Internet Economics I

Burkhard STILLER
Oliver BRAUN
(Hrsg.)

Institut für Informationstechnische Systeme, IIS

Bericht Nr. 2002-04
Juli 2002

Universität der Bundeswehr München

Fakultät für

INFORMATIK

Werner-Heisenberg-Weg 39 • D-85577 Neubiberg

B. Stiller, O. Braun (edts.): Internet Economics I
Technical Report No. 2002-04, July 2002
Information Systems Laboratory IIS, Departement of Computer Science
University of the Federal Armed Forces Munich
Werner-Heisenberg-Weg 39, D-85577 Neubiberg, Germany

Introduction

The Information System Laboratory (Institut für Informationstechnische Systeme, IIS) of the Departement of Computer Science, University of the Federal Armed Forces Munich, Germany started research and teaching during this spring term 2002 (FT02) in the area of communications. One of the closely related topics is addressing the use and application of technology and mechanisms under economic and technical optimization measures.

The topic of Internet Economics is around for some years now. The technology applied to communications and in the domain of electronic businesses has not dramatically changed during recent years, though. However, a number of important extensions of functionality has been developed and further problems have been encountered. Internet communications in general, offering services and selling bandwidth or even QoS-enabled services, all of these tasks are still demanding, but generic and efficient solutions for heterogeneous technologies within today's Internet have not been found yet. Still, or even in contrast to those demands, Internet technology acts as a means to implement new applications, covering inelastic and multimedia applications. Therefore, a review of several challenges and weaknesses of this process of development is required to judge its suitability.

Content

This first edition of the seminar "Internet Economics I" deals with the integration of Internet technology and new ways to support and do business. Starting the talks, a view onto the non-repudiation of electronic transactions is presented, which covers mechanisms, protocols, and important requirements on the hardware. Since the secure access to information in any open and widespread system is essential, the second presentation on security in e-commerce provides an overview on selected mechanisms, protocols, and trading models between seller and buyer.

In addition, peer-to-peer technologies show a modern trend in communications between individuals, which is investigated in the third presentation. Based on the comparison of client/server models with peer-to-peer models, major characteristics are investigated, a potential technology is discussed, and areas of application are considered.

The fourth presentation addresses Wireless Local Area Network technologies based on IEEE 802.11, focussing on technology and its commercial deployment in comparison with other wireless technologies. Security concerns for a commercial application are discussed as well. The final presentation in this Internet Economics seminar addresses mobile services and their future. While Wide Area Network communication services are discussed, device requirements and roaming are outlined. Based on a number of sample applications a brief survey on the market situation is 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, busines 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 a varying audience of students attending the seminar always as well as other interested students and research assistents. Following a general question and answer part, a student-lead discussion debated lively open issues and critical statements with the audience.

Local IIS support for preparing talks, reports, and their preparation by students had been granted Oliver Braun and Burkhard Stiller. Many pre-presentation discussions have provided valuable insights in the emergingly moving field of Internet Economics, 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 people.

This seminar has proven that the effort to be put into place from sudents, supervisors, and the professor are significant in the sense, that addressing a good and satisfactory result requires time and reading. However, those who have succeeded and who have believed that this effort is worthwhile, have achieved a phantastic result in terms of presentations, reports (as included in this Departement of Computer Science Report), and final marks. The presenter's insights into a highly relevant area of research, technology, and speculation - surely adressing the Internet environment - have contributed to advances of understanding many problems. Always highly appreciated, discussions after each of the presentation inter-related many aspects of different talks.

Neubiberg, July 2002

Inhaltsverzeichnis

Kapitel 1

Nicht-Abstreitbarkeit elektronischer Transaktionen

Steffen Mazanek

Inhaltsverzeichnis

1.1 Einleitung

In der heutigen Zeit wird das Internet in einem zunehmenden Maße für kommerzielle Transaktionen benutzt, z.B. in Form von Online-Banking, virtuellen Supermärkten, Online-Auktionen u.ä., denn die neuen digitalen Signaturverfahren erlauben große Distanzen zwischen den Geschäftspartnern, was zu einer enormen Kostenersparnis und Effizienzsteigerung führt.
Daraus resultiert automatisch ein höheres Risikopotential (denn je mehr Vorgänge digital abgewickelt werden, desto mehr entziehen sie sich unserer direkten Kontrolle) und damit einhergehend ein zunehmendes Sicherheitsbedürfnis der Nutzer. Denn auch in diesen Dingen eher unbedarfte Anwender sind durch die Medienberichte über Computerviren, trojanische Pferde, Hackerattacken u.v.m. eingeschüchtert, ja sogar teilweise abgeschreckt. Diese Probleme treten bei der Face-to-Face-Kommunikation nicht auf und es ist meist wesentlich leichter, den Transaktions-Verlauf nachzuvollziehen.
Um den positiven Trend auf diesem Gebiet also aufrechtzuerhalten und den Umgang damit zur Selbstverständlichkeit werden zu lassen, bedarf es einer intensiven Sicherheits- und auch Rechtspolitik sowie deren Unterstützung durch die entsprechenden Technologien und die Ausbildung der Benutzer mit diesen Technologien.

Aber nicht nur für wirtschaftliche Belange sind die neuen Verfahren interessant. Auch die Streitkräfte können davon profitieren und so hat die Bundeswehr als Reaktion auf diese Fortschritte ihren IT-Bereich in den letzten Jahren ausgebaut. Moderne Waffen können jedoch großen Schaden anrichten. Deshalb ist es umso wichtiger, Sicherheitsaspekte gerade auch in diesem Rahmen genau zu untersuchen. Man muss davon ausgehen, dass Angriffe auf Sicherheitslücken zur Informationsgewinnung oder direkten Schädigung beim Militär oft und zielgerichtet erfolgen, sei es durch andere Armeen, Terroristen oder einfach durch einen Selbstbestätigung suchenden Hacker. Dem geneigten Leser sei dazu der Artikel „Die Bundeswehr auf dem Weg ins digitale Schlachtfeld“ [BW] von Ralf Bendrath empfohlen.

In diesem Dokument soll ein integraler Sicherheits-Aspekt herausgegriffen werden: Wie kann man die Leugnung elektronischer Abmachungen und Verträge wirkungsvoll verhindern oder aufklären? Die sich daraus ableitenden Maßnahmen fasst man unter dem Begriff **Non-repudiation** (Nicht-Abstreitbarkeit) zusammen. In der Informatik nennt man festgelegte Regeln und Vereinbarungen, gerade im Bereich der Datenübertragung, häufig **Protokolle**. Nachfolgend sollen nun Anforderungen an Non-repudiation-Protokolle gesammelt und diskutiert werden. Mit Hilfe der dabei gewonnenen Erkenntnisse wird anschließend ein Beispiel-Protokoll untersucht.

1.1.1 Gliederung des Dokuments

Es ist nötig, den Inhalt dieser Arbeit thematisch einzuordnen und abzugrenzen. Dies geschieht in Abschnitt 1.1.2. Ziel ist es, mehrere verschiedene Herangehensweisen aufzuzeigen und gegenüberzustellen. Damit der Detaillierungsgrad nicht darunter leidet, wird an entsprechenden Stellen auf vertiefende Literatur verwiesen. Deshalb wird gleich zu Beginn (1.1.3) auf die zugrundeliegende Literatur eingegangen.

In zweiten Kapitel kann der Leser dann an einem Beispiel die Bedeutung der Non-repudiation-Verfahren nachvollziehen, denn mögliche Problemfelder und Fehlerquellen

werden identifiziert und erklärt. Untrennbar mit dem Thema verbunden ist auch die juristische Sachlage. Deshalb soll der Fragestellung, wer die Beweispflicht innehat, nachgegangen werden. Anhand möglicher Streitszenarien wird dann untersucht, welche Beweise nötig sind und zu welchen Zeitpunkten während der Transaktion ihre Erzeugung stattfinden sollte. Daraus kann der Ablauf eines Non-repudiation-Prozesses abgeleitet werden.

Der Kernpunkt der Arbeit ist es, ein mögliches Protokoll vorzustellen und zu untersuchen. Dies geschieht im dritten Kapitel. Außerdem werden allgemeine Anforderungen an solche Protokolle erklärt und ihre Erfüllung an dem Beispiel überprüft, Verbesserungen und Erweiterungen vorgeschlagen und zum Teil ausgearbeitet.

Da Sicherheit im eCommerce zu einem nicht geringen Anteil von der zur Verfügung stehenden Hardware und dem Betriebssystem abhängt, ist das vierte Kapitel dieser Thematik gewidmet.

Im fünften und letzten Kapitel werden in einer Zusammenfassung die wichtigsten Aspekte nochmals genannt und anschließend mit der persönlichen Meinung des Autors bewertet.

1.1.2 Einordnung in den Sicherheits-Kontext

Es sind sehr viele Arbeiten über das Thema Sicherheit im eCommerce verfasst worden. In wenigen anderen Gebieten tut sich derzeit so viel. Diese Arbeit greift einen kleinen, aber bedeutenden Teil mit hoher Zukunftsrelevanz heraus. Der Rest muss in diesem Rahmen leider weitestgehend unberücksichtigt bleiben. So wird grundlegendes Verständnis für Internethandel, Netzwerke und digitale Signaturen (privater, öffentlicher Schlüssel) vorausgesetzt. Die wichtigsten Begriffe sind im **Glossar** in entsprechender Kürze erklärt und können auch in der einschlägigen Fachliteratur, z.B. „Network Security Essentials“ von William Stallings, nachgelesen werden.
Die Korrektheit von Non-repudiation-Protokollen hängt stark von den allgemeinen Sicherheitsmechanismen in Computer-Netzwerken ab. Teilweise müssen nämlich darüber bestimmte Annahmen getroffen werden. Aber auch dieser Sachverhalt bleibt im Folgenden bis auf die Angabe von Literaturreferenzen außen vor.

1.1.3 Zugrundeliegende Literatur

Bei den meisten der Quellen handelt es sich um Online-Dokumentationen, die im Internet frei verfügbar sind. Den Einstieg in dieses Thema kann die Ausarbeitung von Karl-Fredrik Blixt und Åsa Hagström, „Adding Non-repudiation to Web Transactions“ ([BLIXT]), erleichtern. Rechtliche Grundlagen und Aspekte sind sehr übersichtlich dargestellt in dem Dokument „Non-repudiation in the Digital Environment“ ([McCUL]) von Adrian McCullagh und William Caelli, ein korrektes Protokoll ist in „Evolution of Fair Non-repudiation with TTP“ ([ZHOU1]) zu finden und Möglichkeiten der Effizienzverbesserung dieses Protokolls sind unter anderem erläutert worden in „On the Efficient Implementation of Fair Non-repudiation“ ([ZHOU2]), letztere beide von Jianying Zhou et al. Einige der Definitionen sind angelehnt an den ISO-Standard 13888-1, „Information technology–Security techniques–Non-repudiation“ ([ISO]).

1.2 Hintergrund

Schlüsselbegriffe werden im folgenden eCommerce und Internet sein. Deshalb erfolgt an dieser Stelle eine kurze Begriffsklärung.

Als **Internet** wird heute üblicherweise die Gesamtheit aller Netzwerke und Computer bezeichnet, die über TCP/IP-Verbindungen erreichbar sind. Doch eine genaue Abgrenzung ist aufgrund der Gateways zu anderen Netzen, die teils Internet-Verbindungen, teils andere Mechanismen nutzen, nicht einfach zu ziehen. Da das Internet aus vielen autonomen Netzwerken besteht, gibt es keine Organisation, die das Internet „leitet“ oder „regiert“ ([Net]). In punkto Sicherheit birgt diese Heterogenität enorme Schwierigkeiten, z.B. in Bezug auf Verantwortlichkeiten für Inhalte.

Electronic Commerce ist ein Konzept zur Nutzung von bestimmten Informations- und Kommunikations-Technologien zur elektronischen Integration und Verzahnung von Wertschöpfungsketten und unternehmensübergreifenden Geschäftsprozessen ([eCom]). Manchmal wird für diese Kombination aus Technologie und Handel auch der Begriff **Internet Economics** verwendet.

Ein wesentlicher Nach- aber auch gleichzeitig Vorteil des eCommerce ist, dass Kunden und Verkäufer anonym sind, d.h. dass z.B. ein Web-Shop, betrieben in Amerika, durchaus Abnehmer in Europa haben kann. Bei dieser Art des Handels ist das Währungsproblem wohl noch das geringste. Vielmehr findet diese Internationalität eher unbefriedigende Beachtung durch nationale Verbrauchergesetze. Diese stehen den neuen Anforderungen teilweise hilflos gegenüber. Und so ist es unvermeidlich geworden, dass sich die betroffenen Parteien selbst auf eine Vorgehensweise verständigen müssen, die für beide Seiten sowohl komfortabel und praktikabel als auch fair ist.

1.2.1 Mögliche Probleme - Ein Beispiel

Am besten lassen sich die möglichen Probleme und Streitigkeiten an einem kleinen Beispiel erklären. Ein Kunde $\mathcal{A}$ interessiert sich für ein Angebot des Shops $\mathcal{B}$. Wenn $\mathcal{A}$ jetzt Waren bei $\mathcal{B}$ bestellen möchte, will er sicher gehen, dass $\mathcal{B}$ sich zu dem Auftrag verpflichtet, d.h. ihn weder vergisst noch ignoriert. Andererseits befindet sich $\mathcal{B}$ in ähnlicher Situation, der Verkäufer möchte eine Garantie dafür, dass der Abnehmer zu seiner Anforderung steht und die Produkte abnimmt und bezahlt. Dieser Aspekt wird durch Non-repudiation Schemata verallgemeinert und behandelt.

1.2.2 Pflicht des Beweises, eine rechtliche Streitfrage

Sollte ein Fall von Unterschriftsabstreitung auftreten, muss eine der beiden Parteien die Beweispflicht übernehmen. Damit entsteht ihr natürlich ein Nachteil (Aufwand, Kosten, ungewisser Ausgang, etc.). Der Beweis ist in der Regel schwierig, wenn nicht ein glaubwürdiger Zeuge zur Verfügung steht, wie es bei herkömmlichen handschriftliche Signaturen oftmals der Falls ist, bei digitalen Signaturen jedoch nie. Die Streitfrage ist somit, wer diese lästige Aufgabe übernehmen muss. Nachfolgende Erklärungen sind angelehnt an [McCUL].

Traditionelle rechtliche Regelung

Der vermeintliche Unterzeichner hat immer das Recht die Unterschrift abzustreiten, denn es könnte sich ja wirklich um eine Fälschung handeln oder sie könnte durch skrupelloses Verhalten oder Täuschung durch die zweite oder irgendeine dritte Partei erwirkt worden sein. Generell gilt in diesem herkömmlichen Sinne, dass die Gläubiger nachweisen müssen, dass die Signatur echt und damit rechtens und verbindlich ist.

Kryptotechnische Bedeutung

Zunehmend mehren sich Befürworter der Anschauung, dass mit digitalen Signaturverfahren anders verfahren werden sollte als mit handschriftlichen, denn die Beweispflicht immer dem Gläubiger zu übertragen, ist nicht fair. Natürlich hat sich das zuerst beschriebene Verfahren über lange Zeiträume bewährt und es ist auch gewagt, dererlei Inkonsistenzen in das geltende Recht einzubeziehen.

Nicht umsonst wird Non-repudiation in diesem Kontext sehr streng definiert als **ein Service, der einen Beweis für die Integrität und Originalität von Daten in einer unfälschbaren Weise liefert, der jederzeit nachgeprüft werden kann und auch später mit hoher Versicherung nicht angefochten werden kann** (nach [McCUL]). Der letzte Punkt ist umstritten, da das Recht, die Signatur zu dementieren, verweigert wird, bzw. die Beweislast nun vollständig beim Unterzeichner liegt. Außerdem sind die Begriffe **unfälschbar** und **Beweis** in der digitalen Umgebung nur bedingt verwendbar. Kommen auf diese Weise Abmachungen zu Stande, so müssen sich vorher beide Beteiligten auf diese Geschäftsbedingung verbindlich geeinigt haben und die Abwicklung muss sehr sicher erfolgen. Außerdem wird eine vertrauenswürdige Instanz benötigt, die im Zweifel schlichtet und entscheidet.

Die Frage nach der Beweispflicht lässt sich nur dann auf faire Weise lösen, wenn vertrauenswürdige Computersysteme zur Verfügung stehen, die Schlüsseldiebstahl ausschließen und mobilem Code widerstehen. Dann kann nämlich die traditionelle rechtliche Position aufrechterhalten werden, da der Vertrauende gute Nachweismöglichkeiten dafür besitzt, dass der Geschäftspartner die Signatur wirklich bewusst getätigt haben muss.
Ob die harten Anforderungen in naher Zukunft allerdings erfüllbar sind, ist ungewiss (mehr dazu im Kapitel 1.4).

Internationale Transaktionen

In diesem Abschnitt soll die derzeitige gesetzliche Lage betreffs der Gültigkeit digitaler Signaturen bei multinationalen Verträgen aufgezeigt werden. Im Gesetz zur digitalen Signatur (SigG), dessen Zweck es nach §1 ist, *„Rahmenbedingungen für digitale Signaturen zu schaffen, unter denen diese als sicher gelten und Fälschungen digitaler Signaturen oder Verfälschungen von signierten Daten zuverlässig festgestellt werden können“*, behandelt §15 die Problematik länderübergreifender Gültigkeiten digitaler Signaturen. So heißt es im Absatz 1:

„Digitale Signaturen, die mit einem öffentlichen Signaturschlüssel überprüft werden können, für den ein ausländisches Zertifikat aus einem anderen Mitgliedstaat der Europäischen Union oder aus einem anderen Vertragsstaat des Abkommens über den Europä-

ischen Wirtschaftsraum vorliegt, sind, soweit sie gleichwertige Sicherheit aufweisen, digitalen Signaturen nach diesem Gesetz gleichgestellt.“

Und im Absatz 2 wird noch weiter (geographisch gesehen) verallgemeinert:
„Absatz 1 gilt auch für andere Staaten, soweit überstaatliche oder zwischenstaatliche Vereinbarungen über die Anerkennung der Signaturschlüssel-Zertifikate getroffen sind.“

Länderspezifische Differenzierungen sollten, wo möglich, vermieden werden, denn die Attraktivität des Internets könnte sonst darunter leiden, wenn Benutzer erst Gesetzesbücher lesen müssten, bevor sie eine Transaktion starten. In Kombination mit den ISO-Standards belässt dieses Gesetz weitestgehend die gewünschte Flexibilität, denn diese geforderten Abkommen sind dann nicht schwierig zu treffen.

1.2.3 Repudiation-Szenarien und -Ablauf

In einem Kommunikations-Ablauf, z.B. einem Vertragsabschluss, gibt es die Möglichkeit, dass einer der Beteiligten später die Bedingungen abstreitet, oder noch schlimmer, leugnet, dass der Informationsaustausch überhaupt stattgefunden hat. Die folgenden Szenarien können auftreten ([ISO], [BLIXT]):

Leugnung der Herkunft: Die Beteiligten streiten darüber, ob eine bestimmte Seite die betreffenden Daten erzeugt und gesendet hat, z.B. die Bestellung (repudiation of origin), wichtig wäre also in diesem Zusammenhang ein Beweis für den Empfänger, dass die Nachricht wirklich vom angeblichen Sender initiiert worden ist (evidence of origin).

Leugnung der Übernahme durch eine Zustellungsinstanz: Der ISP (Internet Service
Provider) des Senders oder eine andere Zustellungsinstanz streitet ab, die Daten von diesem übernommen zu haben (repudiation of submission), also benötigt der Sender einen Nachweis, dass seine Nachricht wirklich von dieser angenommen worden ist.

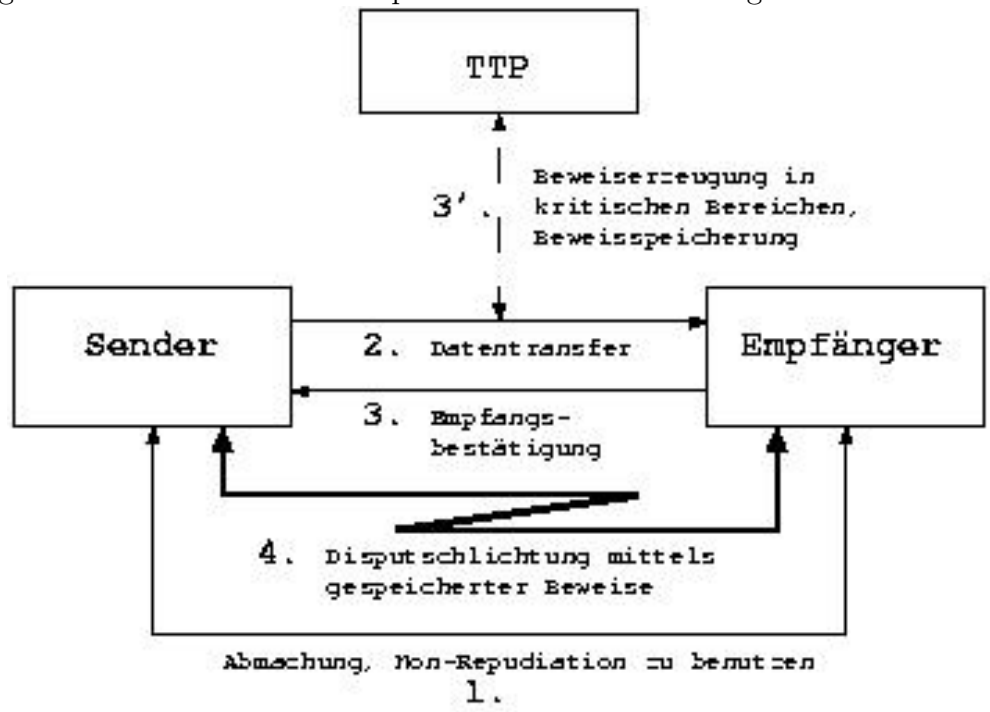
Leugnung der Annahme: Die Parteien widersprechen einander, indem der
Adressat abstreitet, die Daten empfangen zu haben (repudiation of delivery). Hierunter fallen auch Unstimmigkeiten über die Empfangs-Zeit der Nachricht. Es besteht also ein Beweis-Bedarf für den Sender, dass seine Nachricht (der richtigen Person) zugestellt worden ist.

Nichtanerkennung des Inhalts: Ein Beteiligter weist den Inhalt, z.B. einer Bestellung, zurück (repudiation of receipt). Auch hier muss nachweisbar sein, wessen Angaben zutreffen (evidence of receipt).

Natürlich sind auch Kombinationen dieser Szenarien denkbar. Im ISO-Dokument sind noch mehr Fälle aufgelistet und erklärt. Diese können aber unter die genannten subsummiert werden. Jedes dieser möglichen Ereignisse muss also durch ein Non-repudiation-Protokoll verhindert werden.

Ablauf des Non-repudiation-Prozesses

Im Allgemeinen wird eine Kommunikation, die Non-repudiation verwendet, wie im Folgenden beschrieben ablaufen (nach [BLIXT]). Als erstes müssen Sender und Empänger ihr Einverständnis geben, dass sie Non-repudiation verwenden wollen und werden. Dies kann entweder durch eine bestehende Abmachung oder durch den expliziten Wunsch einer der beiden Parteien geschehen. Während der Transaktionen wird jeweils, wenn einer der oben beschriebenen kritischen Abschnitte erreicht ist, ein entsprechender Beweis erzeugt. Dabei muss die Seite, die später diese Aktion abstreiten könnte, die Beweiserstellung unterstützen. Dieser Beweis wird dann dem Empfänger überbracht, von diesem verifiziert und dann abgespeichert (z.B. auch zusätzlich noch durch eine vertrauenswürdige dritte Partei TTP). Sollte später eine Uneinigkeit entstehen, müssen die gesammelten Beweise neu ausgewertet werden und der Disput kann auf dieser Basis gelöst werden.



1.3 Protokolle zur Realisierung von Nicht-Abstreitbarkeit

1.3.1 Anforderungen

Fairness: Von einem guten Protokoll wird erwartet, dass es beide Parteien fair behandelt, d.h. **dass es sowohl dem Sender als auch dem Empfänger einen unanfechtbaren gültigen Beweis liefert, ohne dass eine Partei einen Vorteil erlangen würde im Falle eines unvollständigen Ablaufs in jeder nur denkbaren Weise** (nach [ZHOU1]).

Effizienz: Effizienz ist im Gebiet der Informatik immer ein wichtiger Faktor. Besonders relevant ist dieser jedoch, wenn es um Netz-Belastung geht. Diese Anforderung stellt sich somit als vordergründig dar. Gewisse gute Ideen sind an diesem Aspekt

gescheitert, z.B. die Gradual Exchange Protocols ([ZHOU1]) (allmählich Informationen austauschende Protokolle), die versuchen durch aufwendige Algorithmen, Informationen der Gegenseite in kleinen Stückchen zugänglich zu machen, um so nicht selbst in Nachteil zu geraten. Dieser Ansatz verursacht nicht nur einen gewaltigen Overhead an Datentransfer, sondern er benachteiligt auch den Teilnehmer mit dem schwächeren Rechner, da teilweise ganz erhebliche Berechnungen durchgeführt werden müssen.

