang selbständig terminiert. Dieses Verfahren ist äußerst robust und wahrscheinlich die einzige effektive Strategie in einer hochgradig dynamischen Topologie. Es ist jedoch aufgrund der Tatsache, dass viele Kopien ein und desselben Pakets gleichzeitig durch das Netzwerk geleitet werden und sowohl Bandbreite als auch die Energie der Geräte verschwenden ineffizient und um Nutzdaten auszutauschen nur für die kleinsten Netzwerke zu gebrauchen. Der Anteil der tatsächlich genutzten Bandbreite am gesamten Transfervolumen ist unverhältnismäßig gering. Flooding ist Grundvorraussetzung für on-demand Routing, da zum Aufspüren einer Route Pakete durch das gesamte Netz gesendet werden müssen, aber auch für table-driven Routing, um Aktualisierungen in der Netztopologie jedem Gerät bekanntzugeben. Den Gegenpol Zum Flooding bildet übrigens die Centralized Route Computation Strategie. Hier werden Topologieinformationen an einen zentralen Knoten weitergeleitet, der dann alleine für all Transfers die Routen berechnet. Auch dieses Verfahren eignet sich jedoch nur für sehr kleine Netze.

3.2.5 Link State oder Distance-Vector Routing

Routingstrategieren können nicht nur in der Hinsicht kategorisiert werden wann sie eine Route suchen, sondern auch wie und wie und wie dabei Daten ausgetauscht und in den Knoten gespeichert werden. Die Link-State Methode kommt dabei der in herkömmlichen verkabelten Netzen verwendeten zentralen Pfadberechnung sehr nahe. Jeder Knoten verwaltet ein Modell der Netztopologie mit der Anzahl der benötigten Hops von jedem Knoten zu jedem anderen. Um dieses Modell aktuell zu halten versendet jeder Knoten periodisch an alle anderen Knoten die Anzahl Hops die von ihm zu jedem Ziel benötigt werden. Erhält ein Knoten solch ein Update aktualisiert er gegebenenfalls sein Modell der Netztopologie und berechnet und speichert in einer Liste für jedes Ziel den Nachbarn über den die kürzeste Routen führt. Das Modell der Netztopologie muss bei allen Knoten konsistent bleiben, sonst kann es bei den Routen zur Bildung von Schleifen kommen, jedoch nur für die Zeit, die ein Paket braucht um einmal das Netzwerk zu durchqueren. Im Distance-Vector Verfahren kennt jeder Knoten seine direkten Nachbarn und für jedes Ziel x die minimale Anzahl Hops $D_x(i)$ um dieses über seinen Nachbarn i zu erreichen. Soll ein Paket verschickt werden wird es über einen Nachbarn i geleitet, für den $D_x(i)$ minimal ist. Die so Schritt für Schritt gewonne Reihenfolge an Hops entspricht tatsächlich der kürzesten Route, vorrausgesetzt alle Daten waren akkurat und die Netztopoligie hat sich nicht geändert. Um die Distanzvektoren aktuell zu halten, versendet jeder Knoten periodisch an alle Nachbarn für alle Ziele die minimale Anzahl Hops mit der sie von ihm aus zu erreichen sind. Das oben beschriebene Verfahren wir Distributed Bellman-Ford (DBF) Algorithmus genannt. Verglichen mit der Link-state Methode müssen die einzelnen Knoten weniger

Berechnungen durchführen und viel weniger Routingdaten speichern. Außerdem ist DBF einfacher zu implementieren. Das größte Problem beim DBF Algorithmus ist, dass sich langlebige Schleifen bilden können. Fast alle auf DBF aufbauenden Protokolle lösen dieses Problem jedoch.

3.2.6 Hierarchisch (cluster based) oder Flach

Weiterhin kann man zwischen flachen und hierarchischen Protokollen unterscheiden. Hierarchische Netze werden in so genannte Cluster aufgeteilt, das heißt jeder Knoten wird einem bestimmten Bereich zugeordnet und ein Knoten pro Bereich wird zum Cluster Head, dem einzigen Knoten, der Daten von außerhalb des Clusters empfangen, bzw. Daten an Knoten außerhalb des Clusters senden darf. Wird ein Gerät aktiviert setzt es seinen Status auf *unentschieden*, startet einen Timer und broadcastet eine Hello Nachricht an alle Nachbarn. Befindet sich darunter ein Cluster Head, so antwortet dieser sofort mit einer Hello Nachricht. Sobald das neue Gerät die Antwort auf sein Hello erhält setzt es seinen Status auf *member* und ist von nun an Mitglied im Cluster des Cluster Head, der geantwortet hat. Hat das neue Gerät nach einer bestimmten Zeit (Timeout) noch keine Rückmeldung erhalten prüft es ob sich überhaut andere Knoten in Empfangsreichweite befinden und falls dem so ist ruft es sich selbst zum neuen Cluster Head aus, ist jedoch kein weiteres Gerät in seiner Umgebung behält es seinen *unentschieden* Status und wiederholt die ganze Prozedur. Ein Knoten, der Mitglied (member) in einem Cluster ist, muss über alle anderen Knoten in seinem Cluster bescheid wissen. Mit ihnen kann er direkt kommunizieren. Will er jedoch eine Verbindung zu einem Knoten aufbauen, der nicht Mitglied in seinem Cluster ist muss er die Daten zuerst an seinen Cluster Head schicken, der sie dann zum anderen Cluster weiterleitet, und zwar an dessen Cluster Head. Von hier aus erreichen die Pakete nun endlich ihr eigentliches Ziel. Die Cluster Heads werden somit zu Routern über die Knoten aus verschiedenen Clustern kommunizieren. In flachen Protokollen sind alle Knoten gleichberechtigt. Sie bieten sich für kleinere Netze an, während hierarchische Protokolle in großen Netzen von Vorteil sind.

3.3 Routing Protokolle

3.3.1 Destination-Sequenced Distance-Vector Routing (DSDV)

DSDV ist ein flaches reaktives Routing-Protokoll, das auf dem Distributed Bellman-Ford Algorithmus aufbaut. Jeder Knoten verwaltet zwei Listen, die Informationen für Routen zu allen anderen bekannten Knoten enthalten. In der Next Hop Table wird dabei zu jedem möglichen Ziel jeweils der Nachbar über den die Route führt und in der Distance Table die Anzahl Hops dieser Route gespeichert. Schleifenfreiheit wird durch Sequenznummern gewährleistet. Mittels dieser Sequenznummern werden unterbrochene Routen von neuen Routen unterschieden. Um über die aktuelle Netztopologie informiert zu sein, sind regelmäßige Updates nötig, was zusätzlich Bandbreite belegt. Es gibt zwei Arten von Aktualisierungsnachrichten. Ein vollständiges Update enthält alle Information, es werden

eventuell mehrere Datenpakete benötigt. Die zweite Möglichkeit ist das Mitteilen von Änderungen. Diese Updates passen in der Regel in ein einzelnes Datenpaket (network data unit). Zum Routing wird immer die Route mit der aktuellsten Sequenznummer genutzt. Falls zwei Updates mit der gleichen Sequenznummer einen Knoten erreichen, wird die Route mit der kleineren Metrik genutzt.

3.3.2 Clusterhead Gateway Switch Routing (CGSR)

CGSR ist ein hierarchisches Protokoll, welches auf DSDV aufsetzt. Es benutzt zusätzlich diverse Heuristiken, um das Routing festzulegen. Mehrere Knoten werden zu Clustern zusammengefasst, in denen je ein Knoten die Rolle des Cluster Head übernimmt. Bei sehr mobilen Netzwerken kommt es oft zu aufwendigen Wechseln der Cluster-heads. Deshalb wird ein Least Cluster Change Algorithmus verwendet, der die Cluster Heads nur dann ändert, wenn zwei Cluster Heads in direkten Kontakt geraten oder ein Knoten keinen Kontakt mehr zu irgendeinem Cluster Head bekommt. Das Routing geht von der Quelle über Ihren Cluster Head zu einem Gateway Knoten und von dort aus zu einem anderen Cluster Head. Dieser leitet das Paket dann weiter. Gatway Knoten sind Teilnehmer, die in Kontakt mit 2 oder mehreren Cluster Heads stehen. Für dieses Protokoll sind je Node zwei Listen notwenig, die cluster member table und die routing table.

3.3.3 Global State Routing (GSR)

GSR ist DSDV sehr ähnlich. Es verwendet den Link-State Ansatz. Bei GSR hat jeder Knoten eine Neighbor list, Topology table, Next Hop table und eine Distance table. Die Neighbor list eines Knotens enthällt all seine Nachbarn (in diesem Fall sind Nachbarn alle Knoten an die direkt Daten gesendet werden, auch wenn von ihnen keine Daten empfangen werden können). Die Topology table enthält für jedes Ziel die Link-State-Informationen, die von diesem Ziel gemeldet wurden und einen Timestamp dieser Informationen. in der Next Hop table steht für jedes Ziel an welchen Nachbarn die Daten weitergeleitet werden müssen, und die Distance table enthält Angaben über die Minimalanzahl Hops zu jedem Ziel. Updatenachrichten werden erzeugt sobald ein Link bricht oder neu Entsteht. Wenn ein Knoten eine Updatenachricht erhält, aktualisiert er seine Topology table falls die Sequenznummer dieser Nachricht neuer ist als die der in der Liste gespeicherten Link-State-Informationen des betreffenden Ziels. Danach rekonstruiert der Knoten seine Nexth Hop table und sendet seine Link-State-Informationen an seine Nachbarn.

3.3.4 Fisheye State Routing (FSR)

FSR ist eine Weiterentwicklung von GSR. Der in GSR durch Aktualisierungen der Listen entstehende Overhead verbraucht zu viel Bandbreite. Deshalb werden in FSR nicht mehr alle Updateinformationen an jeden Knoten gesendet. Stattdessen erhalten nähere Knoten häufiger Aktualisierungsnachrichten and weiter entfernte, wodurch das Datenaufkommen

durch Updatenachrichten global erheblich verringert wird. Abbildung x zeigt den Fisheye-Scope eines Knotens. Der Fisheye-Scope wird definiert als Menge aller knoten die mit einer vorgegeben Anzahl von Hops zu erreichen sind. Ein Knoten hält akurate Informationen über alle anderen Knoten in seinem Fisheye-Scope bereit, über alle außerhalb liegenden Ziele jedoch nur ungenaue oder alte Angaben. Wenn ein Paket an einen Knoten jenseits des Fisheye-Scopes geschickt werden soll ist die Route zunächst ungenau, doch je weiter sich das Paket seinem Ziel nähert, umso genauer wird die Route, da dem Ziel näher gelegene Knoten aktuellere Informationen über selbiges besitzen. Letzendlich erreicht das Paket einen Knoten in dessen Fisheye-Scope das Ziel liegt und wird korrekt an dieses weitergeleitet. FSR eigent sich besonders in größeren Ad-hoc Netzwerken.

3.3.5 Ad-hoc On-Demand Distance Vector Routing (AODV)

AODV ist ein flaches Routing-Protokoll, welches auf dem DSDV-Algorithmus beruht. Es minimiert aber die Anzahl der benötigten Broadcasts, indem die Routen nur bei Bedarf bestimmt werden (reaktives On-Demand Verfahren). Es werden also keine kompletten Routing-Tabellen erstellt. Nodes, die sich nicht im Pfad befinden, aktualisieren auch nicht ihre Routing-Informationen und tauschen auch keine Routing-Informationen aus. Wenn ein Knoten Daten senden will, ohne die Route zum Ziel zu kennen, wird ein Routenfindungsprozess gestartet. Dabei werden route request (RREQ) Nachrichten geflutet, bis das Ziel oder ein Knoten mit einer Route zum Ziel erreicht worden ist. Dieser Knoten schickt in Richtung Vorgänger eine route reply (RREP) Nachricht zurück. Dieser schickt dieses zu seinem Vorgänger und speichert die Route zum Zielhost. Bei AODV werden Sequenznummern eingesetzt. Jeder Knoten erhöht seine Sequenznummer mit jedem gesendetem RREQ. Zusammen mit der Adresse des Knotens sind die RREQs damit einzigartig. Falls ein RREQ doppelt ankommt, wird er verworfen. Wenn sich Teilnehmer bewegen, ist eine Neuberechnung der Route notwendig. Bewegt sich die Quelle, kann sie die Berechnung neu anstoßen. Wenn eine andere Station sich bewegt, erkennt das die Vorgängerstation und schickt eine link failure message zu ihrem Vorgänger. Diese wird bis zur Quelle weitergereicht, die eine Neuberechnung veranlassen kann.

3.3.6 Dynamic Source Routing (DSR)

DSR ist ein flaches reaktives Protokoll. Ein Knoten merkt sich in seinem Routen-Cache alle gültigen Pfade. Der Knoten aktualisiert diese Einträge sobald er von neuen Routen Kenntnis erlangt. Doe beiden Hauptphasen des Protokols sind: Routenfindung (route discovery) und Routenpflege (route maintenance). Wenn ein Knoten ein Datenpaket zu einem bestimmten Ziel senden möchte, überprüft es seinen Routen-Cache ob eine solche Route bereitsbekannt ist. Wenn dort ein gültiger Pfad zum Ziel existiert, wir diese Route zum versenden des Datenpakets benutzt. Falls der Knoten jedoch keine solche Route kennt wird der route discovery process gestartet, indem ein Routenanfragepaket geflutet wird. Das Routenanfragepaket enthält die Adressen von Sender und Empfänger und eine einzigartige Identifikationsnummer. Jeder Knoten der somit erreicht wird prüft of sich in seinem Routen-Cache ein Pfad zum gesuchten Ziel befindet. Ist dies nicht der Fall, so

vermerkt er seine eigene Adresse in dem Paket und leitet es an alle seine Nachbarn weiter. Um die Anzahl der Routenanfragepakete zu verringern werden nur solche weiterverarbeitet, die von einem Knoten nicht schon einmal gesehen wurden und somit seine Adresse nicht enthalten. Eine Routenantwort wird erzeugt wenn entweder ein Routenanfragepaket das eigentliche Ziel erreicht oder einen Knoten der einen Pfad zum Ziel kennt. Ein Routdenanfragepaket, das auf einen solchen Knoten trifft enthält schon die Reihenfolge der Hops vom Sender bis zu diesem Knoten.

3.4 Fazit

MANETs ermöglichen neue aufregende Anwendungen, beinhalten jedoch auch eine Menge technischer Herausforderungen. In dieser Arbeit wurde ein kurzer Einblick in die Technologie der drahtlosen Ad-hoc Netzwerke gegeben und besonders auf die Problematik des Routing in ebensolchen eingegangen. Hierfür gibt es verschiedene Strategien, die jeweils ihre Vor- und Nachteile haben. Es kommt immer auf die Eigenschaften des Netzwerks an welche Strategie gut funktionert und welche schlecht. Bei Routingprotokollen entsteht noch viel Entwicklungsbedarf, Es wird versucht die Vorteile verschiedener Routingstrategien in einem Protokoll zu vereinigen.

3.5 Abkürzungen

AODV	Ad-hoc on-demand Distance Vector
DSR	Dynamic Source Routing
GSR	Global State Routing
FST	Fisheye State Routing
MANET	Mobile Ad-hoc Network
CGSR	Cluster Gateway Switch Routing
DSDV	Destination-Sequenced Distance-Vector
RREQ	Route Request
RREP	Route Reply
DBF	Distributed Bellman-Ford

Literaturverzeichnis

[1] Padmini Misra: *Routing Protocols for Ad Hoc Mobile Wireless Networks*
http://www.cis.ohio-state.edu/~jain/cis788-99/adhoc_routing/index.html

[2] J. Macker, S. Corson: *Mobile Ad-Hoc Networks (manet)*
http://www.ietf.org/html.charters/manet-charter.html

[3] NIST: *Project on Wireless Ad-hoc Networks*
http://w3.antd.nist.gov/wctg/manet/

[4] Yuh-Shyan Chen, Yu-Chee Tseng, Jang-Ping Sheu, and Po-Hsuen Kuo:
On-Demand, Link-State, Multi-Path QoS Routing in a Wireless Ad-Hoc Network
http://www.ing.unipi.it/ew2002/proceedings/006.pdf

[5] Tim Kaldewey: *Simulation von Mobile Ad Hoc Networks*
http://www.tk.informatik.tu-darmstadt.de/Lehre/ss02/semtele/Folien/
/Tim%20Kaldewey%20-%20Simulation%20von%20Ad%20hoc%20Netzwerken.pdf

[6] Marco Günther: *Einführung in Ad-Hoc-Netzwerke*
http://archiv.tu-chemnitz.de/pub/2002/0038/data/vortrag_hrz_pp.pdf

[7] Andrea J. Goldsmith, Stephen B. Wicher: *Design Challanges for Energy-Constraint Ad hoc Wireless Networks*
http://www.comsoc.org/livepubs/pci/Private/2002/Aug/wicker.html

Kapitel 4

AAA and Extensions for Wireless Services

Markus Flingelli

Die mobile Nutzung des Internets gewinnt immer größere Bedeutung. Damit Provider mobilen Zugang zum Internet gewähren können, müssen sie in der Lage sein, am Nutzer eine Authentifizierung, Authorisation und Abrechnung (AAA) vorzunehmen. Dies muß sicher und transparent für den Kunden erfolgen. Einen möglichen Ansatz hierfür bietet die AAA-Architektur. Deshalb beschäftigt sich diese Arbeit mit der AAA-Architektur der IRTF. Es werden deren Mechanismen vorgestellt und exemplarisch ein Protokoll näher betrachtet, dabei wird aber deutlich, daß diese Architektur Probleme und Schwächen aufweist. Diese Probleme können durch die sogenannte generische A^x-Architektur behoben werden. Die generische A^x-Architektur wird am Beipiel AAAC erläutert. Im folgenden werden die Grundlagen für mobiles AAA auf Basis von Mobile IP erklärt und an einem möglichen mobilen Szenario verdeutlicht.

Inhaltsverzeichnis

4.1 Einleitung und Motivation

In IP-basierten Netzwerken schreitet die Integration von festen Kommunikationsarchitekturen und mobilen Systemen immer weiter voran. Vor allem seit das Internet sowohl eine öffentliche als auch eine private Kommunikationsplattform für eine Vielzahl von Diensten geworden ist, welche von Firmen, akademischen Einrichtungen und von Privatpersonen benutzt werden. Dies macht es aber für Anbieter notwendig, ihre Angebote im weiten Feld der Dienste einzugrenzen. Damit sich die Investitionen für sie lohnen, brauchen sie für kommerzielle Services, die Möglichkeit, die Nutzer zu authentifizieren, zu authorisieren und abzurechnen. Außerdem gewinnen Sicherheitsaspekte immer mehr an Wichtigkeit, da die Mobilität von Nutzern und Geräten immer mehr an Beliebtheit gewinnt.

Zu den ökonomischen und marktgesteuerten Aspekten, welche die Notwendigkeit von AAA-Systemen (Authentifikation, Authorisation, Abrechnung) motivieren, erfordert die Kommunikationstechnologie, welche sowohl Umwelt als auch technische Basis für AAA-Systeme ist, eine gründliche Erforschung, um eine Weiterentwicklung für zukünftige AAA-Dienste zu gewährleisten. Die Verschiedenartigkeit in Netzwerkkomponenten, ihre Funktionalitäten, die Signalisierungsprotokolle für Ende-zu-Ende Qualitätszusicherung und die Dienstbereitstellung bestimmen die Hauptmerkmale der gegenwärtigen Internettechnologie. Das Netzwerk der Zukunft wird ein Multi-Dienste Internet, bestehend aus vielen kooperierenden Domänen, die Zugangsdienste, Transportdienste, Applikationsdienste und Inhalte anbieten. Das Hauptproblem ist durch weit erweiterte Zugangskontrolle gegeben, welches zur Zeit aus Authentifizierung, Authorisierung und Abrechnung besteht und dadurch komplexer geworden ist. Es ist notwendig, Zugang zu IP-Netzwerken zu authorisieren und Dienste mit QoS-Garantien zu transportieren. Die Entscheidungen für die Authorisation können durch die Vertrauenswürdigkeit, technische (z. B. die Bandbreite) oder finanzielle Aspekte (z. B. Kreditwürdigkeit) geprägt sein.

Die meisten Dienste kosten Geld, z. B. die Telekommunikationsverbindung und der IP-Zugang hängen von der Verbindungszeit oder dem Datenvolumen, der Transport von QoS-Parametern und Inhaltsdienste von der Art des Inhalts ab. Deshalb umfaßt die Abrechnung mehr als das Messen der Zeit, die ein Nutzer mit dem IP-Netzwerk verbunden war. Um allen Anforderungen zu genügen und für die Zukunft gerüstet zu sein, sind Erweiterungen an den AAA-Systemen notwendig und genererische AAA-Dienste werden benötigt [6].

4.1.1 Begriffserklärungen

Um dem Leser die wesentlichen Begrifflichkeiten eindeutig näher zu bringen, werden im folgenden die wichtigsten Begriffe im Zusammenhang mit AAA (in alphabetischer Reihenfolge erklärt):

- **AAA** ist die Abkürzung für *Authentication (Authentifizierung)*, *Authorization (Authorisierung)* und *Accounting (Abrechnung)*.

- Eine **Abrechnung** ist die Zusammenstellung von Informationen, die die Servicenutzung des Kunden dokumentieren, wobei die technisch gemessenen Daten mittels einer Abbildung in monetäre Einheiten umgerechnet werden. Diese kann abhängen vom Verbrauch von Ressourcen oder den ausgehandelten Ressourcenwerten (beispielsweise feste Bandbreite) [6].

- Die **Authentifizierung** ist die Verifikation der Identität eines Subjekts, welches Aktionen ausführt. Das Subjekt kann sowohl ein Servicenutzer als auch ein Serviceanbieter sein [6].

- Die **Authorisierung** ist das Überprüfen, ob ein bestimmtes Subjekt die notwendigen Rechte besitzt, eine Aktion an einem Objekt ausführen zu dürfen [6].

- Ein **Dienst** (Service) definiert eine Menge von Leistungen, die ein Serviceanbieter einem Kunden anbietet. Die notwendige Hardwareausstattung, die vom Anbieter (Provider) kontrolliert wird, stellt diese Dienste dem Nutzer zur Verfügung. Die Dienste reichen vom Zugang zum Netz, über Transportdienste, die es ermöglichen IP-Pakete über das Internet zu versenden, bis hin zu Anwendungs- und Inhaltdiensten [6].

- **Quality of Service (QoS)** bedeutet die Festlegung von Parametern für die Qualität der Übertragung eines Dienstes. In dieser Arbeit handelt es ich im speziellen Fall um die Qualität der Übertragung von Kommunikationsdiensten, eben der Übertragung von Informationen mittels Protokollen. Die Qualität wird z. B. durch die Übertragungsrate, den Jitter und die Fehlerrate festgelegt. Die Parameter der Übertragung sind Bestandteil des Vertrags zwischen Nutzer und Anbieter [11].

4.2 Das traditionelle AAA

Im folgenden wird ein Überblick über traditionelle AAA-Mechanismen und Erweiterungen gegeben.

4.2.1 AAA Mechanismen

Authentifizierung bezeichnet die Überprüfung der Identität eines Subjekts. Die Authentifizierungsmechanismen können so klassifiziert werden:

- Wissensbasierte Authentifizierung nutzt geteilte Geheimnisse, wie zum Beispiel PINs (Personal Identification Numbers) und Paßwörter.
- Kryptographiebasierte Authentifizierung beinhaltet digitale Signaturen, Challenge-Response Mechanismen und Nachrichtenauthentifizierungscodes. Der Nutzer besitzt einen privaten Schlüssel als Charakteristik.
- Authentifizierung, die auf biometrischen Eigenschaften basiert, nutzt Informationen des Subjekts, wie beispielsweise den Fingerabdruck, die Stimme oder die Charakteristik der Augen (beispielsweise ein Iris-Scan).

- Authentifizierung basierend auf sicheren Token funktioniert zum Beispiel so, daß das Subjekt im Besitz einer Smartcard ist. Dieser Mechanismus ist meistens mit dem kryptographischen Mechanismus kombiniert, um die Information vom Token auf das Lesegerät zu übertragen.
- Nicht weit verbreitet sind digitalisierte Unterschriften, welche digitale Bilder handgeschriebener Unterschriften und Charakteristika der Unterschriften sind (z. B. die Ausrichtung der Schrift, den Druck, die Geschwindigkeit und andere Attribute der handschriftlichen Unterschrift).

Die Authentifizierungsstrategie beschreibt, ob Authentifizierung notwendig ist und welche Mechanismen und Algorithmen unter welchen Einschränkungen genutzt werden sollen [5].

Authorisierung ist definiert als die Überprüfung, ob ein Subjekt eine Aktion an einem Objekt ausführen darf oder nicht. Es gibt zwei Hauptklassen der Authorisierungsmechanismen:

- Authentifizierungsbasierte Mechanismen erfordern die Authentifizierung des Subjekts als Vorbedingung für die Authorisierung. Die Information über die Authorisierungsentscheidung ist in einem Objektesystem gespeichert, ähnlich der Zugangskontrolliste des Betriebsystems in der Form „Nutzer S darf eine Aktion A an dem Objekt O vornehmen“.
- Berechtigungsbasierte Mechanismen nutzen vertrauenswürdige Informationen, welche beim Subjekt während des Authorisierungsprozesses bereitgehalten werden.

Authorisierungsstrategien definieren die Aktionen, die ein Subjekt an einem Objekt vornehmen darf. Es gibt eine positive (erlauben) und eine negative (verbieten) Authorisierungsstrategie. Formal kann dies folgendermaßen definiert werden:

O ist eine Menge von Objekten,
S ist eine Menge von Subjekten und
A ist eine Menge von Aktionstypen;
die Authorisierungsregel ist ein Tripel (s, o, a) mit
$a \in S$, $o \in O$, $a \in A$
$f : S \times 0 \times A \rightarrow \{True, False\}$

Wenn die Bedingung „if f(s,o,a) = True“ erfüllt ist, ist die Authorisierungsentscheidung positiv. Wenn nicht, erhält das Subjekt keine Authorisierung. Die ursprüngliche Definition ist erweitert worden, um Einschränkungen in der Strategie zu ermöglichen. Diese Einschränkungen können aktuelle Objektzustände oder universelle Bedingungen sein. Dies kennzeichnet, daß die Strategieentscheidung von Werten der Objektattribute oder universellen Bedingungen, wie beispielsweise der Zeit, abhängen können. Bei der authentifizierungsbasierten Authorisierung gibt es eine große Ähnlichkeit zwischen den Strategien und den Mechanismen. Bei der berechtigungsbasierten Authorisierung hat die Berechtigung eine ähnliche Form wie die Strategie, wobei die Menge der Objekte nur aus einem Element, nämlich dem Nutzer besteht [5].

Ein Abrechnungssystem hat zwei Hauptaufgaben. Zum einen das Sammeln der Daten durch das Meßsystem und das Versenden der Rechnung an die Kunden. Aus diesem Grund gibt es für das Sammeln und das Versenden unterschiedliche Strategien.

Der Nutzer des Datensatzes kann, abhängig von seinem Ziel, eine Abrechnungsstrategie wählen, welche Information er zu welcher Zeit vom Abrechnungssystem braucht. Diese Strategie kann durch interne Vorgänge ereignisgesteuert sein oder durch ein externes Ereignis (wie das Ende des Monats ist erreicht) oder das Rechnungserstellungssystem greift auf den Abrechnungsdatensatz zu. Die Strategie kann auch pflichtgesteuert sein: „Wenn ein neues Preisschema eingeführt wird, müssen neue Abrechnungsinformationen gesammelt werden."

Eine Aufzeichnungsstrategie beschreibt, welche Information durch das Meßsystem aufgezeichnet und zum Abrechnungssystem transportiert werden muß. Diese Strategien sind durch Signalisierungsereignisse gesteuert, außer wenn statische Meßgeräte benutzt werden, welche die Daten für alle Datenströme in einer festen Körnigkeit sammeln. Offensichtlich beeinflußt die Abrechnungsstrategie die Meßstrategien [5].

4.2.2 AAA Protokolle

Im folgenden wird auf Protokolle eingegangen, die in der IETF (Internet Engineering Task Force)-AAA-Arbeitsgruppe diskutiert werden. Diese Arbeitsgruppe beschäftigt sich hauptsächlich mit den Anforderungen, die AAA-Protokolle erfüllen müssen [1]. Der Schwerpunkt dieser Protokolle liegt auf dem Zugang zum Netzwerk. Authentifizierungsprotokolle für den Netzwerkzugang operieren im allgemeinen zwischen Authentifizierungsserver und Servicenutzer. Das Netzwerkzugangsgerät wie ein Einwahlserver agieren als Relais zum Authentifizierungsserver. Einige der Protokolle, die im Zusammenhang mit AAA von Bedeutung sind, sind folgende:

- RADIUS (Remote Authentication Dial In User Service)
- Diameter
- COPS (Common Open Policy Server)
- SNMP (Simple Network Management Protocol)

Da das Protokoll RADIUS heutzutage die größte Verbreitung hat und auch in vielfachen Implementationen verfügbar ist, wird es an dieser Stelle exemplarisch behandelt. Die übrigen Protokolle sind in [5] beschrieben. Das RADIUS Protokoll wurde entworfen, um Authentifizierungs-, Authorisierungs- und Konfigurationsdaten zwischen dem Netzwerkzugangsserver, welcher einen RADIUS-Client beendet und einem bestimmten RADIUS-Server zu transferieren. Dieser RADIUS-Server hält die Information zum Authentifizieren und Authorisieren des Nutzers. Außerdem kann er anderen RADIUS-Servern als Client dienen. Ursprünglich wurde RADIUS entwickelt, um Einwahlverbindungen zu unterstützen, heutzutage wird es zu verschiedenen Zwecken eingesetzt. RADIUS verwendet verschiedene Authentifizierungsprotokolle, beipsielsweise PAP (Passwort Authentication

Protocol) und CHAP (PPP Challenge Handshake Authentication Protocol) . Es gibt Erweiterungen wie Abrechnungsinformationen, die beispielsweise Start und Stop zu einem RADIUS-Abrechnungsserver übertragen. Es gibt auch einige Mängel, die es nicht erlauben, ein geeignetes generisches AAA-Protokoll zu sein. Diese sind unter anderem, die begrenzte Größe der Attribute, die limitierte Sitzungskontrolle, eine geringe Fehlertoleranz und die fehlende Unterstützung einer Ende-zu-Ende Sicherheit [6].

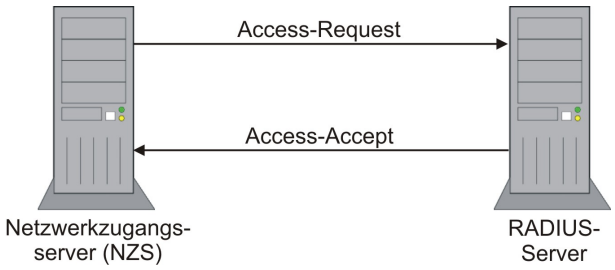

Abbildung 4.1: Authentifizierung eines NZS mit Telnet an einem RADIUS-Server

Das Bild 4.1 zeigt die Authentifizierung eines Netzwerkzugangsserver (NZS) mit Telnet an einem RADIUS-Server . Der NZS schickt ein UDP-Paket(User Datagram Protocol [4]) an den RADIUS-Server (Access-Request). Dieses Paket enthält unter anderem den Benutzernamen und ein 16 Byte langes Sessionpaßwort. Dieses Sessionpaßwort wird erzeugt, indem zuerst auf die Konkatenation des jeden der beiden Rechnern bekannten Paßwortes und der Zufallszahlenfolge der Hashalgorithmus MD5 (die Beschreibung des Algorithmus kann in [8] nachgelesen werden) angewendet wird (der Hashalgorithmus MD5 bildet eine beliebig lange Folge von Bits auf einen 128 Bit langen Wert ab. Es wird für unmöglich gehalten, daß zwei unterschiedliche Zeichenfolgen den gleichen 128 Bit langen Wert aufweisen [8]). Das Ergebnis dieser Operation wird XOR mit der Zufallszahl verknüpft. Eine Verschlüsselung des Paßwortes ist notwendig, da bei Verbindungen von Telnet die Daten unverschlüsselt übertragen werden. Der RADIUS-Server kann anhand des übertragenen Benutzernamens und des Sessionpaßwortes feststellen, ob der anfragende Server das geheime Paßwort zur Sessionpaßwortberechnung benutzt hat. Wenn der RADIUS-Server eine Übereinstimmung feststellt, schickt er in einem UDP-Paket ein Access-Accept an den NZS [7].

4.2.3 Die IRTF AAA-Architektur

Die IRTF (Internet Research Task Force)-Forschungsgruppe *Authentication Authorisation Accounting ARCHitecture Research Group (AAAARCH)* strebt eine Definition einer Architektur und eines Modells für AAA an. Es wird ein strategiebasierter Ansatz gewählt, wobei sowohl Mechanismen als auch Protokolle diskutiert werden. Die Ergebnisse sollen konform sein zu der Arbeit der IETF Strategie-Framework-Gruppe. Die IETF definiert eine Strategie als eine Gruppierung von Strategieregeln, bestehend aus Bedingungen und Strategieaktionen [6], [11].

AAA-Komponenten

Die regelbasierte Maschine (Rule-based Engine RBM) ist eine zentrale Komponente, welche die Strategiebedingungen, die für eine Strategieentscheidung notwendig sind, auswertet und eine Strategieaktion je nach Entscheidung ausführt. Die Strategien werden in so genannten Strategiebehältern (Policy Repositories) aufbewahrt. In der IRTF Forschungsgruppe gilt der Hauptaugenmerk den Authorisierungsstragien für Serviceanfragen und den dazugehörenden Abrechnungsstrategien. Die Strategieaktionen müssen von verschiedenen Komponenten je nach Art der Aktion durchgesetzt werden. Die meisten Aktionen, die zu einer Serviceanfrage gehören, werden von der Serviceausstattung (Service Equipment) erledigt, welche alle Arten von Netzwerkelementen umfaßt. Andere Aktionen wie Hilfsdienste (im Speziellen die Abrechnung) werden von den AAA-Servern getrennt oder integriert in die Serviceausstattung erledigt [5].

AAA-Services

Die Grundlage dieser AAA-Architektur basiert auf der Annahme einer Multi-Domain Internet-Topologie. In jeder administrativen Domäne muß sich mindestens ein AAA-Server befinden. Verteilte AAA-Server bieten den Nutzern Authentifizierungs-, Authorisierungs- und Abrechnungsdienste an. Der Authorisierungsgsdienst ist als Prozeß definiert, der die Authorisierungsentscheidung fällt. Er gewährt oder lehnt die Nutzeranfrage für Dienste ab und stellt in einer authorisierten Sitzung die Serviceausstattung und das Logging zur Verfügung. Die Authentifizierung des Nutzers ist Bestandteil des Authorisierungsprozesses und die Authentifizierungsinformation wird in der Authorisierungsanfrage übermittelt. Abrechungsdienste erfassen relevante Abrechungsinformation entsprechend der Authorisierungsentscheidung und den Ressourcenverbrauch der Sitzung.

Sichere und vertrauenswürdige Beziehungen sind zwischen verschiedenen AAA-Servern notwendig, um AAA-Dienste anbieten zu können. Der Nutzer baut mit einem Vertrag ein Vertrauensverhältnis mit dem gewünschten Service Provider auf, seiner Nutzerheimorganisation (NHO). Diese NHO betreibt einen AAA-Server, genauso wie die Fremdorganisation (FO), von welcher der Nutzer einen Dienst in Anspruch nehmen will. Die FO kann einem Nutzer trauen, wenn die Kette der Vertrauenswürdigkeit zwischen dem weiterleitenden Proxy-AAA-Server, dem Nutzer und der NHO aufgelöst werden kann. Deswegen ist die Authentifizierung zwischen Peer-AAA-Servern Teil von ihren Diensten [5].

Die AAA-Architektur

Die im Bild 4.2 dargestellten Komponenten sind Teil der AAA-Architektur. Die RBM befindet sich in einem AAA-Server. Der AAA-Server empfängt von der Serviceausstattung über ein applikationsspezifisches Modul (ASM) oder von anderen AAA-Servern eine Serviceanfrage. Einerseits wird eine empfangene Serviceanfrage innerhalb des AAA-Servers auf in den Strategiebehältern befindlichen Strategien hin untersucht. Es ist unter Umständen notwendig, andere AAA-Server oder die Serviceausstattung zu Rate zu ziehen, um die Strategiebedingungen auswerten zu können. Dies wird zuerst beim Senden der

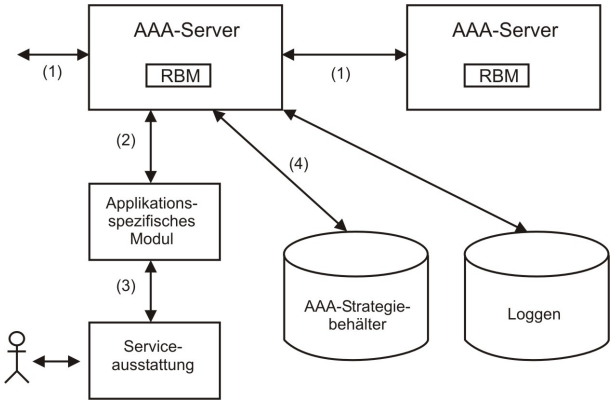

Abbildung 4.2: AAA-Architektur und Schnittstellen [6]

Anfrage zu anderen AAA-Servern und zum zweiten mal durch das ASM erledigt. ASM werden zusätzlich benötigt, um die Strategieaktionen durchzusetzen. Außerdem konfigurieren sie die Serviceausstattung, um einen Dienst bereitstellen zu können. Andererseits werden Strategieaktionen vom AAA-Server selbst in Anspruch genommen. Er speichert den Sitzungszustand, die für die Abrechnung notwendigen Datensätze und loggt die Aktionen.

Die Protokolle, die in dieser Architektur verwendet werden, beinhalten

- (1) ein spezielles AAA-Protokoll, welches vermutlich von der Arbeitsgruppe der IRTF standardisiert wird,
- (2) ein spezielles API (Application Programming Interface) oder wiederum ein AAA-Protokoll,
- (3) abhängig von der Implementation des Strategiebehälters entweder ein API oder das LDAP (Leight-weight Directory Access Protocol) und
- (4) ein anwendungsspezifisches Protokoll [5].

Probleme und Schwächen

Zu AAA wurde für eine ausgewählte Anzahl von Mechanismen und Algorithmen, Vorschläge für Protokolle und Erweiterungen gemacht. Aber oft wurden die Arbeiten isoliert für gekürzte Aufgaben und begrenzte Szenarios durchgeführt, beispielsweise die Verbindungskontrolle durch einen Netzwerkzugangsserver oder Inhaltsübermittlungskontrolle durch ein Rechnungserstellungssystem. Die Erweiterung des existierenden AAA-Systems zieht besonders die Implementation von mobilen Szenarien und das Roaming in Betracht, damit neue integrierte Anforderungen unterstützt werden können. Diese Anforderungen beruhen auf Protokollen wie beispielsweise RADIUS. Diese Erweiterung wirft Probleme auf, da diese abhängig ist von unteren Technologien wie IPv6 oder Mobile IP. Zur Lösung dieser Probleme wurde das Moby Dick (Mobility and Differentiated Services in a Future IP Network) Projekt in das Leben gerufen, um diese Probleme zu lösen. Im Abschnitt 4.3.5 wird auf Moby Dick näher eingegangen.

Die IRTF AAA-Architektur versucht, diese Einschränkungen durch das Bauen von generischen Servern und applikationsspezifischen Modulen zu lösen. Soweit die Vorschläge bis jetzt diskutiert sind, können nicht alle Probleme in Bezug auf AAA-Dienste gelöst werden.

- Die Funktionen der Strategieentscheidung und der Strategieumsetzung sind nicht klar getrennt. Ein AAA-Server macht die Strategieentscheidung bei der Authorisation, aber er erzwingt auch die Abrechnungsstrategie durch die Durchführung der Abrechnungsaufgaben.
- Die Erweiterbarkeit von Funktionen jenseits von AAA, wie die Verifikation der in Rechnung gestellten Kosten ist kompliziert, da die Komponenten nicht generisch definiert sind. Viele Durchführungsfunktionen sind im AAA-Server selbst oder im ASM angesiedelt.
- Die Entwicklung von AAA-Diensten bleibt schwierig, da diese für den Transport und die Aufrechterhaltung der Verbindung entwickelt wurden.
- Die Funktionalität der ASM wurde nicht vollständig definiert. Es agiert als eine Schnittstelle für diese Aufgaben, welche dem AAA-Server nicht generisch zugewiesen werden können.
- QoS relevante, Hand-Over- oder Paging-Hilfsdienste wurden nicht in Betracht gezogen.

Aus diesem Grund wurde eine erweiterte Architektur vorgeschlagen. Das Ziel dieses Vorschlages ist es, A^x-Dienste möglichst generisch zu definieren und eine A^x-Architektur zu schaffen, die es Diensten ermöglicht, in der Unterstützung verschiedenster Nutzerdienste auf unterschiedlichen Ebenen in verschiedenen heterogenen Netzwerkkomponenten und Dienstprotokollen eingesetzt werden zu können [6].

4.3 Die generische A^x-Architektur

Die Voraussetzung für eine generische Architektur beinhaltet eine ausführliche Beschreibung der in A^x-Dienste involvierten Komponenten und die Identifikation von Interaktionsschemata zwischen ihnen. Die Arbeit an A^x wendet drei grundlegende Konzepte für ein Framework der A^x-Architektur an.

- Diensttrennung (erweiterter AAA-Standpunkt),
- Einteilung in Dienstebenen und
- Strategieparadigmen (Nutzen von existierenden Arbeiten).

4.3.1 Die Diensttrennung

Serviceprovider bieten Kunden Dienste an und müssen verteilte Systeme im Internet betreiben. Dieses Managementaufgabe beinhaltet die Konfiguration der Netzwerkgeräte und die Bereitstellung von Protokollmechanismen. Alle existierenden A^x-Funktionen sind Teil dieser Aufgaben und können als Provider für interne Dienste angesehen werden. Sie können von den Diensten getrennt werden, die sie Kunden anbieten. Deshalb werden die Nutzerdienste von den A^x-Diensten der entsprechenden Serviceausstattung in entsprechenden Phasen (siehe Abbildung 4.3) getrennt.

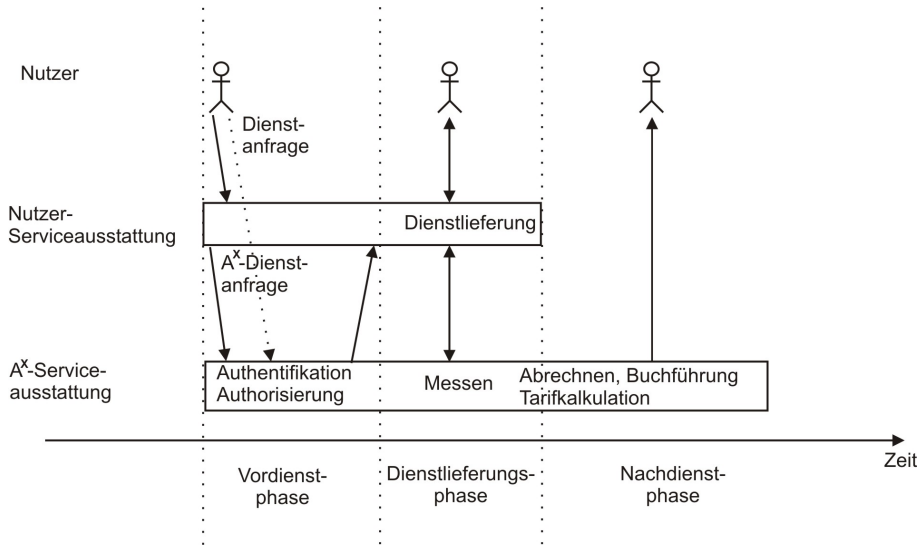

Abbildung 4.3: Dienstinteraktion[6]

Die Trennung der Dienste erlaubt die Definition einer gemeinsamen Schnittstelle für generische A^x-Dienste, welche unabhängig von Netzwerkgeräten oder Protokollen zur Anfrage und Gewährung von Nutzerdiensten ist [6].

4.3.2 Die Aufteilung der Dienstebenen in ein Internetdienstmodell

Das geschichtete Dienstmodell von Internetdiensten ist als Framework definiert, welches aus vier horizontalen Schichten besteht und nur einen eingeschränkten Blick auf den Netzwerkzugang und den Transport gewährt. Die unterste Schicht 1 betrifft die Internetverbindung, Schicht 2 den Transport, Schicht 3 die Anwendung und Schicht 4 den Inhalt. Neben der Einteilung in Schichten, wird eine vertikale Aufteilung in einen Singalisierungs- und Datenpfad, wie Tabelle 4.1 zeigt, vorgenommen.

Die horizontale Aufteilung definiert Dienstklassen mit ähnlichen Charakteristika und A^x-Anforderungen. Die vertikale Einteilung unterstützt die Identifikation, zu welchem Zeitpunkt Hilfsdienste notwendig sind. Wenn beispielsweise ein volumenbasiertes Schema angewendet wird, erfolgen die Authentifizierung und Authorisierung im Kontrollpfad, während die Abrechnung die benötigten Informationen dem Datenpfad entnimmt [6].

Ebene	***Kontrollpfad***	***Datenpfad***
Inhalt	RTSP	News, Videos
Anwendung	HTTP, h.245, SIP	Videokonferenz, IP-Telefonie, Java Applets
Transport	RSVP, RTCP, ICMP	TCP, UDP, RTP
Verbindung	DHCP, ICMP	SONET, SDH, DWDW

Tabelle 4.1: Generische Struktur der Aufteilung [5]

4.3.3 Strategieparadigmen

Die Trennung der Strategie von der Implementation ermöglicht dynamische Änderungen des Managements der Systeme und die Modifikation dessen Verhaltens. Es erlaubt auch eine Wiederverwendung von Strategien in verschiedenen heterogenen Umgebungen, besonders in unterschiedlichen administrativen Domänen. Aus diesen Gründen und Vorteilen wird ein Strategieparadigma angewendet, um eine A^x-Architektur zu schaffen, welche die Strategiedarstellung nicht betrachtet.

Die Strategien werden durch das Strategiemanagementhilfsprogramm geändert, welches eine Inkonsistenz- und Konfliktprüfung vornimmt. Diese sind auf den Strategiebehälter (SB) oder direkt auf den Strategieentscheidungspunkt (SEP) durch Konfiguration verteilt. SEP treffen Entscheidungen, indem sie Strategien mit anderen Daten und potentiellen anderen Strategien auswerten. Wenn eine passende Strategie gefunden wurde, wird die Entscheidung dem Strategieumsetzungspunkt (SUP) gesendet, welcher diese in Konfigurationsdaten übersetzt [6].

4.3.4 A^x-Architektur

Die Abbildung 4.4 zeigt die generische A^x-Architektur, die aus Modulen und Diensten besteht.

Module und Interaktion

Die notwendigen Module der A^x-Architektur können von dem Basisschema abgeleitet werden. Alle verschiedenen A^x-Strategien werden im Strategiebehälter gespeichert. Um Strategien auswerten zu können, wird der SEP als Modul benutzt: ein einzelner SEP für jede verschiedene Art von A^x-Strategie oder eine integrierte. Das instantiierte Design spiegelt diese Abhängigkeiten zwischen den verschiedenen Strategien wieder. SUP definieren Module der A^x-Architektur. Diese sind in der Serviceausstattung angesiedelt, entweder in der Nutzer-Serviceausstattung, um vom Nutzer angeforderte Dienste zu gewähren oder in der A^x-Serviceausstattung, um A^x-Dienste anzubieten.

Authorisationsstrategien werden normalerweise durch eine Entscheidung mit spezifizierten Dienstparametern umgesetzt. Diese Entscheidung beschreibt das Verhalten der Serviceausstattung gemäß der Nutzeranfrage. Die Authorisierungsstrategien werden durch

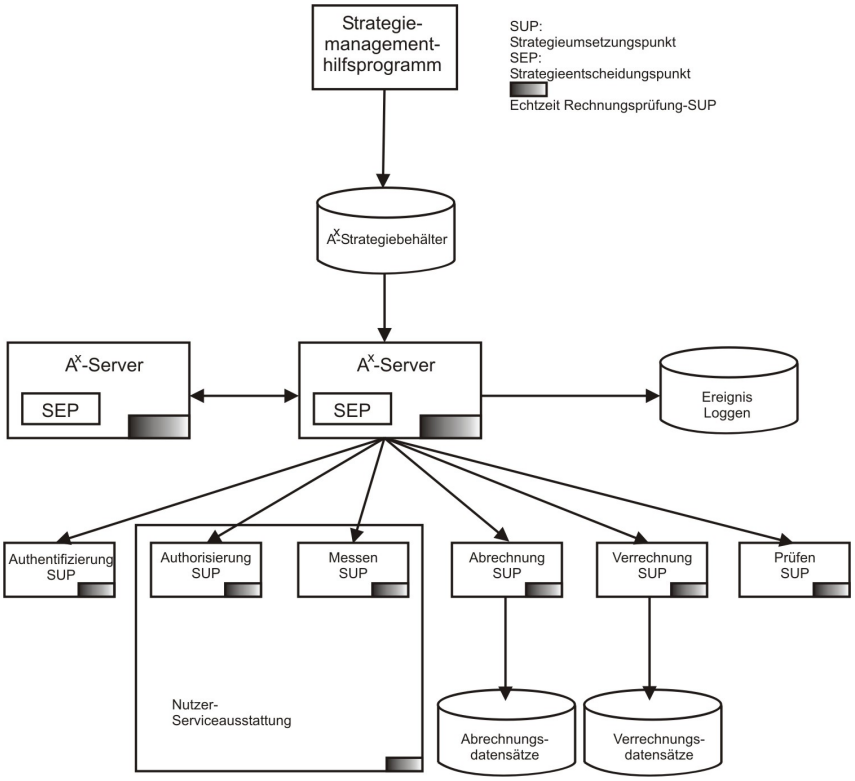

Abbildung 4.4: Generische A^x-Architektur [6]

ein spezielles Modul, welches notwendige Authentifizierungsinformation der Identitäten gespeichert hat, umgesetzt. Der SUP der Meßstrategien ist regulär in der Nutzer-Serviceausstattung oder einer Erweiterung von dieser plaziert. Die entsprechenden Daten können dem Abrechnungsmodul übermittelt werden oder in oder außerhalb der Serviceausstattung gespeichert werden. Für Abrechungs- und Verrechnungszwecke werden die SEP in speziellen Modulen gehalten, welche auf den gemessenen Daten durch Aggregation oder anderen Mechanismen operieren. Die Ergebnisse der Operation werden in der Abrechungs- und Verrechnungsdatenbank gespeichert. Die Speicherstelle der SUP für Prüfungsstrategien hängt von der reservierten Strategie ab. Für Echtzeitprüfung wird die Umsetzung in allen Modulen, die den Dienst anbieten, durchgesetzt. Für eine Prüfung nach Dienstgebrauch ist ein spezieller SUP notwendig.

Aus diesem Grund sind folgende Komponenten einer A^x-Architektur zwingend notwendig, neben dem A^x-Server, dem Strategiemanagementhilfsprogramm und dem Ereignisloggen:

- A^x-SEP als Hauptteil des A^x-Servers
- A^x-Strategiebehälter
- Authentifizierungsmodul
- Authorisierungsmodul in der Nutzer-Serviceausstattung
- Meß-SUP-Modul in der Nutzer-Serviceausstattung

- Abrechnungs- und Verrechnungs-SUP mit zusätzlicher Datenbank für Abrechnungs- und Verrechnungsdatensätzen

- Prüfungs-SUP abhängig von den Prüfstrategien platziert innerhalb jedes anderen Moduls oder als unabhängiges Modul

In dieser generischen Architektur (siehe Abbildung 4.4) werden die Module isoliert dargestellt. Ein einzelnes Modul instantiiert einen einzelnen A^x-SUP, während nur Hauptbeziehungen beschrieben werden. Nach einer gründlichen Untersuchung der genauen Abhängigkeiten zwischen verschiedenen angewandten A^x-Strategien , kann diese Architektur implementiert werden, hauptsächlich getrieben durch die reservierte Leistung und die Sicherheitsfrage. Schließlich werden zusätzliche Elemente für die Implementation der Umsetzung, wie das Ereignisloggen oder Sitzungsverzeichnissse, gebraucht [6].

A^x-Dienste

A^x-Dienste werden der Nutzer-Serviceausstattung oder anderen A^x-Servern im Falle eines Servicezugangs einer fremden Domäne zur Verfügung gestellt. Ein A^x-Server zieht den Strategiebehälter zu Rate, um eine Strategieentscheidung bei einem angeforderten A^x-Dienst zu treffen. A^x-Dienste werden in A^x-Servern durch die Umsetzung von Strategien in unterschiedlichen SUP durchgeführt.

Wenn beispielsweise ein Nutzer eine VoIP-Anwendungsdienst (Voice over IP) anfordert, wird der VoIP-Server eine Authentifizierungs- und Authorisierungsanfrage an den A^x-Server senden. Diese Anfrage muß die Identität und den Berechtigungsnachweis des Nutzers und den angeforderten Dienst enthalten, welcher eine eventuelle QoS-Spezifikation enthält. Abhängig von der Authentifizierungs- und Authorisierungsstrategie können unter Umständen weitere Informationen benötigt werden. Wenn die Anfrage authorisiert worden ist, konfiguriert der A^x-Server die jeweiligen SUP und sendet eine positive Antwort zum VoIP-Server. Alle anderen A^x-Dienste werden umgesetzt, wenn der angeforderte Dienst authorisiert worden ist. Die festgelegte Verrechnungsstrategie und das vertraglich festgesetzte Tarifschema bestimmen die Meß- und die Abrechnungskonfiguration. Für VoIP ist die effektive Verbindungszeit der wichtigste Meßparameter. Das Prüfen stellt sicher, daß der VoIP-Dienst so geliefert wird wie spezifiziert und ein Angriff auf das VoIP-System und die A^x-Infrastruktur festgestellt werden kann.

Wie Abbildung 4.3 zeigt, werden einige Dienst wie Authentifizierung und Authorisation nur einmal während der Signalisierungsanfrage geliefert. Andere Dienste, wie Messen und Echtzeitprüfung, werden kontinuierlich während der Verbindung durchgeführt. Schließlich gibt es Dienste (Abrechnung und Verrechnung), die nach der Dienstlieferung durchgeführt werden können [6].