Unabhängigkeit: Protokolle werden autonom genannt, wenn jede beteiligte Partei ab einem gewissen Zeitpunkt einseitig die Transaktionen komplettieren kann, ohne Fairness einzubüßen. Autonome Protokolle sind am kompliziertesten aufgebaut, denn sie müssen Sub-Protokolle für jeden Abbruchzeitpunkt bereitstellen. Ein solches Protokoll ist in [ZHOU1] beschrieben. Deshalb sei an dieser Stelle darauf nicht weiter eingegangen.

Benutzbarkeit: Nur bedienungsfreundliche Protokolle werden eingesetzt und akzeptiert. Deshalb muss darauf geachtet werden, dass das Non-repudiation-Protokoll eine leicht bedienbare Schnittstelle zulässt (Details der Implementierung sollten verborgen werden können...). Auf diesen Aspekt wird jedoch im Weiteren nicht näher eingegangen.

1.3.2 Third Party Protocols

Kennzeichnend für diese Art von Protokollen, die sich am ehesten durchgesetzt haben, ist das Vorhandensein einer Trusted Third Party, kurz TTP (Vertrauenswürdige Dritte Partei), die Beweise über Transaktionen sammelt, aufbewahrt und im Bedarfsfall zugänglich macht und auf diese Weise vorurteilslos (im Sinne der vorher getätigten Abmachungen) schlichtet. Es gibt mehrere Einsatzmöglichkeiten für das TTP-Prinzip, z.B. zur Schlüssel- und Identitäts-Zertifizierung, als Time-Stamping-Server, Beweishüter, Zustellungsagent und „vollstreckendes“ Organ.
Natürlich setzt der Einsatz von TTPs voraus, dass diese extrem gut gegen unauthorisierte Zugriffe von außen abgesichert sind und dass deren Signaturen versteckt aufbewahrt werden. Ansonsten wird natürlich dieses Paradigma ad absurdum geführt.

On-Line TTP

Am Anfang der Entwicklung mussten die TTPs noch sehr aktiv an den einzelnen Schritten des Protokolls teilnehmen. Das Problem dabei war, dass auf diese Weise die TTP schnell zum Bottleneck (Engpass) wurde, was natürlich dem Effizienzanspruch widersprach. Daraus resultierte die Notwendigkeit, die Einbeziehung der TTP weitestgehend zu reduzieren, also die Entwicklung der TTP von einer Zustellungsbehörde hin zu einem selten beanspruchten Notar.

Off-Line TTP

Die Teilnahme der TTP kann z.B. dadurch verringert werden, dass sie als letzte Instanz hinzugezogen wird und vorher sich die beteiligten Parteien versuchen, selbst zu einigen. Dieser Ansatz ist dann effizienter, wenn die beiden Beteiligten in der Regel fair spielen. Wie ein On-Line- zu einem Off-Line-Protokoll ergänzt werden kann, ist im Abschnitt 1.3.5 kurz dargestellt.

Time-Stamp-Server TTP

Zeitstempel sind ein wichtiges Hilfsmittel, denn auch der Eintreff-Zeitpunkt von Nachrichten spielt eine entscheidende Rolle.
Zur Motivation ein kleines Beispiel: Ein Kunde möchte seine Bestellung ungeschehen machen, veröffentlicht schlauerweise seinen privaten Schlüssel und führt daraufhin einen Schlüssel-Rückruf (Ungültigkeitserklärung) durch, um jetzt argumentieren zu können, dass ja diese Bestellung genausogut von einem anderen hätte stammen können und so macht er natürlich der vertrauenden Partei, die die Beweispflicht hat, das Leben schwerer. Mit Zeitstempeln kann dieser Fall verhindert werden, indem einfach der Zeitpunkt des Rückrufs mit dem der Bestellung verglichen wird.
Welche Voraussetzungen muss also ein TSS (Time Stamping Server) erfüllen? Wenn die Vertrauenswürdigkeit über einen längeren Zeitraum aufrechterhalten werden soll, muss der Zeitschlüssel eine entsprechende Länge besitzen. Außerdem müssen natürlich sowohl der Privatschlüssel als auch die Zeit der Uhr des TSS (darf nicht zurücksetzbar sein) sehr sicher aufbewahrt werden können und Zeitstempel dürfen nicht vergeben werden, ohne die entsprechenden Zugriffsrechte.
Non-repudiation-Protokolle sollten die Zeit mitberücksichtigen, deshalb ist eine TSA (Trusted Time Stamping Authority) erforderlich.
In der nachfolgend vorgestellten Lösung werden die Nachrichten jedoch nicht mit Zeitstempeln versehen. Dies wurde aus Gründen der Übersichtlichkeit weggelassen, kann aber sehr einfach ergänzt werden. Eine konkrete Implementierung dafür kann der geneigte Leser in [BLIXT] nachvollziehen.

1.3.3 Notation

Die folgenden Formalisierungen legen die Grundlage für das Protokoll und können somit für mathematische Beweise, z.B. den Korrektheitsbeweis verwendet werden. Begriffe können im Glossar nachgeschlagen werden.

X, Y	Konkatenation zweier Nachrichten X und Y
$H(X)$	Bild der Nachricht X bezüglich der Einweg-Hashfunktion H
$eK(X)$, $dK(X)$	Ver-, bzw. Entschlüsselung der Nachricht X mit Schlüssel K
$sS_{\mathcal{A}}(X)$	$\mathcal{A}$'s digitale Signatur zur Nachricht X unter Verwendung des privaten Signatur-Schlüssels $S_{\mathcal{A}}$
$\mathcal{A} \rightarrow \mathcal{B} : X$	$\mathcal{A}$ sendet Nachricht X an $\mathcal{B}$
$\mathcal{A} \leftrightarrow \mathcal{B} : X$	$\mathcal{A}$ ruft Nachricht X von $\mathcal{B}$ ab, z.B. via FTP

Das nachfolgend vorgestellte Protokoll ist aus [ZHOU1] übernommen. Es liefert nach vollständigem Ablauf einen EOO (Evidence of Origin, Beweis der Herkunft) und ein EOR (Evidence of Receipt, Beweis des Empfangs), benutzt eine TTP und ist im Abbruchfall so fair, dass keinem Teilnehmer ein Nachteil entsteht. Die Grundidee ist es, die mit einem Schlüssel K kodierte Nachricht M an den Empfänger zu versenden, sich diese Zusendung bestätigen zu lassen und daraufhin den Schlüssel K bei einer TTP für den Empfänger zu hinterlegen.

1.3.4 Ablauf der Transaktionen

Folgende Voraussetzungen müssen erfüllt sein, damit das Protokoll korrekt arbeitet:

- Alle Beteiligten vertrauen der TTP.
- Die Non-repudiation-Strategie muss allen Beteiligten bekannt sein.
- Die Beteiligten besitzen die nötigen Schlüssel (ihren eigenen privaten und die öffentlichen der Anderen).
- Der Schlichter muss die erzeugten Beweise verifizieren können (denkbare Schlichter wären z.B. eine/die TTP oder ein Gericht).
- Die Übertragungskanäle sind nicht permanent unterbrochen.

Folgende Schritte sind nun notwendig, um eine Nachricht M von $\mathcal{A}$ nach $\mathcal{B}$ zu senden und dabei gegen alle genannten Eventualitäten abgesichert zu sein.

1. $\mathcal{A}$ legt einen Schlüssel K für seine Nachricht M fest
2. $\mathcal{A} \rightarrow \mathcal{B} : f_1, \mathcal{B}, L \equiv H(M,K), C \equiv eK(M), EOO_C \equiv sS_{\mathcal{A}}(f_1, \mathcal{B}, L, C)$,
   $\mathcal{A}$ sendet also zuerst ein Flag, das die Art der Nachricht anzeigt (bzw. um welchen Protokollschritt es sich handelt), dann den Empfänger (als Nachricht kodiert), dann die Projektion von M konkateniert mit K in den Ergebnisraum der Hashfunktion H zum Testen der Integrität (siehe Glossar), dann die mit K verschlüsselte Nachricht M und schließlich seine digitale Signatur (siehe Glossar) für diese Nachrichtenfolge als Beweis des Ursprungs für $\mathcal{B}$
3. $\mathcal{B}$ empfängt, prüft und speichert Nachrichtenfolge, insbesondere den Beweis EOO_C, die Überprüfung erfolgt mittels des öffentlichen Schlüssels von $\mathcal{A}$ (bereitgestellt durch eine Public Key Infrastructure PKI)
4. $\mathcal{B} \rightarrow \mathcal{A} : f_2, \mathcal{A}, L, EOR_C \equiv sS_{\mathcal{B}}(f_2, \mathcal{A}, L, C)$, d.h. $\mathcal{B}$ sendet wieder ein spezielles Statusflag, dann den Empfänger, jetzt $\mathcal{A}$, anschließend L (dient in dieser Situation dazu, es rechnerisch aufwendiger zu machen, einen Schlüssel $K_{wrong} \neq K$ zu finden mit $L = H(M,K) = H(M,K_{wrong})$ und $M = dK(C) = dK_{wrong}(C)$) und letztlich seine digitale Signatur zu dieser Nachrichtenkette
5. $\mathcal{A}$ muss nun EOR_C verifizieren und abspeichern
6. $\mathcal{A} \rightarrow TTP : f_5, \mathcal{B}, L, K, sub_K \equiv sS_{\mathcal{A}}(f_5, \mathcal{B}, L, K)$, in diesem Schritt schickt $\mathcal{A}$ der TTP den Schlüssel K, natürlich wieder mit entsprechender digitaler Signatur, diese dient der TTP als Beweis der Schlüsselübergabe (f_5, da Konsistenz in den Bezeichnungen gewahrt werden soll und f_3 und f_4 im ersten Verbesserungsvorschlag für die Zwischenschritte genutzt werden)
7. TTP gibt mit Lese-Rechten das Tupel
   $(f_6, \mathcal{A}, \mathcal{B}, L, K, con_K \equiv sS_{TTP}(f_6, \mathcal{A}, \mathcal{B}, L, K))$ öffentlich frei, z.B. mittels FTP

8. $\mathcal{A}$ und $\mathcal{B}$ rufen dieses Tupel von der TTP ab, in Formeln
   $\mathcal{A}, \mathcal{B} \leftrightarrow f_6, \mathcal{A}, \mathcal{B}, L, K, con_K$, damit hat $\mathcal{A}$ con_K, und folglich den Beweis dafür, dass $\mathcal{B}$ Zugriff auf den Nachrichtenschlüssel K und damit die Nachricht M erhalten kann und $\mathcal{B}$ benutzt den Schlüssel K, um M auszurechnen und speichert con_K als Nachweis, dass der Schlüssel wirklich von $\mathcal{A}$ stammt.

$\mathcal{B}$ muss dafür Sorge tragen, regelmäßig bei der TTP nachzufragen, ob K schon verfügbar ist, ansonsten verliert es sein Disputrecht! Ein unabhängiger Schlichter kann mögliche Unklarheiten leicht überprüfen, indem er sich die von $\mathcal{A}$ und $\mathcal{B}$ gesammelten Beweise anschaut.
In diesem Protokoll kommt der TTP eine eher kleine Rolle zu und der Netzverkehr, der über sie abgewickelt wird, ist auch eher vernachlässigbar, da ja die eigentliche möglicherweise lange Nachricht sie nicht passiert, sondern nur Schlüssel begrenzter Länge und die Hashabbildung der Nachricht, die in der Regel viel kleiner sein wird als das Original. Diese Hashabbildung hat den weiteren positiven Effekt, dass die TTP nichts über den Inhalt der Nachricht weiß, also die Diskretion verbessert wird (noch dazu, da auch andere, nicht vertraunswürdige, Personen Zugriff auf diese öffentliche Freigabe haben). Weiterhin ist die TTP nicht Zustellungsinstanz, muss also nicht ständig versuchen, $\mathcal{A}$, bzw. $\mathcal{B}$, zu erreichen, sondern $\mathcal{A}$ und $\mathcal{B}$ haben die Verpflichtung selbständig die sie betreffenden Informationen abzufragen.
Dieses Protokoll hat auch den Vorteil, dass die Fairness nicht durch die (Un-)Zuverlässigkeit des Kommunikationskanals beeinflusst wird (es sei denn, dieser ist permanent unterbrochen), eine sehr wünschenswerte Eigenschaft!
Ein Korrektheitsbeweis (basierend auf Fallunterscheidungen) für dieses Protokoll kann in [ZHOU3] nachgelesen werden.

1.3.5 Verbesserungen

Obwohl das oben genannte Protokoll die erwartete Funktionalität aufweist, gibt es einige mögliche Verbesserungen, auf die hier im Detail eingegangen werden soll.

Effizienz

Im Allgemeinen sollte die Kommunikation zwischen den Verhandlungspartnern ohne Komplikationen ablaufen. Diese Annahme wird in der folgenden Modifikation des besprochenen Protokolls ausgenutzt. Kern dabei ist es, eine Aufteilung in Haupt- und Recovery-Protokoll vorzunehmen.
Die ersten zwei Schritte verlaufen zunächst analog:

1. $\mathcal{A} \rightarrow \mathcal{B} : f_1, \mathcal{B}, L, C, EOO_C$
2. $\mathcal{B} \rightarrow \mathcal{A} : f_2, \mathcal{A}, L, EOR_C$

Jetzt gibt $\mathcal{A}$ das Schlüsselgeheimnis K frei. Das hört sich im ersten Moment unfair an, diese Situation kann aber durch $\mathcal{A}$ im Notfall noch entschärft werden.

3. $\mathcal{A} \rightarrow \mathcal{B} : f_3, \mathcal{B}, L, K, EOO_K \equiv sS_{\mathcal{A}}(f_3, \mathcal{B}, L, K)$

Jetzt gibt es zwei Möglichkeiten, der Normalfall wird sein, dass $\mathcal{B}$ $\mathcal{A}$ eine Bestätigung „freiwillig“ sendet:

4. $\mathcal{B} \rightarrow \mathcal{A} : f_4, \mathcal{A}, L, EOR_K \equiv sS_{\mathcal{B}}(f_4, \mathcal{A}, L, K)$

So ist die Übermittlung zur Zufriedenheit beider Teilnehmer abgeschlossen.
Sollte jedoch nach einer gewissen Zeit diese Bestätigung noch ausstehen, wendet sich $\mathcal{A}$ an die TTP und die letzten Schritte des ersten Protokolls werden ausgeführt.

3'. $\mathcal{A} \rightarrow TTP : f_5, \mathcal{B}, L, K, sub_K$

4'. $\mathcal{A}, \mathcal{B} \leftrightarrow f_6, \mathcal{A}, \mathcal{B}, L, K, con_K$

Damit ist die Ausgewogenheit wieder hergestellt, da $\mathcal{B}$ nun verpflichtet ist, die nötigen fehlenden Informationen von der TTP abzurufen, sofern diese $\mathcal{B}$ nicht eh schon vorgelegen haben.
Wie schon erwähnt, spart dieser Algorithmus Ressourcen der TTP. Aber auch hier spielen Zeitfaktoren eine entscheidende Rolle: Wie lange soll $\mathcal{A}$ nach dem dritten Schritt warten?

Signaturrückruf–Verkettetes Beweisen

Wie schon geschildert, ist es manchmal wichtig, Signaturen für ungültig zu erklären, da es sich ja um eine Fälschung handeln könnte (Schlüssel gestohlen und ausgenutzt o.ä.). Davon sollen aber nur die Signaturen betroffen sein, die nach dem Rückruf getätigt worden sind. Durch die Verwendung einer TSA (On-Line-Zeitstempel-Service) ist es, wie bereits erwähnt, möglich, dies herauszufinden.

1. $\mathcal{U} \rightarrow TSA : sS_{\mathcal{U}}(X)$

2. $TSA \rightarrow \mathcal{U} : T_g, sS_{TSA}(sS_{\mathcal{U}}(X), T_g)$

Dabei ist $\mathcal{U}$ ein Nutzer und T_g ist der Zeitpunkt der Generierung der Signatur.
Das Problem dabei ist leider, dass damit wieder Abhängigkeiten entstehen (Verfügbarkeit TSA) und die Netzbelastung zunimmt, ein hoher Preis.
Aber auch dafür gibt es Lösungsvorschläge (beschrieben in [ZHOU2]). So könnten die Signatur-Schlüssel in zwei Kategorien eingeordnet werden, widerrufbar (als Langzeitschlüssel) und nicht widerrufbar (temporär), wobei letztere mit Hilfe des widerrufbaren Schlüssels erzeugt werden und zum eigentlichen Unterzeichnen der Nachrichten dienen.
Weiterhin wird in [ZHOU2] das Prinzip des verkettenen Beweisens vorgestellt. Der Vorteil dabei ist, dass $\mathcal{A}$ und $\mathcal{B}$ die TSA nicht kontaktieren müssen, damit EOO und EOR den Zeitstempel bekommen. Erreicht wird dies dadurch, dass die Non-repudiation-Beweise EOO, EOR und con_K miteinander verknüpft werden und in dem Moment Gültigkeit erlangen, in dem die TTP con_K generiert.
Dieses Vorgehen ist aber meiner Meinung nach auch problematisch, denn in dem vorgestellten Protokoll wird die Verknüpfung dadurch realisiert, dass jeweis die Beweise der Vorschritte mitgeschickt werden. Dies führt zu einer höheren Netzbelastung.
Dennoch setzt dieses Protokoll einen der vielversprechendsten Ansätze um, denn es löst die oben genannten Problem zuverlässig, wie in [ZHOU2] auch nachgewiesen wird.

Eine spieltheoretische Annäherung

Ein großes Problem haben diese Protokolle jedoch noch. Es ist absolut nicht offensichtlich, dass sie in jedem denkbaren Fall den Anforderungen gerecht werden, gerade auch, wenn es diverse Subprotokolle und Handlungsalternativen gibt. Dafür sind umfangreichere und auch durch viele Fallunterscheidungen unübersichtliche Beweise nötig. Prinzipiell nicht so schlimm, aber ein gegebenes Protokoll zu erweitern, stellt eine nur schwer lösbare Aufgabe dar. Sehr schnell können sich subtile Fehler einschleichen.
Deshalb verfolgt Kremer einen anderer Ansatz für die Spezifikation und Verifikation solcher Protokolle. Dieser setzt z.B. keine genau festgelegte Reihenfolge der Aktionen voraus. In seinem Paper [GameVer] erfolgt eine spieltheoretische Untersuchung von Nonrepudiation-Protokollen. Kerngedanke ist es, die beteiligten Transaktionspartner (auch die TTP) und die Kanäle als Spieler zu betrachten, die mehrere verschiedene Strategien verfolgen könnten. Beispielsweise kann Fairness für $\mathcal{A}$ wie folgt ausgedrückt werden: *eine Koalition von $\mathcal{B}$ mit allen Kommunikationskanälen hat keine Strategie, die $\mathcal{B}$ EOO verschafft und $\mathcal{A}$ keine Möglichkeit lässt, EOR zu erlangen.* Diese Herangehensweise zieht von Anfang an explizit Fehlverhalten, ja sogar feindliches Verhalten, in Betracht. Und Anforderungen können einfach als Existenz gewisser Strategien festgelegt werden. Die Modellierung der Kommunikationskanäle und der TTP als Spieler räumt mehr Analysefreiräume ein, denn soll ein Wohlverhalten vorausgesetzt werden, z.B. Vertrauen zur TTP, muss nur dafür gesorgt werden, dass der entsprechende Alternativenvorrat zu jedem Zeitpunkt einelementig ist, d.h. eine deterministische Strategie vorliegt.
In dem Dokument ist sogar nicht nur die intensive Untersuchung des Protokolls beschrieben, sondern auch eine winzige Schwachstelle ausgemacht worden (diese kann allerdings nur dann Schaden verursachen, wenn $\mathcal{B}$ mit dem Kommunikationskanal zusammenarbeitet).

Anhand dieser Beispiele wird deutlich, wie viele mögliche Ansätze es gibt und dass zwischen den einzelnen Zielsetzungen Abhängigkeiten, aber auch Konflikte bestehen (Effizienz $\leftrightarrow$ Sicherheit), die noch behandelt werden müssen.

1.4 Hardware- und Betriebssystem-Anforderungen

Die wesentlichsten Verbesserungen der Hardware sind zweifelsohne im Bereich der vertrauenswürdigen Computersysteme von Nöten (TCS Trusted Computing Systems). TCS müssen hohen Qualitäts- und Sicherheitsansprüchen genügen, wenn wichtige Geschäftsbeziehungen mit ihrer Hilfe abgewickelt werden sollen. Sie sollten genau in Einstimmung mit ihrer dokumentierten Spezifikation arbeiten und unerlaubte Aktivitäten verhindern. Es ist von Vorteil, die Computersysteme nach Vertrauenswürdigkeit zu klassifizieren. Dies ist z.B. in den „Common Criteria“ [ComCrit] geschehen und wurde daraufhin zum ISO-Standard (ISO 15408) erklärt. Das Problem dabei ist aber, dass nur auf bekannte Sicherheitsrisiken und -lücken hingewiesen werden kann. Ein weiterer kritischer Punkt ist darin zu sehen, dass heutzutage sehr viele nicht vertrauenswürdige Systeme mit noch weniger vertrauenswürdiger Software vernetzt sind und so ein totaler Schutz während On-Line-Transaktionen wahrscheinlich gar nicht realisierbar ist ([McCUL]). Es wird angestrebt, hoch vertrauenswürdige Subsysteme in diese einzubauen, z.B. in Form von zusätzlichen Pin-Pads o.ä., um größtmögliche Sicherheit zu erreichen.

Die TCSEC (Trusted Computer System Evaluation Criteria) beschreiben sechs fundamentale Anforderungen:

- wohldefinierte Sicherheitsstrategie
- Zugriffskontrolle für kritische Objekte
- Benutzer-Identifizierung
- Informationen müssen so aufbewahrt werden, dass sicherheitsrelevante Handlungen später nachvollzogen werden können, also der Verantwortliche ausfindig gemacht werden kann
- Hardware und Softwaremechanismen können unabhängig bewertet werden
- dauerhafter und kontinuierlicher Schutz

ITSEC (Information Technology Security Evaluation and Certification) [ITSEC] ging sogar noch einen Schritt weiter und definierte eine siebenstufige „Vertrauenshierarchie“ für Informationssysteme, von der niedrigsten Stufe E0 mit unzureichenden Versicherungen bis hin zu E6 mit formalen Beweisen für das genau spezifizierte Sicherheitsmodell.
Nur mit einem vertrauenswürdigem System können Probleme wie der Einsatz von mobilem Code oder Schlüsseldiebstähle eingedämmt werden (es wäre vermessen, anzunehmen, dass ein E6-System total unangreifbar ist, da die Spezifikationen immer noch unzulänglich sein können). Sind derartige Systeme an einer Transaktion beteiligt, fällt es auch nicht schwer, bestehendes Recht für den eCommerce zu übernehmen, denn dann sind Beweise wesentlich leichter zu beschaffen.

1.5 Zusammenfassung und Bewertung

Da Sicherheit im eCommerce eine der Herausforderungen der nächsten Jahre sein wird, ist es von großer Bedeutung, Probleme wie die Abstreitung von Nachrichten zu lösen.
In dieser Arbeit ist ein Protokoll vorgestellt worden, das die wichtigsten Anforderungen an ein Non-repudiation-Protokoll erfüllt. Diese sind Fairness, Sicherheit und im weiteren Sinne auch Effizienz. Außerdem sind einige Verbesserungsvorschläge erwähnt worden. In der Praxis sind Protokolle wie dieses durchaus anwendbar, da vor dem Normal-Nutzer die Details der Implementierung, die recht kompliziert werden können, verborgen sind.
In letzter Zeit ist übrigens ein boomender Markt für professionelle Lösungen in diesem Bereich entstanden, es gibt inzwischen viele Trust-Center o.ä. Dieser Fakt wirft natürlich die Frage auf, wie vertrauenswürdig diese Angebote wirklich sind. Der Kunde sollte unbedingt auf einer Akkreditierung durch die Regulierungsbehörde bestehen, denn nur so haftet das Trust-Center mit hohen Summen dafür, über viele Jahre zuverlässig seine Aufgabe zu erfüllen und alle Unterlagen aufzuheben.
Heutzutage sollte niemand mehr blauäugig elektronische Abmachungen tätigen (egal in welcher Richtung). Denn moderne Signaturverfahren sorgen für sehr gute Aussichten im Falle eines Gerichtsverfahrens, denn obwohl das Gericht Beweise „frei würdigen“ kann, so muss es doch Fälschungssicherheit angemessen berücksichtigen, die durch Infrastrukturen, Verfahrensregeln und natürlich die Technik der „qualifizierten“ digitalen Signatur nach dem deutschen Signaturgesetz gewährleistet werden. Dies bedeutet, dass in einem

privatrechtlichen Streitfall eine gesetzeskonforme elektronische Willenserklärung als unverfälscht akzeptiert wird, sofern keine schwerwiegenden, darzulegenden Gründe dagegen sprechen („Anscheinsbeweis“). Bei rechtlich ungeregelten Signaturen sind hingegen aufwendige Prozesse von Nöten, deren Ausgang ungewiss ist ([TC]).