4.3.5 Die generische A^x-Architektur am Beispiel des AAAC

Differentiated Services Architektur

Bei Differnetiated Services (DS) handelt es sich um eine Erweiterung des IP-Protokolls. In IPv4 bezeichnet das DS-Feld das Type-of-Service-Feld des IP-Headers, in IPv6 die Verkehrsklasse [3]. DS stehen auch für eine QoS-Strategie. Diese Architektur stellt keine Ende-zu-Ende-Dienste bereit, sondern teilt den IP-Verkehr in wenige Klassen ein. Die IP-Pakete werden gemäß ihrer Klassenzugehörigkeit bzw. ihres Eintrags im Type-of-Service-Feld bearbeitet [11].

Moby Dick

Moby Dick ist ein Projekt von Telekommunikationsgesellschaften[1], Herstellern[2], Forschungseinrichtungen[3] und Universitäten[4], das die Entwicklung, die Implementierung und das Testen von Ende-zu-Ende Kommunikationskomponenten in QoS-basierten Mobile IPv6 Netzwerken zum Ziel hat. Weiterhin sollen folgende Eigenschaften unterstützt werden [9]:

- Nahtlose vertikale (Netzwerke) und horizontale (Technologie, z. B. UMTS und Ethernet) Hand-Over-Mechanismen,
- QoS-Mechanismen basierend auf der IETF Differentiated Services Arbeitsgruppe für QoS-fähige Ende-zu-Ende Verbindung,
- AAA-Mechanismen basierend auf der IRFT AAAArch Arbeitsgruppe und der IETF AAA Arbeitsgruppe für QoS-fähigen mobilen Netzwerkzugang und
- Berechnungsmechanismen.

AAAC für die QoS-Infrastruktur

Im folgenden wird die generische A^x-Architektur am Beispiel der AAAC-Architektur (*Authentication, Authorisation, Accounting, Charging*) für eine QoS-Infrastruktur erläutert. In Moby Dick wird die DS-Architektur für das Gewähren von QoS verwendet. Die AAAC-Architektur befaßt sich mit dem Dienstangebot, der Kontrolle des Dienstzugangs und der Abrechnung der angebotenen QoS-Eigenschaften. Das AAAC-System braucht eine Schnittstelle zur DS-Architektur durch ein spezielles ASM. Zwei unterschiedliche Modelle werden in diesem Abschnitt beschrieben.

[1] T-Systems Nova Berkom (Deutschland), Inovação (Portugal)

[2] Centre de Recherche de Motorola (Frankreich), NEC (Großbritannien)

[3] Frauenhofer Gesellschaft FOKUS (Deutschland), EURECOM (Frankreich), Forschungszentrum Telekommunikation Wien Betriebs-GmbH (Österreich)

[4] Eidgenössische Technische Hochschule Zürich (Schweiz), Universität Carlos III von Madrid (Spanien), Universität Stuttgart (Deutschland), Universität Krakau (Polen)

- Deterministischer Ende-zu-Ende QoS (Bandbreitenbroker) und

- wahrscheinlicher QoS mit Netzwerkdimensionierung.

Im ersten Modell wird angenommen, daß der interdomäne QoS-Setup durch einen Bandbreitenbroker (BB) eingerichtet wurde. In dem BB-Modell wird ein Pfad für eine Nutzersitzung gemäß seinen QoS-Anforderungen reserviert, die bei der Anfrage übermittelt werden. Der erste BB, der die Ressourcenallokationsanfrage (RAA) von einem Nutzer erhält, nimmt mit dem AAAC-System über das ASM Verbindung auf. Die Authorisierungsanfrage kapselt die Daten, die zur Authentifizierung und Authorisierung benötigt werden, von der RAA. Das AAAC-System leitet die RAA über ein AAA-Protokoll zum AAAC-System des nächsten nah gelegenen BB weiter, welcher die Authorisierung durchführt. Dies wird wiederholt, bis die endgültige Domäne erreicht ist. Wenn jeder Authorisierung eine Ressourcenallokationsantwort (RAAn) folgt, wird sie durch das AAA-Protokoll zum ersten BB zurückgeschickt, welcher die RAAn zum Nutzer weiterleitet. In jedem AAA-System wird die RAAn zum BB über das ASM weitergeleitet, damit der BB Netzwerkelemente instantiieren kann. Zusätzlich muß der BB oder ein anderes System konfiguriert werden, um die Abrechnungsinformationen für jede Domäne zu sammeln.

Für das wahrscheinliche QoS-Modell markiert die Domäne, in der sich der Nutzer gerade befindet, die angeforderten Flüsse (Flows) mit dem entsprechenden QoS-Parametern abhängig von der Nutzeranfrage. Eine Ressourcenmanagementeinheit verteilt die Ressourcen auf eine drahtlose Zelle gemäß der Nutzerverträge, was in den Nutzerprofilen eines jeden Nutzer gespeichert ist. Jeder Anbieter am Ende-zu-Ende-Pfad hat ein gewisses Maximum von verfügbaren Serviceklassen. Der angebotene QoS hängt von der Dimensionierung der Netzwerkressourcen ab, beispielsweise wie groß die angebotene Quantität gegenüber der Größe der gegenwärtigen Nachfragen aller Nutzer in jeder Domäne ist. Solange die Anforderung kleiner ist als der angebotene Betrag, wird jeder Nutzer seinen gewünschten QoS erhalten. Die Beträge jeder Serviceklasse in einer Domäne sind den Nutzerverträgen angepaßt. Trotzdem werden diese Anpassungen weniger häufig sein, als Änderungen an den Netzwerkbedingungen.

In einem einfachen Fall kann das QoS-Angebot statisch für einen Nutzer und abhängig von seinem Netzwerkzugang sein. Das SLA (Service Level Agreement) definiert, welchen QoS ein Nutzer bekommt und dieser muß nach dem Netzwerkzugang authorisiert werden. In diesem Fall findet keine weitere QoS relevante AAAC-Kommunikation für die ersten zwei As (Authentifizierung und Authorisation) statt. Die Authorisierungsantwort muß für den Netzwerkzugang Parameter für den QoS enthalten und das AAAC-System muß das Abrechnen übernehmen. Die QoS-Sitzung muß unmittelbar nach dem Verbindungsende freigegeben werden [2].

4.4 Szenario für einen mobilen Nutzer

4.4.1 Grundlagen

Das Mobile IP-Protokoll wird dazu verwendet, um die Mobilität eines IP-Hosts über IP-Subnetze zu gewährleisten. Die Mobile IP-Arbeitsgruppe hat folgende Anforderungen definiert, die der Interaktion von AAA und Mobile IP dienen:

- Bessere Skalierung von Sicherheitsaspekten,
- Mobilität über administrative Grenzen der Domänen hinweg und
- dynamische Übergabe der Heimagenten (Home Agents).

Im folgenden wird ein „Multi-Domain“ Modell für Mobile IP entwickelt.

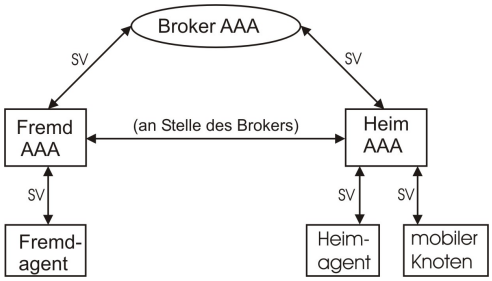

Abbildung 4.5: Mobile-IP Sicherheitsmodell [10]

Das Bild 4.5 zeigt ein neues AAA-Sicherheitsmodell für Mobile IP. Jedes Netzwerk erhält in diesem Modell mobile Knoten (MK) und einen AAA-Server (AAA). Jedes mobile Gerät hält eine Sicherheitsverbindung (SV) mit einem AAA-Server in seinem Heimnetzwerk. Die beiden administrativen AAA-Domänenserver (Fremd- und Heim-AAA-Server) können entweder untereinander eine Sicherheitsverbindung aufrechterhalten oder können diese über einen dazwischen geschalteten Broker betreiben.

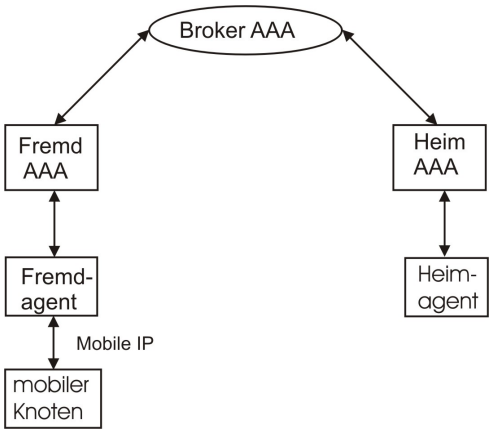

Abbildung 4.6: Drahtlose IP-Architektur für Mobile IP AAA [10]

Das Bild 4.6 zeigt ein Beispiel eines Mobile IP-Netzwerks, welches AAA enthält. In diesem Netzwerk, in dem AAA und Mobile IP integriert ist, wird vorausgesetzt, daß jeder Agent eine sichere Verbindung zwischen sich und seinem lokalen AAA-Server betreibt. Außerdem sind der Fremd-AAA-Server und der Heim-AAA-Server mit einer sicheren Verbindung über den AAA-Servers des Brokers verbunden. Ferner muß jeder mobile Knoten eine sichere Verbindung zu seinem Heim-AAA-Server unterhalten. Ein mobilen Knoten erscheint in diesem Beispiel in einem fremden Netzwerk und benötigt eine Registrierung vom Fremdagenten. Solange der Fremdagent keine Sicherheitsverbindung mit dem Heimagenten unterhält, sendet er eine AAA-Anfrage an seinen lokalen AAA-Server. Diese Anfrage beinhaltet die Authentifizierungsinformation und die Mobile IP-Registrierungsanfrage. Der mobile Knoten kann aus zwei Gründen nicht dirket mit dem Heim-AAA-Server kommunizieren:

- Er hat keinen Zugang zum Netzwerk. Die Registrierungsanfrage wird vom mobilen Knoten gesendet, um Zugang zum fremden Netzwerk zu erhalten.
- Der mobile Knoten hat unter Umständen keine IP-Adresse und fragt eventuell nach, damit ihm eine von seinem Heimprovider zugewiesen wird.

Der Fremd-AAA-Server wird entscheiden, ob die Anfrage lokal durch den Gebrauch des Network Access Identifier (NAI) des mobilen Knotens erfüllt werden kann. Der NAI hat das Format *user@realm* und identifiziert auf Grund des *realm* den Heim-AAA-Server des mobilen Knotens. Sollte keine Sicherheitsverbindung zwischen dem Fremd-AAA-Server und dem Heim-AAA-Server des mobilen Knotens zustande kommen, leitet er die Anfrage an seinen Broker weiter. Wenn der Broker eine Sicherheitsverbindung mit dem Heim-AAA-Server unterhält, kann er die Anfrage weiterleiten, sonst wird eine Fehlermeldung an den Fremd-AAA-Server gesendet.

Sobald der Heim-AAA-Server eine AAA-Anfrage erhält, authentifiziert er den Nutzer und beginnt die Authorisierungsphase. Die Authentifizierungsphase beinhaltet folgendes:

- Das dynamische Generieren der Sitzungsschlüssel (Session Keys), welche an die mobilen Geräte verteilt werden.
- Das dynamische Zuweisen eines Heimagenten (dies ist optional).
- Das dynamische Zuweisen einer Heimadresse ist auch optional und kann ohnehin auch vom Heimagenten erledigt werden.
- Ebenfalls ist das Zuweisen von QoS-Parametern für den mobilen Knoten optional.

Der Heim-AAA-Server sendet nach erfolgter Authentifizierung unverlangt eine AAA-Anfrage an den Heimagenten, welche die Information der originalen AAA-Anfrage und die Authorisierungsinformation, die der Heim-AAA-Server generiert hat, enthält. Der Heimagent stellt die Registrierungsanfrage wieder her und verarbeitet diese. Dann erzeugt er eine Registrierungsantwort und sendet diese dem Heim-AAA-Server in einer AAA-Antwort mit. Diese Nachricht wird durch den Broker zurück zum Fremd-AAA-Server und schließlich zum Fremdagenten weitergeleitet.

Der AAA-Server speichert Sitzungszustandsinformationen, welche auf der Authorisierungsinformation basieren. Wenn sich ein mobiler Knoten zu einem anderen Fremdagenten innerhalb der Fremddomäne bewegt, kann eine Anfrage an den Fremd-AAA-Server schnell erledigt werden, um die Schlüssel zurückzugeben, welche bei dem vorherigen Fremdagenten zugeteilt wurden. Das minimiert einen zusätzlichen Round Trip durch das Internet und ermöglicht eine problemlose Übergabe [10].

4.4.2 Beispiel eines mobilen Szenarios

Die Abbildung 4.7 zeigt ein mögliches Szenario, wie es ein Manager vorfinden kann. Dieser betreibt zu Hause ein kleines Netzwerk und eines seiner Kinder hat einen nicht vernetzten Rechner. In der Firma gibt es auch ein Netzwerk. Wenn er sich auf Konferenzen befindet, braucht er Zugang zu seinem Firmennetzwerk. Mit der AAA-Architektur für Mobile IP ist dies jetzt kein Problem mehr. Wie die Authentifizierung und Authorisierung erfolgen, wurde im vorherigen Abschnitt ausführlich erläutert.

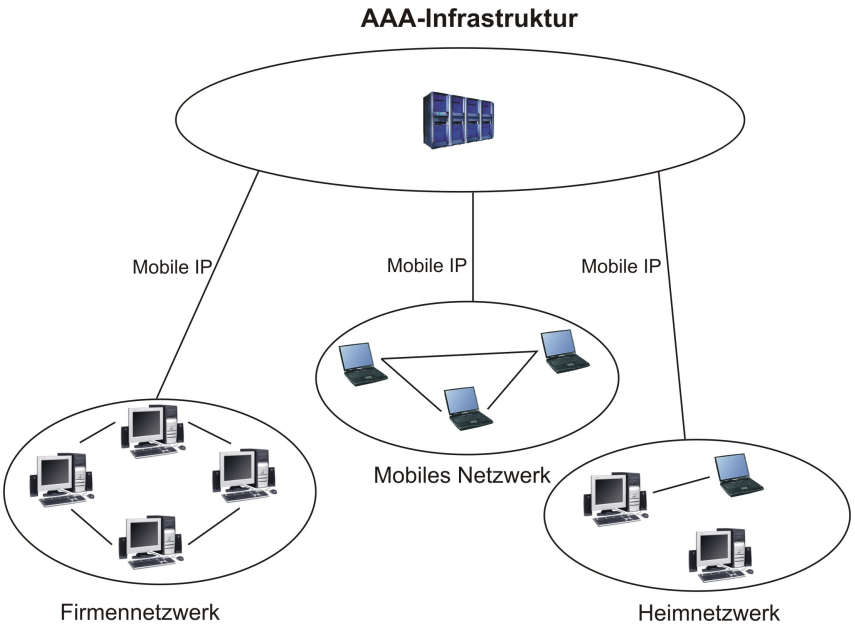

Abbildung 4.7: Drahtlose IP-Architektur für Mobile IP AAA

4.5 Zusammenfassung und Ausblick

Die Nachfrage nach AAA-Diensten und Diensten jenseits von AAA wird immer größer, um die kommerzielle Entwicklung der Dienste zu ermöglichen, die von einem zukünftigen IP- und eventuell Multi-Dienst-Netzwerks angeboten werden. Die A^x-Unterstützungsdienste beinhalten Prüfung, Preisfestsetzung, Verrechnung und Buchführung. Da sich gegenwärtige AAA-Architekturen, Protokolle und Implementationen nicht mit heterogenen Anwendungsszenarien beschäftigen und viele Anforderungen an verschiedene Ebenen der Dienste nicht unterstützt werden, wurde die generische A^x-Architektur vorgeschlagen.

Der generische A^x-Ansatz berücksichtigt diese Aspekte und unterscheidet klar zwischen Hilfsdiensten und Nutzerdiensten. Sie bringt die Vorteile der strategiebasierten Managementarchitektur voll zum Ausdruck, da die Entscheidungspunkte von den Umsetzungspunkten getrennt werden. A^x-Dienste können von einem spezialisierten A^x-System angeboten werden. A^x-Dienste außer Messen können von einem Provider einem anderen wegen der zukünftigen Trennung basierend auf A^x angeboten werden. Aus diesem Grund können Provider gemäß ihres Geschäftsplans ihre Systeme konstruieren [6].

In zukünftigen Arbeiten sollen die Abhängigkeiten zwischen verschiedenen A^x-Strategien und den Gesamtstrategien untersucht werden. Im Besonderen sind Auswirkungen auf die Authentifizierungs- und Authorisierungsstrategien von Interesse. Zusätzlich wird die genaue Erforschung der QoS-Unterstützungspläne durch die A^x-Architektur überprüft, um einen homogenen und integrierten Ansatz zu ermöglichen. Außerdem wird die vorgestellte A^x-Architektur zu einem kompletten Modell einer generischen A^x-Architektur erweitert, welche ein vorgeschlagenes A^x-Protokoll, A^x-Datentypen und eine A^x-Nachrichtensequenzdarstellung umfaßt [5].

Abkürzungungsverzeichnis

AAA	Authentifizierung, Authorisierung, Abrechnung
AAAC	Authentification, Authorization, Accounting, Charging
API	Application Programming Interface
ASM	anwendungsspezifisches Modul
BB	Bandbreitenbroker
CHAP	PPP Challenge Handshake Authentication Protocol
COPS	Common Open Policy Server
DHCP	Dynamic Host Configuration Protocol
DS	Differentiated Services
DWDW	Dense Wavelength Division Multiplexing
FA	Fremdagent
FO	Fremdorganisation
HA	Heimagent
HO	Heimorganisation
HTTP	Hypertext Transfer Protocol
ICMP	Internet Control Message Protocol
IETF	Internet Engineering Task Force
IP	Internet Protocol
IRTF	Internet Research Task Force
LDAP	Leight-weight Directory Access Protocol
MK	mobiler Knoten
NAI	Network Access Identifier
NHO	Nutzerheimorganisation
PPP	Point-to-Point Protocol
QoS	Quality of Service
RAA	Ressourcenallokationsanfrage

RAAn	Ressourcenallokationsantwort
RADIUS	Remote Authentication Dial In User Service
RBM	regelbasierte Maschine
RSVP	Ressource Reservation Setup Protocol
RTCP	Real Time Transport Control Protocol
RTP	Real Time Protocol
RTSP	Real Time Streaming Protocol
SB	Strategiebehälter
SDH	Synchronous Digital Hierarchy
SEP	Strategieentscheidungspunkt
SIP	SMDS Interface Protocol
SLA	Service Level Agreement
SMDS	Switched Multimegabit Data Service
SNMP	Simple Network Management Protocol
SONET	Synchronous Optical Network
SUP	Strategieumsetzungspunkt
SV	Sicherheitsverbindung
TCP	Transport Control Protocol
UDP	User Datagram Protocol
UMTS	Universal Mobile Telecommunication System
VoIP	Voice over IP

Literaturverzeichnis

[1] B. Aboba, D. Mitton. Authentication, Authorization and Accounting (AAA). http://www.ietf.org/html.charters/aaa-charter.html, 10 2002.

[2] Hasan, Davider Singh, Sebastian Zander, Moritz Kulbach, Jürgen Jähnert, Burkhard Stiller. The Desing of an Extended AAAC Architecture. http://www.ist-mobydick.org/publications/wp4-mobsum-2002.pdf, 2002.

[3] K. Nichols, S. Blake, F. Baker, D. Black. Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers. http://www.ietf.org/rfc/rfc2474.txt, 12 1998. RFC 2474.

[4] J. Postel. User Datagram Protocol. http://www.ietf.org/rfc/rfc768, 8 1980. RFC 768.

[5] Christoph Rensing, Hasan, Martin Karsten, Burkhard Stiller. A Survey on AAA Mechanisms, Protocols and Architecture and a Policy-bases Approach beyond: A^x. Technical Report 11, Institut für Technische Informatik und Kommunikationsnetze Eidgenössische Technische Hochschule Zürich, 5 2001.

[6] Christoph Rensing, Hasan, Martin Karsten, Burkhard Stiller. AAA: A Survey and a Policy-based Architecture and Framework. To appear: IEEE Network Magazine, 2002.

[7] C. Rigney, A. Rubens, W. Simpson, S. Willens. Remote Authentication Dial In User Service (RADIUS). http://www.ietf.org/rfc/rfc2138.txt, 4 1997. RFC 2138.

[8] R. Rivest. The MD5 Message-Digest Algorithm. http://www.ietf.org/rfc/rfc1321.txt, 4 1992. RFC 1321.

[9] Davinder Singh, Sebastian Zander. Moby Dick Project: Diameter in a Mobile QoS-enabled Network. http://www.interlinknetworks.com/images/resource/-Moby_Dick_Project.pdf, 4 2002.

[10] J. Vollbrecht, P. Calhoun, S. Farrell, L. Gommans, G. Gross, B. de Brujin, C. de Laat, M. Holdrege, D. Spense. AAA Authoriziation Application Examples. http://www.ietf.org/rfc/rfc2905.txt, 8 2000. RFC 2905.

[11] A. Westerinen, J. Schinzlein, J. Strassner, M. Scherling, B. Quinn, S. Herzog, A. Huynh, M. Carlson, J. Perry, S. Waldbusser. Terminology for Policy-Based Management. http://www.ietf.org/rfc/rfc3138.txt, 11 2001. RFC 3198.

Kapitel 5

Wireless and Location-based Internet Services

René Baldauf

Durent recent years, the technology sector was minted by a extremely high boom. Better and faster production procedures made it possible to build even smaller devices, which are because of mass production on a price level that addresses the common run of mankind. These circumstances, for example, lead to the situation that nearly everybody has a mobile telephone in these days. This provides mankind with the possibility to communicate with nearly everybody around the whole world. But this is just the beginning. With more and more powerful mobile devices the question for other mobile services than telephony raise. The Internet is a service that boomed in a similar way the last years. It is an integral part of our all day life. The combination of these technologies leads us to mobile Internet services, a new group of applications, which will probably boom in the same way the next years.

A part of these services are "location-based Internet services", a completely new section of applications. They use the attained mobility of users to provide very special services. These services determine the user's actual position and exploit these location information to deploy unique applications. Such services were completely unknown in times, when no mobile devices exist. That's why there are a lot of different systems, approaches and opinions. The standardization of location-based Internet services must, therefore, have highest priority, if these applications are thought to become an important part of our all day life.

The present work addresses these services. It introduces classifications for these services, visualizes actual localization techniques and shows gained standards for mobile application design.

Inhaltsverzeichnis

5.1 Introduction

Wireless technology offers the possibility to communicate with nearly everyone around the world. In addition it is possible to provide every imaginable information to the user, wherever he is. By using the different location information a completely new type of services can raise, as discussed in the scenario below. These services are the topic of this report.

- **Scenario:** Imagine, you're on your way to an important meeting. The navigation system shows you the way, because you don't know the exact way. This is already a location-based service, that is nowadays widely spread and that has proved itself in all day life. But the imagination of location-based services goes further on. For example, the fuel stand of the car could automatically be determined. If the system ascertains, that the fuel doesn't last to reach the destination, the system shall show gas stations on the way - eventually compared by prices - and lead the driver there.

The present seminar faces the topic "wireless and location-based Internet services". The scenario above is only one out of a nearly endless list of imaginable applications. Heise.de says, that LBS - short for location-based services - next to UMTS are the coming killer technologies for mobile communications[9]. This chapter gives a definition of LBS and shows with the help of market research, which potential lies within this raising market. The second chapter discusses further examples and classifies LBS. The third and fourth chapter introduce concrete techniques for mobile device localization and the implementation of applications for mobile devices.

5.1.1 Defining wireless and location-based Internet services

Before we discuss LBS in the following chapters, this chapter shall clarify, what LBS are. "Internet services" are all services provided by the Internet, like eMail services, file transfer via FTP, NEWS services and many more. These services are nowadays an integral part of all day life in western countries. As the mobility of Internet users grows, because of PDAs and high-performance mobile telephones, Internet services shall also grow and support the users mobility, but also introduce completely new perspectives.

LBS have the ability to locate a mobile device,
and provide services based on these information.

a) Wireless ...

As long as we use Internet services with a wired network connection, we profit from advantages like high bandwidths and high reliability. For mobile devices the use of this network isn't possible. They had, therefore, to use wireless networks, which operate with radio signals. These connections have to face various problems, but offer also unique advantages:

- **Connectivity:** In wired networks, phenomenons like highly dithering bandwidths and total loss of connection are nearly unknown. For mobile networks these are serious problems. The radio signal, the carrier for the data transmission, is negatively influenced by weather, buildings and other objects. Because of this, the signal can be bend, broken and changed in runtime. This leads to a loss of quality, resulting in lower bandwidths and perhaps total loss of connection. Of course, the user shouldn't notice these circumstances. Special system architectures and the WAP protocol are an attempt to solve these problems. They are discussed in section 1.4.

- **Bandwidth:** A normal desktop PC has nowadays an Internet connection with a relatively high bandwidth. This is averaging at 100 Mbit/s in company or university LAN's, at home with ADSL after all up to 6 Mbit/s or at least constant 64 kbit/s with ISDN. The runtime for packets is within the range of milliseconds, eventually up to 100 ms for transatlantic connections. The situation in wireless networks is completely different. The mobile technology is nowadays far away from reaching similar bandwidths to wired networks. GPRS has nowadays 115,2 kbit/s, HSCD 56,7 kbit/s and GSM after all 9,6 kbit/s. The runtime for packets is here within the range of seconds[5].

- **Substitute for wired connections:** Because of technical and economical aspects, it isn't always possible and suggestive to use wired networks. Trade fairs, for example, need a highly flexible communication infrastructure, which has to be reconfigurated for each event. Therefore, a wireless connection fits much better than a wired one. Historical buildings can be provided with an network connection with a few antennas, avoiding further wires[5].

- **Immunity against natural disasters:** Wireless communication systems are nowadays the only systems that can survive natural disasters, like earthquakes, tidal waves and hurricanes. Traditional telephone and data networks, but also wireless networks that operate with base stations, collapse in these situations. Herein satellite-based and adhoc networks offer the possibility to keep communication alive[5].

b) Location-based ...

The location information of a mobile device is used in various ways in wireless networks. Adjacent services, for example, which work absolutely hidden from the user, are scarcely offer the possibility to use mobil devices. Such adjacent services offer among other things the functionality to keep connection, when the user changes between different cells or networks. Beside these, there are a lot of services, which use the location information to offer services. These services are aware to the user.
In principle, there are 2 possibilities to get location information of a mobil device.

- **User input:** For several applications it is enough, that the user provides the location information by input via keyboard, speech or similar input devices. These information don't have to fit the actual position of the mobile device. They can be used together with automatic localization, to provide routing applications, for example. The actual position can hereby easily be spotted, but it isn't possible to spot the wanted destination. This additional information is provided by the user.

- **Automatic localization:** The second possibility is to spot the location of a mobile device automatically. There are many different techniques, which strongly differ in their basic approaches, accuracy and realization. In section 1.3. the most important techniques and their advantages and disadvantages shall be introduced.

5.1.2 Market research on LBS

With the raging spread of PDA's and, above all, mobile telephones, a completely new market has raised. Location-based services were entirely unknown and senseless in the times of desktop personal computers, as they weren't mobil because of their wired network connection and therefore location information only correspond to their fixed position. In the near future, LBS will probably as certain as nowadays internet connection at home. Hence LBS comply to today's time spirit of the mobile individual.
Market research confirms such and similar statements impressively. A research of IDC (Integrated Data Communications, Inc.) points out, that nearly two thirds of the US citizens are interested in LBS[1]. Thereby, services like emergency services and roadside assistance are at first place[2]. Many of the queried persons said, that they are also willing to pay for these services or receive advertising on their mobile device to lower costs[1]. In Europe, for example, after a research of Mori ordered by AirFlash, a big part of mobile telephone users, interested in LBS, are willing to pay monthly up to 14 EUR for such services[8]. The Strategis Group prognosticate a market size of about 3,9 million US$ in 2004 for the USA only[2]. The ABI (Allied Business Intelligence, Inc.) talks of an income of over 40 billion US$ in 2006. This corresponds to a growth rate of over 81%[1]. For the asian markets the number of mobile Internet users is thought to be 216,3 million in 2007. That are more than ten times more than in 2000[1].
These research results shall show, which great expectations are lying in this market. It remains to hope, that the problems, associated with these services, are successfully solved. Above all it has to succeed to create standards for mobile networks and the Internet. The WAP protocol, which will be introduced below, is a first big step into this direction.

5.2 Applications

There are plenty of possible location-based services and applications, which are entirely different among themselves. This situation makes a graduation in different classes useful and necessary, if we are willing to bother LBS seriously. This section introduces different classifications and shows further examples for location-based services, where they are useful.

5.2.1 Push/Pull services

- **Pull services:** Pull services include all services that the user can fetch. They work like the traditional client/server-model. The client sends a request (pull), which the server answers. Services are, for example, WWW, FTP, NEWS and similar ones[5].

- **Push services:** In opposition to the pull services, where the server answers a client request, within the push services the server sends its content without any request. These services have the advantage, that they produce fewer network traffic as they do not need a client request. This concept indeed don't fit for all applications. Services are, for example, advertising, weather and stock exchange information and similar services, which are thought to be renewed frequently, without a client request[5].

5.2.2 Classification of LBS

The classification of LBS into push and pull services is a very technical approach. This isn't always useful. The next classification separates the LBS into groups of services with different tasks. This groups are not thought to be orthogonal as the classification into push and pull services. For example there are services which can be grouped into adjacent services but also into supporting services. This belongs to a readers point of view.

- **Adjacent services:** This group contains above all services, which are needed to allow the hand-over of a mobile device into another cell or radio network. This technique, already used in traditional telephony service, should also be deployed for internet services. This should among others allow, that services like eMail, file transfer via FTP and other services don't collapse during a cell change, because of changing logical addresses[5]. MobileIP, for example, is a protocol, which offers this functionality. It has been introduced in this seminar series as well[11]. Adjacent services as a rule work invisible to the user, that means the user isn't aware of a cell or network change. His applications and internet services work further on without any interruption.

- **Location-aware services:** This group consists of services, that gather information about the environment. With these information, special services can be provided. An application, for example, is an extended printer service. If the user is in his home office and wants print out a document, the document is send to the standard printer. But what happens, if the user is on a business trip with his laptop and wants to print a document? To print the document on the home printer is rather senseless. But without any information about the environment, this is the only possibility. Therefore a hotel, for example, could provide a printer service, which is spotted by the location-aware service. The document can than be printed on the hotel printer. On the other hand, the location-aware service can send user information to the hotel service to allow an accounting of the used service[5].

- **Security services:** As we have seen above, there are many services that communicate with other stations and exchange data with them. This, indeed, raise the

question for privacy and anonymity. Therefore services have to be created that use the services provided by the environment, but on the other hand don't allow the environment to gather exact information about whereabouts and user behavior. This shall prevent the misuse of personal data for advertising, for example. Furthermore services can be ranged in this group, which activate various filters depending on different locations and times, to keep unimportant events away from the user. For example, it isn't desired, that an unimportant video conference is send to the users mobile device during an important discussion. Nevertheless, urgent eMails have to reach the user further on[5].

- **Information services:** This group contains services, which the user directly perceive. In detail, these are all services that provide the user with information, like WWW, eMail services but also localization services for gas stations, super markets, etc. [5].

- **Supporting services:** Next to the services mentioned above there are a lot of services, which operate as far as possible hidden from the user. These services shall, above all, support the mobility of the user. Examples are the caching of condition information or cache content. This, for example, allows to save internet pages and makes it possible, that the user can continue his works, although there is no internet connection[5].

As seen, there are a lot of services being hidden from the user. They support the user in his mobility, protect his privacy and allow the use of other services beyond firmless bandwidths and loss of connection to the network.

5.2.3 Classification of user-aware LBS

The next classification applies above all for the information services, the services that the user is aware of. This doesn't mean, that special services from other groups don't fit into this classification. These classifications are not absolute, but offer a concept - borders between single groups are indistinct and belong to a reader's point of view.

- **Security services:** As market research shows - see also section 5.1.2 -, these services are on first place among mobil device users. They shall provide the functionality, that there is quick and reliable help for the user in an emergency or dangerous situation. The creation of such services represents the birth of LBS[2]. An initiative from the FCC (The Federal Communication Commission) required all network operators, to provide emergency centers (911 numbers) with the position of a mobile device within an accuracy of 125 m. Nowadays these services are the widest spread LBS in the USA[10]. Another example for emergency services is roadside assistance.

- **Transaction services:** These services shall provide the possibility to easily pay bills, for example, with the help of a mobile device. They had therefore to gather information about the environment, to spot, which services are provided at the current location. An example is the printer service, discussed above. Another example

could be a 24h roadhouse. the reception probably isn't occupied all the day. With the help of a special service, the user could check in, by sending his user information to the hotel. The hotel system can than check the information and hand out the keys, if they are correct.

- **Information services:** Services of this group shall provide the user with plenty of various information. Examples are the current cinema or theater schedule, information about special events like fairs, partys and so on - short, all what is nowadays called lifestyle. Services that provide traffic information or weather and stack exchange data, for example, also fit into this group.

- **Tracking services:** Services of this group provide the user with the location of various objects. They are useful for commercial and private use. An example for a commercial use is the fleet management (ships, trucks, trains). Every object of the fleet has an mobile device, which can be spotted. Based on these information, the manager can make an effective planning for his capacities. A possibility for private use is, for example, a packet tracking system, which can show the user where his packet is in the very moment.

5.2.4 Further examples for LBS

The first part of this section discusses a special location-based Internet service, called "friends around". These services are mostly the first steps of some companies to introduce LBS to the customers - especially young ones. The second part discusses 3 examples for LBS, which doesn't operate on WAP, but using different techniques.

a) Friends around

A special application for tracking services is the display of friends in your vicinity - the so-called "friends around". These services show friends with mobile devices, which are in a definite radius according to the own location. If a friend is found, you can send him a SMS and make a date for dinner or cinema, for example. Such services shall leverage LBS to the same popularity like SMS. The Swiss company SwissCom has deployed such a service - called friendZone[9].

b) LBS and WAP

Nowadays, plenty of existing LBS base on the WAP protocol. This protocol was especially developed for such applications, as traditional internet protocols like TCP/IP and HTTP hardly suit for mobile purposes. But there are also other concepts for providing LBS to the user:

- **CityScout:** The company Viag Interkom provides a LBS system, which operates on SMS. The user therefore sends a SMS with a search string to the number 3463 (f-i-n-d). The system locates the position of the mobile device and replies with a SMS which contains the nearest address corresponding to the search string. On demand, the user can receive more SMS with further addresses[9].

- **Berlin CityGuide:** The company mecorno wants to provide internet content, so far only accessible via WAP, completely via speech portals. An example is the Berlin CityGuide, which supplies the user with information about locations and events. To use the Berlin CityGuide, simply call 030/52 00 51[9].

- **i-mode:** i-mode is a packet-based service for mobile devices, provided by the leading Japanese supplier for mobile technology NTT DoCoMo. It doesn't use the WAP protocol and his language WML, like most of the leading companies, but a simplified version of HTML - called C-HTML or CWML[9][10].

5.3 Techniques for mobile device locating

A prerequisite for location-based Internet services is the technique to determine the position of a mobile device. The network operator has therefore to trespass the thin red line between costs and profit. Simple localization techniques are mostly reasonable and easily to deploy, but offer sometimes only insufficient accuracy to set up applications like emergency services and roadside assistance. More accurate systems are often very expensive or hard to deploy because of legal provisions. In addition these techniques often need hardware changes on the mobile device, wherewith only new customers can be reached at the moment.
This section provides an overview about the most important localization techniques. It therefore also elaborates advantages and disadvantages of these techniques. The second part faces the MLC - short for Mobile Location Center - , an interstage, that allows the network operator to deploy the techniques, presented in part one, within economical aspects.

5.3.1 Localization techniques

COO - Cell Of Origin

To maintain the connection to a network, a mobile device has to register at a network base station. Each of these base stations covers a certain area - also called a cell. If the mobile device crosses the cell border, it automatically unregisters at the current base station and registers at the new base station. The COO positioning system uses this technique to spot the location of the device. It simply checks at which network base station the mobile device is registered. The position of the base station is ascertained and considered to be the location of the mobile device. For this reason the accuracy of COO depends on the base stations cell size. In urban areas where base stations are densely concentrated the

accuracy of COO may be as close as 150 m. Instead, in rural areas, where base stations are less densely concentrated, the ascertained position and the exact one may differ up to 30 km. Therefore, COO can hardly be used for emergency service, roadside assistance or similar LBS.
Although COO positioning is not as precise as other methods, it offers unique advantages:

- The location of a mobile device can be spotted very quick (generally in about three seconds).
- The COO uses existing network technology. It doesn't need any hardware changes or modifications on the mobile device or the base stations. For this reason, COO is very reasonably and easily to deploy.

COO is nowadays the only technology that is widely deployed in wireless networks.

TOA - Time Of Arrival

The TOA positioning system uses three network base stations to locate a mobile device. It measures the runtime of a signal from the device to each base station. With the help of these values the position can be triangulated. To use the TOA technique the network has to be synchronized, using GPS or atomic clocks at each base station.
Therefore, the TOA positioning has to face various problems:

- The cellular network has to be synchronized to use TOA, which is not provided in asynchronous GSM networks for example. For this reason, TOA is expensive to be deployed by network operators.
- TOA offers only little advance in position accuracy, but needs a far longer time of response (generally up to 10 seconds).

On the other hand, it offers only little advantages:

- no modification of the mobile device is needed.

AOA - Angle Of Arrival

The AOA positioning system tries to locate a mobile device by determine the angle of a received signal relative to the cells sites. Therefore a complex 4-12 antenna array is needed at each cell site. The antennas of these arrays can in principle work together and determine the angle from which the signal originated. When at least two several antenna arrays work together the location of the device can be estimated.
The AOA positioning has to face various problems. For these reasons it is hard to deploy.

- The AOA systems suffers from distortion - above all in urban areas - of the wavefront of the cellular signal, caused by multipath and other environmental factors.
- The angular error of the antenna array can translate into a significant error in lateral distance if the mobile device is far from the cell sites.
- Adding antenna arrays to the cell sites is a logistical and aesthetic dilemma. On one hand it is very expensive to set up these antenna arrays, on the other hand the communities are increasingly regulating the built of network base stations.

E-OTD - Enhanced Observed Time Difference

The E-OTD positioning system uses a technique similar to the TOA system. It places location receivers, overlaid on the cellular network, as a location measurement unit - short LMU - at multiple sites geographically dispersed in a wide area. Each of these LMU has an accurate timing source. When a signal from at least three base stations is received by a mobile device and the LMU, the time differences of arrival of the signal from each base station at the device and the LMU are calculated. The location of the mobile device can now be estimated from the differences in time.
In comparison to other systems E-OTD has a few disadvantages:

- The E-OTD system has a longer time of response - typically around five seconds
- The software of the mobile device has to be modified to work with the E-OTD positioning system. For this reason E-OTD cannot be used to provide LBS to existing customer bases.

The E-OTD technique has also advantages:

- E-OTD systems offer a greater accuracy than COO - in fact, between 50 and 125 m.

AGPS - Assisted Global Positioning System

GPS is a satellite system to locate mobile GPS devices up to an accuracy of 10 m for civil purposes and up to 1 m with special military equipment. The GPS is owned and operated by the U.S. Department of Defense but is available for general use around the world. The GPS system consists of 24 synchronized satellites that orbit the earth. From any point on Earth at least four satellites are above the horizon. Each satellite continuously broadcast a radio signal. A mobile GPS device can triangulate its position by receiving the radio signal of at least 3 satellites. With the fourth radio signal, the mobile GPS device can also calculate its height or its moving speed. Assisted GPS combines this technique with others, like COO or E-OTD, to receive better results in urban areas.
Although GPS is a very accurate positioning system it has some disadvantages:

- To estimate the users position via GPS he has to be in sight of at least 3 satellites. This can be difficult indoors or in densely built areas.
- To use GPS the user needs a GPS-equipped mobile device. This mobile devices are nowadays still relatively expensive.

5.3.2 The MLC - The Mobile Location Center

As showed in section one, location-based Internet services have high potential and are the base for many hopes and expectations. Despite these facts some network operators are rather hold back in there implementation. Deterrent the plenty of various techniques and the missing standards are taking effect. Questions like "Which platform should be used?" and "Which applications should be implemented first?" originate uncertainness and prevent "solo runs", because of misgiving to be incompatible to coming standards.
At this position the mobile location center provides a solution. It separates the localization of a mobile device from the application. Many services work quite well with the accuracy provided by COO. Network operators can implement more accurate localization techniques step by step, without waiting for 100% market saturation before they can provide new services. Because the MLC works inside the communication network, it has the ability to change and adapt the possibilities of mobile devices. Whether the user calls a number, sends a SMS or starts a request via the WAP-protocol, the MLC simply sends the position information to the application.
The MLC provides the network operator with the ability to experiment with new applications, better localization techniques, new content and further mobile devices. He can react faster and economically on market demands and develop the system when new technologies become available.

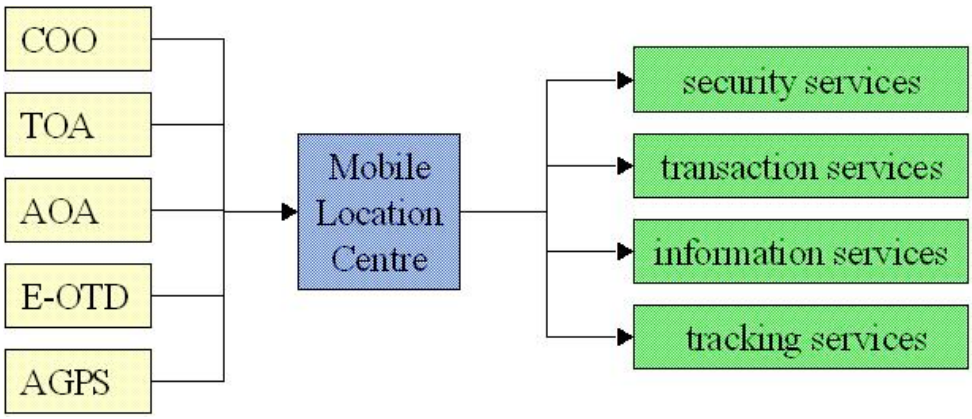

Abbildung 5.1: The MLC divides localization technique and applications

5.4 Concepts for supporting mobile Internet

The traditional Internet is based on the TCP/IP protocol. The WWW as an integral part of the Internet uses the HTTP protocol and its language HTML. Beside there are a lot

of other languages, which all go their own ways. Generally, these are Java applets, Flash animations, ActiveX elements and many others. They all have in common, that they have high expectations to the display - mostly true color with a resolution of at least 800 x 600 pixel. These structures work quite well for the traditional Internet - although there are also problems here - but they nearly entirely fail for mobile Internet applications. Because of this, the WAP protocol was created, to pave the way of the Internet to mobile devices.

5.4.1 HTTP/HTML and their problems with mobile Internet access

a) HTTP - HyperText Transfer Protocol

HTTP is a stateless, simple application level protocol for data transmission between WWW servers and WWW clients. Hereby stateless means, that all HTTP transactions are completely independent from each other. The result of this design is a very easy implementation without the necessary of a complex state machine. On the other hand, this design prevents nearly every possibility to implement a session management. A HTTP transaction consists of a HTTP request, which is send to a WWW server by a client. This server answers by sending a HTTP reply. HTTP therefore originate in using a save communication protocol - mostly TCP for traditional Internet. One problem is, that HTTP in version 1.0 sets up a TCP connection for every transaction. These are, for example, for a WWW page with text, 5 graphical elements like buttons and 2 bigger pictures already 8 TCP connections.
In addition HTTP supports simple mechanisms for caching from data between client and server. This prevents unnecessary data transfer, because content that hasn't changed needn't to be loaded again. With miscellaneous tags in the head of a WWW page it is possible to ease/influence caching. Possible are hereby, for example, tags containing information about the expiration date of WWW pages or tags that prevent caching, because the page contains dynamic content.
Because of these features there are various problems in wired networks, but even more in wireless networks:

- **Bandwidth and delays:** HTTP isn't designed for connections with little bandwidth and possible delays. Queries are mostly very large and show high redundancy, because of stateless implementation. In addition, the data is send uncompressed and in plain text (ASCII). As shown above, HTTP/1.0 creates a new connection for each object. This means a save 3 way connection establishment, the data transfer and a save connection abort. These expenses are also necessary for little objects like graphical elements, what creates corresponding overhead. HTTP isn't designed for such a communication scheme with short requests and maybe longer replies. But this scheme is typical for transactions. Further problems create the Slow-Start mechanism of the TCP protocol. At the beginning this protocol sends with a lower data rate, because the transfer window hasn't open completely. But before TCP can use the full bandwidth, the transfer is possibly ended. The effect is, that TCP never leaves the Slow-Start phase and creates new delays. Further delays are created

because of DNS queries. Many WWW pages contain objects, whose links has to be translated in their logical address.

- **Caching:** The problem isn't the caching, but the structure of WWW pages. Many pages contain objects like counters, date and time displays, which can't be cached. Furthermore, many companies place advertising on Internet pages and use the clicks on these objects for market research. But this system don't work, if the pages are cached, wherefore the caching is oppressed. In addition, many pages are nowadays dynamically created via PHP or ASP. Hereby the problem is, that the lower pages are completely dynamically created, wherefore the entrance to these pages happens only on the start page. Authentification mechanisms also hinder the caching, because the authentification happens between Client (Browser) and WWW server an not with the help of the Cache.

- **Send with POST:** Hereby arise problems, when WWW servers don't accept the send of highly delayed data. An application between client and server that caches content during a loss of connection and later with Internet connection simulates the client and sends the data also don't solve the problem.

b) HTML - HyperText Markup Language

HTML is a language that is mainly used to describe pages for the WWW. The language was designed for a normal desktop pc with a wired Internet connection. It therefore ignores every heterogeneities between personal computers and mobile devices. These are the Internet connection and, on the other hand, the great difference in processor power, capacity and display possibilities. HTML therefore don't provide the possibility to create WWW pages for such different requirements. In addition WWW pages are nowadays designed regarding optical aspects.
Further problems create extensions like Java, ActiveX, etc. . Many of these extensions are unknown to the browser, wherefore plugins are needed. But these plugins only exist for common platforms like Linux or Windows, but not for the special operating systems used for mobil devices. But even if these plugins exist, the problem isn't solved. The browser nevertheless isn't able to show a true-color video on a monochrome display. Similar problems exist for picture maps and other effects used for WWW pages.

5.4.2 Proprietary solutions

The problems introduced with TCP, HTTP and HTML have lead to proprietary solutions, which are discussed below. Some of these basic approaches had slipped into the creation of the WAP protocol and his language WML.

- **Scaling of pictures:** Pictures are not send to the mobil device in original size, but scaled to a lower size, fewer colors or down to the picture title. The user can than decide, if he wants to load the original. In addition techniques to create picture cut-outs, blowups and detail studies can be used.

- **Content transformation:** Many documents exist in formats like Postscript or PDF, that are not readable by many mobil devices. Here it is possible to transfer the content before the transfer into a readable format for mobile devices.

- **Content extraction/semantic compression:** To avoid the loading of a full web page, it can be useful to extract headlines or make a text summary and transfer this data to the mobile device. The user can than decide, if he wants to load the whole page. Problems occur, if tags a misused for layout aspects. In addition the compression of arbitrary texts is very difficult.

- **Use of special protocols and languages:** There are various efforts to replace HTTP and HTML by more suitable protocols and language. Examples are the HDTP (Handheld Device Transport Protocol) and HDML (Handheld Device Markup Language) from Phone.Com (formerly Unwired Planet). This protocol and his language slipped in the WAP protocol and WML.

- **Push techniques:** To avoid traffic, a server can automatically send data to the user. This can also prevent the effort to set up a new connection for every object. The push of data, indeed, is only useful for special applications like news, stock exchange values, and weather, which don't need a interaction with the user.

- **HTTP/1.1:** HTTP/1.1 solves some problems of the first version. It therefore allows the reuse of connections, so that connections could be used for several requests and replies. The caching mechanism was enhanced, so that is also possible to cache replies. In addition there exist mechanisms to transfer compressed data, to verify the integrity of mails and a authentification between client, server and cache.

5.4.3 System architectures

The classical system architecture of the WWW is a client/server structure. Without extensions, the consequence of a click on a link is the download of the data. For long, low-bandwidth connections with frequently loss of connection the cache as extension plays an elementary role. Within occasional loss of connection they are the only possibility to support the browser. This section introduces various architectures and extensions, which shall support the browser and mobil devices:

The first architecture consists of a mobile client with a web browser. This browser has a integrated cache as an extension. This offers the user the ability to disconnect from the network and nevertheless work with the saved content. The cache thereby can't load data forward, but save only received content.

A second possibility is to use an accompanying application, which supports the browser. This application can support the forward load of data, the caching and the use without connection. The problem is, that the Internet connection isn't transparent to the browser, because he has two different ways for the server access.
Because of this, an accompanying application isn't use, but a client proxy. With this

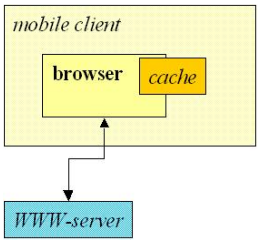

Abbildung 5.2: Browser with cache as extension

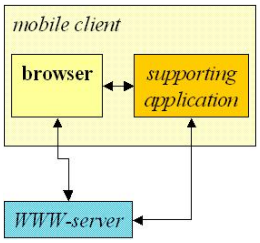

Abbildung 5.3: Browser with accompanying application

proxy, the connection to the network is transparent to the browser and the proxy can load and cache data with various strategies. Examples for this system are CaudWeb, TeleWeb and the Weblicator[5].

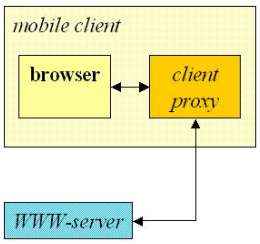

Abbildung 5.4: Browser with client proxy

Another possibility is to implement the proxy on the network side. This proxy can support the mobile client from the wired network, by using adaptive mechanisms for content transformation and extraction, for example. The network proxy can also cache or forward load content. This is sensible for wireless connections with high failure rates. Examples are TranSend and Digestor[5].

The use of two proxies is also useful. The connection between client and server is still transparent and the proxies can use this configuration to coordinate their forward load and caching. A possibility, for example, is, that the client proxy sends the network proxy user information. With these information, the network proxy can adjust his loading and caching strategy. WebExpress is an example for this architecture[5]. The last architecture builds a network subsystem between client proxy and network proxy. Therefore TCP and HTTP can replaced by more suitable protocols and data can transferred compressed, whereby the data transfer is optimized. Mowgli is a system that follows this concept[5].

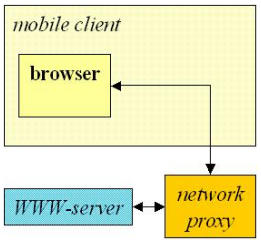

Abbildung 5.5: Network with network proxy

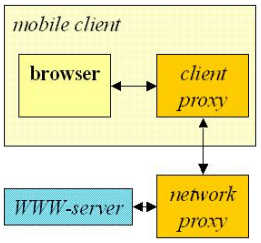

Abbildung 5.6: use of client and network proxy

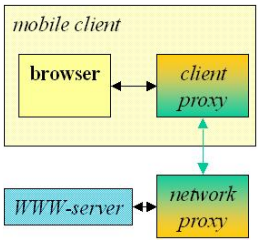

Abbildung 5.7: use of client and network proxy with a network subsystem

5.4.4 WAP - The Wireless Application Protocol

In June 1997 the WAP forum was created by Ericsson, Motorola, Nokia and Phon.Com (formerly Unwired Planet). The intention was, to prevent the creation of plenty of incompatible solutions and the provision of Internet content and other data services for mobile devices. Therefore a protocol family should developed, that provides the communication within different network technologies like GSM, CDPD, UMTS. The developed solutions hereby shall have the following attributes:

- **Interoperability:** All mobile devices and applications of various manufacturers shall work together within different networks.
- **Scalability:** The used protocols and services shall grow with the number and the demands of the users.

- **Efficiency:** The implementation shall guarantee a service quality, that is possible in wireless networks.
- **Reliability:** The solutions shall represent a consistent and predictable platform for the implementation of new services.
- **Security:** The integrity of data and users has to be secured. In addition devices and services have to be protected.

Nowadays the WAP forum consists of more than 300 members of software industry, various device manufacturers, network operators and computer telecommunication companies. These members work together in different workgroups facing different components of the protocol family. In April 1998 the first specification was released - the version 1.0. In May 1999 followed the version 1.1. and finally in November 1999 the version 1.2. The architecture shall now be introduced:

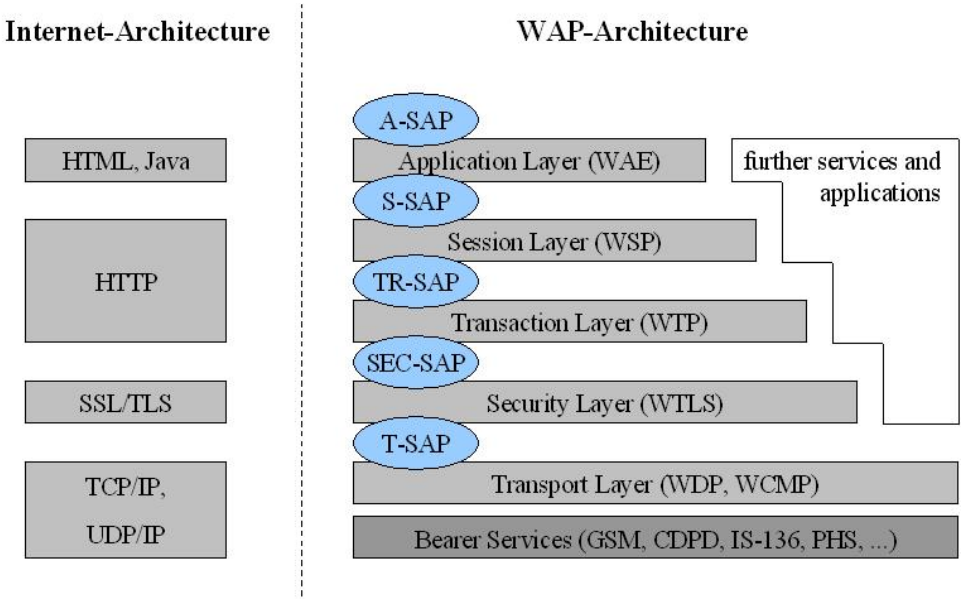

Abbildung 5.8: WAP architecture compared with the traditional Internet architecture

- **Bearer services:** The base of the WAP protocol family are bearer services like GSM, CDPD, IS-136, PHS. WAP thereby doesn't use a special bearer service, but it tries to use existing and include coming ones. Between bearer services and the transport layer there is no special access point defined, because the adaption depends on the bearer service.
- **Transport layer:** The transport layer owns the two protocols WDP (Wireless Datagram Protocol) and WCMP (Wireless Control Message Protocol). They therewith provide the upper layers of the WAP architecture a service access point (T-SAP), which provides, independent from the underlying bearer services, a datagram-oriented transfer service. The intrinsic communication is done transparently with one of the underlying bearer services.