1.5.1 Untersuchung einer heutigen Anwendung eines Nichtabstreitbarkeitsverfahrens

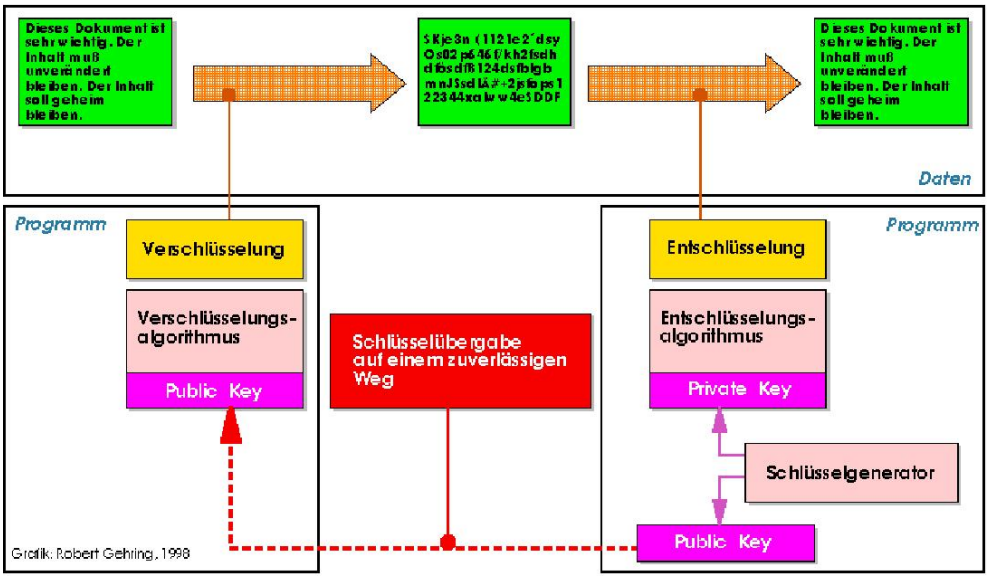
Ich habe mich im Rahmen dieser Arbeit auch bei Kreditinstituten informiert, inwieweit Non-repudiation-Protokolle z.B. beim Online-Banking schon zum Einsatz kommen. Dabei musste ich feststellen, dass Sicherheit zumeist auf andere (unelegantere) Weise zu garantieren versucht wird. So benutzen die meisten Banken das PIN/TAN-Prinzip, bei dem der Nutzer eine PIN festlegen muss, eine Liste von TANs (Transaktionsnummern) zugesendet bekommt und bei elektronischen Transaktionen jeweils eine benutzt und abstreicht.
Dies ist kein Verfahren mit Zukunftsperspektive, denn diese Nummern könnten ja wirklich leicht gestohlen werden. Ich denke, dass biologische Merkmale in nächster Zeit verstärkt eingesetzt werden sollten, denn nur auf diese Weise ist eine 100prozentige Identifikation möglich. Aber derartige Technologien sind im Moment noch nicht für die Massenproduktion ausgereift und finanzierbar. **Non-repudiation-Protokolle sind dem PIN/TAN-Prinzip meiner Meinung nach vorzuziehen.**

Die Anwendung von Non-repudiation-Verfahren hilft nicht nur bei der Beweisführung, nein, es dürfte auch eine wirkungsvolle Abschreckung darstellen, so dass leichtfertige Vertragsabschlüsse hinter dem Schutz der Anonymität nicht mehr ohne weiteres durchführbar sind.

1.6 Glossar

Im folgenden werden die gebräuchlichsten Begriffe im Bereich eCommerce, speziell auf Non-repudiation bezogen, in alphabetischer Reihenfolge erläutert (Definitionen nach [ISO], Bilder und ein paar Erklärungen aus [CRYPTO]).

Asymmetrisches Verschlüsselungs-Verfahren: Der Sender besitzt einen **öffentlichen Schlüssel**, der von außen frei zugänglich ist, und einen **privaten Schlüssel**, den nur er kennt. Dieses Schlüsselpaar hat die besondere Eigenschaft, dass mit dem einen Schlüssel entschlüsselt werden kann, was mit dem anderen Schlüssel verschlüsselt wurde. Kodiert also dieser Sender seine Nachricht mit seinem geheimen Schlüssel, können andere diese Nachricht lesen (dank öffentlichem Schlüssel) und wissen gleichzeitig, dass diese Nachricht wirklich vom Sender stammt (**Authentifizierung**). Der andere Fall ist, dass jemand ein Geheimnis an diesen schicken will. Dieses verschlüsselt er einfach mit dem öffentlichen Schlüssel der betreffenden Person, und nur diese Person kann (dank ihrem Privatschlüssel) diese Nachricht hinterher sinnvoll entschlüsseln.



Authentifizierung: Beweis der eigenen Identität.

Digitale Signatur: Daten, angehängt an oder eine kryptographische Transformation von einer Daten-Einheit, die dem Empfänger erlaubt, die Quelle und die Integrität zu überprüfen und die gegen Fälschung, z.B. durch den Empfänger, schützt ([ISO]). Digitale Signaturen können unter sicheren Voraussetzungen in elektronischen Medien (u.a.) die Funktionen von Unterschriften übernehmen. Mit ihrer Hilfe kann dem Unterzeichner nachgewiesen werden, daß er „eigenhändig“ ein vorliegendes Dokument „unterzeichnet“ hat.

- Es kann nachgewiesen werden, daß die Signatur zum Dokument gehört.
- Es kann nachgewiesen werden, daß der Inhalt des Dokuments unverändert ist.
- Wenn eine Zeitmarkierung vorgesehen ist, kann nachgewiesen werden, daß das Dokument zu einem bestimmten Zeitpunkt unterzeichnet wurde.

mögliches Vorgehen: Hash-Wert der Nachricht wird mit privatem Schlüssel des Senders kodiert und an die Nachricht angehängt.
Wenn außerdem technisch sichergestellt ist, daß das unterzeichnete Dokument vor der Unterzeichnung vollständig wahrgenommen wurde, erfüllen digitale Signaturen die rechtlichen Anforderungen an Unterschriften. Letztere Informationen stammen aus [CRYPTO].

eCommerce: Electronic Commerce ist ein Konzept zur Nutzung von bestimmten Informations- und Kommunikations-Technologien zur elektronischen Integration und Verzahnung von Wertschöpfungsketten und unternehmensübergreifenden Geschäftsprozessen (Def. nach P. Koter). Dieser Begriff umfasst also mehr als nur das reine Verkaufen über das Medium Internet.

Hash-Funktion: Eine Funktion, die beliebige Bit-Strings abbildet auf Bit-Strings fester Länger und dabei den folgenden Ansprüchen genügt:
Es ist auch mit Hilfe von Computern nahezu unmöglich,

- eine Eingabe zu finden, die auf eine gegebene Ausgabe abgebildet wird

- zu einer gegebenen Eingabe eine zweite (natürlich von der ersten verschiedene) zu finden, die auf die gleiche Ausgabe abgebildet wird.

Hashfunktionen werden hauptsächlich benutzt, um die Integrität von Daten zu überprüfen.
Dabei wird folgendermaßen vorgegangen (zitiert aus [CRYPTO]):
Zuerst wird die zu hashende Information in Blöcke passender Länge zerteilt. Sollte die Länge der Information kürzer sein, als für die Hashfunktion benötigt, wird sie durch Hinzufügen willkürlicher Zeichen „verlängert“. Jeder einzelne Block durchläuft dann die Hashfunktion und wird auf eine feste Länge komprimiert. So werden z.B. aus 512 Bytes (Zeichen) Information 128 Bytes (128 Zeichen) Hashwert erzeugt. Dieser resultierende Hashwert wird der Hashfunktion zusammen mit dem nächsten Informationsblock wieder zugeführt und mit dessen Hashwert verknüpft (z.B. durch XOR). Dieser Vorgang wird solange wiederholt, bis alle Blöcke der ursprünglichen Information abgearbeitet sind. Im Ergebnis erhält man einen einzelnen Hashwert, den sogenannten „message digest“, auch „digitaler Fingerabdruck“ genannt. Zu jeder Information, die eingegeben wird, erzeugt die Hashfunktionen einen anderen „Fingerabdruck“.

Integrität (von Daten): Die Eigenschaft, dass Daten nicht in einer unauthorisierten Weise verändert oder zerstört worden sind.

Schlüssel: Eine Sequenz von Symbolen, die Operationen kryptographischer Transformation kontrollieren, z.B. Ver- und Entschlüsselung, Erstellung kryptographischer Prüffunktionen, Berechnung und Verifizierung digitaler Signaturen.

privater: siehe asymmetrisches Verschlüsselungs-Verfahren

öffentlicher: siehe asymmetrisches Verschlüsselungs-Verfahren

Sicherheits-Zertifikat: Eine Menge sicherheitsrelevanter Daten, ausgestellt von einer Sicherheits-„behörde“ oder einer vertrauenswürdigen dritten Partei, zusammen mit Sicherheits-Informationen, die benutzt werden, um die Integrität und den Ursprung der Daten zu authentifizieren.

Symmetrisches Verschlüsselungs-Verfahren: Es werden für die Ver- und Entschlüsselung einer Nachricht Kopien ein- und desselben Schlüssels eingesetzt. Wichtig ist, dass die Schlüsselübergabe auf sicherem Weg erfolgt.

Vertrauen: Eine Beziehung zwischen zwei Elementen, eine Menge von Aktivitäten und eine Sicherheitsstrategie in der Element x y vertraut dann und nur dann, wenn x glaubt, dass y sich auf eine wohldefinierte Art und Weise verhalten wird, die nicht die gegebene Sicherheitsstrategie verletzt.

Vertrauenswürdige dritte Partei (TTP Trusted Third Party): Eine Institution, der im Hinblick auf sicherheitsbezogene Aktivitäten vertraut wird.

Zeitstempel: Ist nur dann einsetzbar, wenn er von einer TSA (Trusted Time Stamp Authority) erstellt werden kann. Enthält Informationen über Zeit und Datum einer Aktion.

Literaturverzeichnis

[BLIXT] Karl-Fredrik Blixt, Åsa Hagström.
Adding Non-Repudiation to Web Transactions. 1999
http://www.it.isy.liu.se/∼asa/publications/NonRepudiation.html

[BW] Ralf Bendrath.
Die Bundeswehr auf dem Weg ins digitale Schlachtfeld. 2000
http://www.heise.de/tp/deutsch/special/info/8326/1.html

[ComCrit] *The Common Criteria Homepage.*
http://www.commoncriteria.org/

[CRYPTO] Robert Gehring.
Sicherheit im elektronischen Geschäftsverkehr - Rechtsverbindlichkeit durch digitale Signaturen. 1998
http://ig.cs.tu-berlin.de/ap/rg/1998-06

[eCom] P. Koter.
Marketing-Management. Analyse, Planung, Umsetzung und Steuerung. 1999
http://www.online-marketingmix.de/definition-ecommerce.htm

[GameVer] Steve Kremer, Jean-François Raskin.
Formal Verification of Non-Repudiation Protocols – A Game Approach. 2000
http://www.ulb.ac.be/di/scsi/skremer/research.html

[ISO] ISO/IEC 13888-1.
Information technology–Security techniques–Non-repudiation, Part 1: General. 1997

[ITSEC] *Information Technology Security Evaluation and Certification Scheme.*
http://www.cesg.gov.uk/assurance/iacs/itsec/index.htm

[McCUL] Adrian McCullagh, William Caelli.
Non-Repudiation in the Digital Environment.
http://www.firstmonday.dk/issues/ossie5_8/mccullagh/

[Net] Klaus Schliessl.
Das Internet. 1995
http://www.fmi.uni-passau.de/internet/Internet.html

[TC] D-Trust.
Rechtliche Auswirkungen des deutschen Signaturgesetzes im Überblick.
http://www.d-trust.de/internet/content/sigg-ausw-ueberblick.html

[ZHOU1] Jianying Zhou, Robert Deng, Feng Bao.
Evolution of Fair Non-repudiation with TTP.
http://homex.coolconnect.com/user/jyzhou/publications.html

[ZHOU2] Cheng-Hwee You, Jianying Zhou, Kwok-Yan Lam.
On the Efficient Implementation of Fair Non-repudiation.
http://www.acm.org/sigcomm/ccr/archive/1998/oct98/ccr-9810-you.pdf

[ZHOU3] Jianying Zhou, Dieter Gollmann.
A Fair Non-repudiation Protocol.
http://homex.coolconnect.com/user/jyzhou/publications.html

Kapitel 2

Security in E-Commerce

Nico Krebs

Inhaltsverzeichnis

2.1 Einleitung

Durch seine sehr starke Verbreitung bietet das Internet Herstellern, Händlern und Vertretern ganz neue Möglichkeiten seine Produkte anzubieten. Man erreicht heute über das Internet theoretisch fast jeden Haushalt weltweit. Dieses Handeln mit elektronischen oder materiellen Gütern und die Bereitstellung von Dienstleistungen im Internet wird

als E-Commerce bezeichnet. Dabei wird der gesamte Bereich nochmals in zwei Bereiche unterteilt. Mit B2B (Business to Business) werden Handelsbeziehungen zwischen zwei Unternehmen bezeichnet. Und B2C (Business to Customer) bezeichnet die Handelsbeziehungen zwischen Unternehmen und ihren Kunden. Das Internet als Medium des Handels hat jedoch einen Nachteil. Mangelhafte oder fehlende Sicherheitsstandards machen es unsicher. Um dennoch eine sichere Kommunikation über ein unsicheres Medium führen zu können, bedarf es bestimmter Verfahren und Techniken.

2.1.1 Ein Modell für den Verkaufsprozess im Internet

Der überwiegend von kleineren Anbietern im Internet betriebene Handel kann durch ein sehr einfaches Modell beschrieben werden (siehe Abbildung 1). In diesem Modell, dass einen Verkaufsprozess modellieren soll, gibt es mindestens einen Verkäufer und einen Kunden. Der gesamte Prozess des Verkaufens kann in verschiedene Phasen aufgeteilt werden. Der Prozess beginnt mit einer Werbungsphase. Hier wird dem Kunden ein Angebot gemacht. Dies geschieht zum Beispiel, indem er sich einen Katalog des Verkäufers im Internet anschaut. Hat der Kunde nun den Wunsch, etwas davon zu erwerben, beginnt die Bestellphase. Er wird in der Regel ein Formular auf der Webseite ausfüllen und zum Verkäufer schicken. Der Verkäufer wird die Bestellung bearbeiten und zusammenstellen. Nun beginnt die Lieferphase. Nachdem der Kunde die Ware erhalten hat, beginnt die Zahlungsphase. Er wird zum Beispiel die mitgelieferte Rechnung begleichen.

Abbildung 2.1: Ein einfaches Modell für den Verkaufsprozess mit seinen vier Phasen.

Diese vier Phasen kann man nochmals in zwei Bereiche teilen. Die Werbungsphase und die Bestellphase finden online, also über das Internet statt, die Lieferphase und die Zahlungsphase dagegen offline. Relevant für unsere Betrachtungen sind daher die Werbungsphase und die Bestellphase. In der Werbungsphase gelangt der Kunde auf die Webseite des Verkäufers und kann dort Preisinformationen über bestimmte Artikel einsehen. Wird der gewünschte Artikel von mehreren Verkäufern angeboten, so wird der Kunde zunächst einen Preisvergleich machen. Aber was würde geschehen, wenn ein konkurrierender Verkäufer A die Webseite eines anderen Verkäufers B derart manipuliert, dass B´s Preise nun höher sind? Das ist nachteilig für den Kunden und den Verkäufer B. Der Kunde muss jetzt mehr bezahlen und dem Verkäufer B fehlt der Umsatz. Im Falle einer Verringerung

der Preise würde der Kunde darauf bestehen, den Artikel vom Verkäufer B zu dem niedrigeren Preis zu erhalten.
Auch in der Bestellphase gibt es Sicherheitslücken. So könnten beispielsweise Bestellungen unter Vortäuschung falscher Identitäten getätigt werden. Dies würde sich dann erst nach der Lieferung herausstellen. Wer kommt in diesem Fall für die Versandkosten auf? Wie sollte man beweisen, dass man nichts bestellt hat? Es wäre auch möglich, dass Bestellungen auf dem Weg durch das Internet abgefangen und manipuliert werden.
Das Problem liegt hier jeweils bei den Identitäten. Der Kunde möchte sicher sein, dass er wirklich die Angebote sieht, die der Verkäufer auf seiner Webseite anbietet und der Verkäufer möchte gern die Sicherheit haben, dass der Kunde auch wirklich der ist, für den er sich ausgibt. Aber wie kann der Kunde den Verkäufer davon überzeugen, dass er der richtige ist?

2.1.2 Ziel und Gliederung dieser Arbeit

Ziel dieser Arbeit ist es, ausgewählte Sicherheitsmechanismen und Verfahren vorzustellen. Stärken und Schwächen dieser Verfahren sollen aufgezeigt werden, um sie den verschiedenen Anwendungsgebieten des E-Commerce zuzuordnen.

Zunächst werden an dem in ersten Abschnitt beschriebenen Modell die notwendigen Kommunikationsschritte und die darin enthaltenen Sicherheitsrisiken aufgezeigt. Es werden Institutionen, Techniken und Verfahren vorgestellt, die diese Sicherheitslücken schließen können. Der dritte Abschnitt behandelt 3 Protokolle aus den Bereichen des sicheren Bezahlens und des sicheren Datentransfers. Dabei soll auf die Stärken und Schwächen eingegangen werden, um sie aufgrund ihrer Eigenschaften bestimmten Bereichen des E-Commerce zuordnen zu können. Im vierten Abschnitt werden etwas komplexere Handelsmodelle vorgestellt, die auf den selben zuvor vorgestellten Voraussetzungen für eine sichere Kommunikation aufbauen.

2.2 Ausgewählte Mechanismen für eine sichere Kommunikation

Als Kommunikation wird der Austausch von Informationen bezeichnet. Dabei kann es sich zum Beispiel um Angebote, Bestellungen oder Kreditkarteninformationen handeln, die über das Internet versandt werden. Das Internet selbst stellt aufgrund seiner mangelnden Sicherheitsstandards ein unsicheres Medium dar. Als ein Netzwerk, dass auch nach größeren Ausfällen noch funktionsfähig sein sollte, wurde es im Auftrag des US-Militärs entwickelt. Um das gewährleisten zu können, wurde das Internet dezentral organisiert. Informationen werden in Pakete zerteilt und über viele dem Sender oft unbekannte Stationen weitergeleitet. Durch die Öffnung des Netzes kann sich heute leicht ein Rechner an das Internet anschließen und selbst Anbieter von Diensten werden, die Pakete weiter leiten. Dabei könnten die Pakete vor dem Weiterleiten gelesen oder verändert werden. Bestimmte Mechanismen machen es jedoch möglich eine sichere Kommunikation über ein unsicheres Medium aufzubauen. Abhörsicherheit gewährleistende Verschlüsselungsverfahren, Verfahren der Identitätsüberprüfung und das Verfahren der digitalen Signatur,

welches der Nachricht einen bestimmten Absender zuordnen und die Unverfälschtheit der Daten bestätigen kann, sollen in diesem Abschnitt vorgestellt werden.

2.2.1 TTP - die Schiedsrichter

Will eine Person einer unbekannten Person über das Internet glaubhaft machen, dass er der ist, für den er sich ausgibt, dann benötigt er bewiesenermaßen einen Schiedsrichter, also eine dritte Partei. Eine trusted third party (TTP) ist eine solche dritte unparteiische Institution. Eine Form einer TTP ist die certificate authority (CA). Eine ihrer Aufgabe ist es, Identitäten zu prüfen und zu bestätigen. Treffen nun zwei sich zunächst unbekannte Parteien aufeinander, so könnten diese jeweils behaupten zu sein, wer sie sind und auch eine CA angeben, die das bestätigen kann. Beide haben also einen Zeugen für die Echtheit ihrer Identität. Es scheint zunächst so, als würde das eigentliche Problem nur verlagert. Warum sollte man diesem Schiedsrichter trauen können? Es gibt einige CAs, die den gesetzlichen Anforderungen genügen. Zum Beispiel gibt es in Deutschland das Signaturgesetz [SigG]. Darin sind Rahmenbedingungen für den Betrieb einer CA (§4) und der Vergabe der Zertifikate geregelt (§5). Kann eine CA nachweisen, dass sie den gesetzlichen Anforderungen genügt, so bekommt sie ein Gütesiegel von der zuständigen Behörde (in Deutschland von der Regulierungsbehörde für Telekommunikation und Post (RegTP)). Einer auf dieser Weise akkreditierten CA kann also vertraut werden. Durch dieses Vertrauen, was beide Personen gegenüber der CA haben, kann auch das Vertrauen direkt zwischen den Personen aufgebaut werden.

2.2.2 Zertifikate

Eine Aufgabe der CA ist es, Identitäten zu überprüfen. Wurde eine in digitaler Form vorliegende Identität gemäß den gemachten Angaben überprüft, so wird diese Identität von der CA unterschrieben. Eine solche Identität nennt man dann Zertifikat.

Zertifikate haben ein bestimmtes Format. Das am weitesten verbreitete Format ist X.509v3 [ITU88]. Das Zertifikat enthält somit ganz bestimmte Angaben, wie den Namen des Inhabers, eine laufende Nummer, den Gültigkeitszeitraum, den Namen der ausstellenden CA, den Namen des Staates in dem die CA niedergelassen ist, die Unterschrift der CA, die Zertifikatsart und die Zertifikatsklasse. Das X.509v3-Format erfüllt somit auch die gesetzlichen Forderungen [SigG].

Zertifikatsarten

Zertifikate sind aufgrund ihres einheitlichen Formates eigentlich gleichwertig und sollten nicht in ihrer Art zu unterscheiden sein. Jedoch ist eine CA für die von ihr ausgestellten Zertifikate haftbar. Deshalb kann sie die vergebenen Zertifikate auch nur für bestimmte Anwendungsbereiche freigeben. Diese Bereiche sind beispielsweise die sichere Übertragung von E-Mails, die Clientauthentifizierung, die Unterschrift unter einem Programm oder die Serverauthentifizierung.

Zertifikatsklassen

Je nach geleisteten Aufwand den die CA erbringt um die Inhalte der Zertifikate zu überprüfen, stellt sie Zertifikate in unterschiedlichen Klassen aus. Anhand der Klasse eines vorgelegten Zertifikats kann auf einfache Weise die Vertrauenswürdigkeit der angegebenen Inhalte abgeschätzt werden. Genau vorgeschrieben ist die Vergabe der einzelnen Klassen nicht. Dennoch erfolgt die Vergabe der Zertifikatsklassen bei den verschiedenen CAs ungefähr gleich. Zum Beispiel bietet TC Trustcenter, eine in Deutschland tätige CA, Zertifikate in den Klassen 0 bis 4 an. In ihren Zertifizierungsrichtlinien ist festgelegt, wie die Daten in den einzelnen Klassen überprüft werden [TC Richtlinien].
Die Klasse 0 ist nur für Testzwecke gedacht und der angegebene Inhalt wird keiner Prüfung unterzogen.
Zertifikate der Klasse 1 werden ausgestellt, nachdem die angegebene E-Mail-Adresse überprüft worden ist.
Für die Klasse 2 ist zusätzlich eine Kopie des Ausweises einzureichen.
Die persönliche Identitätsfeststellung, zum Beispiel in einer Postfiliale, ist für die Klasse 3 notwendig.
Die sicherste Stufe stellt die Klasse 4 dar. Hier wird die Identitätsprüfung bei einer Meldebehörde vorgenommen.

Bei Zertifikaten für Unternehmen tritt an die Stelle des Ausweises ein Auszug aus dem Handelsregister. Für den E-Commerce werden Zertifikate der Klasse 3 empfohlen um das notwendige Vertrauen zu schaffen.

Wenn unser kleines Modell sicher sein soll, so brauchen also Verkäufer und Kunde jeweils ein Zertifikat (siehe Abbildung 2). Der Verkäufer sollte eine Zertifikatsart haben, mit der er seinen Server authentifizieren kann. Denn er möchte, dass der Kunde sicher ist, dass er auch wirklich die Webseite sieht, die der Verkäufer erstellt hat. Damit der Kunde auch dem Zertifikat trauen kann, sollte es mindestens eines der Klasse 3 sein. Der Kunde sollte seine Identität auch nachweisen können. Er benötigt ebenfalls ein Zertifikat der Klasse 3 und abhängig von der Art des Verbindungsaufbaues ist eine Zertifikatsart notwendig, die ihn als Client authentifiziert oder die ihm das Schreiben sicherer E-Mails ermöglicht. Ebenfalls wird das Vorhandensein einer CA notwendig. Sie muss zunächst die Zertifikate ausstellen und später die Möglichkeit bieten Zertifikate zu überprüfen.

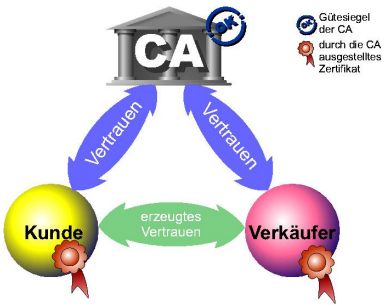


Abbildung 2.2: Kunde und Verkäufer besitzen ein Zertifikat, dass von einer vertrauenswürdigen CA ausgestellt wurde.

Alle Beteiligten haben jetzt eine überprüfbare Identität. Aber wie kann man sicher sein, dass diese nicht gefälscht ist oder jemand einfach die Identität eines anderen kopiert und selbst verwendet? In der realen Welt können sich Personen mit ihrem Personalausweis ausweisen. Das funktioniert, da der Personalausweis einige nicht kopierbare Merkmale besitzt und somit nicht gefälscht werden kann. Im Internet liegen Zertifikate jedoch als digitale Daten vor, die beliebig oft kopiert und verändert werden können. Um die eindeutige Zuordnung zwischen Zertifikat und Person auch im Internet zu gewährleisten, benötigt man spezielle Verschlüsselungsverfahren.

2.2.3 Verschlüsselungsverfahren

Mathematisch umkehrbare Verfahren, bei denen lesbare Daten so verändert werden, dass sie nicht mehr für einen Unberechtigten lesbar sind, nennt man Verschlüsselungsverfahren. Diese kommen zur Anwendung, wenn man eine geheime Nachricht versenden möchte und sich nicht sicher sein kann, ob diese auf dem Weg bis zum Empfänger von anderen Personen gelesen wird. Dem Empfänger muss es möglich sein, die Verschlüsselung rückgängig zu machen. Bei den Verfahren die hier vorgestellt werden kommt es nicht auf die Geheimhaltung der benutzten Verfahren, sondern auf die Geheimhaltung der dabei benutzten Schlüssel an. Solche Schlüssel sind digitale Daten, wie zum Beispiel Zahlen oder Buchstabenfolgen.

Symmetrische Verschlüsselungsverfahren

Bei symmetrischen Verschlüsselungsverfahren benutzen Sender und Empfänger den gleichen Schlüssel. Beide Seiten müssen also einen Schlüssel besitzen, der nur ihnen bekannt ist. Es ist notwendig, für jeden Kommunikationspartner einen Schlüssel sicher zu verwahren. Neue Kommunikationspartner müssen sich zunächst auf einen gemeinsamen Schlüssel einigen. Der Austausch dieser Informationen oder des Schlüssels muss aber auf einem sicheren Weg erfolgen.

Asymmetrische Verschlüsselungsverfahren

Werden zur Ver- und Entschlüsselung zwei verschiedene Schlüssel benutzt, nennt man ein solches Verschlüsselungsverfahren asymmetrisch. Ein weiteres theoretisches Merkmal, das bei der Definition von asymmetrischen Verfahren genannt wird, ist, dass sich keiner der beiden Schlüssel aus dem jeweils anderen herleiten lassen darf. Diese Forderung kann aber nicht erfüllt werden. Daher beruht die Unmöglichkeit einer solchen Schlüsselberechnung darauf, dass der Aufwand der Berechnung viel größer ist, als die derzeit vorhandenen Großcomputer mit den besten Algorithmen in vertretbarer Zeit bewältigen können. Dabei spielt die Länge des verwendeten Schlüssels eine große Rolle. Längere Schlüssel sind wesentlich schwerer zu berechnen, da sie wesentlich mehr Möglichkeiten der Kombination bieten, die ausprobiert werden müssen. Aber auch die Entwicklung der Technologie hat einen Einfluss auf die durch die Schlüssellänge gegebene Sicherheit. Galt noch vor 1998 eine Schlüssellänge von 512 Bit als sicher, so waren es bis 2000 schon 768 Bit. Heute stellen 1024 Bit die untere Grenze dar [geeignete Kryptoalgorithmen]. Die Möglichkeit, eine sichere Verbindung mit einem unbekannten Kommunikationspartner aufzubauen, ist hier ein wichtiger Vorteil. Nehmen wir an, A und B wollen eine sichere Verbindung aufbauen. Dazu müsste A ein asymmetrisches Schlüsselpaar erzeugen und einen Schlüssel zu

B senden. Der versendete Schlüssel wird als öffentlicher Schlüssel bezeichnet. Der sicher verwahrte Schlüssel wird dagegen als der private Schlüssel bezeichnet. B ist nun im Besitz des öffentlichen Schlüssels von A. Nachrichten, die B mit diesen Schlüssel verschlüsselt können nur mit dem privaten Schlüssel von A wieder entschlüsselt werden. Eine unbefugte Person C kann also nicht mit einem abgefangen öffentlichen Schlüssel Kenntnis vom Inhalt einer verschlüsselten Nachricht erlangen. Um auch Nachrichten von A zu B sicher zu übertragen, muss auch B ein solches Schlüsselpaar generieren und ebenfalls den öffentlichen Schlüssel zu A senden. Sobald beide Seiten jeweils den öffentlichen Schlüssel der anderen haben, ist eine sichere Verbindung aufgebaut.

Ein Vergleich zwischen symmetrischen und asymmetrischen Verfahren zeigt deutliche Vor- und Nachteile auf. Ein Vorteil der symmetrischen Verfahren ist die Geschwindigkeit. Sie benötigen um die gleiche Sicherheit wie asymmetrische Verfahren zu bieten wesentlich kleinere Schlüssel, die weniger Rechenaufwand bedeuten. So entsprechen 128 Bit bei symmetrischen Verfahren der gleichen Sicherheit, wie 1024 Bit bei asymmetrischen Verfahren. Ein entscheidender Vorteil der asymmetrischen Verfahren ist, dass eine sichere Verbindung auch mit einem unbekannten Kommunikationspartner aufgebaut werden kann.
Asymmetrische Verfahren können auch mit nur einem Schlüsselpaar zur sicheren Kommunikation verwendet werden. Es muss jedoch je ein Schlüssel bei den Kommunikationspartnern sicher verwahrt werden. Hierbei entsteht der gleiche Nachteil, wie schon beim symmetrischen Verfahren. Irgendwann muss einer der Schlüssel übertragen werden. Man kommt also nicht um den Aufwand herum, zwei Schlüssel für einen Kommunikationspartner zu verwalten, wenn man mit einer unbekannten Person sicher kommunizieren möchte. Nachteilig ist auch, dass die Generierung der Schlüssel sehr rechenintensiv ist. Es gibt derzeit einige asymmetrische Verschlüsselungsverfahren. Der am weitesten verbreitete Algorithmus ist RSA. Da der RSA-Algorithmus recht einfach nachvollzogen werden kann, sollte dieser stellvertretend etwas genauer betrachtet werden.

RSA Der Algorithmus RSA wurde bereits 1977 entwickelt und ist bis heute der am weitesten verbreitete Algorithmus. Die Abkürzung RSA steht für die Namen seiner Entwickler Ron Rivest, Adi Shamir und Len Adleman [RSA2]. Da dieser Algorithmus asymmetrisch arbeitet, gibt es zwei verschiedene Schlüssel. Der private Schlüssel setzt sich zusammen aus zwei großen natürlichen Zahlen d und n, der öffentliche Schlüssel besteht aus den Zahlen e und n. Damit A eine Nachricht M zu B schicken kann, muss A zunächst die Nachricht verschlüsseln. Er betrachtet den Inhalt der Nachricht M als Zahlen und berechnet $C = M^e \ mod \ n$. Diese gleichlange Nachricht C ist jetzt nicht mehr lesbar. Sie kann nur mit dem privaten Schlüssel wieder entschlüsselt werden. Daher kann diese Nachricht auch nur noch B entschlüsseln, da er der einzige ist, der den privaten Schlüssel besitzt. Die Nachricht C kann so sicher über das Internet übertragen werden. Erhält B dann die verschlüsselte Nachricht C, so berechnet er, ähnlich wie zuvor A, $M = C^d \ mod \ n$. Er erhält so die unverschlüsselte Nachricht M (siehe Abbildung 3).

Schlüsselgenerierung Die Zahlen für d, e und n müssen jedoch ganz bestimmte Bedingungen erfüllen, damit die Verschlüsselung umkehrbar ist. Die Zahl n muss das Produkt aus frei gewählten Primzahlen p und q sein. Die Benutzung von Primzahlen, also Zahlen, die nur durch sich und durch 1 teilbar sind, findet man auch bei fast allen anderen

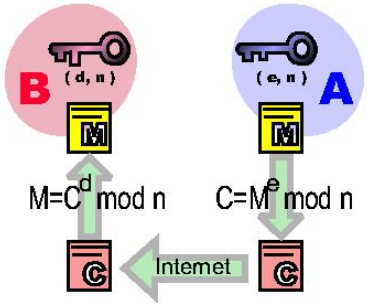


Abbildung 2.3: Eine schematische Darstellung des RSA.

asymmetrischen Verfahren. RSA benutzt noch weitere mathematischen Grundlagen, die auch schon zur Zeit Eulers bekannt waren. Es ist die Eulersche Phi-Funktion, die hier Anwendung findet. Diese Funktion berechnet zu einer natürlichen Zahl x die Anzahl der natürlichen Zahlen, die kleiner x und relativ prim zu x sind. Relativ prim zu x sind dabei alle Zahlen, die mit x nur die 1 als gemeinsamen Teiler haben, also teilerfremd sind. Zum Beispiel sind nur die Zahlen 1,5 und 7 kleiner als 12 und relativ prim zur 12. Die Phi-Funktion liefert also den Wert 3. Es gibt auch eine implizite Berechnungsvorschrift, die hier jedoch aufgrund einer besonderen Eigenschaft nicht gebraucht wird [RSA1]. Die Phi-Funktion angewendet auf eine Primzahl x liefert nämlich immer $x - 1$. Aus der Definition einer Primzahl folgt ja, dass es keine kleinere Zahl gibt, durch die sie teilbar ist. Somit müssen alle Zahlen, die kleiner sind als die Primzahl relativ prim zu dieser sein. Für e muss nun folgende Bedingung gelten:
$(0 < e < \phi(n))$ und $GGT(\phi(n), e) = 1$.
Um $\phi(n)$ zu berechnen, kann man eine weitere Eigenschaft der Phi-Funktion heranziehen: Die Multiplikativität. Ist nämlich n ein Produkt aus zwei teilerfremden Zahlen p und q, so gilt:
$\phi(n) = \phi(p) \cdot \phi(q)$.
Unsere Zahlen p und q sind Primzahlen und somit teilerfremd. Zudem gilt für die Primzahlen p und q:
$\phi(p) = p - 1$ und $\phi(q) = q - 1$.
Also:
$\phi(n) = (p-1) \cdot (q-1)$. Die Zahl e muss also eine natürliche Zahl kleiner $\phi(n)$ sein und darf außer der 1 keinen weiteren Teiler mit $\phi(n)$ gemeinsam haben. Unter allen Zahlen, die dieser Bedingung genügen, kann dann e frei gewählt werden. Für d gilt neben dieser noch eine weitere Bedingung:
$e \cdot d \; mod \; \phi(n) = 1$. Das Produkt aus e und d muss nach der Division durch $\phi(n)$ einen Rest von 1 aufweisen. Bei M muss noch beachtet werden, dass gilt:
$(0 < M < n)$.
Bei einer Implementierung von RSA würde dann nicht M nach oben hin begrenzt sein, sondern vielmehr gilt der theoretische Maximalwert einer Nachricht M als untere Grenze für n. Damit $M = 0$ ausgeschlossen werden kann, könnte man zum Beispiel M vor der Verschlüsselung immer um 1 erhöhen. Nach der Entschlüsselung muss diese 1 wieder abgezogen werden.

Ein Beispiel Die Primzahlen p und q werden zunächst frei gewählt:
$p = 3$ und $q = 11$.
Daraus ergibt sich für n:
$n = p \cdot q = 3 \cdot 11 = 33$ und für
$\phi(n) = \phi(p) \cdot \phi(q) = (p-1) \cdot (q-1) = 2 \cdot 10 = 20$.
Für die Wahl von e und d kommen alle teilerfremden Zahlen von $\phi(n)$, also von 20 in Frage: 1, 3, 7, 9, 11, 13, 17 und 19.
Gewählt sei jetzt: e = 13.
Für d bleibt jetzt nur noch eine Möglichkeit: $d = 17$.
Denn: $d \cdot e = 13 \cdot 17 = 221$ und $221\ mod\ 20 = 1$.
Für die Schlüssel ergibt sich somit:
Privat : $(d, n) = (17, 33)$ und
Öffentlich : $(e, n) = (13, 33)$.
Für Nachrichten M muss zusätzlich gelten: $(0 < M < 33)$.
Sei nun $M = 5$. Dann ergibt sich für C:
$C = M^e\ mod\ n = 5^{13}\ mod\ 33 = 1220703125\ mod\ 33 = 26$.
Erhält nun der Empfänger dieses C und kennt d und n, so berechnet er:
$M = C^d\ mod\ n = 26^{17}\ mod\ 33 = 113382731538515072555417 6\ mod\ 33 = 5$.

Man erkennt recht gut, dass schon bei der Wahl sehr kleiner Werte sehr große Zwischenergebnisse auftreten. RSA arbeitet heute schon mit einer Schlüssellänge von mindestens 1024 Bit [geeignete Kryptoalgorithmen]. Daran lässt sich auch die Ursache für den großen Geschwindigkeitsunterschied zwischen symmetrischen und asymmetrischen Verfahren festmachen.

2.2.4 Digitale Signaturen

Digitale Signaturen sind vergleichbar mit den Unterschriften des täglichen Lebens. Es sind Daten in digitaler Form, die anderen digitalen Daten angefügt werden. Wie normale Unterschriften dienen sie auch der Authentifizierung, also der Beglaubigung der Echtheit [SigG].
Die Verfahren, die eine digitale Signatur erzeugen können, sind die asymmetrischen Verschlüsselungsverfahren. Einige von ihnen eignen sich nicht zum Verschlüsseln von Nachrichten. Zu dieser Gruppe gehört zum Beispiel der DSA - digital signature algorithm. Der DSA setzt nämlich voraus, dass sowohl der Sender, als auch der Empfänger die Nachricht bereits kennen. Der DSA soll lediglich dazu dienen, die Herkunft zu bestätigen, also die Unterschrift zu überprüfen. Es wird nicht näher auf DSA eingegangen, da der bereits vorgestellte RSA ebenfalls zum Erzeugen einer digitalen Signatur verwendet werden kann.
Dazu ist zunächst zu klären, was eine Hash-Funktion ist. Weit verbreitete Hash-Funktion sind zum Beispiel MD4 und MD5. Zwei bekannte Vertreter der MD4-Familie sind RIPEMD-160 und SHA-1 [geeignete Kryptoalgorithmen]. Diese Funktionen bilden eine große Datei auf einen, im Vergleich dazu, sehr kleinen 160 Bit großen Hash-Wert ab (MD5 auf 128 Bit). Eine solche Funktion ist natürlich nicht umkehrbar. Es ist sogar erwünscht, dass aus dem Hash-Wert die originale Nachricht nicht mehr hergeleitet werden kann. Eine weitere Eigenschaft der Hash-Funktion ist, dass es fast unmöglich ist, eine zweite Nachricht zu generieren, die den gleichen Hash-Wert besitzt. Soll nun eine Nachricht unterschrieben werden, so wird zunächst mit der Hash-Funktion der Hash-Wert berechnet (siehe Abbil-

dung 4). Dieser Wert wird dann mit dem privaten Schlüssel verschlüsselt und zusammen mit der lesbaren Nachricht zum Empfänger übertragen. Der verschlüsselte Hash-Wert ist jetzt die Unterschrift des Senders unter der Nachricht. Diese Unterschrift wurde sicher vom Sender geleistet, denn nur er ist im Besitz des privaten Schlüssels. Der Sender erhält nun die lesbare Nachricht gemeinsam mit der Unterschrift. Er berechnet zunächst selbst den Hash-Wert dieser Nachricht mit der selben Hash-Funktion. Daraufhin entschlüsselt er den verschlüsselten Hash-Wert, also die Unterschrift, und vergleicht beide Werte. Sind sie identisch, so kann er sich sicher sein, dass diese Nachricht wirklich vom richtigen Sender stammt und nicht mehr nach dem Versenden verändert wurde. [Non-Repudiation]
Dieses Verfahren funktioniert jetzt genau entgegengesetzt zum oben vorgestellten RSA-Algorithmus. Man nutzt jetzt die Tatsache, dass nur der Sender seinen privaten Schlüssel kennt. Alle, die jetzt eine solche Nachricht erhalten, können mit dem öffentlich verfügbaren Schlüssel die Unterschrift verifizieren.

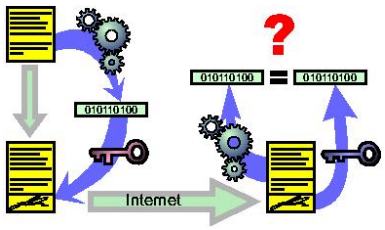


Abbildung 2.4: Zeigt den Ablauf des Signierens (links) und der Verifikation (rechts).

In unserem kleinen Modell könnte die Signatur zum Beispiel dazu dienen, dass der Kunde seine Bestellung unterschreiben kann. Somit wäre der Verkäufer sicher, dass der Kunde die bestellte Ware wirklich haben möchte und dass diese Bestellung nicht mehr nachträglich verändert wurde. Für den Fall, dass der Kunde später abstreitet jemals diese Bestellung gesendet zu haben, dient die Unterschrift des Kunden als Beweismittel. [Non-Repudiation] Genauso kann auch der Verkäufer seinen über das Internet verbreiteten Katalog unterschreiben um sicherzustellen, dass der Kunde die Angebote auf Echtheit prüfen kann. Auch der Kunde kann diese Unterschrift als Beweismittel verwenden, wenn der Verkäufer nicht bereit ist, die angebotene Ware zu den im Katalog beschriebenen Konditionen zu liefern.

2.3 Sichere Protokolle

Damit zwei Systeme kommunizieren, also Daten austauschen können, müssen sie die gleiche Sprache mit ganz bestimmten Regeln sprechen. Eine solche Sprache wird Protokoll genannt. Das Internet basiert auf mehreren Protokollen wie z.B. FTP, HTTP, TCP, IP.
Sichere Protokolle sollen mit Hilfe der bereits vorgestellten Techniken eine Kommunikation ermöglichen, die nach bestimmten Gesichtspunkten sicher ist. Eine Kriterium könnte die Abhörsicherheit sein, ein anderes die Unverfälschtheit der versandten Daten. (Ein Protokoll, dass die Abstreitung vom Versandt bestimmter Daten unmöglich werden lässt, wurde bereits in einer anderen Arbeit vorgestellt. [Non-Repudiation])
In diesem Abschnitt sollen exemplarisch 3 Protokolle vorgestellt werden. Die gewählten Protokolle unterscheiden sich in ihrer Arbeitsweise und haben daher auch bestimmte Stär-

ken und Schwächen.
Als ein Vertreter, der die Abhörsicherheit gewährleisten kann, soll der von Netscape entwickelte SSL (Secure Socket Layer) dienen. XML (Extensible Markup Language) soll als ein Vertreter der Protokolle vorgestellt die im B2B-Bereich zur Anwendung kommen. Und schließlich soll noch auf das von VISA und MasterCrad entwickelte SET (Secure Electronik Transaction) eingegangen werden, da es im Bereich der Online-Bezahlung zur Anwendung kommt.

2.3.1 Das SSL-Protokoll

SSL (Secure Socket Layer) wurde von Netscape entwickelt, um eine abhörsichere Verbindung zwischen einem Client und einem Server, also zwischen dem System des Anwenders und dem des Anbieters aufzubauen. Dabei operiert SSL direkt auf TCP/IP (Transmission Control Protocol / Internet Protocol) [SSL]. Betrachtet man SSL im OSI-Referenzmodell, so ist es der Transportschicht zugeordnet. XML und SET dagegen operieren auf der Anwendungsschicht (siehe Abbildung 5).

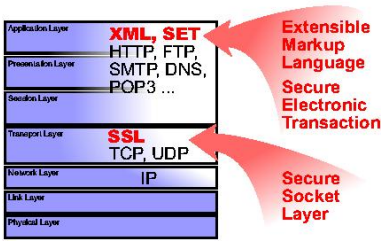


Abbildung 2.5: Zeigt die Einordnung der Protokolle in das OSI-Referenzmodell. Darstellung basiert auf einer Darstellung von Netscape [SSL].

Diese vergleichsweise niedrige Einbindung im gesamten Kommunikationsprozess bringt natürlich Vorteile mit sich. Aufgrund der Tatsache, dass sich SSL zwischen TCP/IP und der eigentlichen Applikation einblendet, kann eine TCP/IP-Anwendung (z.B. Telnet oder HTTP) ohne Veränderungen über eine SSL-Verbindung sicher betrieben werden.

Das beschriebene RSA-Verfahren ist gut geeignet für eine sichere Übertragung über das Internet. Allerdings ist es, so wie alle asymmetrischen Verfahren, sehr langsam. Ein gutes Protokoll muss jedoch schnell sein. Wie oben schon erwähnt, gibt es neben den asymmetrischen Verfahren noch das viel schnellere symmetrische Verfahren. SSL (Secure Socket Layer) ist ein Protokoll, dass gute Geschwindigkeiten ermöglicht. Dabei kombiniert es die Eigenschaft von asymmetrischen Verfahren, eine sichere Verbindung mit einem unbekannten Kommunikationspartner aufzubauen, mit der Schnelligkeit von symmetrischen Verfahren [SSL]. Im folgenden soll der Ablauf des Verbindungsaufbaus grob beschrieben werden (siehe Abbildung 6).
Der Client (das System des Benutzers) sendet als erstes die Versionsnummer seiner SSL-Version, die zur Verfügung stehenden Verschlüsselungsverfahren und einen aus Zufallszahlen bestehenden Datenblock. Der Server (das System des Anbieters) sendet daraufhin seine SSL-Versionsnummer, die nun zu verwendenden Verschlüsselungsverfahren, sein Zertifikat und den signierten Datenblock zurück. Der Client kann jetzt mit dem Zertifikat die Signatur überprüfen. Verläuft diese Prüfung positiv, so berechnet er aus den zur Verfügung

stehenden Daten ein Premaster Secret. Dieses wird nun mit dem öffentlichen Schlüssel verschlüsselt und zum Server gesendet. Der Server kann mit seinem privaten Schlüssel das Premaster Secret wieder entschlüsseln. Beide Seiten haben nun das Premaster Secret und berechnen daraus den Master Secret. Daraus wird im nächsten Schritt der Session Key generiert. Der Session Key ist für die symmetrische Verschlüsselung geeignet und hat eine Länge von bis zu 128 Bit. Der Client sendet nun dem Server, dass er von jetzt an alle Daten mit dem Session Key verschlüsselt. Danach sendet der Server eine ähnliche Nachricht und die sichere Verbindung steht.

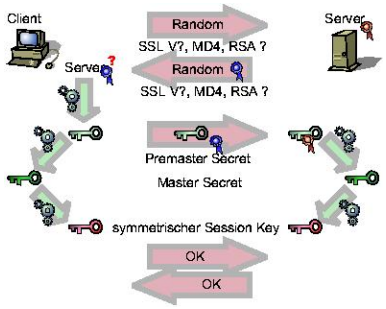


Abbildung 2.6: Der schematische Ablauf des SSL-Verbindungsaufbaus (Handshake).

SSL bietet darüber hinaus auch die Möglichkeit der Identitätsüberprüfung des Clients. Will der Client eine solche Verbindung, so sendet er diesen Wunsch mit den ersten Daten. Der Server wird darauf das Zertifikat vom Client anfordern. Dieser sendet zusammen mit dem Premaster Secret sein Zertifikat und je nach gewählten Verfahren die Signatur von Daten, die beiden Seiten bekannt ist. Um dieses erweiterte Protokoll nutzen zu können, benötigt der Client jedoch ein eigenes Zertifikat.
Auf diesen Weg wurde mit Hilfe der asymmetrischen Verschlüsselung ein symmetrischer Schlüssel sicher übertragen. Ein Protokoll, das nur auf asymmetrischer Verschlüsselung basiert, ist sicher auch realisierbar. Jedoch sind symmetrische Verschlüsselungsverfahren wesentlich schneller (ca. 100 mal so schnell [TC]) als asymmetrischen Verfahren. Für einen Server, der mehrere Clients gleichzeitig bearbeiten können muss, ist dieser Geschwindigkeitsvorteil entscheidend.