- **Security layer:** The security protocol WTLS (Wireless Transport Layer Security) bases on the TSL protocol (Transport Layer Security; formerly SSH), which is used for secure data transfer in WWW. The WTLS protocol was optimized for wireless networks with relatively low bandwidths. It provides functionality for data integrity, privacy, authentification and a certain protection against DoS-attacks (Denial of Service attacks). The services are provided at the SEC-SAP.

- **Transaction layer:** Above the security layer lies the transaction layer with its protocol WTP (Wireless Transaction Protocol). The transaction layer provides a simple transaction service. Services thereby effectively can process secure and insecure requests and asynchronous transactions. The service access point is the TR-SAP.

- **Session layer:** The session layer and the WSP (Wireless Session Protocol) provide the functionality, that is needed to deploy connection-oriented services. Therefore a special service (WSP/B) was defined for the WWW. This service provides the HTTP/1.1 functionality and further more services for persistent sessions, the parking and resume as well as the migration of sessions. In addition there are plenty of more services. The service access point for the session layer is the S-SAP.

- **Application layer:** The application layer - also WAE (Wireless Application Environment) - is the highest layer of the WAP architecture. It offers a framework for various Internet and telephony applications. It therefore owns plenty of protocols and services, which are provided at miscellaneous access points. These are among others various script languages, page description languages, access points for telephony applications and various content formats.

An application hasn't completely to follow the WAP architecture. WAP doesn't force an application to use the whole protocol stack, but it can also used parts of this architecture. So, for example, a simple service, that doesn't need a session management, can directly set up on the WDP.

5.5 Summary

5.5.1 Localization techniques

As we've seen in section 3 there are a lot of different localization techniques. In these days the Cell of Origin technique (COO) is the only one that is widely deployed. Its accuracy is quite well for most of today's application. These applications are mostly the first steps of service providers within location-based Internet services. When they have become more popular and service providers have attained more experience with them, new applications will be deployed, which will probably also need more accurate localization techniques. To provide this accuracy, the network operators will probably install the Enhanced Observed Time Difference technique and later update to Assisted GPS technique. Time of Arrival and Angle of Arrival will probably not set up[4], because of their various problems and the little advantage in accuracy they offer.

GPS and Assisted GPS is nowadays also used for tracking and route planning. The problem is, that GPS devices are still relatively big and expensive. When they become smaller and cheaper they probably will be integrated in mobile telephones or PDAs. The provided accuracy of up to 7 m will bring a completely new quality for location-based services.
Their is also research for other localization techniques. One other is to measure the field strength received by the mobile device of the radio signal that is broadcasted by the network base stations. But this technique has still big problems with environmental factors and is therefore far from implementation.

5.5.2 Applications

In these days WAP is the only standard that is widely deployed and used to provide Internet content for mobile devices. It was developed and deployed for this type of application and faces therefore a lot of problems that occurred, for example, with TCP and HTTP. These problems where discussed in section 4. In spite of them NTT DoCoMo with its i-mode successfully uses the HTTP protocol and a special language called C-HTML to provide Internet content. But NTT DoCoMo also said, that it will eventually become compatible to WAP sooner or later[10].
Other techniques, like providing Internet content via voice portals or SMS (see also section 2) can also be very successful as they use common technique and offer greater comfort. It is much easier to tell the system your wishes than typing them with a little mobile telephone keyboard and using shortcuts and keywords. The system can also answer your question via voice and avoid so to display plenty of information on a little screen - probably monochrome. Such techniques are very comfortable, but it is hard to believe that all internet content can provided in this way. WAP also includes the possibility to establish such services with its WTA - Wireless Telephony Application.

References

[1] M. Prasad: Location Based Services;
http://gisdevelopment.net/application/lbs/lbs002pf.htm, September 2002.

[2] V. Wentworth, T. Wrappe: Location Based Services: Oppurtunities & Technologies (slides); http://www.qualcomm.com/brew/ brewtimes/ppt/wentworth_wrappe1.pdf, September 2002.

[3] IEEE Personal Communications, Volume: 7 Issue: 1, Februar 2000, System support for mobile, adaptive applications, Noble, B., Page(s): 44-49

[4] Location Interoperability Forum; TD 201, V.3.0.0; http://www.locationforum.org, Ferbruar 2002

[5] Schiller, Jochen: Mobilkommunikation, 2000

[6] An Introduction to mobile Positioning;
http://bestgsm.virtualave.net/mobposit/mobposit.html

[7] Mobile Positioning; http://www.mobilepositioning.com

[8] Location-Based Services; http://www.mobileinfo.com/LocationBasedServices/resources.htm

[9] heise.de; http://www.heise.de/mobil/artikel/2002/03/04/lbs/

[10] WhatIs.Com; http://www.whatis.com

[11] Wagenbrenner, Stefan: MobileIP; Seminar: Mobile Systems, Talk No. 2

Kapitel 6

VoIP in Wireless Environments

Carsten Schwede

As the Internet grows and grows, it is no longer used to just offer information users research and retrieve, but to connect people all over the world and provide communication in the global village. This is usually done by e-mail, newsgroups, chats or instanst messaging, whereas the basic and more comfortable voice conversation still takes place using standard telephones. This report will introduce protocols and procedures allowing people to place telephone calls and to practice voice communication over Internet Protocol-based networks, in short VoIP. As Internet Procotol itself does not provide real-time communication, Real-time Transport Protocol that does is presented as well as some Internet Protocol-basics like Transmission Control Protocol and User Datagram Protocol. Two of the most common protocols used to implement real-time conversation are H.323 and Session Initiation Protocol. These are described in detail and sample call connection procedures are explained. Furthermore, the implementation of VoIP in wireless environments is introduced as in those the Quality-Of-Service is especially endangered. As it is mandatory for the success of VoIP to achieve a Quality-Of-Service being compatible with the one of todays telephone networks, impacts on Quality-Of-Services are presented as well as solution to compensate drawbacks of packet-based transmissions with real-time requirements.

Inhaltsverzeichnis

6.1 Introduction

As the Internet becomes more and more popular, it becomes a favoured medium of communication. Although social scientists speak already of a communication revolution and influences of e-mail, newsgroups and instant messaging in the real world, one of the oldest forms of communication seems to remain untouched by the Internet hype: voice conversation.
This report gives an insight of todays realizations of voice communication over IP-based networks, in short VoIP.
The usage of already existent network resources enables VoIP to offer voice connections at more favourable prices than telephone providers that realize voice connections over designated wires that have to be maintained. VoIP can be easily integrated in present IP applications and provides even higher degrees of integration. Including voice links on web sites would enable customers surfing an e-shop to speak to the shopownwer directly through the web site to have questions answered. In this case the shopowner wouldn't even need to posess a computer, VoIP could realize the connection between the computer of the websurfer and the usual telephone of the shopwner. It is even possible to connect standard telephones and still benefit from VoIP, as VoIP uses gateways to transmit voice between IP-based and normal telephone networks. Thereby VoIP develops big potentials, market researchers forecast a strong accretion as well as many shares in the telecommunication market since several years. But for some reasons VoIP is not as popular as it should be according to the prognose.
This report will not examine the readiness of todays VoIP products for the telecommunication market, but show up which problems VoIP has to face and how these should be solved to make VoIP competitive. The main goal of VoIP developments at the present time is to measure up with the Quality-Of-Service of public telephone services. Along with protocols and procedures that cover the connection setup und termination, methods to improve VoIPs Qualitiy of Service are presented. This becomes especially important if VoIP is to be used in wireless environments, as these show characteristics that heavily impact VoIPs Quality-Of-Service.

6.2 VoIP in General

Voice communication is usually done by transmitting the voice data over designated wires with constant bandwiths. VoIP realizes the conversation using IP-based networks, but problems arise as IP has not been built for real-time transmission. To transmit data with IP, data has to be divided to packets that have to be rearranged after arrival.
Two main protocols that realize the data transport in packets using IP are TCP and UDP. TCP is used if reliable data transportis required as packets are retransmitted if they are not acknowledged by receiver and additional header informations provide better error recognition. UDP does not guarantee this but provides smaller headers, and needs lesser time and bandwith [3][4].
IP can not gurantee the same packet order at departure and arrival, since IP packets may use different routes for every transmission. To provide continous stream of data that is

needed for voice communication, it is therefore necessary to recover the packet order after arrival and detect packet loss. To make real-time data packets more important than other data like e-mail in the flow of packets, an identification and priorization of packets by routers would make the flow of communication stream more effective. Instead of detecting packet loss by acknowledging every packet, a sequence number included in every packet would expose those missing and increase efficiency. Both is introduced by the following protocol.

Real-Time Transport Protocol

The Real-time Transport Protocol (RTP) was specified by the Intenet Engineering Task Force (IETF) in RFC 1889 [5].
Within the header RTP includes a sequence number to provide the recovery of packet order as well as a payload type to identify included media. The overall RTP header has a length of at least 16 Byte, the sequence number needs 16 Bits, whereas the payload type field has a length of 7 Bits. The content of the payload typee field is specified by the Internet Assigned Number Authority (IANA) and corresponds to the codec used for encoding the payload. For example 0 for G.711 or 4 for G.723.1. The range of 96 until 127 is designated to dynamic use, whereas 32 - 95 is reserved [6]. Besides sequencing and payload identification RTP provides the following functions:

- Frame Identification, to mark beginning and ending of audio and video frames
- Source Identification, to determine originator of frame in multicast session
- Intramedia Synchronization, to synchronize audio and video packets by time stamp

The recovery of packet order and the deriving frame sequence is done using the sequence number and the timestamp. The sequence number starts with an initial random value and increments by one for every RTP packet sent, whereas the timestamp does not need to be monotonic. If one frame needs more than one packet to be transmitted, these packets will have different sequence numbers, but the same timestamp. If more than one frame is transmitted in one packet, the time between those has to be calculated based on more information in the payload, as RTP only provides the timestamp for the first frame.

In addition to RTP RFC 1889 defines the following protocol [5].

Real-Time Transport Control Protocol

Real-time transport control protocol (RTCP) provides control services for RTP. During RTP sessions participants send RTCP messages from time to time to exchange information. These may be addressing information like name or e-mail address as well as information about connection quality. This enables communication partners to switch codec according to available bandwith.

RTCP has different message typees in order to achieve this:

- Sender report
- Receiver report
- Source description
- Bye
- Application-defined RTCP packet

With these the following services are provied:

- Identification, information about participants
- QoS Feedback, like number of lost packets, jitter and Round Trip Time
- Session Control
- Intermedia Synchronization

6.3 H.323

The H.323 standard is a technology used for transmission of real-time audio, video, and data communications over packet-based networks [8]. It is developed by Internet Telecommunications Union - Telecommunication Sector (ITU-T) and specifies protocols components, and procedures necessary to achieve multimedia communication using networks with packet-based transport protocol [7].

6.3.1 Components of H.323

The H.323 standard specifies four kinds of components, which provide point-to-point and point-to-multipoint multimedia communication services:

1. terminals
2. gatekeepers
3. gateways
4. multipoint control units (MCUs)

All components may be integrated in one device, although treated seperately.

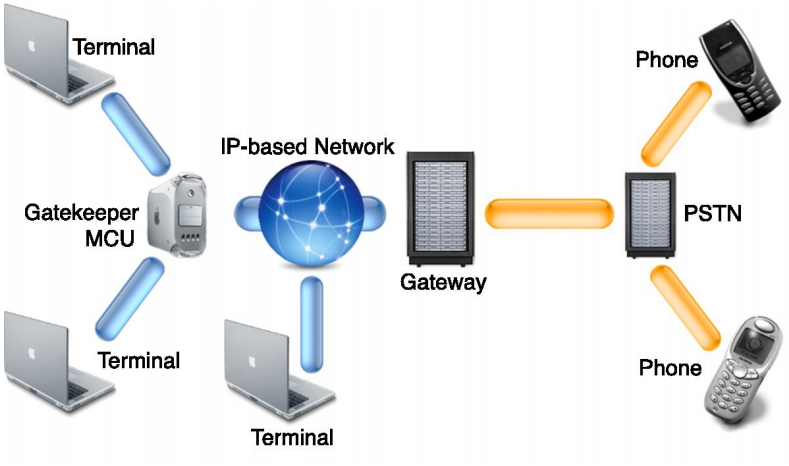

Abbildung 6.1: H.323 components in real network

Terminals

A terminal in H.323 is what could be best described as a multimedia endpoint on which H.323 and the appropriate multimedia applications are implemented. Every terminal has to support at least:

- RTP/RTCP, used as transport protocol for media
- RAS, used for the registration and admission control with gatekeeper
- H.245, used for the exchange of terminal capabilities and the setup of media channels
- H.225, used for the call signaling and the call setup
- G.711 audio codec, used for encoding and decoding audio samples

Other capabilities like video, T.120 data-conferencing and the use of MCU are optional and will not be described further.

Gatekeepers

A gatekeeper is optional within a H.323 network, but has to be used if present. It provides important services like addressing, authorization and authentication of terminals and gateways, bandwidth management, call-routing, and accounting. The H.323 standard defines services the gatekeeper must provide and others which are optional. Gatekeeper are useful to increase performance in network, as they can perform routing descision based on network and gateway load.

Functions the gatekeeper must provide are:

Address Translation

The gatekeeper is used to translate aliases or E. 164 telephone numbers to corresponding network addresses. Therefore it is possible to address a H.323 terminal from within the network with an alias, or with an E.164 telephone number from without.

Admission Control

The gatekeeper uses RAS messages to control the admission of the endpoints in H.323 networks. These are admission request (ARQ), admission confirm (ACF) and admission reject (ARQ). It is possible to reject admission (ARQ) by criteria like available bandwith, or to accept all admission requests.

Bandwidth Control

A very important functions a gatekeeper provides is the support for bandwith control. This is again managed by RAS messages, namely: Bandwith request (BRQ), bandwith confirm (BCF) and bandwith reject (BRJ). H.323 terminals are hereby offered an option to request more bandwith for conferences or higher audio-video quality.

Zone Management

All terminals, gateways and MCU that are registered with the gatekeeper belong to the same zone. All of the functions listed above can be used by H.323 components in its zone.

Optional functions include:

Call-Control Signaling

The gatekeeper may allow the direct exchange of H.225 call-signaling messages between endpoints, but can also route call-signaling messages between these on his own.

Call Authorization

The gatekeeper may accept or reject a call when receiving call-signaling messages from endpoints, based on access-based or time-based restrictions.

Call Management

The gatekeeper may reroute calls for load balancing, therefore the gatekeeper may track active H.323 calls to gather information about used bandwith and routing.

Gateways

A gateway is used to connect incompatible networks. H.323 gateways do not only provide access to other non-H.323 packet-based networks, but also to public switched telephone networks (PSTN) like ISDN. On each side of the networks the gateway connects, it appears as an endpoint terminal of the corresponding network, as it implements the appropriate protocols for call setup, call release and media format.

For example, to perform audio-video communication between H.323 and ISDN network, the gateway would make use of the G.711 audio and H.261 video codec as media formats. These are used by ISDN and supported by H.323 and conversion is not necessary. On the ISDN side, the gateway would run ISDN-specific protocols for call setup and release, whereas on the H.323 it would run H.245 control signaling for exchanging capabilities and H.223 call signaling for call setup and release. H.225 registration, admissions and status (RAS) is used as well for registration with gatekeepers, if present.

Due to this an ISDN network would recognize the gateway as an ISDN terminal, whereas H.323 terminals would think of the gateway as a H.323 terminal endpoint.

Terminals do not differentiate between gateways and other terminals, whereas gatekeepers do know which H.323 endpoints are gateways instead of terminals as this is part of the information in the RAS process between terminal/gateway and gatekeeper.

To provide compatibility to non-H.323 networks the gateway implements call setup und call release protocols according to the networks have to be translated, as well as the conversion of the actual media.

A gateway is optional in communication between H.323 terminals.

Multipoint Control Units

The MCUs provide audio-video conference support with three or more H.323 terminals. An MCU consists of a mandatory Multipoint Controller (MC) and an optional Multipoint Processor (MP). The MC is used for negotiation of terminal capabilities with H.245 and provides conference control. The MP is used for processing media stream including conversion of different codecs and bitrates and provides video multicast capabilities.

6.3.2 Protocols Specified by H.323

H.323 standard is independent of the network and packet protocol, but specifies other protocols in order to provide communication services. H.323 uses UDP instead of TCP to transmit media streams for the following reasons. TCP ackknowledges every packet and would send lost packets again, this would lead to delayed packets which had to be discarded, as they are unusable in the stream of communication. Another disadvantage of TCP is it has to set up connections before transmitting data. UDP does not need this and is therefore faster and provides a 8 Bytes smaller header causing smaller overhead than TCP.

- H.225 registration, admission, and status (RAS)
- H.225 call signaling
- H.245 control signaling

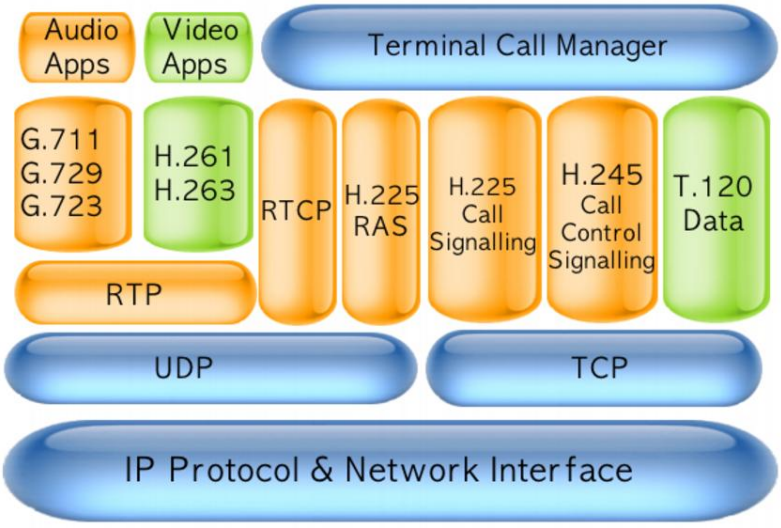

Abbildung 6.2: H.323 Protocol Stack

H.225 Registration, Admission, and Status

Registration, admission, and status (RAS) is used for the communication between endpoints and the gatekeeper. RAS provides messages that can be divided into Requests (xRQ), Confirmations (xCF) and Rejects (xRJ). The following procedures defined in RAS make use of them:

Gatekeeper Discovery

This process defines how endpoints determine which gatekeepers are present and which one they use. If the gatekeeper address is not known to the endpoint, the terminal starts a discovery process by multicasting a gatekeeper request (GRQ). Gatekeepers in reach may respond with a gatekeeper confirm (GCF) or gatekeeper reject (GRJ). If more than one gatekeeper replies with a GCF the endpoint may choose which gatekeeper to use, if no gatekeeper replies, the endpoint may rerequest. Included in GCF is the transport address of the corresponding gatekeepers RAS channel.

Endpoint Registration

If the endpoint is aware of a present gatekeeper that confirmed the GRQ, the endpoint needs to register with the gatekeeper. This is done by the endpoint sending a Registration Request (RRQ) including its transport and alias addresses to the gatekeepers RAS channel transport address, that was discovered in the gatekeeper discovery process. The gatekeeper may respond with a Registration Confirm (RCF) and is now able to resolve the endpoints alias to its transport address, but may also respond with a Registration Reject (RRJ). If an endpoint wishes to unregister with the gatekeeper it sends an Unregistration Request (URJ) to gatekeeper, which responds with an Unregister Confim (UCF).

Endpoint Location

A gatekeeper or an endpoint that wants to determine the contact information of one endpoint by its alias may send a Location Request (LRQ). The endpoints corresponding gatekeeper responds with a Location Confirmation (LCF) message including contact information about the endpoint.

Admissions and Bandwith

Admission Request (ARQ) is sent by endpoints in order to ask gatekeeper for call connection, specifying which bandwith the endpoints wishes to use. Admission Confirm (ACF) is sent by gatekeeper if the gatekeeper allows connection. If gatekeeper cannot provide the requested bandwith it may reduce it and inform endpoint in ACF. Admission Reject (ACF) may be sent by gatekeeper if endpoint is not allowed to perform call connection or if bandwith is unavailable. Bandwith can be requested during call connection using Bandwith Change Request (BRQ). If the gatekeeper is able to provide the desired bandwith, a Bandwith Change Confirm (BCF) is sent, whereas Bandwith Change Reject (BRJ) is responded if desired bandwith cannot be assured.

H.225 Call Signaling

To set up connections between endpoints H.225 call signaling is used. Call signaling is done by exchanging H.225 protocol messages over reliable transport channels (with packet acknowledgement, like TCP). H.225 call signaling messages are directly exchanged between two endpoints if no gatekeeper is present, otherwise the gatekeeper may allow a direct exchange (direct call signaling) during the ACF or route the messages on its own (gatekeeper-routed call signaling).

H.245 Control Signaling

H.245 control signaling is used to exchange H.245 control messages between communicating endpoints. H.245 control messages are used to open and close unidirectional logical channels in RTP/RTCP over which the media is transmitted. They can also include information about one terminals transmit and receive capabilities, e.g. the ability to receive and process incoming media streams. H.245 control signaling also provides Conference Control by appointing MC in case the call is extended to a conference.

6.3.3 Sample Connection Procedure

Theses are the steps involved in setting up a H.323 call, the establishment of the media transport and the call release. This example assumes two H.323 terminals Alice and Bob that are connected to their corresponding gatekeepers Alice-GK and Bob-GK, which both allow direct call signaling.

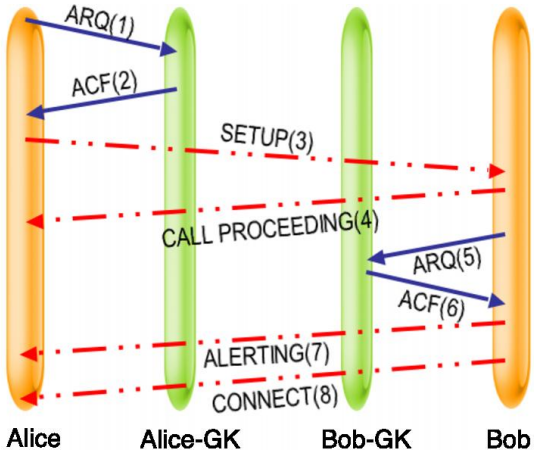

Abbildung 6.3: H.323 Sample Call Setup Procedure

1. In order to register with her gatekeeper Alice-GK, Alice sends an ARQ message to Alice-GK requesting direct call signaling to Bob.

2. Alice-GK looks up if Alice is allowed to perform calls, if direct call signaling can be allowed and if bandwith is sufficient. If so, it will resolve the network address of Bob by his alias and send it along with the ACF to Alice, indicating that direct call signaling is allowed.

3. Alice, now knowing Bobs network address is now able to request a connection with Bob by sending him H.225 call signaling setup message.

4. Bob responds with a H.225 call proceeding message to Alice.

5. As bandwith may be insufficient in Bobs network, or Bob may be unallowed to place phone calls, he has to register with his gatekeeper Bob-GK, therefore he sends ARQ message to Bob-GK.

6. If Bob-GK has no objection, he sends ACF to Bob.

7. Since Bob is now allowed to establish a connection in his zone, he alerts Alice by sending a H.225 alerting message.

8. Alice confirms the establishment of the connection by sending a H.225 connect message to Bob.

The call is now established.

9. Now that the call is established, Alice and Bob need to negotiate which ports to use, a H.245 control channel is established and Alice sends a H.245 TerminalCapabilitySet message including her terminals capabilities to Bob.

10. Bob acknowledges Alice capbilities by sending a H.245 TerminalCapabilitySetAck message.

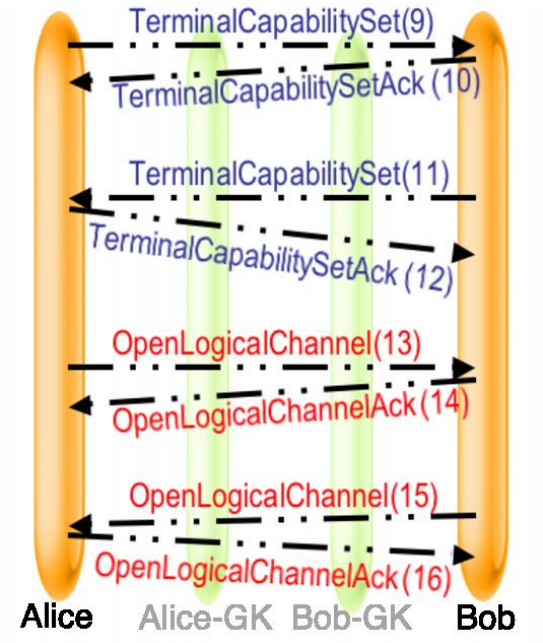

Abbildung 6.4: H.323 Sample Call Setup Capability Exchange and Channel Opening

11. Bob sends his capabilities to Alice by sending a H.245 TerminalCapabilitySet Message.

12. Alice acknowledges Bobs capabilities by sending a H.245 TerminalCapabilitySetAck message.

13. With this information Alice can open a media channel to Bob by sending a H.245 openLogicalChannel message, in which the transport address of the RTCP channel is included.

14. Bob acknowledges the opening of the channel by sending a H.245 openLogicalChannelAck message, which include the RTCP address received from Alice before, as well as the RTP address Bob wants Alice to use for sending the RTP media stream.

15. Bob can now open a media channel to Alice by sending a H.245 openLogicalChannel message, in which the transport address of the RTCP channel is included.

16. Alice acknowledges the opening of the channel by sending H.245 openLogicalChannelAck message, which include the RTCP address received from Bob before, as well as the RTP address Alice wants Bob to use for sending the RTP media stream.

The bidirectional media stream is now established, and RTCP messages are exchanged.