SSL bietet als ein Protokoll, dass für die sichere und schnelle Datenübertragung entwickelt wurde, die gewünschte Abhörsicherheit und die erforderliche Schnelligkeit. Dadurch kann SSL überall dort eingesetzt werden, wo es auf die schnelle Verarbeitung der Daten ankommt oder wo sehr große Datenmengen übertragen werden müssen. Eine weitere Stärke von SSL ist die Einbindung unterhalb der Anwendungsschicht. Diese Stärke kann aber auch zum Nachteil werden. SSL bietet nämlich nur die Identifizierung und eine sichere Verbindung. Was darüber hinaus benötigt wird, muss von der jeweiligen Anwendung selbst bereitgestellt werden. Dabei wird meist auch eine zusätzliche Verwaltung von Zertifikaten und Schlüsseln nötig. Aufgrund dieser Eigenschaften findet SSL verstärkt im B2C-Bereich Anwendung. Dort kommt es vor allen darauf an, dass der Kunde so einfach wie möglich sicher mit dem Händler in Verbindung treten kann. Dazu gehört, dass er seine bisherigen Anwendungen weiterhin hutzen kann, die Datenübertragung nicht zu lange dauert und er seine Daten abhörsicher übertragen kann.

2.3.2 XML

XML (Extensible Markup Language) stellt eine der vielen Alternativen zu SSL dar. Dabei ist XML eine Metasprache, die viele verschiedene Sprachen definieren kann. Die durch XML definierten Sprachen, die Aspekte der Sicherheit beinhalten, werden unter dem Begriff XML-Security zusammengefasst. Sie finden Anwendung bei großen Unternehmen, die XML zur Kommunikation nutzen. Bei diesen Kommunikationen sind jedoch viele weitere Sicherheitsaspekte zu berücksichtigen als das bei SSL möglich ist. Zum Beispiel bietet SSL keine Möglichkeit verschiedene Zugriffsrechte zu vergeben. Der Client wird entweder akzeptiert oder eben nicht. XML dagegen kennt viele verschiedene Zugriffsrechte (zum Beispiel für den Zugriff auf Datenbanken [Thuraisingham]). Ein weiterer Vorteil von XML ist die Möglichkeit einzelne Elemente zu verschlüsseln. Somit können bestimmte Teile offen übertragen werden, die nicht geheim gehalten werden müssen. Andere Teile hingegen könnten nur für bestimmte Personen lesbar sein. Durch diese Trennung kann Zeit bei der Übertragung gewonnen werden. Dabei stützt sich XML auch wieder auf die bereits vorgestellten Verschlüsselungsverfahren ab. Digitale Signaturen, asymmetrische Verschlüsselung und Hash-Funktionen sind auch hier vorhanden. Als ein Beispiel für eine Erweiterung sei die Hash-Funktion DOMHASH erwähnt [Xavier]. Diese kann neben der normalen Hash-Wertberechnung auch Hash-Werte über Verzeichnisstrukturen berechnen. Das kann sehr hilfreich sein, wenn man gegenseitig große Datenstrukturen abgleichen muss.

Ein Vorteile von XML ist die Vergabe von verschiedenen Zugriffsrechten, der besonders dann zur Geltung kommt, wenn es eine Vielzahl von Zugriffsberechtigten für unterschiedliche Bereiche gibt. Eine weitere Stärke von XML ist die Möglichkeit Elemente einzeln und verschieden zu verschlüsseln. XML als Metasprache bietet darüber hinaus noch den Vorteil der grenzenlosen Erweiterbarkeit. Zudem sorgt XML für eine echte Transparenz, die auf der Lesbarkeit des übermittelten Datenformats beruht. Ein entscheidender Nachteil von XML ist, dass andere, bereits bestehende Anwendungen nicht über XML arbeiten können, so wie das bei SSL möglich ist. Der private Anwender, in unserem Fall der Käufer, wird daher SSL vorziehen. Dieses viel feiner abgestufte Sicherheitskonzept findet hauptsächlich Anwendung beim Austausch von Informationen zwischen zwei Unternehmen (B2B - Business to Business). Beispielsweise wenn der Hersteller einen Zwischenhändler mit dem Vertrieb seiner Ware beauftragt oder wenn ein Unternehmen bestimmte Kundendaten mit einem Kreditinstitut abgleicht.

2.3.3 SET

Das Protokoll SET (Secure Electronic Transaction) wurde von MasterCard und VISA entwickelt um Geschäfte sicher über das Internet abzuwickeln [MasterCard]. Dabei müssen neben Kundendaten und Bestellungen auch Kreditkarteninformationen sicher übertragen werden. Dabei handelt es sich immer nur um sehr kleine Datenpakete, so dass SET keine Kombination symmetrischer und asymmetrischer Verschlüsselung bieten muss. SET ist also nicht geeignet für den sicheren Austausch großer Datenmengen. SET lässt sich auch in keine der beiden Kategorien B2B oder B2C stecken, da es entwickelt wurde, um Zahlungen zu garantieren. Dazu muss das Protokoll neben einer Kunde-Verkäufer-Verbindung noch eine weitere Verkäufer-Bank-Verbindung aufbauen.
Der Kunde benötigt zum Einkaufen ein Zertifikat, das er von der Bank bekommt, von der

er auch seine Kreditkarte hat. Der Verkäufer benötigt ebenfalls ein Zertifikat von seiner Bank. Wurde eine Verbindung ähnlich wie bei SSL aufgebaut, so kann der Kunde online kaufen (siehe Abbildung 7). Dazu sendet er seine Kreditkarteninformationen verschlüsselt an den Verkäufer. Die Daten sind jedoch nur für die Bank bestimmt und können auch nur dort wieder entschlüsselt werden. Der Verkäufer leitet also diese Daten dann ungelesen an seine Bank weiter. Nachdem die Banken untereinander Verbindung aufgenommen haben und die Zahlung bewilligt wurde, sendet der Verkäufer dem Kunden die Bestätigung der Transaktion. Bei dem Datentransfer zwischen den Banken werden die bereits zwischen den Banken bestehenden Verbindungen genutzt (Banknet).

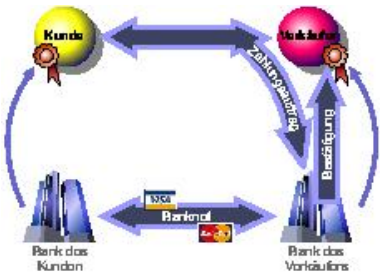


Abbildung 2.7: Diese Darstellung zeigt den Ablauf einer SET-Transaktion.

Vergleicht man SET mit SSL, so erkennt man, dass beide sicher Daten übertragen können. Während SSL dafür entwickelt wurde um eine Verbindung aufzubauen, über die große Datenmengen ausgetauscht werden kann, sollte SET Zahlungen garantieren können. Dafür reichen jedoch schon sehr kleine Datenpaket aus. Dadurch, dass der Verkäufer innerhalb weniger Sekunden eine Zahlungsbestätigung erhält, kann er die Ware sofort risikolos versenden. Dieser Vorteil macht sich besonders bemerkbar, wenn es sich um digitale Waren handelt, die er sofort übertragen kann. Diese Übertragungen können dann jedoch nicht mehr mit SET erfolgen. Ein weitere Vorteil ist, dass der Verkäufer nur an die verschlüsselten Zahlungsdaten des Kunden kommt. Der Kunde braucht also keinen Missbrauch durch den Verkäufer zu befürchten.

2.4 Weitere Modelle

Im zweiten Abschnitt wurde ein einfaches und nicht gesichertes Verkaufsmodell vorgestellt. Außerdem wurden im dritten Abschnitt Protokolle vorgestellt, mit deren Hilfe Sicherheit in den verschiedenen Bereichen des E-Commerce gewährleistet werden kann. Mit diesen Protokollen lässt sich jedoch noch viel mehr konstruieren. Hier sollen jetzt weitere Modelle des Internethandels aufgezeigt werden, um daran zu zeigen, dass auch diese mit den vorgestellten Verfahren und Protokollen sicher gemacht werden können.

2.4.1 Zwischenhändlermodelle

Diese Modelle kennen außer dem Käufer und dem Verkäufer noch einen Zwischenhändler [Hauswith]. Die einzelnen Modelle unterscheiden sich darin, welche Aufgaben der Zwischenhändler übernimmt und welche beim Verkäufer verbleiben. Solche Aufgaben sind die Angebotsunterbreitung, die Verhandlung mit dem Kunden und der Vertragsabschluss,

die Lieferung, der Zahlungsvorgang und die Kundenbetreuung. Letzteres ist jedoch für Sicherheitsbetrachtungen weniger relevant.
In diesen Modellen müssen einige sichere Verbindungen aufgebaut werden. Zum einen möchte der Kunde seine Daten sicher zum Zwischenhändler und zum Verkäufer übertragen. SSL ist hier am besten geeignet, da der Verbindungsaufbau automatisch vom Browser übernommen wird. Zum anderen muss der Zwischenhändler die Kundendaten, Verkaufzahlen, Bestellungen oder ganze Verkaufskataloge sicher mit dem Verkäufer abgleichen. Diese hier notwendige Verbindung ist eine B2B-Verbindung. Natürlich könnte auch hier SSL benutzt werden. Jedoch bieten Protokolle wie XML auch gleich die Möglichkeit des Zugriffs auf Datenbanken. Für den Fall, dass online bezahlt werden soll, braucht der Kunde noch eine sichere Verbindung, die geeignet ist um Zahlungsinformationen zu übermitteln. Und abhängig davon, mit wem die Zahlung abgewickelt wird, braucht entweder der Zwischenhändler oder der Verkäufer noch eine sichere Verbindung zu einer Bank. Hier kann ebenfalls wieder auf SSL zurückgegriffen werden. Zum Beispiel, wenn es sich um Lösungen handelt, die sich auf das Homebanking abstützen. SET würde hier eingesetzt werden, wenn es darum geht Zahlungen zu garantieren und um den gesamten Prozess des Verkaufens zu beschleunigen.
Mit den im dritten Abschnitt vorgestellten Protokollen lässt sich also ein solches Modell leicht so konstruieren, dass es sicher ist (siehe Abbildung 8).

Abbildung 2.8: Zeigt ein Zwischenhändlermodell mit den Einsatzmöglichkeiten der verschiedenen Protokolle. Der senkrechte, gestrichelt dargestellte Pfeil deutet die Möglichkeit der unterschiedlichen Aufgabenverteilung an.

Weitere Varianten von Zwischenhändlermodellen entstehen, wenn ein Kunde mehrere Verkäufer über einen Zwischenhändler erreichen kann. Zum Beispiel bei Preisagenturen, die dann den günstigsten Verkäufer vermitteln. Ebenso denkbar sind Modelle, bei denen mehrere Käufer über einen Zwischenhändler ein Produkt erwerben wollen. So können zum Beispiel diese Käufer dann in den Genuss eines Mengenrabattes kommen.

2.4.2 Das Vermittlermodell

Als Vermittlermodell wird in dieser Arbeit ein Modell bezeichnet, dass einen Vermittler und viele Kunden kennt. Dabei werden die Kunden jeweils aneinander vermittelt und der eigentliche Handel findet unter ihnen statt (siehe Abbildung 9). Es gibt also keinen richtigen Verkäufer. Ein Beispiel hierfür sind Online-Auktionen. Damit sich ein Kunde vermitteln lassen kann, muss eine Verbindung zwischen dem Vermittler und dem Kunden aufgebaut werden. Hierzu eignet sich wieder SSL. Sind nun zwei Kunden sicher mit

dem Vermittler verbunden und sollen aneinander vermittelt werden, so muss zwischen den Kunden eine neue Verbindung aufgebaut werden. Abgesehen davon, dass eventuell noch Vermittlungsgebühren zu zahlen sind, schließt sich hier wieder das ganz einfache Modell aus Abschnitt 2 an. Da jedoch keiner der beiden Kunden ein professioneller Verkäufer ist, kann hier keine Rede von B2C sein. Daher werden auch die zur Verfügung stehenden Verfahren, wie SSL, SET usw. nicht zur Anwendung kommen. Denkbar ist hier jedoch der Vertragsabschluss mittels signierter E-Mails.

Abbildung 2.9: Ein Modell mit vielen Kunden und einem Vermittler. Ein Problem stellt die sichere Verbindung zwischen den vermittelten Kunden dar.

2.5 Zusammenfassung

An einem sehr einfachen, aber gerade von kleineren Anbietern benutzten Verkaufsmodell wurden Verfahren und Techniken vorgestellt, mit denen der Handel im Internet sicher gemacht werden kann. Die drei vorgestellten Protokolle haben gezeigt, dass es möglich ist, Sicherheit in vielen Bereichen zu erzeugen. Und schließlich konnte an den komplexeren Modellen gezeigt werden, dass die Sicherheit beim Handeln im Internet gewährleistet werden kann.

Diese Arbeit zeigt ein Idealbild, nicht jedoch die Realität. In der Wirklichkeit stellen sich nämlich noch weitere Fragen. Was soll der Verkäufer mit einem Kunden machen, der ein Zertifikat der Stufe 2 hat? Genauso stellt sich die Frage, was mit einem abgelaufenen Zertifikat geschieht. Und was passiert, wenn das Zertifikat von einer unbekannten oder von einer nicht akkreditierten CA ausgestellt wurde? Ebenso ist es möglich, dass der Kunde gar kein Zertifikat besitzt. Lehnt man ihn als Kunden ab, so schickt man ihn sicher nicht zur nächsten CA, sondern zur Konkurrenz.
Bei der Verwendung von SSL stellt sich auch die Frage nach der Länge des symmetrischen Schlüssels. So gibt es noch Browser, die aufgrund von damaligen Ausfuhrbeschränkungen nur eine schwache Verschlüsselung bieten (40 Bit). Der heutige Standard beim Homebanking ist jedoch die starke Verschlüsselung mit 128 Bit. Banken hatten damit ein großes Problem. Sie wussten, dass 40 Bit nicht sicher genug waren. Aber sie wussten auch, dass sie noch keine 128 Bit fordern konnten. Dieses Problem wird es aber immer wieder geben, denn auch bei RSA und DSA gibt es regelmäßige Anpassungen der Schlüssellängen. Wie soll also mit Kunden verfahren werden, die noch mit einer älteren Software arbeiten?
Der Verkäufer muss sich auch die Frage stellen, welche Zahlungsmöglichkeiten er seinen Kunden bieten möchte. Bietet er nur eine Zahlung per Rechnung oder Nachname, so

verliert er die Käufer aus dem Ausland. Bietet er nur die Zahlung mit Kreditkarte an, so schränkt er seinen Kundenkreis sehr ein. Um wirklich konkurrenzfähig zu sein, muss er also viele Zahlungsarten anbieten. Und solange es genug Anbieter im Internet gibt, die noch herkömmliche Zahlungsarten anbieten, wird der Kunde nicht die Notwendigkeit sehen, ein Zertifikat zu erwerben. Ein anderer wichtiger Punkt ist die Kostenfrage. Alle angesprochenen Zertifikate gibt es nicht kostenlos. Abbildung 10 zeigt die Preise für die verschiedenen Zertifikate.

Zertifikat/CA	VeriSign	Thawte	TrustCenter
SSL Server (pro Jahr)	$ 349	$ 125	130 €
weitere Jahr	$ 249	$ 100	130 €
Versicherungssumme	$ 100.000	-	-
SSL Server 128 Bit (pro Jahr)	$ 895	$ 300	-
weitere Jahr	$ 895	$ 300	-
Versicherungssumme	$ 250.000	-	-
persönliches Zertifikat	-	$ 25	62 €
weitere Jahr	-	$ 0	62 €

Quellen: [VeriSign] [Thawte] [TC] Stand: 05.2002

Abbildung 2.10: Ein Überblick über die finanziellen Aufwendungen für SSL-Zertifikate.

Auch SET hat seinen Preis. Der Kunde zahlt für die Beantragung des Zertifikates bis zu 7,5 EUR, bei einigen Banken einen jährlichen Beitrag von bis zu 5 EUR und für die Erneuerung alle 3 Jahre weitere 7,5 EUR. Auf der Händlerseite fallen deutlich mehr Kosten an. Für die Einrichtung der nötigen Software müssen 150 - 200 EUR gezahlt werden. Zudem werden monatliche Kosten von 13 - 30 EUR fällig. Für das Zertifikats zahlt der Händler für die einmalige Ausstellung weitere 1000 EUR und einen jährlichen Beitrag von 125 EUR. Zudem fallen noch Transaktionsgebühren von 3,5% des Kaufpreises an. (Quellen : Sparkasse, Volksbank)
Es wird deutlich, dass es die hier vorgestellte Sicherheit nicht kostenlos gibt. Es stellt sich die Frage, wie viel man bereit ist, für die eigene Sicherheit auszugeben. Ein Kunde, der über das Internet einkaufen möchte, erwartet heutzutage schon eine sichere Verbindung [Gollman]. Für den Verkäufer heißt das, dass er ein SSL-Server-Zertifikat braucht. Die Kosten für den Privatanwender sind schon sehr gering und dennoch viel zu hoch. Der Kunde ist in der Regel nicht dazu bereit, noch zusätzliche Kosten zu tragen. Dagegen ist man eher dazu bereit, ein etwas höheres Risiko in Kauf zunehmen.
Die Bereitschaft und die Notwendigkeit Sicherheit zu praktizieren steigt mit dem Preis der angebotenen Ware. Daniel W. Manchalla stellt in Ë-Commerce Trust Metrics and Modelsëine Lösung vor, die diesem Sachverhalt gerecht wird [Manchalla]. So soll die Identität des Kunden nur überprüft werden, wenn die Ware einen gewissen Wert übersteigt. Und für den Fall, dass die vorherigen Zahlungen eines Kunden bekannt sind, könnte man bei guter Zahlungsmoral diese Grenze nach oben verschieben. Diese Methode ist gedacht für sehr große Unternehmen, die einen großen Kundenstamm und ein breites Angebot haben. Dabei zielt man auch auf die Einsparung der nötigen Systemressourcen ab.

Diese Arbeit bezog sich nur auf die Sicherheit der Kommunikation zwischen den Handelspartnern und nicht auf die Sicherheit der benutzten Systeme. Die längsten Schlüssel nutzen nichts, wenn die Systeme, auf denen private Schlüssel gespeichert sind, unsicher sind. Unsichere System stellen einen Angriffspunkt für Hacker und Viren da. Neben den

angesprochenen Verschlüsselungsverfahren sind also noch weitere Sicherungen, wie eine Firewall oder Virenscanner nötig [Ahuja].

Abschließend möchte der Autor feststellen, dass es theoretisch möglich ist, Handel im Internet sicher zu betreiben. In der Praxis stößt man jedoch schnell an die Grenzen der Praktizierbarkeit. Der Kostenaufwand für die Sicherheit ist noch recht hoch und lohnt sich erst bei großen Umsatzzahlen. Zudem ist auch die Bereitschaft des Kunden noch sehr gering, für seine Sicherheit Geld auszugeben. Bestätigt wird der Kunde in seinem Verhalten durch Verkäufer, die bereit sind das volle Risiko zu tragen. Ändern wird das wohl erst, wenn der wirtschaftliche Schaden über dem erwirtschafteten Gewinn liegt oder es weltweit gesetzliche Bestimmungen gibt, die einen Sicherheitsstandard für den E-Commerce vorschreiben. Letzteres wird jedoch nur schwer realisierbar sein, da die regionalen Gesetze und Bestimmungen noch viele Besonderheiten besitzen.

Literaturverzeichnis

[Ahuja] V. Ahuja.
Building trust in electronic commerce.
IT Professionel, Volume 2, Issue May-June 2000, Page(s): 61-63
ftp://ftp.tik.ee.ethz.ch/pub/lehre/iteco/SS02/material/00846215.pdf

[geeignete Kryptoalgorithmen] *Geeignete Kryptoalgorithmen gemäß § 17 (2) SigV.*
Bundesanzeiger Nr. 213 - Seite 18.638
http://www.iid.de/iukdg/algorith.html

[Gollman] D. Gollman.
E-commerce security.
Computing & Control Engineering Journal, Volume 11, Issue: 3, June 2000, Page(s): 115-118
ftp://ftp.tik.ee.ethz.ch/pub/lehre/iteco/SS02/material/00850785.pdf

[Hauswith] M. Hauswirth, M. Jazayeri, M. Schneider.
A phase model for e-commerce business models and ist application to security assessment.
System Sciences, 2001. Proceedings of the 34th Annual Hawaii International Conference, 2001. Page(s): 4146-4155
ftp://ftp.tik.ee.ethz.ch/pub/lehre/iteco/SS02/material/00927285.pdf

[ITU88] ITU-T Recommendation X.509
Information Technology − Open Systems Interconnection − The Directory: Authentication framework. 1988 (ISO/IEC 9594-8)
http://www-t.zhwin.ch/it/ksy/Block08/ITU/X509_4thEditionDraftV8.pdf

[Manchalla] D. W. Manchalla.
E-Commerce Trust Metrics and Models.
IEEE Internet Computing, Volume 4, sue 2, March-April 2000, Page(s): 36-44
ftp://ftp.tik.ee.ethz.ch/pub/lehre/iteco/SS02/material/00832944.pdf

[MasterCard] *SET Secure Electronic TransactionTM – Setting the Stage for Safe Internet Shopping - an enticing concept.*
http://www.mastercardintl.com/newtechnology/set

[Non-Repudiation] Non-Repudiation - Nicht-Abstreitbarkeit elektronischer Transaktionen
Steffen Mazanek
Eine andere Seminararbeit des Seminars.

[RSA1] *Vernetzte Informationssysteme.*
http://caladan.wiwi.uni-frankfurt.de/IWI/Veranstaltung/Vis_ws9798/visscript

[RSA2] Neal Koblitz.
Algebraic Aspects of Cryptography.
Springer-Verlag

[SigG] *Signaturgesetz - Gesetz über Rahmenbedingungen für elektronische Signaturen.*
Bundesgesetzblatt Nr. 22 vom 21. Mai 2001
http://www.bundesanzeiger.de/bgbl1f/findex01.htm

[SSL] Netscape.
Introduction to SSL.
http://developer.netscape.com/docs/manuals/security/sslin/contens.html

[TC] *TC TrustCenter.*
http://www.trustcenter.de

[TC Richtlinien] *TC TrustCenter Zertifizierungsrichtlinien.*
Version vom 1. Oktober 1999
http://www.trustcenter.de/legal/policy/policy_de/r_de.pdf

[Thawte] *Thawte* http://www.thawte.com/

[Thuraisingham] B. Thuraisingham, C. Clifton, A. Gupta, E. Bertino, E. Ferrari.
Directions for web and E-commerce applications security.
The MITRE Corporation. This Paper appers in: Enabling Technologies: Infrastructure for Collaborative Enterprises, 2001. WET ICE 2001. Proceedings. On page(s): 200-204, June 20-22, 2001, ISSN: 1080-1383
ftp://ftp.tik.ee.ethz.ch/pub/lehre/iteco/SS02/material/00953414.pdf

[VeriSign] *VeriSign* http://www.verisign.com/

[Xavier] E. Xavier.
XML based security for e-commerce applications.
8th IEEE Symposium and Workshop on Engineering Coputer-Based Systems Vienna, Virginia, U.S.A., 17 - 20 April 2001, Page(s): 10-17
ftp://ftp.tik.ee.ethz.ch/pub/lehre/iteco/SS02/material/00922398.pdf

Kapitel 3

Incentives for Peer-to-peer Networks

Sebastian Kühne

Inhaltsverzeichnis

3.1 Einleitung

Ursprünglich als Netzwerk vieler gleichberechtigter Knoten geschaffen, hat sich das Internet mit der wachsenden Kommerzialisierung und der massiven Zunahme von Dial-up-Verbindungen zu einer „Zweiklassengesellschaft “entwickelt. So dominieren heute Client/

Server-Architekturen seine Servicestruktur. Vom E-Mail-Dienst über Suchmaschinen, Firmenpräsenzen und Online-Shops, überall trifft der Anwender auf die gewohnte Zugriffsform: Mittels Browser holt er sich Informationen und andere Dienstleistungen von zentralen ständig erreichbaren Servern. Aber spätestens seit dem Aufbau der Musiktauschbörse Napster und deren weitreichender Verbreitung ist eine alte Idee wieder in das Bewußtsein der Internet-User und -Macher gelangt: Peer-to-peer (P2P). Viele unabhängige Knoten, die direkt und gleichberechtigt miteinander agieren.