17. Bob sends media stream to Alice using RTP.

18. Alice sends media stream to Bob using RTP.

19. Alice sends RTCP messages to Bob.

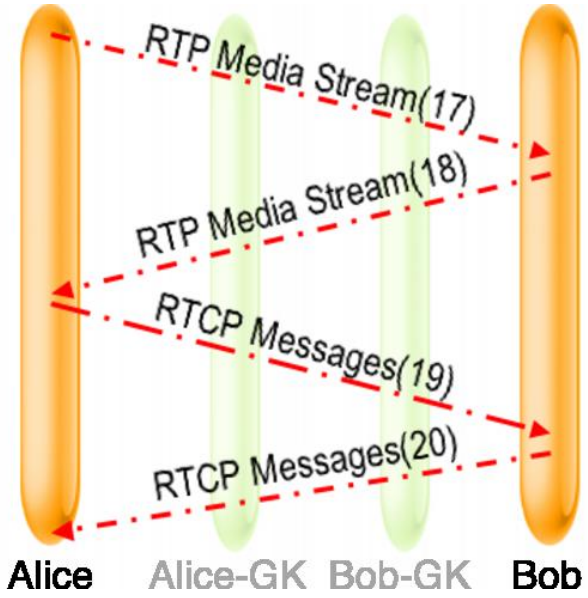

Abbildung 6.5: H.323 Sample Call

20. Bob sends RTCP messages to Alice.

These procedures repeat as long as the call connection continues, during the call connection BRQ may be send by Alice and Bob to their gatekeepers to adapt to bandwith requirements and to change to another codec they have negotiated using the exchange of RTCP messages.

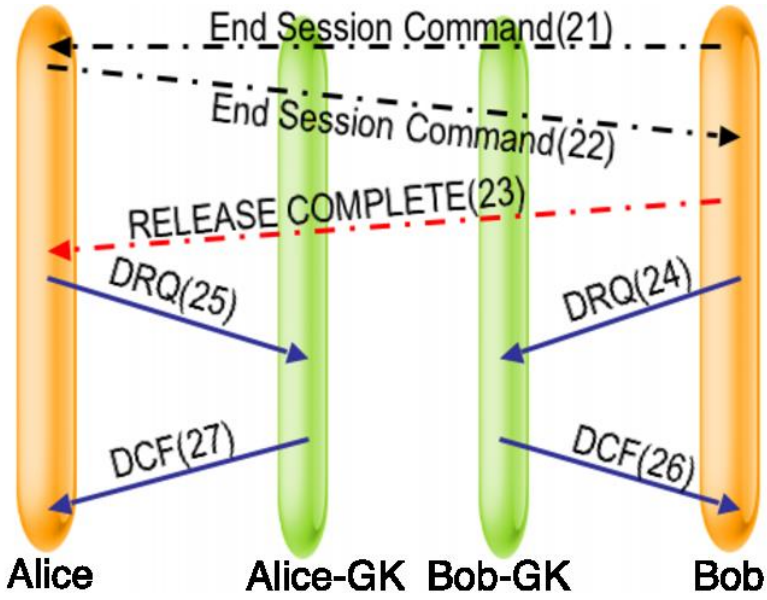

Abbildung 6.6: H.323 Sample Call Release

21. If Alice wishes to terminate the call, she initiates a call release by sending a H.245 EndSessionCommand message to Bob.

22. Bob may confirm the release by sending a H.245 EndSessionCommand message to Alice.
23. Alice now makes the call release complete by sending a H.225 release complete message to Bob.
24. Bob disengages with Bob-GK by sending RAS DRQ to Bob-GK
25. Alice disengages with Alice-GK by sending RAS DRQ to Alice-GK
26. Bob-GK comfirms the disconnect by sending DCF to Bob
27. Alice-GK confirms the disconnect by sending DCF to Alice

The call is now terminated.

Alice-GK and Bob-GK may be one and the same gatekeeper if both reside in the same H.323 zone.

6.4 SIP

SIP (Session Initiation Protocol) is specified in RFC 2543 and is a protocol developed by the IETF (Internet Engineering Task Force) for establishing VoIP connections between two or more users. It is modeled upon text-based Internet protocols like HTTP and SMTP and makes use of the client-server architecture. Requests are generated by the client and sent to the server, which sends a response back to client after processing the request. Therefore SIP does not rely on reliable protocols like TCP, since every request is acknowledged by response. Identification of users is made simple by e-mail like addresses (user@company.com) and the usage of Domain Name System (DNS) [9].
SIP provides protocol mechanism that enables end systems and proxy servers to provide the following services:

- User location, to determine the end system used for communication
- User capabilities, to determine media parameter
- User availability, to determine if called party user is willing to accept call
- Call Setup, to inform called party user of incoming call and establishing call parameters between users
- Call handling, to control the transfer, forwarding and termination of calls
- Number resolving, to resolve numbers to aliases and vice versa
- Terminal capability negotiation, to exchange information of terminals capabilities to provide better Terminal-type selection

SIP is able to initiate conferences using MCUs.

6.4.1 Components of SIP

SIP consists of the following two components:

1. User Agents
2. Network Server

User Agents

A user agent is an endpoint and can be divided into a client and a server part. The client part is called User Agent Client (UAC) and is used to initiate SIP requests whereas the server part is called User Agent Server (UAS) and is responsible for returning responses.

Network servers

- registration server
- proxy server
- redirect server

The registration server receives updated information regarding the current location of one user. Proxy server forward requests to the next-hop sever with more information about the location of the called user, in contrast to the redirect server which does not forward the request but determines and returns the address of the next-hop server.

6.4.2 Protocols

Session Description Protocol

Session Description Protocol (SDP) is specified in RFC 2327 and is used to provide information about the sessions media stream. Transport protocols, media typees and format, as well as bandwith and contact information are exchanged using SDP and are used for negotiation between endpoints [10].

Session Announcement Protocol

Session Announcement Protocol (SAP) is specified in RFC 2974 and is used if conferences are announced. This is done by a SAP announcer that periodically multicasts announcements packets [11].

6.4.3 Messages and Responses

SIP defines the following messages and responses for initiating sessions [9].

- INVITE invites a user to a call
- ACK is used for reliable exchange of INVITE messages
- BYE terminates the connection between users or declines a call
- CANCEL terminates requests or searches for one user
- OPTIONS returns information about call capabilities
- REGISTER registers the users current location to SIP registration server
- INFO is used for signalling during initiated session
- 1xx Informational; e.g. 100 Trying, 180 Ringing
- 2xx Succesful; e.g. 200 OK, 202 Accepted
- 3xx Redirection; e.g. 302 Moved Temporarily
- 4xx Request Failure; e.g. 404 Not Found, 482 Loop Detected
- 5xx Server Failure; e.g. 501 Not Implemented
- 6xx Global Failure; e.g. 503 Decline

6.4.4 Sample Initiation Procedure

When Alice wants to call Bob, Alice has to initiate the call with an INVITE request. This single request contains the following information that enables Bob to immediately establish a connection with Alice:

- Alice user address
- TCP/UDP port Alice wishes to use
- Audio codec Alice wishes to use

As we need one connection for each direction, Bob sends an 200 OK Response, included are the following information:

- Bob user address
- TCP/UDP Port Bob wishes to use
- Audio codec Bob wishes to use

Alice is now able to establisch a connection to Bob and responds with an ACK response.

The initiation procedure is completed.

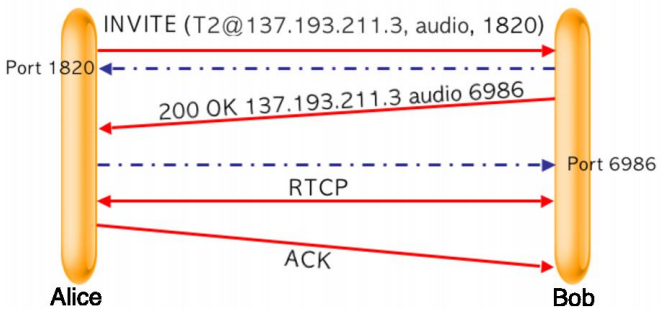

Abbildung 6.7: SIP Sample Session Initiation

6.5 H.323 vs. SIP

It should become clear that SIP is structured more simply and needs lesser information exchange to achieve a call connect. This applies only to the first version of H.323, H.323v2 provides faster call setups as fast as SIP. Signaling in H.323 was built like signaling procedures in PSTN for compatibility reasons. Therefore H.323 is more complex but more complete, and offers built-in compatibility to many other networks as well. H.323 provides complete conference capbilities with audio, video and data communication as well as control mechanism that allow synchronization of these. SIP is able to perform audio and video communication but lacks control procedures. This makes H.323 as one all-in-one package not very modular, whereas SIP can be extended in every manner. Application building should be easier with SIP as it uses plain-text messages for information exchange in contrast to H.323 that encodes messages in binary format, but can achieve lesser message sizes. SIP supports multicast messages and uses DNS, whereas H.323 has to maintain it's own database to accomplish user addressing [12].

	H.323	**SIP**
Intention	Complete package providing multimedia communication over IP-based networks	Modular package defining the initiation of sessions
Message Format	Binary	ASCII
Adressing	H.323 alias, E.164 number	user@host.com
PSTN Compatibility	PSTN-like signaling procedures and protocols provided	PSTN-like signaling procedures have to be self-implemented
Conferencing	Audio, video and data conferences supported	Audio, video conferences supported
Encryption	Provided by H.235	SSL, PGP, S/MIME, and other

6.6 Audio Codec Overview and MOS

An audio codec is responsible for the encoding and decoding of audio signals.
The compression ratio of one codec depends on its frame size, as more voice data is acquired, the compression gets more efficient but frame size adds to the overall delay.
As sound quality is very subjective, a value has to be introduced that makes audio codecs comparable regarding their voice quality.
This is achieved by the Mean Opinion Score (MOS). The MOS of one codec is determined by the following steps. A preselected voice-sample is encoded with codec and played to a statistical mixed group. Every member of this group rates the quality of the encoded sample based on a scale from 1 to 5:

1. Means bad audio quality and the member was unable to understand.
2. Means poor audio quality and the member was able to understand with considerable effort.
3. Means fair audio quality and the member was able to understand with moderate effort.
4. Means good audio quality and the member was able to understand with attention.
5. Means excellent audio quality and the member was able to understand even if completely relaxed.

The MOS is calculated as the averages of the groups ratings.

The following table lists popular codecs and their corresponding bandwith needs, frame sizes and associated MOS.

Codec	**kbit/s**	**frame size**	**MOS**
G.711	64	0.125 ms	4.1
G.726	32	0.125 ms	3.85
G.728	16	0.625 ms	3.61
G.729	8	10 ms	3.7
G.723.1	5.3	30 ms	3.65

6.7 Adaptions for Wireless VoIP

Although VoIP application can be deployed in wireless environment without any adaption, QoS becomes even more important as wireless networks show characteristics that impact QoS. Those of most wireless network protocols like 802.11b are Collision Avoidance (CA) and Packet Acknowledgement [13]. Unfortonately these affect VoIP in a harmful way, as they cause additional delays. As these procedures are integrated in the network protocol (OSI Layer 2) they are out of VoIPs range of control.

6.7.1 Quality-Of-Service

PSTNs have evolved to provide optimal service for time-sensitive voice applications and although the voice quality is not perfect user have become accustomed to the PSTN level of quality. IP network were supposed to support non-real-time application like file transfer, e-mail or WWW. These services may require high bandwith but are not sensitive to delay or delay variation. As users of traditional telephone networks are used to the PSTN level of quality, VoIP has to offer the same to become a coequal solution and to be able to integrate into present PSTNs as smooth as possible. To compete with PSTN, VoIP has to cope with three major impacts on its QoS [14].

Delay

Delay can be devided to the following three.

1. accumulation delay
2. processing delay
3. network delay

The first delay is caused by the need to collect frames of audio samples that can be compressed by the audio codec. The codec can compress more efficiently if more audio frames have been collected. The frame size is determined by the audio codec, current codecs frame sizes ranges from 0.125 ms (G.711) to 30 ms (G.723.1).

Processing delay describes the delay caused by:

1. encoding audio samples and collecting into packet (5 ms - 30 ms)
2. extracting and decoding encoded audio samples from packet(5 ms - 10 ms)

This delay depends on the algorithm used in the codec, the computing performance.
Network delay is caused by the distance between sender and receiver, the physical medium and properties of the network in use, its bandwith, routing capabilities, bit error rates. Wireless LANs do have considerably higher bit error rates compared to other LAN relalizations as can be seen in the following table.

LAN	**Bit Error Rate**
10 Base T	10^{-10}
10 Base 5	10^{-8}
10 Base 2	10^{-7}
Wireless	10^{-3} - 10^{-5}

If the overall delay exceeds 50 ms echoing is perceivable, and above 250 ms talker overlap becomes a problem, i.e. stepping on the other talkers speech.

Delay Jitter

If every packet that leaves the sender would need a fixed time to be transmitted to the receiver, the packets would have the same order at arrival as at departure. Since the delay is mutable the packet order might be disturbed. The variability in arrival time of packets is called Delay Jitter.
To achieve a continous stream of packets, incoming packets need to be buffered in a Delay Jitter Buffer. If the buffer is filled up, the packets included are being sorted and put out. As the first arriving packet is not forwarded until the buffer is full, this packet will be delayed according to the size of the buffer.
Measurements showed that Delay Jitter Buffer should buffer less than 30 ms.

The overall delay should be less than 150 ms.

Packet Loss

Especially in wireless networks packets might get lost under peak load, handoffs, or during periods of traffic congestion.

There are three procedures that deal with the problem of packet loss.

1. Packet Replay
2. Packet Extrapolation
3. Packet Redundancy

Packet Replay simply replays the last received packet. This can only be done if lost packets occure occasionaly and not in a row.

Packet Extrapolation tries to guess information in the missing packet by analysing packets received before. Fourier Transform is used to determine frequency spectrum and corresponding amplitude to extrapolate waveform. This procedure is experimental, as Fourier Transformation means intensive numerical calculations and cannot be used to regenerate packets lost in a row.

Packet Redundancy include information of n-th packet in the (n-1)-th packet. This information can either be an identical copy of the following packet or a re-encoded copy with lowered bitrate to decrease packet size.

Measurements showed that packet loss should be less than 1%, but meaning can be understood at packet loss up to 10%.

A procedure that reduces the overall number of packets is silence suppression. During the call a basic noise level is measured. If volume level drops down and reaches noise level, no audio date is sent during that period and the receipient is informed that sender is not speaking. The receipient may now create artificial noise corresponding to the noise level measured before. As human conversation is mainly half-duplex (i.e. one person speaking at a time), silence suppression can approximately reduce bandwith requirements to 50%.

6.7.2 Overhead

The Overhead describes the size of the headers in one packet related to its payload. This report pointed out that a lot of protocols are needed to provide the desired services. The following overview shows a sample protocol setup with its corresponding header sizes.

1. Ethernet 802.11b DSSS (24 Bytes)
2. IPv4 (20 Bytes)
3. UDP (8 Bytes)
4. RTP (12 Bytes)

The overall header size sums up to 64 Bytes, which results in an overhead of 24%, an average payload size of 200 Bytes is assumed.

Because there is a high degree of redundancy in the header fields of consecutive packets the headers can be compressed. Algorithms that provide header compression do maintain a context based on the last succesfully decompressed header. Compressed headers do only include changes to this context. If packets get lost, the context cannot be properly updated and the decompression of following headers fails. Therefore mechanisms are required, that enable the compression algorithm to install a context, to detect when it's out of date, and to repair the context if needed.

Algorithms that provide header-compression are among others:

6.7.3 Compressed RTP

Compressed RTP is standardized by IETF in RFC 2508 and can compress the IPv4/UDP/RTP header fields from 40 Bytes down to a minimum of 2 Bytes. CRTP relies on an upstream link over which it can send requests in order to update headers. While the context is out of date, all packets received are lost since header cannot be decompressed [15].

Simulations showed that the packet-loss rate for CRTP is about three times higher than without. Therefore another algorithm was inventend:

6.7.4 Robust Checksum-based Header-compression

Ericsson focused on local context repair and developed Robust Checksum-based Header-compression (ROCCO). Included in the ROCCO header is a checksum A generated by the original uncompressed header. After decompression, ROCCO will calculate a checksum B of the decompressed header and compare checksum A with checksum B. Furthermore ROCCO includes information in headers on how header fields have changed. ROCCO will try to reconstruct header with modified header field if reconstruction fails. ROCCO supports different compression profiles for different channel conditions and RTP streams and can hereby adjust header compression [16].

6.8 Conclusion

This report presented a short overview over VoIP at the present time. Protocols like H.323 and SIP were introduced. By now, it is not possible to predict which protocol will prevail. SIP becomes more and more popular, but H.323 settled its slow call setup by newer versions. H.323 might be the answer for those looking for a built-to-run solution with audio, video and data communication and compatibility to PSTNs. Developers that need clear and modular structures and expandibility might be better off with SIP. As for now, the question is not which protocol to choose, as both do not provide the same quality as PSTN due to network reasons and the nature of IP. As more and more networks develop that offer more and more bandwith, the focus is to set on real-time capabilities. The reason VoIP still cannot compete with PSTNs is not the audio quality itself but the QoS degraded by delays and jitter. As wireless network in particular suffer these problems the fast evolution of wireless network is to be observed, but the popularity of VoIP will rise instantly if its QoS becomes equal to PSTNs. Nevertheless VoIP points out interesting possibilities that cannot be neglected by todays telephone providers. As soon as they stop thinking of VoIP as a menace and instead support the developement VoIP wil take a big step forward. The convergence of electronic communication services to one single layer is a process that can be observed since some years and will not stop, until television, audio-video conferencing and todays Internet are just ones and zeros transmitted through one and the same network.

References

[1] Cisco Systems
Voice-over-IP Overview
(*http://www.cisco.com/univercd/cc/td/doc/product/access/acs-mod/1700/1750/1750voip/intro.htm*)

[2] IEC
Voice and Fax over Internet
(*http://www.iec.org/online/tutorials/vfoip/topic01.html*)

[3] M. del Rey
Transmission Control Protocol, RFC 793
(*http://www.ietf.org/rfc/rfc793.txt*)

[4] J. Postel
User Datagram Protocol, RFC 768
(*http://www.ietf.org/rfc/rfc768.txt*)

[5] H. Schulzrinne, S. Casner, R. Frederick, V. Jacobsen
RTP: A Transport Protocol for Real-Time Applications, RFC 1889
(*http://www.ietf.org/rfc/rfc1889.txt*)

[6] Internet Assigned Number Authority
RTP Payload typees (PT) for standard audio and video encodings
(*http://www.iana.org/assignments/rtp-parameters*)

[7] OpenH323
A Primer on the H.323 series standard
(*http://www.cis.ksu.edu/ deep/h323/home.html*)

[8] IEC
IEC: H.323
(*http://www.iec.org/online/tutorials/h323/topic01.html*)

[9] M. Handley, H. Schulzrinne, E. Schooler, J. Rosenberg
SIP: Session Initiation Protocol, RFC 2543
(*http://www.ietf.org/rfc/rfc2543.txt*)

[10] M. Handley, V. Jacobsen
SDP: Session Description Protocol, RFC 2327
(*http://www.ietf.org/rfc/rfc2327.txt*)

[11] M. Handley, C. Perkins, E. Whelan
Session Announcement Protocol, RFC 2974
(*http://www.ietf.org/rfc/rfc2974.txt*)

[12] Packetizer, Inc.
H.323 versus SIP: A Comparison
(*http://www.packetizer.com/iptel/h323-vs-sip/*)

[13] Kanoksri Sarinnapakorn
IEEE 802.11b High Rate Wireless Local Area Networks
(*http://alpha.fdu.edu/ kanoksri/IEEE80211b.html*)

[14] Cisco Systems
Quality-Of-Service: QoS
(*http://www.cisco.com/warp/public/784/packet/oct02/pdfs/qos.pdf*)

[15] S. Casner, V. Jacobson
Compressing IP/UDP/RTP Headers for Low-Speed Serial Links, RFC 2508
(*http://www.ietf.org/rfc/rfc2508.txt*)

[16] L.-E. Jonsson, M. Degermark, H. Hannu, K. Svanbro
RObust Checksum-based header COmpression (ROCCO)
(*http://www.ludd.luth.se/users/larsman/rocco/drafts/draft-ietf-rohc-rtp-rocco-01.txt*)

List of abbreviations

CRTP	Compressed RTP
DNS	Domain Name System
IANA	Internet Assigned Number Authority
IETF	Internet Engineering Task Force
IP	Internet Protocol
ISDN	Integrated Services Digital Network
ITU-T	International Telecommunications Union - Telecommunication Sector
MC	Multipoint Controller
MCU	Multipoint Control Unit
MP	Multipoint Processor
OSI	Open Systems Interconnection
PSTN	Public Switched Telephone Network
QoS	Quality-Of-Service
RAS	Register, Admission and Status
RFC	Read For Comment
ROCCO	Robust Checksum-based Header-compression
RTCP	Real-time Transport Control Protocol
RTP	Real-time Transport Protocol
SAP	Session Announcement Protocol
SDP	Session Descrition Protocol
SIP	Session Initiation Protocol
TCP	Transmission Control Protocol
UDP	User Datagram Protocol

Kapitel 7

Security in Wireless Networks

Mohamed Kallel

Diese Arbeit behandelt die Sicherheit von drei Funksystemen: Wireless LAN *(IEEE 802.11),* Bluetooth *und GSM.*

IEEE 802.11 zeichnet sich durch das zusätzliche Sicherungsprotokoll Wireless Equivalent Privacy *(WEP) aus. WEP war dazu gedacht, im Wireless LAN eine ähnliche Sicherheit zu bieten wie es in drahtgebundenen LANs aufgrund der festinstallierten Kabel der Fall ist. Wir werden sehen, dass WEP große Sicherheitslücken hat und keinen verlässlichen Schutz bietet. Bluetooth als typisches* Wireless Personal Area Network *(WPAN) dient in erster Linie als Kabelersatz und ist für eine Kommunikation von Computerperipherie und Kleingeräten entworfen worden. Das Sicherheitverfahren in Bluetooth ist auf zwei Schlüsseln basiert: dem* Link key *und dem* Encryption Key*. Aber dies hat auch seine Sicherheitslücken und für Anwendungen die ein besonderes Maß an Sicherheit verlangen, besteht noch Bedarf an Nachbesserung. Das GSM-Netz (*Global System for Mobile communication*) hat sich in den letzten zehn Jahren zum Standard der drahtlosen Kommunikation entwickelt. Die Einführung der GSM-Technologie sollte keine zusätzlichen Sicherheitsrisiken gegenüber dem Festnetz mit sich bringen. Dafür werden die Algorithmen A3, A5 und A8 benutzt, um für die Sicherheit von GSM zu sorgen. trotzdem gibt es bei GSM auch Sicherheitslücken.*

Die Sicherheitsverfahren, die Sicherheitslücken, die Gegenmaßnahmen und Verbesserungen dieser drei Funksysteme werden wir in dieser Arbeit näher betrachten.

Inhaltsverzeichnis

7.1 Einleitung

Wireless Networks haben immer mehr Bedeutung in unserem Leben. Sie bieten mehr Mobilität sowie mehr Bequemlichkeit als normale Netzwerke.
Nehmen wir beispielsweise ein WLAN. Anstatt zum Beispiel in einem Firmengebäude in die Wände Löcher für die Kabel zu bohren, reicht es, eine zentrale Funkstation, *Access Point*, an das Strom- und lokale Netz anzuschließen, in den Rechner die Funknetz-Karte einzubauen und die Software zu installieren. Dieser *Access Point* übernimmt dann die Versorgung mit der drahtlosen Netzanbindung und dient als eine Ethernet-Brücke. Er schickt die Kommunikationen zum passenden Netz, entweder das verdrahtete Netz, oder das drahtlose Netz.
Es genügt also, sich in ausreichender Nähe zum *Access Point* zu befinden, um eine Kommunikationsverbindung aufzubauen. Schon die Strasse vor dem Firmengebäude ist oft ausreichend versorgt. Aus dieser Tatsache muss das WLAN gesichert werden, weil die Informationen frei durch die Luft gesendet werden und zum Einbruch ins WLAN eine Antenne und eine ausreichende Nähe zum *Access Point* reichen. Ein Beispiel ist der Parkplatzangriff (s. Abb. 7.1), in dem der Angreifer auf dem Parkplatz der Firma sitzt und Zugang zum internen Netzwerk der Firma hat. Solche Angriffe können fatale Folgen für die Firma haben.

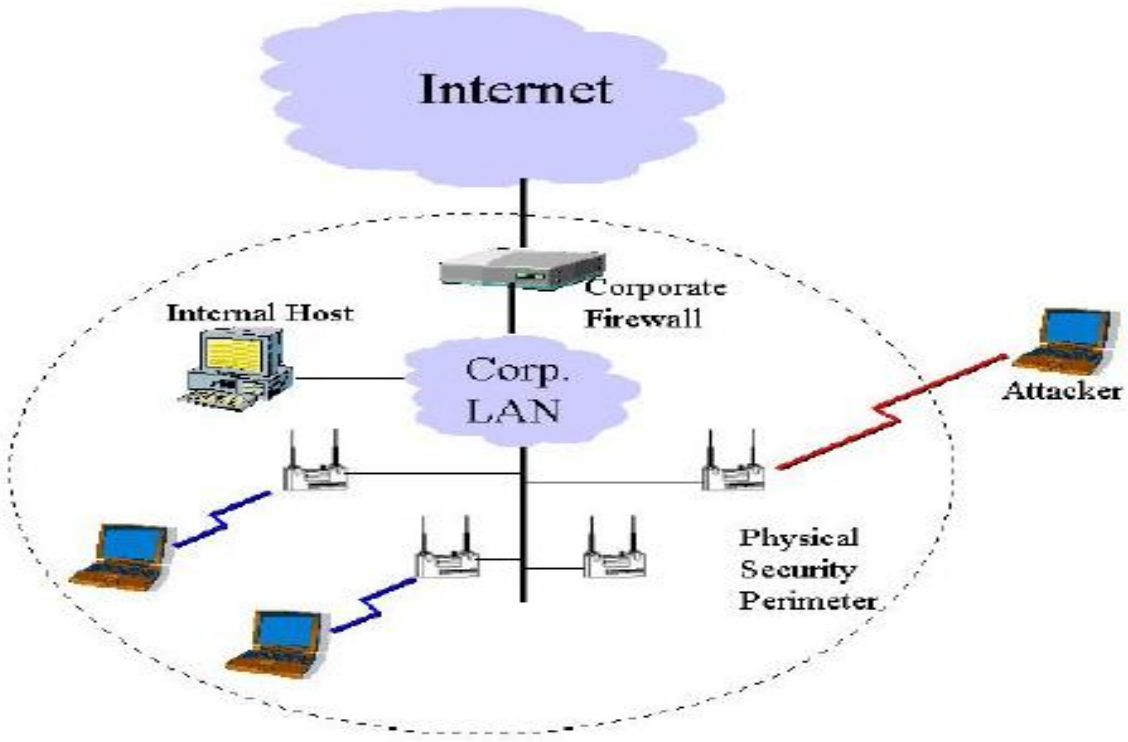

Abbildung 7.1: Der Parkplatzangriff [8]

Um die verschiedenen *Wireless Networks* gegen solche Angriffe zu schützen, wurden Sicherheitsverfahren entwickelt. Diese Verfahren sollen folgende Ziele erreichen [5]:

Authentisierung: Nur die Personen, die autoriziert sind, bekommen Zugang zum Netzwerk.
Das Authentisierungprotokoll, welches häufig verwendet wird, ist das *Challenge-Response* Verfahren. Dieses Verfahren läuft folgendermaßen ab: Zwei Geräte-das Erste will eine Verbindung zum Zweiten bauen. Das zweite Gerät spielt jetzt die Rolle des Prüfers und schickt einen *Challenge* zum Ersten, der diesen *Challenge* beantwortet und eine Antwort züruckschickt. Der Prüfer wertet die Antwort aus und entscheidet, ob das erste Gerät Zugang erhält.