Mittlerweile basieren zahlreiche Applikationen auf einer Peer-to-peer-Architektur (P2P). Da wären zum Beispiel die verschiedenen Instant Messenger Services (ICQ, AOL Instant Messenger, etc.) und Filesharing Services (z.B. Gnutella und Freenet). Es sind jedoch auch schon wieder Angebote verschwunden. Der vielleicht populärste Dienst Napster wurde durch eine Initiative der Musikindustrie, die den Schutz ihrer Urheberrechte einklagte, gestoppt. Derlei Rückschläge sind dem aktuellen Hype um die Peer-to-peer-Idee zwar kaum abträglich, jedoch zeigt gerade dieses Beispiel den Bedarf an der Diskussion der Fähigkeiten und Eigenschaften von P2P-Systemen. Welche technischen, rechtlichen oder sozialen Probleme können aus der Existenz von P2P-Netzwerken entstehen oder durch sie gelöst werden? Nur wenige Stichworte sind: Explosion des Verkehrs- und Datenvolumens, Schutz von Urheberrechten und die Verbreitung von unerwünschtem Material.

Im Zuge dieser Seminararbeit sollen insbesondere die technischen und ökonomischen Aspekte der P2P-Technologie beleuchtet und dadurch eine Erklärung für den derzeitigen Hype gefunden werden. Welchen Nutzen können Firmen oder auch einzelne Personen daraus ziehen? D.h., welche ökonomischen Anreize lassen sich finden, daß beliebige, weltweit verteilte, Benutzer ihre Rechner als P2P-Peers zur Verfügung stellen? Mit dem Schwerpunkt auf das kommerzielle Potential von P2P gerichtet soll das Bild eines möglichen zukünftigen P2P-Internets gezeichnet werden. Dazu werden im folgenden zuerst die Alternativen auf technischer und konzeptioneller Ebene verglichen und anhand von Beispielen vorgestellt (Kapitel 2). Um die Eigenschaften von P2P-Netzwerken effektiv nutzen zu können, ist eine formale Definition ihrer Bestandteile und vor allem ihrer Sprache nötig. JXTA, vorgestellt in Kapitel 3, basiert auf einer solchen Definition. Schließlich befasst sich Kapitel 4 mit der konkreten Anwendung von P2P-Technologie in geschäftlichem und privatem Umfeld.

3.2 Netzwerkarchitekturen

Als Grundlage für die Diskussion ihrer Fähigkeiten werden im Folgenden die beiden Architekturen konzeptionell und anhand von Beispielen vorgestellt und schließlich miteinander verglichen.

3.2.1 Client/ Server

Client/Server ist die klassische Architektur. Ein zentraler Server stellt Dienstleistungen bereit, die von vielen Clients in Anspruch genommen werden. Dabei sind die Clients voneinander unabhängig und merken in der Regel nichts von der Existenz der anderen. Die Aufgaben eines Clients bestehen allein aus der Formulierung und dem Senden von Kommandos an den Server und dem Empfangen der daraus resultierenden Ergebnisse. Daraus ergeben sich die Aufgaben des Servers: Empfang von Kommandos für einen Dienst,

Ausführen des gewünschten Dienstes und Senden der Ergebnisse. Sowohl der Begriff des Clients als auch der des Servers sind nicht jeweils an eine eigene Maschine gebunden. Die Bezeichnungen beziehen sich in der Regel auf einzelne Prozesse, die die obigen Charakterisierungen erfüllen. So können mehrere Server oder auch Clients parallel auf einer Maschine existieren.

Als kurzes Beispiel sollen hier nur einige der gängigsten Dienste kurz erwähnt werden. Mit einem Browser als Client-Applikation kann ein User die unterschiedlichsten Dienste von Servern im WWW abrufen. Mit dem universellen Interface, das gängige Browser mit der Darstellung von HTML-Seiten und unterstützten Scriptsprachen bieten, können zum Beispiel Anfragen an Suchmaschinen formuliert und die Ergebnisse direkt angezeigt werden. Die einzige Voraussetzung ist, daß die Adresse des Servers bekannt ist, der den gewünschten Dienst zur Verfügung stellt. Trotz der Universalität der Browser ist es keiner Anwendung vorenthalten, als Client für einen speziellen Dienst zu wirken. E-Mail-Clients und Newsreader zum Beispiel sind hinlänglich bekannt. Es ist wohl nicht zuletzt der Vielfältigkeit der Browser, die vom gemeinen User oft mit dem eigentlichen Internet untrennbar unifiziert werden, zu verdanken, daß sich die C/S-Architektur so stark durchgesetzt hat.

3.2.2 Peer-to-peer

Neben einer allgemeinen Einführung wird hier auch eine Einteilung von P2P-Netzwerken nach technischen und nutzenspezifischen Eigenschaften vorgenommen.

Es existiert keine formale Definition für P2P. Die Peer-to-peer Working Group [W2] beschreibt es so: *„Put simply, peer-to-peer computing is the sharing of computer resources and services by direct exchange between systems. These resources and services include the exchange of information, processing cycles, cache storage, and disk storage for files.*

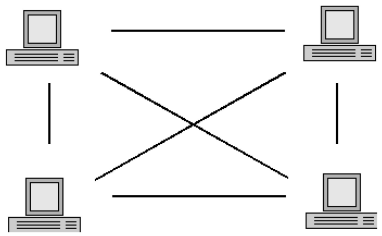
Die Grundlage eines P2P-Netzwerkes ist also die Nutzung der Ressourcen der Peers zu einem gemeinsamen Zweck. Ein Peer vereint die klassischen Aufgaben von Client und Server in sich. Die Kunst beim Entwurf eines solchen Netzwerkes besteht darin, die anfallenden Aufgaben mit möglichst hoher Effizienz zu verteilen. Dabei sind unter Umständen die verschiedensten Eigenschaften des Peers zu berücksichtigen, wie etwa die Qualität und Bandbreite der Anbindung, verfügbarer Speicherplatz und Rechenleistung. Je nach Zweck des Netzwerks kann jedoch auch so etwas wie die persönliche Zuverlässigkeit des zum Peer gehörenden Users eine Rolle spielen.

Um die Nützlichkeit von P2P-Netzwerken für bestimmte Anwendungsgebiete besser einordnen zu können hat die Gartner Group [21] fünf Modellgruppen eingeführt: *Atomistisch*, *benutzerorientiert*, *datenorientiert*, *Web Mk2* und *berechnungsorientiert*. Das einfachste atomistische Modell könnte man auch als pures P2P bezeichnen. Es besteht einfach aus einzelnen Peers, die miteinander kommunizieren, ohne daß das Netzwerk einen Mechanismus zur Verfügung stellt, der diese Verbindungen vermittelt. Ein Computerspiel, bei dem sich Spieler durch schlichte Eingabe der IP-Adresse in ein anderes Spiel einlinken können, würde solch ein „Netzwerk“erschaffen. Benutzer- und Datenorientierte Netzwerke stellen Mechanismen zur Verfügung, die aufgrund von Benutzer- oder Dateneigenschaften Verbindungen vermitteln. Beispiele dafür sind Instant Messenger und Filesharing Netzwerke wie sie im folgenden Kapitel ausführlich beschrieben werden.

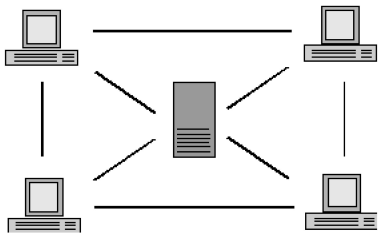
Eine Kombination der ersten drei Modelle mit bestehender Webarchitektur ist Web Mk2. Im Zentrum dieses hoch komplexen Modells steht eine höchst personalisierte brow-

serartige Peer-Software, die das immense Datenaufkommen speziell auf den jeweiligen Benutzer zugeschnitten darbieten muß. Beim berechnungsorientierten Modell steht die gemeinsame Rechenleistung des Netzwerks im Vordergrund, die zur Bearbeitung eines oder weniger Prozesse nutzbar gemacht werden soll. Gartner selbst weist bei der Beschreibung der Modelle oft Servern bestimmte Aufgaben zu. Etwa zur Verwaltung der User- oder Datenindizierung im benutzer-/datenorientierten Modell. Dieses ist jedoch nicht zwangsläufig nötig wie sich auch im Folgenden am Beispiel Gnutella zeigen wird.

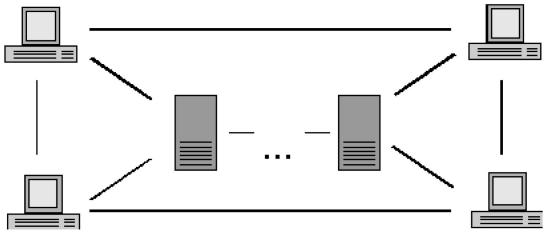
Eine andere Einteilung, die zur Verfeinerung der obigen dienen kann, wird von Theodore Hong [18, Kap. 14] vorgenommen. Hier wird eine grobe Unterscheidung von *zentral koordiniert*, *hierarchisch* und *völlig dezentral* getroffen. Im ersten Fall wird für die Koordination des Netzwerks ein Server benutzt. Im zweiten existieren mehrere Koordinatoren, die jeweils ihren eigenen Zuständigkeitsbereich besitzen. Völlig dezentrale Netzwerke wie das Gnutella-Netz koordinieren sich selbst.



Völlig dezentral



Zentral koordiniert



Hierarchisch

Für die Betrachtung von P2P-Netzwerken bezogen auf ihre Funktionalität sind die Einteilungen der Gartner Group völlig ausreichend, aber eine technische Diskussion kommt ohne eine detailiertere Einteilung, wie sie Hong vornimmt, nicht herum.

3.2.3 Beispiele für P2P-Netzwerke

IP

IP ist zwar keine eigenständige Netzwerkarchitektur, aber das Protokoll auf dem alle Dienste im Internet basieren müssen. Das ist der P2P-Idee jedoch nicht abträglich, denn IP ist pures P2P. Es regelt einzig den direkten Datentransfer zwischen Rechnern, die eine weltweit eindeutige IP-Adresse besitzen. Die Voraussetzungen für P2P-Netzwerke im Internet sind also durch dessen Fundament gegeben. Auch die in der Regel untrennbar mit IP verbundenen Zuverlässigkeitsprotokolle TCP und UDP sind nicht hinderlich.

Gnutella

Mit Gnutella [1] wird keine spezielle Applikation verbunden. Es ist als Protokoll spezifiziert, das von zahlreichen Anwendungsprogrammen implementiert wird. Ironischerweise werden diese als Gnutella-Clients bezeichnet, obwohl sie vom klassischen Client nicht weiter entfernt sein könnten. Es sind Clients für alle gängigen Betriebssysteme erhältlich.

Zielsetzung: Gnutella ist als Filesharing-Netzwerk konzipiert und fällt somit in den Bereich der datenorientierten Modelle. Es ist nicht wie etwa Napster auf einen bestimmten Datentyp spezialisiert. Die Suche nach Daten erfolgt in Echtzeit und völlig anonym. Keine Anfrage kann zu ihrem Ursprung zurück verfolgt werden und es ist höchst unwahrscheinlich, daß ein Treffer bereits veraltet ist bevor die Daten empfangen werden können. Der Download hingegen ist nicht völlig anonym. Jeder Knoten weiss, wer seine Daten empfängt. Dennoch ist es nicht möglich mit dieser Information so etwas wie User-Profile zu erstellen, da Quellen gewöhnlich ständig wechseln. Das Gnutella-Netzwerk ist gegenüber Störungen jeglicher Art kaum anfällig. *Gnutellanews.com [W3] : „Gnutella is designed to survive a nuclear war [..] [and] can withstand a band of hungry lawyers“.*

Technik: Mit *„The client is the server is the network“*faßt Gene Kan [18, Kap. 8] das technische Konzept zusammen. Es existiert keine ständige Verbindung zwischen den einzelnen Knoten. Das ganze Netzwerk basiert auf einem Message-Broadcast-System. Neue Knoten werden in das Netzwerk eingefügt indem der User Adressen von aktiven Knoten im Chat oder von Host Lists erfährt und diese in seinem Client ausprobiert bis ein Knoten antwortet. Von da an übernimmt das Netzwerk alles weitere. Der gefundene Knoten überträgt eine Liste der ihm bekannten aktiven Knoten und innerhalb kurzer Zeit organisieren Analysealgorithmen die Position des neuen Knotens im Netz. Diese können von Client- zu Client-Software unterschiedlich sein. Jedoch bringt die Orientierung an der Anbindungsgeschwindigkeit der Knoten das Netz in eine effiziente Form. So bauen Knoten hauptsächlich Verbindungen mit Knoten ähnlicher Geschwindigkeit auf. Das erzeugt eine leistungsfähige Backbone bestehend aus schnellen Knoten. Durch die relativ zufällige Wahl des Eintrittspunktes eines neuen Knotens entstehen Cluster unterschiedlicher Größe innerhalb des Netzes. Es ist so auch keine Garantie gegeben, daß das Gnutella-Netz immer vollständig zusammenhängend ist. Erweiterungen der Analysealgorithmen bewirken zusätzlich bei einer zu großen Anzahl von Knoten in einem Cluster eine Aufteilung in zwei oder mehrere neue Cluster. Wird das Eintreten von neuen Knoten zusätzlich durch Passwörter kontrolliert, lassen sich auf einfachste Weise auch private Gnutella-Netzwerke einrichten.

Die anonyme Suche nach Inhalten ist durch ein erstaunlich einfaches System realisiert. *„Gnutella is the game Telephone [Stille Post]“*[W3]. Der suchende Knoten schickt eine Suchanfrage an seine benachbarten Knoten, diese wieder an ihre, usw.. Dabei merkt sich jeder Knoten nur den Knoten von dem er die Anfrage erhalten hat, um Ergebnisse aus seinem eigenen Datenbestand oder der weitergeleiteten Anfrage genau auf diesem Weg wieder zurück zu schicken. Kein Knoten kann also sagen, ob der Knoten von dem er die Anfrage erhalten hat der Ursprung der Suche ist oder nur ein Vermittler. Um die offensichtlich große Gefahr eines stark erhöhten Datenverkehrs oder der mehrfachen Bearbeitung der selben Suchanfrage zu beseitigen, wurden zwei Mechanismen eingebaut. Jede Anfrage erhält einen 128-bit breiten einzigartigen Identifier. So erkennt ein Knoten eine bereits bearbeitete Anfrage und kann sie ignorieren. Weiterhin bekommt jede Anfrage eine

[1] http://www.gnutella.com

Time-to-live-Number. Jeder Knoten der eine Anfrage bearbeitet zählt die TTL um eins herunter. Hat diese den Wert Null erreicht, wird die Anfrage nicht mehr weitergeleitet.

Eine besondere Eigenschaft des Suchmechanismus bezeichnet Gene Kan als *„Distributed Intelligence“*. Es bleibt der Client-Software überlassen wie sie Suchanfragen interpretiert. Das führt dazu, daß eine Anfrage wie „3+1*4“bei vielen Clients keine Ergebnisse produziert. Ein algebraisch orientierter Client würde jedoch das Ergebnis 7 liefern. In besonders kleinen Clustern ist mit einem Client, der die eintreffenden Anfragen auflistet, sogar ein Chat mittels Suchanfragen möglich.

Der eigentliche Datentransfer findet im Gegensatz zur Suche direkt von der Quelle zum Empfänger statt. Ein Suchergebnis enthält zu diesem Zweck die Adresse der Quelle. Der Empfänger kann so eine direkte Verbindung aufbauen. Dieses Vorgehen ist zwar nicht vollständig anonym, bedenkt man aber, daß das Netz ständig seine Form ändert, ist es unmöglich ein Profil über das Downloadverhalten eines bestimmten Users zu erstellen.

Probleme: Gnutella lebt von der aktiven Teilnahme der Knoten. Die Qualität des Netzes häng unmittelbar von den zur Verfügung gestellten Daten ab. Da kein Mechanismus existiert, der User zwingt oder wenigstens animiert Daten zum Download freizugeben, existieren eine Menge Free Rider. Sie belasten das Netz mit Suchanfragen und Downloads ohne die Datenmenge zu vergrößern und erhöhen so indirekt die relative Anzahl der Zugriffe auf andere Knoten.

Durch die begrenzte Lebensdauer einer Suchanfrage besitzt jeder Knoten einen bestimmten Horizont, bis zu dem er in das Netzwerk eindringen kann. Es ist also gut möglich, daß ein gewünschter Inhalt im Netz vorhanden ist, jedoch nicht von der Suchanfrage erreicht wird. Nun stelle man sich einen Knoten vor, der von einer ungewöhnlich großen Zahl Free Rider umgeben ist. Eine solche Konstellation würde den Horizont noch zusätzlich einengen, da die Anfrage nur an Lebenszeit verliert ohne Ergebnisse produzieren zu können.

Eine technische Eigenart (Problem oder Vorteil) ist, daß keine Inhalte unerwünschter Natur aus dem Gnutella-Netz verbannt werden können. Das verhindert die Zensur zum Beispiel von Minderheitenmeinungen aber ermöglicht auch die Verbreitung von illegalem Material wie Kinderpornografie. Die „(Gnutella) Wall of Shame“[2] ist ein Beispiel für einen möglichen Lösungsansatz derartiger Probleme.

Eigenschaften	**Technik**	**Probleme**
Filesharing (Datenorientiert)	Protokoll	Free Rider
Anonyme Suche	Vollständig dezentral	Suchhorizont
Quasi anonymer Download	Messag-Broadcast-System	Zensur unmöglich
Geringe Störungsanfälligkeit	Manuelles Einfügen neuer Knoten	
private Netzwerke	Algorithmen organisieren Struktur	

Überblick: Gnutella

Groove

Zielsetzung: Groove [3] ist ein kommerzielles Produkt für Unternehmen jeder Größe. Eine im Funktionsumfang stark abgespeckte Version ist für den Privatgebrauch jedoch frei

[2] www.zeropaid.com/busted

[3] http://www.groove.net

erhältlich. Es handelt sich um eine Netzwerkarchitektur mit P2P-Technologie zur Realisierung von computergestützter Gruppenarbeit (CSCW). Man könnte es also als User Centered einordnen. Zwar stellt Groove hauptsächlich eine Plattform für die Entwicklung von individualisierten Unternehmenslösungen dar, jedoch ist auch ein Paket von Standardanwendungen auf dessen Basis verfügbar. Ziel ist eine direkte Vernetzung von Endgeräten auch über die Grenzen von Firewalls und NAT-Routern (Network Adress Translation) hinweg. Das Groove-Netzwerk ermöglicht die Bildung von Arbeitsgruppen unter anderem zur gemeinsamen Bearbeitung und Betrachtung von Dokumenten. Zum Grundfunktionsumfang gehören auch Chat- und Voice-Chat-Tools.

Technik: Groove unterstützt sowohl asynchrone als auch synchrone Gruppenarbeit. Eine Arbeitsgruppe in Groove besitzt einen gemeinsamen Workspace. Jeder Client hält eine lokale Kopie des gesamten Workspace. So ist es dem User möglich auch ohne Netzwerkverbindung (asynchron) auf den gemeinsamen Daten zu arbeiten. Jeder Client synchronisiert sich automatisch mit den anderen sobald wieder eine Verbindung verfügbar ist. Die Daten im Workspace werden beim Client in einem komprimierten und verschlüsselten XML-Object-Store gehalten. Die XML-Technik erlaubt unter anderem bei Änderungen an Dateien nur die Informationen zur Änderung an die anderen Clients zu verschicken statt der gesamten Datei. Michael Hurwicz [W1] benutzt den passenden Begriff Deltas für diese Änderungsinformationen. Das Verfahren hat bei großen Dateien eine enorme Bandbreitenersparnis zur Folge.

Groove verschlüsselt alle Daten sowohl im gemeinsamen Workspace, auf den Relay Servern (s.u.) und beim Versand mit einer 192-Bit Verschlüsselung auf der Basis einer PKI. Mit öffentlichen/privaten Schlüsseln wird auch die Identifikation der Clients sichergestellt, die jeder eine eigene Identität besitzen.

Um die Notwendigkeit zu Beseitigen, daß zwei Clients gleichzeitig online sein müssen um sich zu synchronisieren, existiert ein zentraler Relay Server (bei www.groove.net), über den die Synchronisation optional ausgeführt werden kann. Für große Unternehmen besteht die Möglichkeit eigene Relay Server einzurichten. Neben der Synchronisation fallen dem Relay Server noch drei weitere Aufgaben zu: Awareness, Fanout und Transparency. Der Awareness-Service des Servers dient zur Erfassung des Online/Offline-Status der Clients, den Mitglieder eines Workspace jeweils voneinander beim Server abfragen können. Ein Client versendet die Deltas beim Ändern einer Datei jeweils direkt an alle anderen Clients in seinem Workspace. Ist die Verbindung eines Clients so schlecht, daß ein direktes Versenden zu zeitaufwendig wäre, schickt er nur eine einzelne Kopie des Deltas zum Relay Server. Dieser produziert dann einen Fanout des Deltas (sendet Kopien an die übrigen Clients). Transparency beschreibt in diesem Fall, daß ein User nicht bemerkt, daß zum Beispiel einige User in seinem Workspace an Rechnern hinter einer Firewall sitzen. Eine Firewall würde eine direkte P2P-Verbindung wie sie zum Versenden der Daten benutzt wird verhindern. Um eine Kommunikation dennoch möglich zu machen „verpackt"der Relay Server die Daten in HTTP, für das Firewalls normalerweise durchlässig sind. Der Client hinter der Firewall entpackt die Daten dann wieder.

Für die Anwendung in Unternehmen existieren noch zwei weitere Server-Arten. Der Enterprise Integration Server dient dazu Groove-Applikationen in bereits vorhandene zentrale Dienste zu integrieren. Der Enterprise Management Server bietet einen zentralen Punkt für die Administration des Groove-Netzes. Von dort können zentral Lizenzen vergeben, die Nutzung des Netzwerkes überwacht und Identitäten für Clients geschaffen werden.

Probleme: Daten von kleinen Firmen und Privatleuten, die keine eigenen Relay Server betreiben werden häufig über den zentralen Server von Groove verschickt oder sogar dort gespeichert. Trotz der recht starken Verschlüsselung ist so ein gewisses Vertrauen gegenüber dem Anbieter erforderlich. Bei einer hohen Inanspruchnahme des zentralen Relay Servers kann offensichtlich an dieser Stelle schnell ein Bottleneck entstehen.

Eigenschaften	**Technik**	**Probleme**
CSCW-System (Benutzerorientiert)	Implementationsgrundlage	Relay Server:
Synchrone u. asynchrone Arbeit	Zentral koordiniert	Bottleneck
unterstützt Firewall und NAT	192 bit Verschlüsselung	Datensicherheit
Datensicherheit	Workspace als Replikate	
	Synchronisation durch Deltas	

Überblick: Groove

3.2.4 Gegenüberstellung

Eine Schwäche von C/S-Systemen ist die bedingte Fähigkeit, Daten bereit zu stellen, für die ein erhöhter Bedarf besteht. Besonders temporäre Bedarfsspitzen sind kaum in ausreichender Qualität bedienbar. Eine Erweiterung der Verbindungsbandbreite des Servers, nur um eventuelle Spitzen bedienen zu können, wäre unwirtschaftlich. Intelligente P2P-Systeme dagegen können erhöhten Bedarf an bestimmten Daten und daraus resultierende Verkehrskonzentrationen an betroffenen Knoten erkennen. Durch kontrollierte Lastbalancierung kann der Verkehr auf das Netz verteilt und so insgesamt gemindert werden. Hierbei werden besonders gefragte Daten automatisch repliziert und an verschiedenen Stellen im Netz verfügbar gemacht. Durch Replikate auf Knoten in Nord Amerika und Europa kann zum Beispiel die Nutzung teuer Überseeverbindungen vermieden werden. Bestimmte P2P-Strukturen erlauben neben diesem aktiven ein natürliches und dadurch automatische Lastbalancierung. So kann ein Knoten, der von einem anderen Daten empfangen hat, diese zusätzlich selbst zur Verfügung stellen.

Durch die Möglichkeit beziehungsweise Eigenschaft Replikate von Daten an den verschiedensten Stellen im Netzwerk zu halten entsteht in P2P-Netzwerken eine hohe Datenredundanz. Ausfälle von einzelnen Knoten oder sogar Clustern beeinträchtigen so das restliche Netz kaum oder gar nicht. C/S-Systeme kommen durch einen Ausfall des Servers hingegen völlig zum Erliegen und sind dadurch natürlich auch verwundbarer gegenüber feindseligen Angriffen.

Ein großes Plus für P2P-Netzwerke sind ihre Sucheigenschaften. Ergebnisse einer Datensuche wie in Gnutella realisiert verweisen so gut wie nie auf bereits tote Verbindungen. Wohingegen die zentrale Datenbank einer Suchmaschine nur bedingt aktuelle Links enthält. Die Aktualität von Daten kann besonders davon profitieren, daß sie in P2P-Netzwerken dort gespeichert werden wo auch der Verantwortliche für sie arbeitet. So können zum Beispiel Versionskonflikte vermieden werden.

P2P-Netzwerke bergen jedoch auch Schwierigkeiten in sich. Eine verteilte Datenhaltung und eine relativ lockere physische Struktur machen ein reines P2P-Netzwerk nur schwer oder kaum administrierbar. Zugriffsrechte, die sich auf bestimmte Daten beziehen, können nicht wie beim Server auf zentrale Verzeichnisse bezogen werden. Das Logging von Vorgängen müsste entsprechend der Struktur des P2P-Netzwerkes angepasst und entsprechend komplexer werden.

Für Aufgaben bei denen es besonders auf die Sicherheit von Daten ankommt sind P2P-Netzwerke nur bedingt geeignet. Wenn Daten auf vielen verschiedenen Maschinen liegen, muss auf allen diesen Maschinen auch der nötige Sicherheitsstandart (Firewall, NAT, Virenscanner) erreicht werden. Das würde ungleich mehr Aufwand erfordern, als die Sicherung von Daten auf einem Server.

Aber besonders in Firmen standardmäßig genutzte Einrichtungen wie Firewalls, NAT-Router und Proxy-Server erschweren den Aufbau eines P2P-Netz-werkes, da sie die direkte Kommunikation über die Grenzen des lokalen Netzes hinaus einschränken. Ein P2P-Netzwerk muß also Mechanismen zur Verfügung stellen, um diese Grenzen zu umgehen. Das wirft natürlich wieder Fragen nach dessen Vertrauenswürdigkeit und somit neue Sicherheitsfragen auf.

Letztlich kann die Funktionalität von P2P-Netzwerken stark von der Qualität der Verbindung der einzelnen Knoten abhängen. Netzwerke, die Nachrichten von Knoten zu Knoten weiterreichen, würden von langsamen Knoten als ganzes gebremst werden. Ein Server wird durch langsame Clients dagegen sogar entlastet.

3.3 JXTA - Eine gemeinsame Sprache und Implementationsumgebung für P2P-Netzwerke

Das Ziel von JXTA [16] ist ein großes gemeinsames P2P-Netz, in dem alle Anwendungen die gleiche Sprache sprechen. Es stellt grundlegende Werkzeuge zur Erstellung von P2P-Applikationen zur Verfügung und definiert grundlegende P2P-Komponenten und eine Sprache für P2P-Netzwerke.

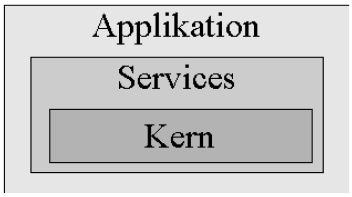
3.3.1 Zielsetzung und Struktur

Da die Struktur eines P2P-Netzwerkes stark durch das Verhalten der individuellen Knoten bestimmt ist, muss ein P2P- im Gegensatz zu einem C/S-Client eine Fülle zusätzlicher Aufgaben übernehmen, die je nach Netzwerk unterschiedlich ausfallen. Deshalb geht mit einem P2P-Netzwerk auch immer eine individuelle Applikation einher, die ihre eigenen Protokolle für die Datenübertragung zwischen den Peers definiert. Verschiedene P2P-Applikationen verstehen sich also in der Regel nicht. Dieses kann für den Anwender umständlich werden, wenn er verschiedene P2P-Netzwerke nebeneinander benutzen oder gar deren Inhalte kombinieren möchte. JXTA kann als gemeinsame Sprache unterschiedlicher P2P-Applikationen dienen.

JXTA ist nicht als eigenständige P2P-Software konzipiert, sondern bietet eine Implementationsgrundlage für verschiedenste P2P-Applikationen in Form eines Frameworks aus Java-Klassen, die Protokolle auf XML-Basis implementieren. Diese Protokolle regeln jede denkbare Art der Kommunikation innerhalb eines P2P-Netzwerkes. Die Abstützung auf Java macht JXTA plattformunabhängig.

Eine auf JXTA basierende Anwendung lässt sich in drei Schichten einteilen, die aufeinander aufbauen: Core Layer, Service Layer und Applications Layer. Die Core Layer besteht aus einem Großteil des JXTA-Frameworks und ist bei allen auf JXTA basierenden Anwendungen identisch. In ihr sind die grundlegenden Elemente eines P2P-Netzwerkes definiert (s.u.). Die Service Layer umfasst Elemente, die bei der Erstellung eines P2P-Netzwerkes nützlich sein könnten, aber nicht unbedingt benötigt werden. Alle diese Elemente müssen

auf Elementen der Core Layer basieren. Die restlichen Elemente des JXTA-Frameworks fallen in diese Schicht. Auch bei der Entwicklung von Anwendungen zählen Module, die grundlegende technik- bzw. architekturnahe Funktionen realisieren, zur Service Layer. Die benutzernahen und anwendungsspezifischen Funktionen der Applications Layer basieren wiederum auf Funktionen der Service Layer.



JXTA-Applikationsstruktur

3.3.2 Grundlegende Elemente eines P2P-Netzwerkes

Im JXTA-Framework sind die grundlegenden Elemente eines P2P-Netzwerkes modelliert.

- **Peer** *Any entity capable of performing some useful work, and communicating the results of that work to another entity over a network, either directly or indirectly.*
- **Peer Groups** dienen zur Aufteilung des Netzes in Interessengemeinschaften. Zwar könnten alle Peers miteinander kommunizieren, aber ein Filesharing Peer hat beispielsweise kein Interesse irgendein Instant Messenger Peer zu entdecken.
- **Network Transport Elements** sind für alle Aspekte des Datentransports zuständig.
  - **Endpoints** sind als Interface-Klassen definiert. Ihre Implementation realisiert die Nutzung eines konkreten Übertragungsmechanismus (TCP/IP, Bluetooth, GSM, etc.).
  - **Pipes** sind die virtuellen Verbindungen zwischen Endpoints.
  - **Messages** sind Container für Daten (Content), die über Pipes von Endpoint zu Endpoint transportiert werden.
- **Services** (eines Peers oder auch einer Peer Group) sind alle Funktionalitäten, die von anderen Peers genutzt werden können.
- **Advertisements** sind die formale Beschreibung aller Services eines Peers oder einer Peer Group.
- **Entity Naming Elements** beschreiben Elemente, die eine eindeutige Identifizierung benötigen (Peers, Peer Groups, Pipes und Content). Dabei wird eine übertragungsartunabhängige Kennzeichnung verwendet. So werden zum Beispiel Peers nicht durch ihre IP-Adresse identifiziert.
- **Protocols** (siehe Kapitel 3.3)
- **Security and Authentication Primitives** bieten u.a. Verschlüsselungsmechanismen für die Datenübertragung.

3.3.3 Die JXTA-Protokolle

Die sechs Protokolle der Core Layer regeln die Kommunikation im Netzwerk. Sie umfassen...

... die Anforderung von Advertisements von/durch Peers (Peer Discovery Protocol).

... die Anforderung entdeckter Dienste (Peer Resolver Protocol).

... die Anforderung von Statusinformationen der Peers (Peer Information Protocol).

... die Organisation von Peer Groups (Peer Membership Protocol).

... den Aufbau von Verbindungen zwischen Peers und den damit verbundenen Austausch von Informationen über eine Pipe (Pipe Binding Protocol).

... das Routing einer Verbindung zwischen Peers (Endpoint Routing Protocol). Also das Finden einer Route, über die eine Verbindung mittels des Pipe Binding Protocol aufgebaut werden kann.

3.3.4 Vor- und Nachteile

Mit Java als Sprache ist das JXTA-Framework plattformunabhängig verwendbar. Der Verzicht auf die Bindung an eine bestimmte Übertragungsart und das Abstützen auf XML für jeglichen Informationsaustausch macht JXTA zwar flexibel und ermöglicht die Kombination verschiedener P2P-Applikationen, aber erhöht den Aufwand, der für die Implementation einer Applikation benötigt wird. Das kann dazu führen das Entwickler sich gegen die JXTA-Plattform entscheiden um Aufwand und somit Kosten zu minimieren.

nach **Eigenschaften**	nach **Technik**
Anwendungsabhängigkeit	Plattformabhängigkeit
Flexibilität (Struktur)	Kompatibilität
Störungsanfälligkeit	Datensicherheit
Bedien-/Administrierbarkeit	

Überblick: Bewertungsgrundlagen für P2P-Netzwerke

3.4 Möglichkeiten für den Einsatz von P2P-Technologie

Für den Einsatz von P2P-Technologie sind zwei Marktsegmente erkennbar, die sich aufgrund ihrer Ansprüche und Zielvorstellungen unterscheiden. Firmen erwarten Zuverlässigkeit und Funktionen, die Kosten sparen und Qualität erhöhen und sichern. Privatleute dagegen legen weniger Wert auf Zuverlässigkeit, schätzen aber die Erweiterung ihrer persönlichen Möglichkeiten sowohl zur Gestaltung ihrer Freizeit als auch zur Erledigung ihrer Alltagsgeschäfte.

3.4.1 Marktsegment: Unternehmen

Die Gartner Group unterscheidet formale und informale P2P-Netzwerke. Für die Nutzung in Unternehmen wird ein formales Modell vorausgesetzt. In diesem ist der Zugriff auf sämtliche Ressourcen überwachbar ist (durch Logbücher, Rechtevergabe, Mitgliedschaft in Arbeitsgruppen). Der Enterprise Management Server ist das zentrale Werkzeug, das Groove zu einem formalen Netzwerk macht. Das informale Gnutella hingegen verfügt über keine derartige Kontrolleinrichtung. Weiterhin werden drei Trends identifiziert, mit denen Unternehmen sich in naher Zukunft auseinandersetzen müssen. Data Explosion, Need for Content Management Solutions und Internetworking and Collaborative Commerce.

Data Explosion

Die Menge an zu verarbeitenden Daten steigt kontinuierlich an. Um diese Daten filtern und mit gängigen Mitteln (Datenbanken) sammeln zu können, müssen sie in strukturierter Form vorliegen. Unstrukturierte Daten sind durch ein solches System also nicht erfassbar. Distributed Content Management Systeme (DCM) auf der Basis von P2P-Technologie könnten solche Daten mit vergleichsweise geringem Aufwand urbar machen. Zwar erfüllt Gnutella nicht die nötigen Voraussetzungen für eine professionelle Anwendung, jedoch könnten dessen Fähigkeiten zur Behandlung heterogener Datenbestände als Vorbild für ein solches System dienen. Die Dezentalisierung würde auch die Last von immer größer werdenden Datenbeständen mindern und die Konzentration von Verkehrslasten am zentralen Punkt im Netzwerk auflösen. Dadurch könnte die Anschaffung von neuen Servern und Netzanbindungen mit höherer Kapazität vermieden oder vermindert und eventuell schon vorhandene Ressourcenüberschüsse (Speicherplatz, Rechenleistung) auf Desktopsystemen ausgenutzt werden, was insgesamt zu einer Kostenersparnis führen würde.

(Web) Content Management Solutions

Mit wachsender Präsenz von Unternehmen im Internet und damit verbundenen Dienstleistungen wächst auch die Komplexität der Webseiten. Bei steigender Komplexität ist ein angemessener Qualitätsstandart ohne Unterstützung durch Software kaum mehr zu realisieren. Bestehende Web Content Management Systeme (WCM) basieren jedoch auf einem zentralen Datenbestand, wodurch die Erreichbarkeit und Aktualisierung der Daten den für C/S-Systeme bekannten Beschränkungen (vgl. Kapitel 2.3.) unterlieg. Bei global operierenden Unternehmen kommt noch die Notwendigkeit hinzu, Daten für die Nutzung in verschiedenen Regionen aufzubereiten. Das führt dazu, daß Replikate von Daten mit dem selben Inhalt an verschiedenen Orten gehalten werden. Dieser Umstand vergrößert die genannten Probleme nochmals. Gartner schlägt vor, bestehende WCM mit DCM zu kombinieren. Also die zentrale Datenbank durch einen virtuell zentralen Datenbestand zu ersetzen, der unter anderem leichter zu aktualisieren ist, da Daten an den Stellen gehalten werden könnten an denen sie ihren Ursprung haben. Lokalisierungen können dann separat auf den zentralen Daten aufsetzen, aber in örtlicher Nähe zu ihrem Einsatzgebiet gehalten werden. Dadurch wird die Gefahr von Inkonsistenzen oder Versionskonflikten gemindert. Verdienstausfälle oder Qualitätsverlust durch nicht erreichbare Dienste oder inhaltlich veraltete Lokalisierungen werden so gemindert.

Internetworking and Collaborative Commerce

Viele Unternehmen arbeiten zur Steigerung von Qualität und Effizienz nicht nur alleine an bestimmten Projekten, sondern teilen sich anfallende Aufgaben mit anderen Unternehmen. DCM könnten diese Zusammenarbeit durch ihre speziellen Fähigkeiten effizienter machen. Zum Beispiel könnte der Informationsaustausch in Projekten von den Sucheigenschaften in P2P-Netzwerken profitieren. Viele Unternehmen unterhalten bereits Portale, über die ihre Mitarbeiter Zugriff auf für sie relevante Daten haben. Diese könnten durch DCM auch für den Zugriff von Firmenexternen und sogar Kunden erweitert werden. Durch entsprechende P2P-Lösungen wäre auch der Austausch von unstrukturierten Daten möglich. Eine formale Plattform wie Groove könnte zur Realisierung genutzt werden. Selbstverständlich kann nicht davon ausgegangen werden, daß zwei Unternehmen immer die gleiche Kommunikationssoftware benutzen. Um die Vorteile einer P2P-Struktur dennoch zu nutzen, könnte ein solches Netzwerk auf einer Lösung wie JXTA aufsetzen.

Insgesamt verspechen Arbeitsabläufe durch P2P-Technologien weniger Unterbrechungsanfällig zu werden. In jedem Fall werden sie angenehmer für die Mitarbeiter, weil anstelle von technischen natürliche Umgangsformen treten (Daten vom Kollegen statt vom Server). Zwar ist der Gewinn für Unternehmen durch Motivationssteigerung schwer messbar, aber vermiedener Arbeitsausfall verrechnet sich sofort positiv.

3.4.2 Marktsegment: Privatpersonen

Ein entscheidender Faktor für die Attraktivität von P2P-Anwendungen im privaten Bereich stellt die allgemein verfügbare Verbindungsbandbreite dar. Denn Attraktivität für die breite Masse bedeutet in der Regel Multimediafähigkeit. Aber Voice- und Video-Streaming von einem 56k Knoten zum anderen bricht keine Qualitätsrekorde und die Übertragung von gängigen Multimediadateien wird zur Geduldsprobe. Mit dem Erscheinen von DSL-Zugängen (Digital Subscriber Line) und den Plänen für Internet aus der Steckdose und über das TV-Kabelnetz scheinen die Weichen für eine breitbandige Zukunft gestellt. Ein Trend bei bereits erhältlichen DSL-Zugängen bremst die P2P-Euphorie jedoch noch. Die günstigsten und somit gängigsten Angebote sind asynchrone Zugänge. Als Produkt der C/S-Kultur bieten sie zwar eine erheblich größere Bandbreite für den Downstream (üblich: 768kbit/s), aber der Upstream ist immer noch auf ISDN-Niveau (128kbit/s). Für P2P-Netzwerke sind aber beide Richtungen annähernd gleich wichtig. Anwendungen der P2P-Technologie, die wirklich auf massenhaften Zuspruch stoßen, werden nicht zuletzt deshalb also noch eine Weile auf sich warten lassen.

Die gängigsten P2P-Anwendungen im privaten Bereich sind heute Filesharing-Netzwerke und Instant Messenger (IM). Erstere sind besonders wegen der Möglichkeit zum Tausch von Musik- und Videodateien beliebt. Die P2P-Technologie wird aber nicht nur zum Vergnügen genutzt. In vielen Ländern ist die freie Meinungsäußerung immer noch gefährlich (z.B. China). Das Internet bietet dort eine Möglichkeit der Zensur zu entgehen. Mittlerweile überwachen Regierungen aber die üblichen. Wer etwa verschlüsselte E-Mail oder andere Daten versendet oder empfängt macht sich schon verdächtig. Außerdem können Adressen mit unerwünschten Inhalten geblockt werden. Derlei Beschränkungen zu umgehen ist zum Beispiel mit dem Red Rover[4] Netzwerk möglich. Gestützt auf ein P2P-Netzwerk von freiwilligen Clients wird hier das unauffällige Empfangen von unerwünschten

[4] http://redrover.org

Inhalten möglich. Andere Systeme wie Publius[5] und Free Haven[6] stellen durch eine dezentrale Netzwerkarchitektur eine Technik zur Verfügung, um Dokumente anonym und praktisch nicht zensierbar bereitzustellen.

3.4.3 Überblick: Angesprochene P2P-Netzwerke und ihre Anwendung

Netzwerk	**Anwendungsgebiet**
AOL Instant Messenger	IM,Chat,Filetransfer
ICQ	IM,Chat,Filetransfer
Napster	Filesharing(mp3)
Gnutella	Filesharing(allgemein)
Groove	CSCW-System(kommerziell)
Publius	Dokumentarchiv anonym, unzensierbar
Free Haven	Dokumentarchiv anonym, unzensierbar
Red Rover	Informationsbroker (Empfang von geblockten Inhalten)

3.5 Schlussfolgerungen

Wie schon am Beispiel der Definition für die fünf Modelle von P2P-Netzwerken der Gartner Group zu sehen ist, ist der Übergang von C/S- zu P2P-Netzwerken fließend. Gartner verwendet selbst in der Definition der P2P-Modelle oft den Begriff Server (vgl. Kap. 2.2.1.). So besteht wohl auch der sinnvollste Einsatz der P2P-Technologie in einer Kombination mit gewohnten C/S-Strukturen. Ein P2P-Netzwerk definiert sich also schon durch die Benutzung von P2P-Technologie für die Lösung bestimmter Probleme, ohne dabei C/S-Strukturen völlig auszuschließen. Im Idealfall können sich so die Vorteile beider Systeme ergänzen. Das ökonomische Potential solcher Lösungen ist entsprechend hoch einzuordnen.

Wenn es gelingt Administrierbarkeit und Sicherheitsaspekte mit P2P-Tech-nologie zu vereinbaren, dann verspricht deren Einsatz in Zukunft nicht nur angenehmere Arbeitsabläufe für Mitarbeiter, sondern durch effizientere Ausnutzung von Ressourcen wie Netzwerkbandbreite und Speicherkapazität und einer erhöhten Ausfallsicherheit der Systeme auch Kostenersparnisse für Unternehmen. So können Mitarbeiter benötigte Daten in der aktuellsten Version direkt von dem dafür zuständigen Kollegen beziehen. Der unnatürliche Umweg über den zentralen Server in der Firma entfällt. Auch synchrone Arbeit an gemeinsamen Dokumenten kann unabhängig von der Erreichbarkeit eines zentralen Servers realisiert werden. Mit gemeinsamen Standards wie JXTA für die P2P-Kommunikation können Daten auch mühelos mit firmenexternen Rechnern und sogar auf einer „ad hoc"-Basis wie etwa im Meeting ausgetauscht werden.

Für private Anwender bietet P2P die Möglichkeit, Verbindungen ohne Mittelsmänner und so ohne die Gefährdung persönlicher Daten direkt untereinander aufzubauen. Dabei beschränkt sich die Datenübertragung nicht nur auf ein bestimmtes Medium wie etwa beim

[5] http://publius.cdt.org
[6] http://freehaven.net

Telefon sondern ist universell nutzbar. So kann der Internet-User auch ohne zusätzlichen finanziellen Aufwand zum Provider werden.

Mit steigender Verfügbarkeit von Breitbandzugängen wird sich das Internet weiter von der reinen Datenquelle zu einem Medium entwickeln, in dem die aktive Partizipation und Interaktion von Usern einen immer größeren Anteil annimmt. Durch ihre Vorteile auf diesen Gebieten werden sich P2P-Architekturen weiter verbreiten. Wenn man sich auf gemeinsame Standards einigen kann, werden sich unterschiedliche P2P-Applikationen in Zukunft bis zu einem gewissen Grad verstehen. Technische Grenzen zwischen Gemeinschaften wie heute etwa den Usern der verschiedenen Instant Messenger oder verschiedenen Unternehmen werden verschwinden.

P2P ist keine komplett neue Idee. Sie ist eine neue Betrachtungsweise. C/S schränkt auf eine zentralisierte Struktur ein. P2P ist flexibler und somit besser geeignet für die Kommunikation in einer Gesellschaft, die sich ständig sowohl auf ökonomischer als auch sozialer Ebene und mit scheinbar immmer größerer Geschwindigkeit weiterentwickelt.

Literaturverzeichnis

[16] Brendon Wilson: JXTA; URL: http://www.brendonwilson.com/projects/jxta/, Mai 2002

[18] A. Oram, N. Minar, C. Shirky, T. O'Reilly: Peer-to-peer: Harnessing the Power of Disruptive Technologies; O´Reilly & Associates, 2001

[19] G. Fox: Peer-to-peer Networks; Computing in Science and Engineering, May/June 2001, pp 75-77

[20] M. Parameswaran, A. Susarla, A. B. Whinston: P2P-Networking: An Information sharing Alternative; IEEE Computer, July 2001, pp 31-38

[21] Gartner Consulting: The Emergence of Distributed Content Management and Peer-to-peer Content Networks; White Paper, January 2001

[W1] Michael Hurwicz: Emerging Technology: Groove Networks: Think Globally, Store Locally, URL: http://www.networkmagazine.com, 07.05.2001

[W2] P2P Working Group, URL: http://www.p2pwg.org, Mai 2002

[W3] Gnutella News, URL: http://www.gnutellanews.com, Mai 2002

Kapitel 4

Wireless LAN — Technologie und kommerzielle Nutzung

Christian Czosseck

Inhaltsverzeichnis

4.1 Einleitung

Wireless LAN, im weiteren auch WLAN genannt, fristete im letzten Jahrtausend eher ein Schattendasein. Technologisch zwar ausgereift, litt es jedoch unter mangelnder Akzeptanz unter den potentiellen Käufern. Da keine einheitliche Norm vorhanden war, blieb es Kind einiger weniger Pioniere, die versuchten, diese Technologie kundenfreundlich zu machen. Mit der IEEE 802.11 Norm verabschiedete nun das Institut of Electrical and Electronical Engineers 1999 einen einheitlichen Standard.

Unter diesem wurden mehrere Übertragungsmedien und -verfahren vereint, um eine einheitliche und einfache Einbindung in die existierende IEEE[1] 802 Normreihe zu ermöglichen.

Diese Abhandlung wird nun IEEE 802.11 von mehreren Seiten beleuchten. Hierbei wird auf die beiden unteren Schichten des OSI-Modells eingegangen. Ein besonderes Augenmerk wird den im Wireless LAN neu eingeführten Sicherheitsschichten gegeben.

Desweiteren wird über den ökonomischen Nutzen dieser Technologie gesprochen. Hierbei wird über das Für und Wider diskutiert und gezeigt, wo sich Wireless LAN nutzbringend einsetzen lässt.

Der erste Teil dieser Ausarbeitung wird sich nun mit der technischen Seite des Wireless LAN beschäftigen. Hierbei werden die Rahmenbedingungen aufgezeigt, in den Wireless LAN geschaffen wurde um danach die beiden Schichten des OSI-Modells zu besprechen, die durch die IEEE 802.11 Norm definiert werden. Dieses wird auch der Sicht eines Technikers geschehen.

4.2 IEEE 802.11: Standard für drahtlose Netze

Nachdem das erste kommerzielle Wireless LAN Produkt im Dezember 1990 auf dem US-Markt verfügbar war, entwickelte sich diese technologische Neuerung eher zäh. Nicht zuletzt war eine fehlende Spezifikation Grund dafür. Die Produkte, die vorhanden waren, waren zueinander inkompatibel. Die Firmen, die diese vertrieben, waren aber auch darauf bedacht, ihre eigenen Produkte zu verkaufen. Nichts desto trotz scheuten viele, gerade kleinere, potentielle Anbieter die Investition, da sie damit rechen mussten, dass ihr Produkt sich nicht durchsetzen würde.

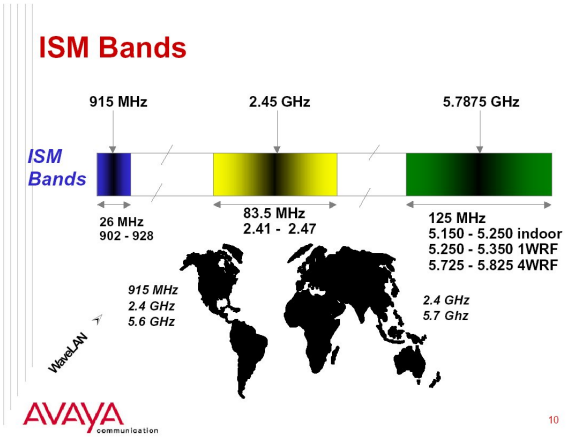


Abbildung 4.1: Die (bedingt) freien ISM-Bänder [6]

Es zeichnete sich ab, das Wireless LAN ein immenses Wachstumspotential haben würde und daher verabschiedete die IEEE im Jahr 1997 nach sieben Jahren Arbeit die IEEE 802.11 Norm in seiner ersten Version. Hiermit wurde ein einheitlicher Standard für die Datenfernübertragung erstellt, der Übertragungsraten von 1 und 2 MBit/s spezifiziert.