Vertraulichkeit: Es soll verhindert werden, dass unautorizierte Personen Zugriff auf Informationen haben, wenn diese im Klartext übertragen werden. Hierzu werden Veschlüsselungsverfahren benutzt, das heißt die Informationen werden nicht als Klartext gesendet, sondern sie werden mit Hilfe eines Verschlüsselungsalgorithmuses verschlüsselt und als Ciphertext übertragen.
Beispiele für Verschlüsselungsalgorithmen sind der E0-Algorithmus von Bluetooth oder RC4-Algorithmus von WEP (das ist die Abkürzung für *Wired Equivalent Privacy* und bezeichnet das Verschlüsselungsverfahren von *IEEE 802.11*, welches eigentlich -WEP ist schon gebrochen worden- ein Abhören des drahtlosen Datentransfers durch Unbefugte verhindern soll).

Datenintegrität: Die Nachricht, die geschickt wird, wird während des Transportes nicht geändert, ohne dass eine solche Änderung nicht festgestellt wird.
Ein Beispiel: in *IEEE 802.11* (das zur Zeit am weitesten verbreitete WLAN-Verfahren) hat jedes Datenpaket eine CRC-32 (*Cyclic Redundancy Check*) Prüfsumme, die zeigt, ob ein Paket während der Übertragung verändert wurde. Vor dem Senden eines Pakets wird seine CRC-32 berechnet und mit dem Paket verschlüsselt. Im Ziel werden die verschlüsselten Daten entschlüsselt, die CRC-32 neu berechnet und mit der CRC-32, die in dem Packet geschickt wurde, verglichen.

Als Vertreter für *Wireless Networks* werden in den folgenden Kapiteln WLAN (*IEEE 802.11*), Bluetooth und GSM genauer betrachtet.

7.2 Sicherheitsverfahren in *Wireless Networks*

7.2.1 Sicherheit im WLAN (*IEEE 802.11*)

Authentisierung: Um Nachrichten in einem WLAN übertragen zu können, muss sich eine Station zunächst bei dem *Access Point* anmelden. Hierfür sind vom *802.11* Standard

die beiden Methoden *Identify-based* und *Challenge-Response* gedacht (s. Abb. 7.2).

Bei der *Identify-based* Methode braucht die Station nur den *SSID*(*Der Service Set Identifier* ist eine alphanumerische Kennung, die im *Access Point* eingestellt ist)einzugeben. Es gibt zwei Arten der Authentisierung:

- **Die *Open System* Authentisierung:** Die Station bekommt den Zugang, wenn sie sich einfach mit *NULL* als *SSID* anmeldet.
- **Die *Closed System* Authentisierung:** Die Station muss sich hier mit einem gültigen *SSID* anmelden, um Zugang zu bekommen.

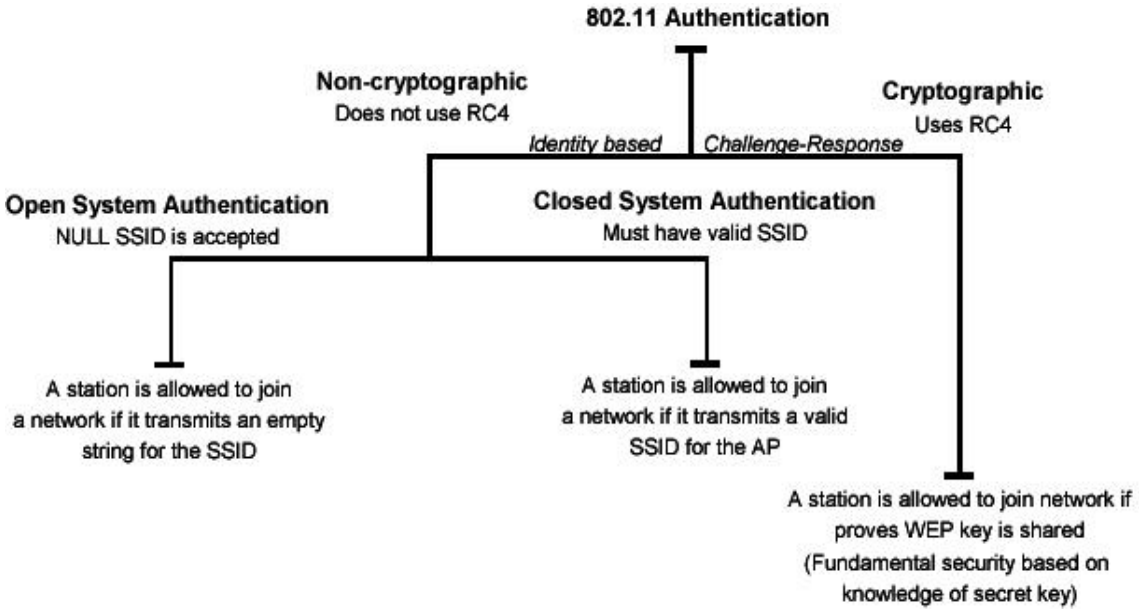

Abbildung 7.2: 802.11 Authentication techniques [1]

Bei der *Challenge-Response* Authentisierung oder auch *Shared key* Authentisierung (s. Abb. 7.3) läuft die Anmeldung einer Station bei einem *Access Point* folgendermaßen ab: die Station fordert eine *Challenge* beim *Access Point* an. Dieser generiert eine zufällige Zahl und schickt sie zur Station. Sie antwortet dann, indem sie diese *Challenge* mit dem gemeinsamen Schlüssel (*Shared key*) nach dem RC4 Algorithmus verschlüsselt. Dieser Schlüssel wird der Station außerhalb des Netzwerks mitgeteilt. Der *Access Point*, der auch diesen Schlüssel hat (deswegen auch der Name *Shared Key*), entschlüsselt die Antwort und verifiziert dann diese mit seiner generierten Zufallszahl und reagiert entsprechend-entweder mit Einlass oder Abweisung.

Abbildung 7.3: Shared-key Authentication Message Flow [1]

WEP-Protokoll: Ein 802.11 Datenpaket besteht aus einem *Header*, gefolgt von Daten (die man übertragen will) und einer CRC-Prüfsumme. Vor der Übertragung wird dieses Datenpaket noch mit WEP verschlüsselt [13].

Es läuft alles wie folgt ab [3] (s. Abb. 7.4):

1. Prüfsumme berechnen:
   Zuerst wird die CRC-Prüfsumme c(M) der Nachricht M berechnet. Der Klartext P sieht dann so aus: P=<M,c(M)>.

2. Verschlüsselung:
   Der Klartext P wird in diesem Schritt mit Hilfe des RC4 Algorithmus' verschlüsselt. Dafür Braucht der *RC4* zwei Angaben: den *Shared Key* K und den *Initialisierungsvektor* (IV).
   Der IV ist ein 24 bit langes Feld. Er erweitert den *Shared Key* und sorgt dafür, dass jedes Paket einen anderen *RC4 Key* bekommt. Er wird unverschlüsselt mit dem Paket geschickt.
   Mit dem IV und dem *Shared Key* erzeugt der *RC4 Algorithmus* eine "unendlich" lange Folge von Datenbits RC4(IV,K), die zur Verschlüsselung vom Klartext P verwendet wird, indem RC4(IV,K) und P mit XOR verknüpft werden. Wir erhalten dann den Ciphertext C :

$$C = P \oplus RC4(IV, K)$$

3. Übertragung und Entschlüsselung:

   Nach der Verschlüsselung wird das Datenpaket geschickt. Ein mal ins Ziel gekommen, wird der selbe RC4(IV,K), der beim Sender generiert wurde, regeneriert und

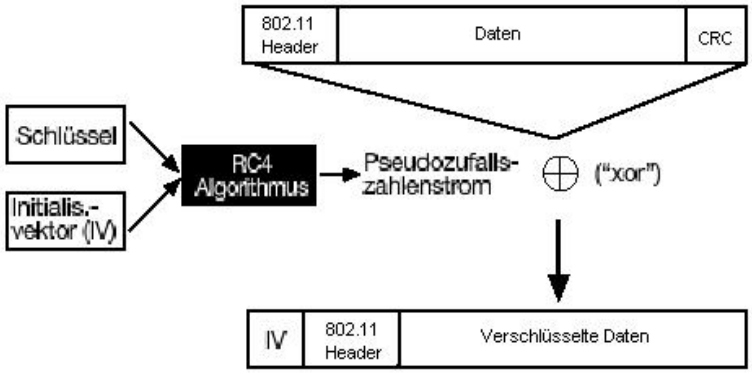

Abbildung 7.4: WEP-Verschlüsselungsverfahren [13]

mit dem Ciphertext C mit XOR verknüpft. So bekommt man den Klartext P', der normalerweise der ursprüngliche Klartext P sein sollte.

$$P' = C \oplus RC4(IV, K) = (P \oplus RC4(IV, K)) \oplus RC4(IV, K) = P.$$

Der entschlüsselte Klartext P' hat auch die Form P'=<M',c'>. Zunächst wird die Prüfsumme c(M') berechnet und mit der empfangenen Prüfsumme c' verglichen. Dies stellt sicher, daß nur Pakete mit einer gültigen Prüfsumme durch den Empfänger angenommen werden.

Aber WEP ist seit ca. einem Jahr gebrochen worden und deshalb in dieser Form unsicher.

7.2.2 Bluetooth Security

In Bluetooth werden drei verschiedene Sicherheitsbetriebsarten unterschieden [1]:

- **Mode 1:** Das ist der unsicherste Zustand, indem ein Bluetooth Gerät keinerlei Sicherheitsprozeduren einleitet. In diesem Zustand erlaubt es anderen Geräten mit ihm eine Verbindung aufzunehmen.
- **Mode 2:** Auf dem *Service-Level* erzwungene Sicherheit.
- **Mode 3:** Auf dem *Link-Level* erzwungene Sicherheit.

Der Unterschied zwischen Mode 2 und 3 ist, dass beim Mode 3 die Sicherheitsprozedur bereits initialisiert wird, bevor der Kanal aufgebaut wird.

In fast allen Bereichen, in denen Bluetooth zur Anwendung kommt, werden private Daten des Benutzers verwaltet oder zum Teil benutzt. Um sie vor Missbrauch zu schützen haben die Entwickler von Bluetooth verschiedene Mechanismen entworfen und in das Bluetooth-Umfeld eingebracht. Die drei Bluetooth Sicherheitsmechanismen sind das Key Management, die Verschlüsselung und die Authentisierung.

Key Management: Bluetooth verwendet unterschiedliche Schlüssel. Der Wichtigste ist der ***link key***. Dieser dient sowohl als Grundlage in die Verschlüsselung, als auch zur späteren Authentisierung zwischen zwei Geräte. Die vier verschiedenen Arten des *link keys* sind 128-bit lang und heißen [5]:

- ***Initialisation key:*** Wenn zwei Geräte sich zum Erstenmal "begegnen", müssen sie einen gemeinsamen Schlüssel vereinbaren, falls eine verschlüsselte Übertragung gewünscht ist: Als Erstes wird ein init key generiert. Dieser wird von beiden Geräten mittels der Funktion E22 aus einem geheimen PIN, der Device Adresse des Geräts und einer Zufallzahl berechnet und anschliessend überprüft. Der PIN ist eine Zahl, die in beide Geräte eingegeben wird und zwischen 8 und 128 Bits lang sein muss. Wird kein PIN gewählt, wird der default PIN 0 verwendet. Die Zufallszahl wird unverschlüsselt übertragen. Die Device Address ist durch vorangehende, unverschlüsselte Kommunikation bekannt [9].
- ***Unit key:*** Dieser Schlüssel wird bei der Installation des Bluetooth-Gerätes einmalig erzeugt und im Speicher des Gerätes hinterlegt und nicht mehr geändert. Er wird mittels der Funktion E21 aus einer Zufallszahl und der Gerät-Adresse gebildet. Die gleiche Funktion E21 wird benutzt, um den *Combination key* zu berechnen.
- ***Combination key:*** Dieser Schlüssel wird während des Initialisierungsprozesses von 2 miteinander kommunizierenden Geräten erzeugt. Diese Schlüsselart kommt zum Einsatz, wenn beide kommunizierende Geräte über genügend Speicher verfügen. Er bietet eine höhere Sicherheit als ein Unit Key.
- ***Master key:*** Dieser Schlüssel wird für Multi-Point Verbindungen generiert, damit alle Geräte den gleichen *Encryption Key* verwenden.

Es gibt noch einen wichtigen Schlüssel: den ***encryption key***. Dieser Schlüssel, der beim Verschlüsselungsalgorithmus benutzt wird, wird mit einem internen Key-Generator (KG) erzeugt. Dieser Key-Generator berechnet die *encryption keys* aus dem *link key*, der Zufallszahl (EN_RAND) und dem ACO-Wert. Der ACO Parameter, eine 96-bit *Ciphering Offset Number*, wurde während des Authentisierungverfahrens erzeugt (s. Abb. 7.5).

Authentisierung: Das Bluetooth Authentisierungsschema benutzt eine *Challenge-Response* Strategie, die durch ein Protokoll prüft, ob die andere Partei den geheimen Schlüssel kennt. Da das Protokoll symmetrische Schlüssel benutzt, beruht eine erfolgreiche Authentisierung darauf, dass beide Geräte sich den selben Schlüssel teilen.

Die Schritte im Authentisierungsprozeß sind die folgenden (s. Abb. 7.5) :

1. **Schritt 1:** Der *Claimant* übermittelt seine 48-bit Adresse (BD_ADDR) an den Prüfer.
2. **Schritt 2:** Der Prüfer sendet eine Zufallszahl (AU_RAND) als *Challenge* zurück.

3. **Schritt 3:** Beide Parteien benutzen die Authentisierungsfunktion E1 mit dieser Zufallszahl, der Geräteadresse und dem aktuellen *link key*, um die Antwort zu berechnen.

4. **Schritt 4:** Der *Claimant* schickt seine berechnete Antwort (*SRES*) zum Prüfer.

5. **Schritt 5:** Der Prüfer vergleicht den *SRES* vom *Claimant* mit seinem berechneten *SRES*.

6. **Schritt 6:** Wenn die zwei 32-bit *SRES*-Werte gleich sind, setzt der Prüfer den Verbindungaufbau fort.

Wenn die Authentisierung fehlschlägt, muss eine gewisse Zeit vergehen bis ein weiterer Versuch gestartet werden kann. Diese Zeit erhöht sich exponential, um wiederholte Versuche zu verhindern.

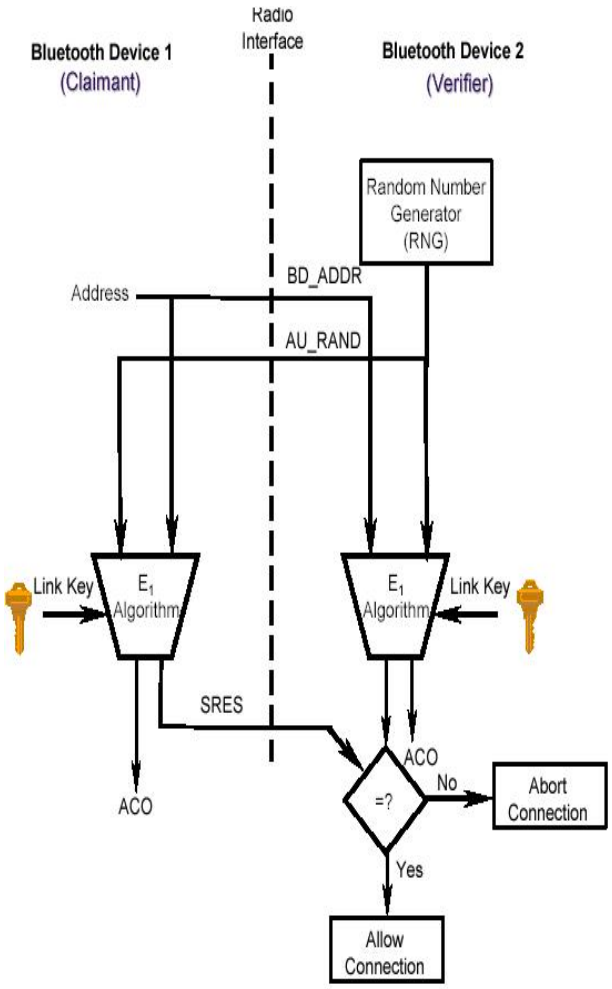

Abbildung 7.5: Bluetooth-Authentisierung [1]

Verschlüsselung: Das Bluetooth-Verschlüsselungsverfahren basiert auf einer *Cipher stream* E0. Ein erzeugter Keystream wird mit den Payload-bits XORed und zum anderen Gerät gesendet. Dieser Keystream wird mit einem Verschlüsselungsalgorithmus erzeugt, der auf linear feedback shift registers basiert (LFSR). Die Verschlüsselungsfunktion nimmt die Master-Adresse (BD_ADDR), die Zufallszahl (EN_RAND), eine Slot-Zahl und den *encryption key*, der das LFSR vor dem Senden jedes Pakets initialisiert, als Eingangsparameter (s. Abb. 7.6).

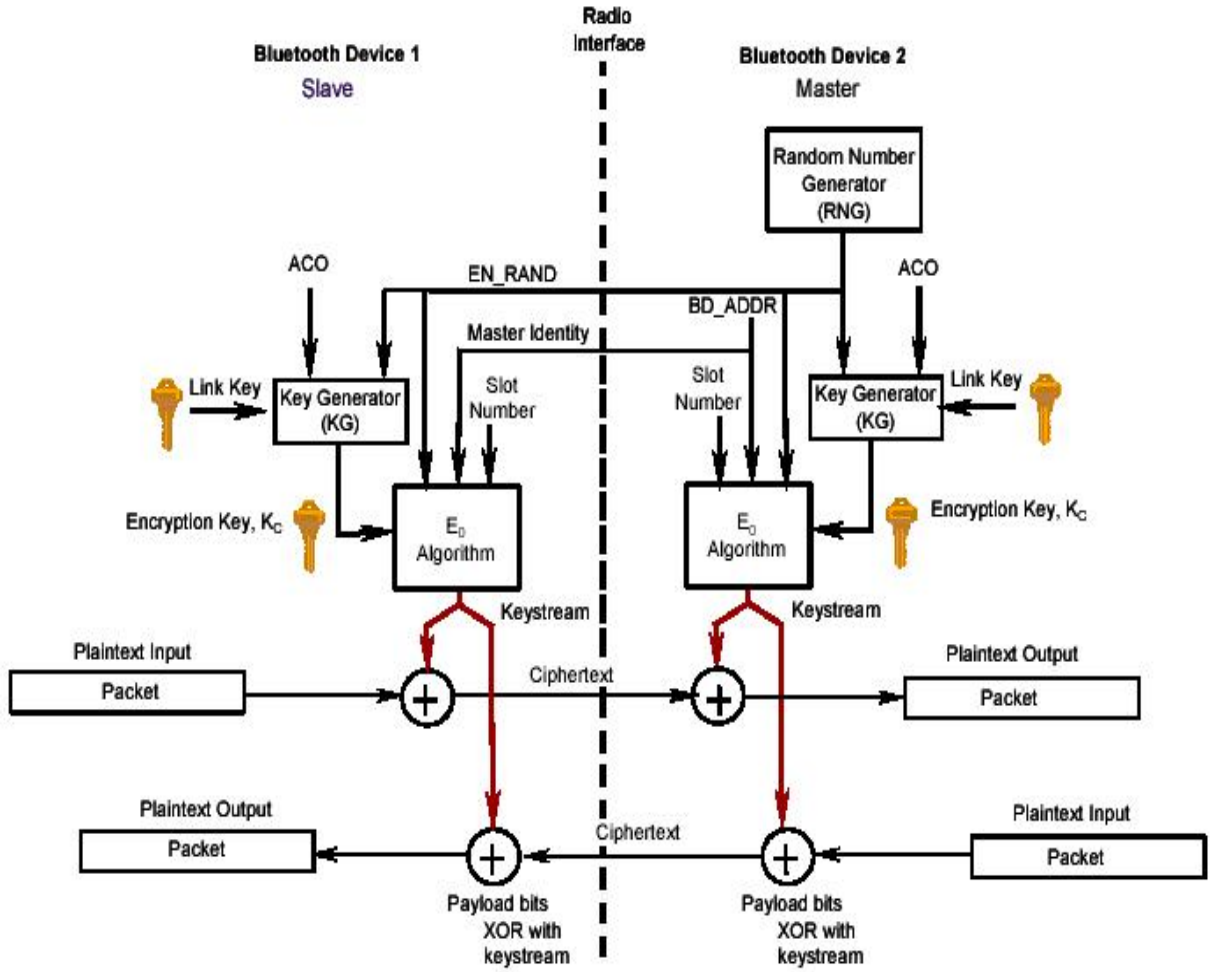

Abbildung 7.6: Bluetooth-Verschlüsselung [1]

7.2.3 Sicherheit von GSM

GSM verfügt über drei zentrale Sicherheitsalgorithmen, diese sind A3, A5 und A8.
Die Algorithmen A3 und A8 sind auf der SIM Karte gespeichert und im Authentication Center des Netzbetreibers vorhanden. Diese Algorithmen sind durch die Spezifikation von GSM nicht vorgegeben und können vom Hersteller selbst festgelegt werden. Der A3 dient zur Authentifizierung des Teilnehmers gegenüber dem Netz, während A8 zur Erzeugung eines 64 Bit Chiffrierschlüssels Kc genutzt wird. Hierzu greifen die Algorithmen auf einen teilnehmerindividuellen Schlüssel Ki zu, der sowohl in der SIM, als auch im Authentication

Center gespeichert ist.
Im Gegensatz zum A3 und A8 Algorithmus ist A5 im Endgerät (Mobile Equipment) implementiert. A5 ist eine französische Entwicklung und ein europaweit standardisierter Stromchiffrieralgorithmus. Vom A5 Algorithmus existieren mehrere Versionen, die sich in ihrer Sicherheit unterscheiden, diese sind: der A5/1 und der A5/2. Mit A5/0 wird eine unchiffrierte Kommunikation bezeichnet [5].

Comp128 ist ein in Deutschland entworfener Algorithmus, mit dem bereits A3 und A8 kombiniert realisiert werden können, falls die Netzbetreiber keinen eigenen Algorithmus entwerfen wollen [5].

Zugangskontrolle:

Der Benutzer weist seine Identität durch die Eingabe einer persönlichen Geheimzahl, der sogenannten PIN (Personal Identification Number), nach. Die PIN ist eine vier- bis achtstellige Zahl. Nach dreimaliger Falscheingabe wird die Karte gesperrt und kann nur mit einer separaten, achtstelligen Geheimzahl, der PUK (PIN Unblocking Key), wieder freigeschaltet werden. Nach zehnmaliger Falscheingabe der PUK wird die Karte als unbrauchbar markiert. Insgesamt soll durch diese Maßnahmen eine unberechtigte Benutzung z.B. durch Diebstahl verhindert werden (Teilnehmerschutz). Speziell für Notfall-Situationen wurde der Notruf-Dienst eingeführt, der auch ohne Eingabe der PIN funktioniert [15].

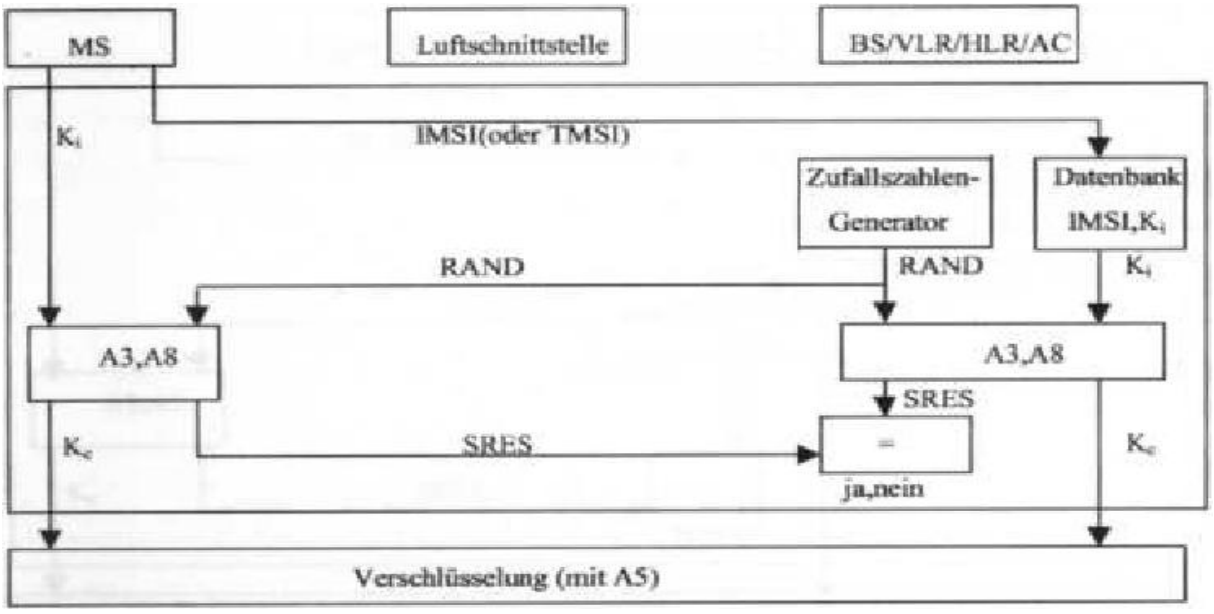

Abbildung 7.7: GSM-Authentisierung [5]

Authentisierung:
In GSM-Netzen erfolgt die Authentisierung mit einem *Challenge-Response*-Verfahren (s. Abb. 7.7). Bevor die Prozedur beginnen kann, muss sich der Teilnehmer zuerst mit seiner speziellen Kennung IMSI (Jeder Mobilfunk-Teilnehmer ist durch eine europaweit eindeutige Mobilfunk-Kennung, der sogenannten IMSI *International Mobile Subscriber Identity* identifizierbar.) beim Netz anmelden. Die IMSI wird nur bei der ersten Kontaktaufnahme unverschlüsselt über die Funkschnittstelle verschickt. Sie ist dem Teilnehmer nicht bekannt.

An das Mobilgerät wird eine 128-Bit Zufallszahl RAND gesendet (*Challenge*). Es nutzt sie und den Ki um mit Hilfe des A3 Algorithmus' eine Antwort SRES (*Response*) zu generieren, die zum Netz zurückgesandt wird. Gleichzeitig wird durch den Algorithmus A8 ein Chiffrierschlüssel Kc generiert, der nach einer erfolgreichen Authentifizierung zur Verschlüsselung genutzt werden kann[12].

Verschlüsselung:
Die Datensicherheit im GSM-System betrifft nicht nur die Sprachdaten, sondern alle zu übertragenen Teilnehmerdaten, d.h. sowohl die Daten der Verkehrs- als auch die der Signalisierungskanäle werden verschlüsselt.

Bevor eine Verschlüsselung vorgenommen werden kann, muss ein Chiffrier-Schlüssel mittels des A8-Algorithmus' generiert werden. Wie bei A3 wird auch bei A8 die zuvor erhaltene Zufallszahl RAND mit dem Schlüssel Ki verknüpft. Das Ergebnis ist diesmal der 64-Bit Schlüssel Kc.

Die eigentliche Verschlüsselung der Daten erfolgt durch den A5-Algorithmus. Der A5-Algorithmus generiert aus dem 64-Bit Chiffrierschlüssel Kc und der 22-Bit Rahmennummer des TDMA-Frames (Rahmennummern der zu verschlüsselnden Daten) einen pseudozufälligen Bitstrom, der nun einfach mit den zu übertragenden Daten mittels der XOR-Operation verknüpft wird [12].

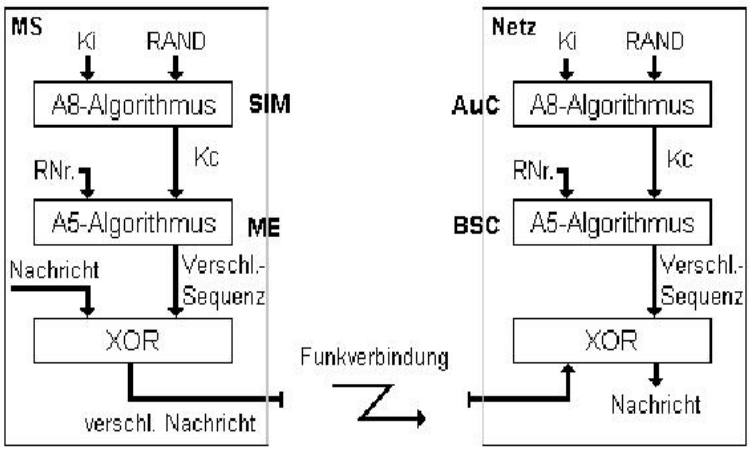

Abbildung 7.8: GSM-Verschlüsselung [10]

Anonymisierung: Wie bei dem Verfahren der Authentisierung bereits beschrieben, muß sich der Teilnehmer zunächst mit seiner IMSI im Klartext anmelden. Damit dieser später nicht immer mit der eindeutigen IMSI angesprochen wird, bekommt er nach erfolgter Anmeldung eine neue temporäre Kennung TMSI (*Temporary Mobile Subscriber Identity*) zugeordnet. Das Netz weist die TMSI der Mobilstation verschlüsselt zu, da zu diesem Zeitpunkt schon der Übertragungsschlüssel Kc berechnet wurde. Diese temporäre Kennung dient nun der Identifizierung der Mobilstation in der jetzigen Location Area. Verlässt der Mobilteilnehmer die aktuelle LA wird ihr automatisch vom neuen VLR (*Visitor Location Register* ist eine lokale Datenbank, in der die für das Management benötigten Daten temporär gespeichert werden können, um so einen effizienten Verbindungsaufbau garantieren zu können) eine neue TMSI zugewiesen. Ebenso kann das VLR nach einer gewissen Zeit der Mobilstation eine neue TMSI zuordnen, auch wenn die aktuelle LA nicht verlassen wurde. Die zur Zeit gültige TMSI und der *Location Area Identifier* (LAI) sind auf der SIM Karte und im VLR gespeichert. Dadurch können abgehörte, aber dennoch verschlüsselte, Gespräche nicht einfach einem Teilnehmer zugeordnet werden. Wie eine neue TMSI generiert wird, ist nicht bekannt. Die Netzbetreiber legen fest, wann und wie oft die TMSI geändert wird. Mittels dieser Anonymisierung auf der Luftschnittstelle bleiben der Teilnehmer und sein Aufenthaltsort anonym [12].

7.3 Angriffe und Sicherheitslücken

Leider sind diese Sicherheitsverfahren nicht perfekt und haben viele Schwachstellen oder besser gesagt Sicherheitslücken. Sie bieten also für unbefugte Teilnehmer die Möglichkeit, eine Reihe von Angriffe zu realisieren.

Angriffe

Alle diese Angriffe kann man wie in Abb. 7.9 in zwei verschiedene Kategorien zusammenfassen [1]:

1. **Passive Attack:**

   ist ein Angriff, in dem ein nicht autorisierter Teilnehmer einfach Zutritt zu dem Netz erhält. Er hört den Netzverkehr auf Nachrichten ab ohne sie zu ändern (d.h. Lauschen). Passive Angriffe können entweder einfaches heimliches Zuhören (*Eavesdropping*) oder die Verkehrsanalyse (*Traffic Analysis*) sein.

   - *Eavesdropping*:
     Der Angreifer überwacht einfach den Verkehr, um den Inhalt der Nachrichten zu sehen. Ein Beispiel dieses Angriffs ist eine Person, die die Übertragung zwischen zwei Workstations in einem LAN abhört.
   - *Traffic Analysis*:
     Bei diesem Angriff handelt es sich um die Überwachung des Datenverkehrs, um wichtige Informationen über die übertragenen Daten herauszufinden.

2. **Active Attack:**

   Bei diesem Angriff fügt ein nicht autorisierter Teilnehmer Änderungen an die übertragene Nachricht an. Aktive Angriffe können folgende Formen (oder Kombinationen davon) annehmen:

   - *masquerading*:
     Der Angreifer verkörpert einen autorisierten Benutzer und gewinnt dadurch bestimmte nicht autorisierte Privilegien.
   - *replay*:
     Der Angreifer überwacht den Datenverkehr (passiver Angriff) und schickt dann Nachrichten als wäre er der berechtigte Benutzer.
   - *message modification*:
     Der Angreifer ändert eine gesendete Nachricht, indem er sie löscht, ändert, neuordnet oder ihr etwas hinzufügt.
   - *denial-of-service*:
     Der Angreifer verhindert oder verbietet den normalen Gebrauch oder das Management von Kommunikationsmöglichkeiten.

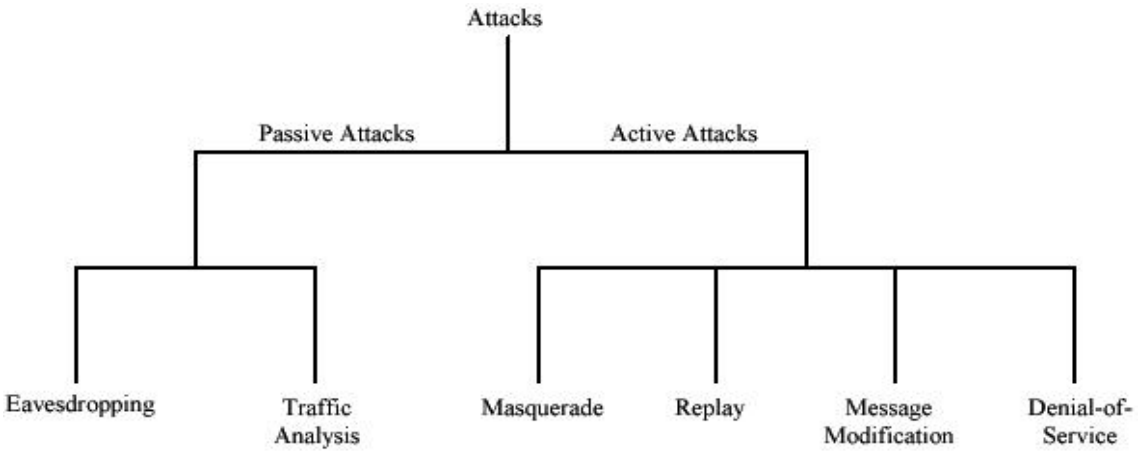

Abbildung 7.9: Angriffe [1]

7.3.1 Sicherheitslücken von IEEE 802.11

1. Deaktivierte Sicherheitseinstellungen:

   Die *Open System*-Authentisierung ist als Standardauthentisierungalgorithmus gesetzt. Das heißt, dass die Hersteller ihre Systeme mit einem deaktivierten *WEP* verschicken.

2. Kurze IVs:

   Der Initialisierungsvektor ist 24-bit lang und wird als plaintext gesendet. Er ändert sich für jedes Datenpaket nach bestimmten Mustern: Das einfachste Verfahren inkrementiert ihn einfach um 1. Je nach Größe des Initialisierungsvektors läuft dieser

Zähler früher oder später über und es können Initialisierungsvektor-Kollisionen entstehen. Eine Initialisierungsvektor-Kollisionen ist die Wiederverwendung der selben Initialisierungsvektoren. Wird der gleiche Initialisierungsvektor mit dem gleichen Shared Key verwendet, entsteht daraus der gleiche WEP-Key. Sobald dieser auf genügend Frames angewendet ist, lässt er sich extrahieren.

3. Kurze Verschlüsselungsschlüssel:

   Ein weiterer Schwachpunkt im Design von WEP ist die mit 40-Bit keineswegs als sicher anzusehende Schlüsselgröße. Aber eine Vergrößerung des Schlüssels wäre keine signifikante Änderung für die Sicherheit von WEP.

4. Gemeinsame Verschlüsselungsschlüssel:

   Viele Benutzer in einem drahtlosen Netz teilen sich möglicherweise für lange Zeit den selben Veschlüsselungsschlüssel. Das heißt, wenn ein Computer (z.B.:ein Laptop) verloren oder gestohlen werden sollte, könnte er eventuell die Sicherheit aller anderen Computer gefährden, die den selben Schlüssel mit ihm teilen.

5. Aktualisierung der Verschlüsselungsschlüssel:

   Verschlüsselungsschlüssel sollten häufig geändert werden, um *Brute-Force* Angriffe zu verhindern. *Brute-Force* gilt als die einfachste und primitivste Methode ein Passwort herauszufinden und ist nur das stupide Durchprobieren aller möglichen Zeichenkombinationen.

6. Schwachpunkt der RC4:

   Der gravierendste Angriffspunkt ergibt sich allerdings aus einer Schwäche des RC4. Ungefähr jeder 2000ste Initialisierungswert ist “schwach“: Das bedeutet, dass eine kleine Anzahl Bits aus dem Initialisierungswert einen überproportional starken Einfluss auf die Anfangswerte des Schlüsselstroms haben. Umgekehrt sind so aus dem Anfang des Schlüsselstroms auch Rückschlüsse auf den verwendeten Initialisierungswert und damit den geheimen Schlüssel möglich. Der mathematische Beweis wurde im August 2001 veröffentlicht.

7. Schwache Datenintegrität [3]:

   CRC-32 ist eine lineare Funktion. Das heißt:

$$CRC(x \oplus y) = CRC(x) \oplus CRC(y).$$

   Nehmen wir an, dass ein Angreifer eine verschlüsselte Nachricht C abfängt, die zu einer Klartext-Nachricht M gehört. Es gilt also

$$C = RC4(IV, K) \oplus < M, c(M) >.$$

   Anstelle der Nachricht M will er eine veränderte Nachricht

$$M' = M \oplus \delta$$

senden, die in verschlüsselter Form C' heißt und sich wie folgt berechnen lässt:

$$\begin{aligned} C' &= C \oplus <\delta, c(\delta)>. \\ &= RC4(IV, K) \oplus <M, c(M)> \oplus <\delta, c(\delta)>. \\ &= RC4(IV, K) \oplus <M \oplus \delta, c(M) \oplus c(\delta)>. \\ &= RC4(IV, K) \oplus <M', c(M \oplus \delta)>. \\ &= RC4(IV, K) \oplus <M', c(M')>. \end{aligned}$$

8. Keine reelle Identifizierung des Benutzers:

   Das heißt nur das Gerät wird identifiziert und deswegen hat auch ein gestohlenes Gerät Zugang zum Netzwerk.

9. Anfälligkeit der identifikationsbasierten Systeme (Benutzung von SSID):

   Viele Anwender denken nun, es reiche aus, den Namen des Service Sets (SSID) zu ändern. Dieser Wert dient aber keineswegs der Sicherheit, sondern lediglich der Unterscheidung einzelner Funknetze. Zwar lassen sich die Clients so einstellen, dass sie nur zu einem WLAN mit einer bestimmten SSID eine Verbindung aufbauen - ein Muss ist das jedoch nicht. Mithilfe geeigneter Treiber ist auch das Ansprechen beliebiger Netze kein Problem. Die Vergabe einer aussagekräftigen SSID hat für den Hacker sogar Vorteile. Er erhält auf diesem Weg zusätzliche Informationen über das Netz, in dem er sich befindet.

7.3.2 Sicherheitslücken von Bluetooth

***Frequency Hopping*:**
Frequency Hopping wurde eingeführt, um Interferenzen mit zum Beispiel Mikrowellen oder Wireless LANs, die im gleichen Bereich arbeiten, zu vermeiden. Es beruht darauf, dass die Funkmodule immer, nachdem sie ein Datenpaket verschickt oder erhalten haben, zu einer neuen Frequenz springen.
Oft wird die pseudozufällige Hopping Sequence, nach der Geräte auf bestimmten Frequenzen senden bzw. empfangen, als Sicherheitsmaßnahme bezeichnet, da ein Angreifer nicht voraussagen kann, nach welchem Muster gesendet wird, und somit nicht in der Lage sein soll, die Kommunikation zu belauschen. Allerdings ist es technisch nicht sehr schwierig die 79 parallelen Kanäle zu überwachen. Zudem ist die Seed des Pseudozufallszahlengenerators einfach in Erfahrung zu bringen, da sie aus der Geräteaddresse und der clock des Masters besteht. Beide Informationen sind nicht geheim und werden in der Initialisierungsphase offen übertragen [9].

***Man in the Middle* Angriff:**
Dieses ist zum Beispiel an einem öffentlichen Platz möglich, wo ein Angreifer den kompletten Prozess des "Pairings" der beiden Bluetooth Geräte abhören kann. Falls keine zusätzliche Sicherung auf der Applikationsebene stattfindet oder der Angreifer eine Man-In-

The-Middle Attacke durchführen kann, könnte er so eine komplette Kopie aller Dokumente erhalten, die der Besitzer der Bluetooth Geräte über die Funkschnittstelle austauscht. Wenn der Angreifer das Opfer nachahmt, kann er Nutz- oder Steuerdaten verändern [13].

Brechen des geheimen Schlüssels:
Wenn der PIN zu kurz oder schwach ist, kann der *link key* durch eine *brute force*-Attacke erraten werden. Dazu belauscht der Angreifer die Initialisierungsphase und erfährt dabei die verwendete Zufallszahl zur Generierung des init keys und die Daten, die zur Verifikation des init keys übertragen werden. Er wählt einen Wert für den PIN, führt selbst offline die Schritte zur Initialisierung und Verifizierung durch und vergleicht den berechneten Funktionswert mit den belauschten Daten. Stimmt die eigene Verifikation mit der Belauschten überein, hat der Angreifer mit sehr hoher Wahrscheinlichkeit den richtigen PIN gewählt. Da alle weiteren Schlüssel auf dem geheimen PIN basieren, kann ein Angreifer nun alle weiteren Kommunikationen entschlüsseln und selbst Nachrichten einspeisen. In einer Variante dieser Attacke initiiert der Angreifer selbst die Kommunikation.

***Denial of Service*:**
Die Elemente eines drahtlosen Netzes sind von einer begrenzten Energiequelle abhängig. Einer der Hauptangriffe von *Denial of Service* besteht darin, die anzugreifende Station mit Arbeit zu überfordern, um ihre Energie so schnell wie möglich zu verbrauchen.

***Location Attack*:**
Bluetooth Geräte können in zwei Modi betrieben werden:

- *Detectable Modus*: In diesem Modus, antwortet ein Bluetooth Gerät auf jede Anfrage eines anderen Gerätes mit seiner eindeutigen *Device Adresse*.
- *Non-Detectable Modus*: In diesem Modus, antwortet das Gerät auf keine Anfrage.

Das bedeutet, dass ein Angreifer den Standort eines Gerätes bestimmen kann, falls dies im *Detectable Modus* arbeitet.

7.3.3 Sicherheitslücken von GSM

***Brute-Force* Angriff:**
Den Stromverschlüsselungsalgorithmus A5 gibt es in zwei Varianten [12] :

- A5/1: ist die strengere Variante. Sie hat eine Schwäche darin, daß sie "nur" 2^{64} Zustände einnehmen kann. Untersuchungen von A5/1 zeigten, dass die letzten 10 Bits immer auf 0 gesetzt wurden. Dies reduziert die Komplexität für die Schlüsselsuche bereits auf 2^{54}. Es wurden Verfahren entwickelt die durch Raten bestimmter Bits die Komplexität auf 2^{45} und 2^{40} senken. Jedoch benötigt ein normaler PC für diese Attacken immer noch über einen Monat, um den Schlüssel zu finden.

- A5/2: ist die leichtere Variante. Für diese Variante wurde ein Verfahren entwickelt, das von der Komplexität 2^{16} ist, und zeigt, daß diese Variante völlig unsicher ist.

***Man in the Middle* Angriff:**
Ein *IMSI-Catcher* simuliert eine Basisstation. Da dieser ein stärkeres Signal sendet als die wirklichen Basisstationen, melden sich alle MS innerhalb seiner Reichweite bei ihm an.
Er kann in zwei Modi arbeiten: Fangen und abhören.
Im Fangmodus kann der *IMSI-Catcher* die IMSI und auch die IMEI abfragen. Während dieser Prozedur ist es nicht möglich, mit einer gefangenen MS Gespräche zu führen oder zu empfangen. Im Abhörmodus kann man abgehende Gespräche des gefangenen MS in Klartext aufzeichnen. Dazu verhält sich der *IMSI-Catcher* zusätzlich gegenüber der echten BS wie die gefangene MS und leitet das Gespräch ans Netz weiter. Das Abhören in Klartext ist deshalb möglich, da die Verschlüsselung durch einen speziellen GSM-Befehl, der vom *IMSI-Catcher* benutzt wird, vom Netzbetreiber abgeschalten werden kann.

Kopieren der SIM Karte:
Es wurden zwei Wege gefunden, um eine SIM Karte zu kopieren [16]:

1. 1998 ist es dem Chaos Computer Club (CCC) gelungen eine D2-Karte zu kopieren, indem man den geheimen Teilnehmerschlüssel Ki aus einer SIM Karte extrahiert hat. Grundlage dafür ist eine Sicherheitslücke in dem verwendeten Algorithmus COMP128.
   Dies geschieht über bestimmte Kombinationen von Anfragen, die, wenn sie ein identisches Ergebnis liefern, bestimmte Bits des Schlüssels verraten. Der Angriff erfordert um die 150.000 Anfragen, die die SIM auffordern, sich zu authentifizieren, und dauert bis zu 12 Stunden. Neben einem entsprechenden Kartenleser, benötigt man ebenso die PIN der entsprechenden Karte.

2. Im Mai 2002 veröffentlichte die Firma IBM eine Mitteilung, derzufolge IBM-Mitarbeiter ein Verfahren gefunden haben sollen, mit dem sie innerhalb weniger Minuten den geheimen Schlüssel Ki anhand des Authentifizierungsvorgangs rekonstruieren können. Dieses Verfahren erfordert zwar keinen Eingriff in die SIM-Karte, benötigt jedoch physikalischen Zugang zum Endgerät, um unter anderem Stromaufnahme und elekromagnetische Abstrahlung der Karte zu messen. Auf diese Weise ist es möglich, SIM Karten zu fälschen und auf Kosten anderer Teilnehmer zu benutzen.
   Details der sogenannten *partitioning attack* wurden von IBM zunächst nicht veröffentlicht.

7.4 Gegenmaßnahmen und Verbesserungen

7.4.1 Der Fall IEEE 802.11

1. *WEPplus*:

WEPplus ist eine Entwicklung von Agere Systems. Die alleinige Aufgabe von WEPplus ist es, statt WEP die Generierung von Initialisierungsvektoren vorzunehmen und damit zu verhindern, dass schwache Paketschlüssel erzeugt werden. Um das Netzwerk vollständig gegen schwache Initialisierungsvektoren abzusichern, ist es jedoch notwendig, alle WLAN Access Points des jeweiligen Netzes mit WEPplus aufzurüsten [5].

2. WEP2:

   Das WEP2 Protocol wurde von der TaskGroup i (TGi) im Rahmen des Standards IEEE 802.11i im Jahre 2001 entworfen und sollte einen Initialisierungsvektor mit einer Länge von 128 Bit beinhalten und ein periodisches Schlüsselupdate durchführen. Der Entwurf der TGi hatte allerdings Schwächen und wurde 2001 wieder verworfen [5].

3. *Virtual Private Networks*:

   Sie verschlüsseln die Kommunikation zwischen VPN-Server und VPN-Clients (so genannte Ende-zu-Ende-Sicherheit) und können zum Beispiel mit einem PPTP-Server (Point-to-Point-Tunneling) realisiert werden (s. Abb. 7.10). Die Verwendung von Protokollen wie IPSec erhöht die Sicherheit ebenfalls [6].

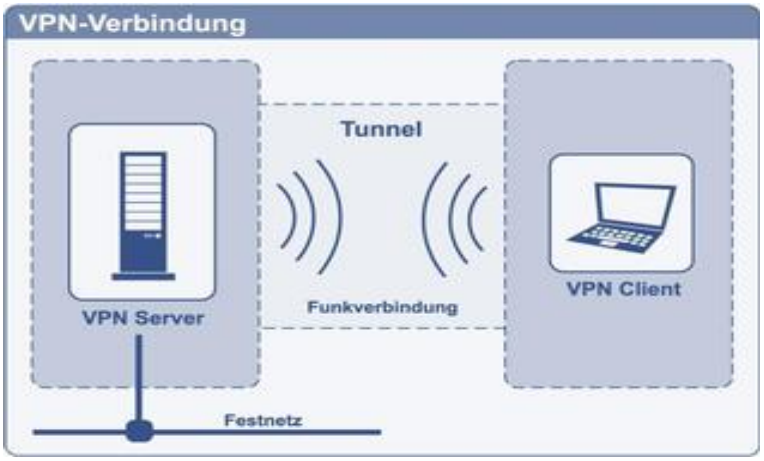

Abbildung 7.10: Virtual Private Networks [6]

7.4.2 Der Fall Bluetooth

***Virtual Private Networks*:** Die Verschlüsselung in den Bluetooth Geräten lässt sich erfolgreich angreifen. Deswegen sollten bei Kommunikation über Bluetooth, soweit möglich, Applikationen verwendet werden, die zusätzliche Sicherheit bieten. Dies wäre unter anderem ein Virtual Private Network (VPN) (s. 7.4.1).

***Non-Detectable Modus*:** Es ist notwendig, Bluetooth Geräte in den Betriebsmodus *Non-Detectable* zu schalten, um eine Erstellung von Bewegungsprofilen zu unterbinden.

Diese Bewegungsprofile sind nur so lange anonymisiert, wie keine Zuordnung der Bluetooth Adresse zu der Person erfolgen kann, die dieses Bluetooth Gerät besitzt. Eine Einleitung eines Kommunikationsvorgangs zwischen zwei Bluetooth Einheiten (Pairing) wird so allerdings erschwert und recht unbequem. Bequemlichkeit hat sich aber schon oft als Grund für verringertes Sicherheitsinteresse herausgestellt.

7.4.3 Der Fall GSM

- COMP128-2:

  Es existiert bereits eine neue Version von COMP128, die das Kartenkopieren nicht zulässt.

- Einige Verbesserungsvorschläge:

  – Gegenseitige Identifizierung: Nicht nur die Mobilstation muss sich identifizieren, sonder auch die Basisstation. Das würde die *Man in the Middle* Angriffe verhindern, da die falsche Basisstation sich nicht authentisieren kann.

  – Die vollständige Ausnutzung des 64-bit Kc würde mehr Sicherheit bieten

  – Man könnte auch die Länge von Kc verdoppeln.

7.5 Zusammenfassung

Wireless Networks haben mehr Vorteile als Normale Netzwerke. Sie müssen aber extra geschützt, weil die Informationen frei durch die Luft gesendet werden. Es reicht ein Antenne, um diese Informationen zu empfangen. Für Unternehmen sind Sicherheitsmängel ihrer Netze unter Umständen ein nicht unerhebliches, finanzielles Problem. Nicht nur die Sicherheit der Unternehmen ist gefährdet, sondern auch sensible Daten der Kunden werden nicht sicher verwaltet und der Datenschutz ist somit nicht gewährleistet. Deswegen müssen *Wireless Networks* sehr gut geschützt.
802.11, Bluetooth und GSM weisen erhebliche Sicherheitsmängel auf. Ein Beispiel ist das Sicherheitsverfahren von 802.11 WEP. Er ist seit einem Jahr gebrochen worden. Deshalb ist er in dieser Form unsicher.
802.11, Bluetooth und GSM erfüllen also nicht die Anforderungen für eine Nutzung in sensiblen Bereichen.
Wir haben die Schwachstellen in den Sicherheitsverfahren der drei Funksysteme demonstriert und einige Angriffsmöglichkeiten beschrieben, die aus diesen Schwachstellen resultieren. Es müssen also zusätzliche Vorsichtsmaßnahmen getroffen werden, um den Netzverkehr zu schützen.

Literaturverzeichnis

[1] T. Karygiannis, L. Owens: *Wireless Network Security 802.11, Bluetooth™ and Handheld Devices*; http://csrc.nist.gov/publications/drafts/draft-sp800-48.pdf, September 2002.

[2] N. Borisov, I. Goldberg, D. Wagner: *Security of the WEP Algorithm*; http://www.isaac.cs.berkeley.edu/isaac/wep-faq.html, Januar 2002.

[3] N. Borisov, I. Goldberg, D. Wagner: *Intercepting Mobile Communications: The Insecurity of 802.11*; http://www.cs.berkeley.edu/~daw/papers/wep-mob01.pdf, Januar 2002.

[4] S. Rubner: *TechReport: Wireless Security*; http://techupdate.zdnet.de/story/0,,t422-s2111737,00.html, ZDNet Juni 2002.

[5] A. Kösling: *Sicherheitsanalyse in drahtlosen Netzen*; http://security.koesling.net/wireless, Mai 2002.

[6] Pc-Welt: *Sicherheit in Funknetzen*; http://www.pcwelt.de/ratgeber/extras/26424/index.html.

[7] Bundesamt für Sicherheit in der Informationstechnick: *Sicherheit im Funk-LAN (WLAN, IEEE 802.11)*; http://www.bsi.de/fachthem/funk_lan/wlaninfo.pdf, Juli 2002.

[8] A. Arbaugh, N. Shankar, J. Wan: *Your 802.11 Wireless Network has No Clothes*; http://www.drizzle.com/~aboba/IEEE/wireless.pdf, März 2002.

[9] S. Keuser: *Sicherheit in mobiler Kommunikation*; http://www.inf.ethz.ch/vs/edu/SS2001/MC/beitraege/08-security-rep.pdf.

[10] J. Pflüger: *GSM - Global system for mobile communication*; http://goethe.ira.uka.de/seminare/rkt/gsm/#ToC7.

[11] T. Funck: *Authentifizierungsprotokolle und Sicherheitsmechanismen im Mobilfunk*; http://wwwhegering.informatik.tu-muenchen.de/Events/Sarntal/Sarntal2000/sicherheit-mobilfunk.pdf, September 2002.

[12] M. Frey: *Sicherheit in GSM-Netzen*; http://medien.informatik.uni-ulm.de/lehre/courses/ss02/sem_mobsec/seminar_mobsec.pdf.

[13] F. Heckel: *Sicherheit bei IEEE802.11*;
http://medien.informatik.uni-ulm.de/lehre/courses/ss02/sem_mobsec/seminar_mobsec.pdf.

[14] K. Siebenrok: *Bluetooth-Security*;
http://medien.informatik.uni-ulm.de/lehre/courses/ss02/sem_mobsec/seminar_mobsec.pdf.

[15] W. Kowalk: *Rechnernetze: Sicherheit in GSM-Netzen*;
http://einstein.informatik.uni-oldenburg.de/rechnernetze/gms-sicherheit.htm.

[16] SecurityServer: *GSM-Sicherheit Cloning von SIM-Karten*;
http://www.infoserversecurity.org/gsm.php.