[1] Institute of Electrical and Electronics Engineers, [12]

Hierfür wurden Spezifikationen des MAC[2]-Protokolls und die zu unterstützenden Verfahren des PHY[3]-Layers festgeschrieben. Die zu unterstützenden Übertragungsverfahren sind zwei Frequenz-Spreizverfahren für HF-Übertragung im 2,4 GHz und 5 GHz Band und ein Infrarotverfahren.

Das 2,4 GHz ISM[4]-Band Netz wurde gewählt, da es auf der ganzen Welt kostenfrei für nicht kommerzielle Nutzung zur Verfügung steht. Dennoch gibt es Einschränkungen, da in einigen Ländern staatliche Restriktionen existieren zum Beispiel in Japan, Frankreich und Spanien. Dieses ist in Tabelle 4.1 auf Seite 69 nachlesenbar. Aber auch die erlaubte Sendeleistung und weitere Spezifikationen unterscheiden sich teils erheblich. Das 5 GHz Netz ist innerhalb der USA frei zugänglich, in Europa liegt es aber unter der Kontrolle der ETSI[5], da hier das patentierte Produkt HiperLAN operiert.

4.2.1 Die ursprüngliche Definition: 802.11

Der IEEE 802.11 entstand, wie schon erwähnt, durch das IEEE, in dessen Umfeld aber auch zwei weitere Gremien agierten. Zum einen war es die **Wireless LAN association** [9], die die Verbreitung des Wireless LAN durch Öffentlichkeitsarbeit und Marketing unterstützte. Zum anderen die **Wireless Ethernet Compatibility Alliance** [10], welche durch scharfe Spezifikationen die Interoperabilität von IEEE 802.11 Geräten durch eine Zertifizierung sicher stellte. Geräte, die dieses Zertifikat erhalten, dürfen das Kürzel **Wi-Fi** tragen, was für Wireless-Fidelity steht.

So entstand der IEEE 802.11 , welcher sich einigen Herausforderungen zu stellen hatte, denn anders als bei Ethernet konnten im Funknetz Probleme entstehen, die so noch nicht vorgekommen waren. Die wesentlichen Punkte, die Wireless LAN erfüllen sollte, lassen sich wie folgt zusammenfassen:

Robustheit der Übertragung, besonders unter Berücksichtigung des *Hidden-Station-Problems*, welches in Abschnitt 4.2.1 näher beschrieben wird.

Multi channel roaming, welches ähnliche Möglichkeiten bieten sollte, wie es von Handy-Funknetzen her bekannt war.

Power Management Funktionen, welche besonders wegen des hohen Energiebedarfs einer durchgehenden Funkverbindung für Notebooks und PDA's ein Problem darstellen.

Automatic rate selection ermöglicht es den Geräten, je nach Verbindungsqualität zwischen 2 und 1 MBit/s zu wählen.

Security WEP Da das Übertragungsmedium Funkwelle wesentlich angreifbarer ist als eine drahtgebundene Variante, kam der Bedarf einer Sicherung der Übertragung schon auf unterster Ebene in Betracht.

[2]Medium Access Layer
[3]Physical Layer
[4]Industrial, Scientific, Medical
[5]European Telecommunications Standards Institute

Einordnung

Die IEEE 802.11 gehört zur Gruppe der IEEE 802, der Spezifikation, welche sich mit Netzwerken, Netzwerkprotokollen und Übertragungsrichtlinien beschäftigt. Im Bild 4.2 werden die gegenseitigen Beziehungen dieser Spezifikationen gezeigt.

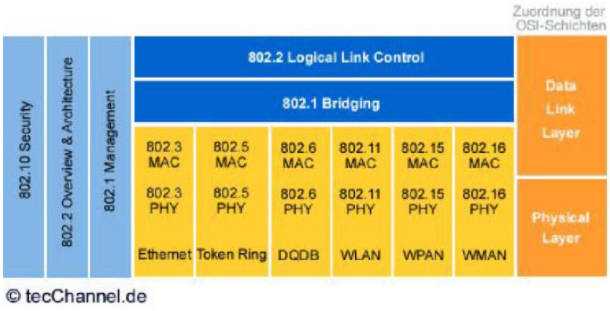


Abbildung 4.2: Einordnung von Wireless LAN [1]

Wie zu erkennen ist, erstrecken sich die Spezifikationen auf die unteren zwei Schichten des OSI-Modells. Auf diese wird im weiteren Verlauf eingegangen. Vorab werden aber erst einige allgemeine Spezifikationen aufgezeigt.

Frequenzspektren Wireless LAN nutzt elektromagnetische Wellen als Träger der Informationen. In der IEEE 802.11 werden drei mögliche Frequenzen definiert.

1. 2.4 GHz ISM, ist frei
2. 5 GHz, nur für 802.11a, nur in den USA frei
3. Infrarot, Wellenlänge

Erwähnenswert ist, dass die IEEE 802.11 die erste Norm ist, die gleichzeitig drei verschiedene Übertragungsmedien (genauer Frequenzen) definiert. Für diese sind sogar noch verschiedene Modellierungsverfahren vorgesehen, auf die später ab Seite 4.2.1 eingegangen wird.

Im weiteren wollen wir uns auf das 2,4 GHz Spektrum beschränken, da es das zur Zeit wichtigste Band ist.

Die Übertragung von elektromagnetischen Wellen bedarf dem Einsatz von Sende- und Empfangsantennen. Diese können in verschiendenen Bauformen existieren. Das liegt in der Wellenlänge des 2,4 GHz Bandes begründet, welche ca. 12,5cm beträgt und lässt sich somit in Antennen von ca. 3cm Größe realisieren (/4-Antenne). Das ist ideal für den Einsatz in PCMCIA-Karten. Die Antenne ist in diesem Fall entweder direkt an der Karte angebracht und ragt dann ca. 2 cm aus dem PCMCIA-Slot hervor oder wird via Kabel am Notebookdeckel angebracht. Neuere Geräte mit integrierter Wireless LAN Fähigkeit bauen die Antenne zumeist im Deckel des TFT ein.

Die zweite Variante sind die sog. **Access Points**. Diese werden benutzt, um die BSS bzw. ESS, auf welche in 4.3.2 auf Seite 84 näher eingegangen wird, aufzubauen. Die folgenden Bilder zeigen einige Beispiele.

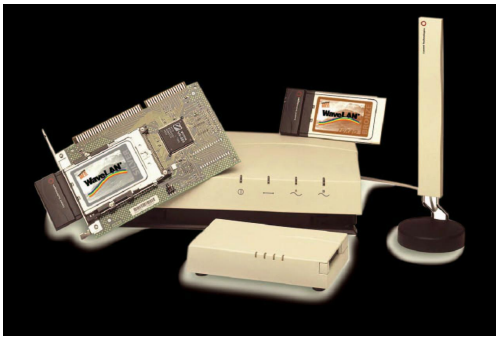


Abbildung 4.4: Beispiele für Wireless LAN Geräte [6]

Physical Layer

Der Physical Layer, die unterste Schicht des OSI-Modells, ist für die eigentliche Bitübertragung zuständig. Somit ergeben sich hier naturgemäß die größten Unterschiede zur drahtgebundenen Kommunikation. Bei der Spezifikation dieser Schnittstelle musste man besonders auf mögliche Probleme einer Luftübertragung eingehen, welche im Wesentlichen sind:

1. Rauschen und Interferenzen durch andere (z.B.: Mikrowellen)
2. Andere nach IEEE 802.11 arbeitende Sendegeräte, denn auch wenn deren Sendeleistung nicht mehr zum Datenaustausch ausreicht, können sie doch störend wirken.
3. Signale anderer Sendegeräte, die nicht nach IEEE 802.11 arbeiten, aber das selbe Frequenzband nutzen.

Aufbau des PHY

In Anbetracht der oben genannten Besonderheit des PHY-Layers und besonders der Tatsache, dass drei voneinander gänzlich unterschiedliche Übertragungsverfahren und -medien genutzt werden, war es nötig in der IEEE 802.11 eine weitere Schicht einzufügen. Hierfür wurde die 1. Schicht des OSI-Modells nochmals unterteilt und zwar in eine

1. Physical Medium Dependant Sublayer (PMD) und eine
2. Physical Layer Convergence Protocol (PLCP).

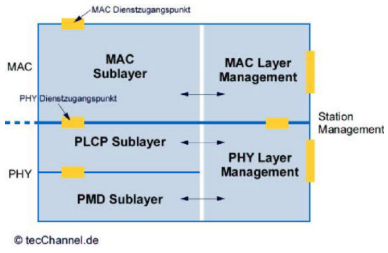


Abbildung 4.5: Aufteilung des PHY Layers bei Wireless LAN [1]

Hierbei übernimmt die PMD die Kodierung und Modulation während die PLCP einer gewöhnlichen PHY-Layer-Schicht entsprechende Funktionen bietet. Insbesondere liefert die PLCP das Clear Channel Assignment Signale (CCA), welches den momentanen Zustand des Mediums angibt.

Schmalband- vs. Breitbandtechnologie Dem Wireless LAN , besonders in seiner momentan verbreiteten Version IEEE 802.11b , steht ein Frequenzbereich von ca. 83,5MHz zur Verfügung.

Um dieses Band nutzen zu können, sind im Wesentlichen zwei mögliche Technologien denkbar:

Schmalbandtechnik Hierbei wird die zur Verfügung stehende Kanalbreite in gleichen Teilen in Kanäle aufgeteilt. Jede einzelne Bandbreite wäre dann ca. 1MHz, was in Europa eine Anzahl von ungefähr 79 Kanälen bedeuten würde. Nachteil ist aber, dass jeder einzelne Kanal sehr störungsanfällig wäre, wenn nicht sogar unvorhersehbar unnutzbar.

Besonders unangenehm ist u.a. die Tatsache, dass Mikrowellen ebenfalls in diesem Spektrum arbeiten. Somit gäbe es eine Diskrepanz auch darüber, wer nun die besseren und wer die schlechteren Frequenzen nutzt.

Breitbandtechnik Die Lösung dieses Problems ist die Nutzung des gesamten verfügbaren Spektrums, bzw. eines wesentlich größeren Teils auf einmal. Damit würden sich die Störungen quasi auf die gesamte Breite aufteilen.

Im Wesentlichen werden und wurden hierfür zwei Verfahren verwandt:

FHSS Frequency Hopping Spread Spectrum und

DSSS Direct Sequence Spread Spectrum.

Während FHSS gerade bei den ersten Produkten große Verbreitung fand, dieses lag zum einen daran, dass es einfacher und somit billiger zu verwirklichen war, wird heutzutage nahezu ausschließlich nur noch DSSS verwandt. Das liegt unter anderem darin begründet, dass DSSS höherer Datenübertragungsraten zulässt.

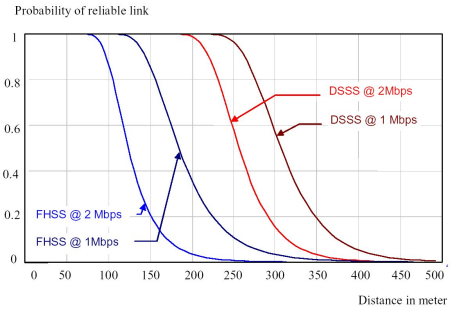


Abbildung 4.6: Vergleich von Reichweite zu Durchsatz zw. FHSS und DSSS [8]

FHSS Mit FHSS ist die einfache Nutzung eines größeren Spektrums realisierbar. Das Prinzip ergibt sich schon aus dem Namen. Der verfügbare Bereich wird wie bei der Schmalbandtechnik in einzelne Kanäle geteilt. Diese haben einen Bereich von ca. 1MHz und ergeben ca. 79 Kanäle in Europa; näheres siehe Seite 4.1. Die IEEE 802.11 sieht vor das diese verfügbaren Kanäle in 3 Gruppen á 26 Mustern zusammengefasst werden. Hierbei wird die Abfolge der verwendeten Frequenzen aus einer Basisfolge berechnet, die einer

Region	**Frequenzband (GHz)**	**Sprungfrequ. (GHz)**	**# Sequenzen**
Europa, USA	2,4000 - 2,4835	2,402 - 2,483	79
Japan	2,4710 - 2,4970	2,473 - 2,495	23
Frankreich	2,4465 - 2,4835	2,447 - 2,473	27
Spanien	2,4450 - 2,4750	2,448 - 2,482	35

Tabelle 4.1: FHSS: Nutzbare Frequenzbereiche aufgeschlüsselt nach Region [1]

Pseudo-Zufallskette von Zahlen im Intervall 0 bis 78 entspricht. Hierbei ist eine minimale Sprungdistanz von 6 Kanälen einzuhalten.

Die Basisfolge wird zwischen Sender und Empfänger ausgehandelt, kann zufällig bestimmt sein oder aber fest eingegeben.

Sei nun diese Basisfolge wie folgt gegeben: $b(i) = \begin{cases} 0, & \text{i=1;} \\ 54, & \text{i=2;} \\ 75, & \text{i=3;} \\ 45, & \text{i=4.} \end{cases}$

Dann bestimmt sich die k-te Frequenzfolge als:

$$f_k(i) = 2402 + (b(i) + k) mod 79[GHz],$$

wobei i immer im Intervall von 1 bis 4 durchlaufen wird. Hierdurch wird sichergestellt, dass man sich zum einen im vorgegebenen Frequenzbereich befindet und zum anderen eine Pseudo-Zufallsreihe erstellt wurde. Im oberen Beispiel ergibt sich:

k/i	1	2	3	4
0	02	56	75	47
1	03	57	76	48
2	04	58	77	49
3	05	59	78	50
4	06	60	**00**	51
⋮	...	...	...	...

Zur Verdeutlichung des Verfahrens und dessen Nutzens folgt eine kleine schematische Simulation, welche aus [1] übernommen worden ist. Hierbei wird angenommen, dass es 2 Sender (FHSS-Sender 1 und 2) und einen Störsender gibt. Letzterer blockiert durchgängig die Frequenz f_3. Durch die Nutzung des FHSS konnte somit die Anzahl von gegenseitigen Störungen und die negative Wirkung des Störsenders minimiert werden. Es ist verständlich, dass bei einer höheren Anzahl von Sendern natürlich häufiger gegenseitige Interferenzen auftreten werden. Dieser Umstand und die Tatsachen, dass zum einen ein Störsender einen gesamten Bereich lahm legt und zum anderen diese Anfälligkeit Störungen gegenüber diese Technologie weniger geeignet mach für höhere Datenübertragungen als 2 MBit/s, wurde in der IEEE 802.11 eine Alternative definiert.

DSSS Im Gegensatz zu FHSS ist DSSS wesentlich aufwändiger und somit teurer zu realisieren. Daher wurde es bis 1999 eher seltener eingesetzt. Als aber der Ruf nach höheren Übertragungsraten laut wurde, wurde mit dem IEEE802.11b, einer Erweiterung des ursprünglichen Protokolls, DSSS zum bis heute genutzten quasi-Standardübertragungsverfahren im Wireless LAN.

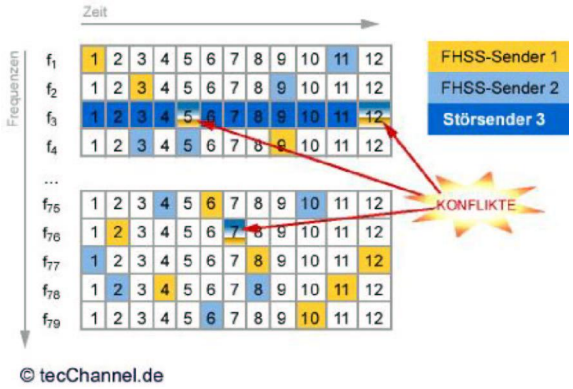


Abbildung 4.7: Störresistent durch das Springen durch die Kanäle [1]

DSSS ist ebenfalls ein Breitbandverfahren, welches aber grundsätzlich anders arbeitet als FHSS. Die zu übermittelnden Daten werden mit einem weiteren Datenstrom von 11bit Länge kodiert, der Mindestlänge bei einem Direct Sequence-Verfahren. Diese Pseudo-Random Numerical Sequence (PN) kann eine beliebige Sequenz sein. Dann muss aber sichergestellt sein, dass sowohl Sender als auch Empfänger dieselbe Sequenz benutzen. Hierbei hat sich herausgestellt, dass der so genannte Barker-Code (10110111000) besonders geeignet ist, da er besonders gute Autokorrelationseigenschaften hat. Er wird auch in anderen Direct Sequence Geräten verwandt.

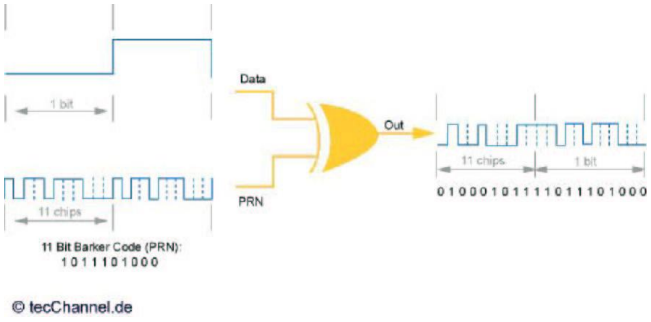


Abbildung 4.8: Verknüpfung des Daten- zum Sendesignal mit der PN-Sequenz [1]

Hierbei wird jedes einzelne Bit des zu übertragenden Stromes mit dem 11bit langem Spreizcode XOR verknüpft. Hieraus ergibt sich aber eine wesentlich höhere Bandbreite von 22MHz:

- Quelle bei 1MHz Übertragungsfrequenz,
- jedes Bit wird mit dem Baker-Code moduliert,
- um diese 11MHz verlustfrei übertragen zu können, wird ein Band von 22MHz Breite benötigt.

Hieraus resultiert, dass der verwendbare Frequenzbereich nur noch drei Kanäle parallel zueinander betreiben kann, ohne dass sich diese untereinander behindern. Insgesamt sind im ISM-Frequenzband bis zu 14 Kanäle möglich, welches in Tabelle 4.2 dargestellt wird.

Wie man in Abbildung 4.9 sieht, ist es möglich zwei weitere Kanäle dazwischen zu betreiben. Dann muss aber bedacht werden, dass diese sich gegenseitig behindern und somit

Region	Frequenzband (GHz)	DSSS-Nutzung	Kanäle	Sendeleistung
USA	2,4000 - 2,4835	2,412 - 2,462	11	1000 mW
Europa	2,4000 - 2,4835	2,412 - 2,472	13	100 mW (EIRP)
Japan	2,4710 - 2,4970	2,484	1	10mW/MHz
Frankreich	2,4465 - 2,4835	2,457 - 2,462	2	100 mW (EIRP)
Spanien	2,4450 - 2,4750	2,457 - 2,472	4	100 mW (EIRP)

Tabelle 4.2: DSSS: Nutzbare Frequenzbereiche aufgeschlüsselt nach Region [1]

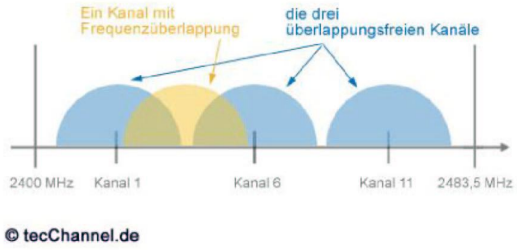


Abbildung 4.9: DSSS: Die drei Haupt- und mögliche Zwischenkanäle [1]

die maximale Übertragungsrate senken. Wird genug Abstand zu einander gehalten, wären aber dennoch hohe Übertragungsraten möglich. Die Bänder erschienen sich dann gegenseitig als Hintergrundstörungen. Das weitere Platzieren von Kanälen führt aber sicherlich zu erheblichen Überlagerungen und somit Fehlern.

Das Direct Sequence-Verfahren, welches auch im Mobilfunk Anwendung findet, hier werden aber verschieden PN's gleichzeitig im selben Frequenzband verwendet, hat aber auch noch einen anderen Effekt: Es ist wesentlich unanfälliger gegenüber Störungen und somit leistungsfähiger, was man in Abbildung 4.6 sieht. Nun wollen wir und dieses verbesserte Störungsanfälligkeit mit Hilfe des folgenden Bildes 4.10 verdeutlichen.

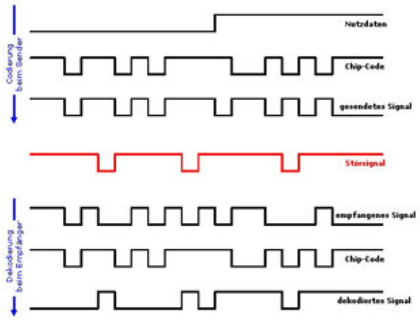


Abbildung 4.10: Fehlerbeseitigungseigenschaft des DSSS [2]

Der oberste Strom ist der eigentliche Datenstrom, mit der die darunter liegenden PN-Sequenz XOR-Verknüpft wird. Dieser liegt, wie schon beschrieben, im Verhältnis 11:1 vor. Das dann eigentlich gesendete Signal werde dann durch ein willkürliches Störsignal beeinflusst und somit ergibt sich das empfangene Signal. Dieses wird dann mit der selben PN wieder XOR verknüpft und ergibt das unten angegebene dekodierte Signal. Da die Länge des Chip-Cods bekannt ist und das Signal synchronisiert übertragen wird, kann bis zu einer Anzahl von 5 (unter 11) Störsignalen noch sicher das Originalsignal rekonstruiert werden.

Infrarot Technologie Wird nicht mehr benutzt, da unpraktikabel im Alltagsgebrauch (direkte Sichtverbindung), zu geringe Leistung und Reichweite.

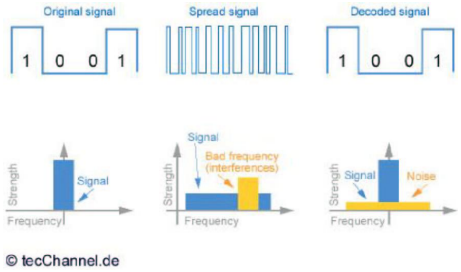


Abbildung 4.11: DSSS ist gegenüber Hintergrundstörungen unanfällig [1]

MAC Layer

Die MAC-Schicht, die zweite im OSI-Modell steht für *Medium Access Control*. Sie ist für die einheitliche Verwendung des gewählten Übertragungsmediums durch die oberen Schichten zuständig und regelt somit die Übertragung von Informationen. Hierbei wird sichergestellt, dass die Information am anderen Ende wirklich ankommt und richtig übertragen wurde. Für den zu übertragenden Inhalt ist sie aber nicht zuständig.

Die MAC-Schicht des IEEE 802.11 hat große Ähnlichkeiten zu seinem Bruder im 802.3 (Ethernet). Dennoch gibt es besondere Änderungen, die sich gerade wegen des verwendetem Mediums Funkwelle ergaben. Die IEEE 802.11 verwendet für die Übertragung einen CSMA/CA[6]-Algorithmus (im Unterschied zu 802.3 dort ist es ein CSMA/CD[7]), der für folgendes steht:

Carrier Sense Jeder Teilnehmer des Übertragungsmediums überwacht dieses durchgängig und passt seine eigenen Aktionen dem momentanen Zustand an.

Multiple Access Mehrere Teilnehmer teilen sich das selbe Medium, hier im besonderen die selbe Frequenz.

Collision Avoidance Da bei Funkübertragungen eine Kollision nicht feststellbar ist (vgl. 4.2.1), wird versucht, diese gänzlich zu verhindern, was mehr oder weniger gut gelingen kann.

Um sicherzustellen, dass ein Datenpaket ohne eine Kollision, die dieses ja zerstören würde, übertragen wird, musste man sich ein Verfahren ausdenken, was die Anzahl der Kollisionen minimiert.

Das Problem, vor dem man stand, war, dass eine Kollision nicht sicher erkannt werden konnte. Diesen Sachverhalt verdeutlicht das sog. Hidden-Station-Problem am besten.

Hidden Station Problem Es gibt zwei klassische Beispiele für dieses Phänomen. Beim ersten seien zwei Rechner durch ein Hindernis derart getrennt, dass sie keine Verbindung aufbauen können. Ein dritter Rechner habe Kontakt zu beiden (vgl. Bild 4.12). Nun kann jeder der beiden von einander getrennten Rechnern ein Paket zu dem dritten schicken. Beide werden, wie in 4.2.1 auf Seite 76 näher beschrieben, das Medium auf eine laufende Übertragung in überprüfen. Tun das beide gleichzeitig, werden beide einen freien Raum finden, da sich sich ja gegenseitig nicht wahrnehmen. Nun beginnen beide zu senden und

[6]Carrier Sense Multiple Access / **Collision Avoidance**
[7]Carrier Sense Multiple Access / **Collision Detection**

die Pakete werden nahezu gleichzeitig ankommen, womit beide zerstört sind. Da aber nur auf Empfängerseite geprüft werden kann, ob eine Übertragung fehlerfrei war, wird dem Sender das nicht auffallen.

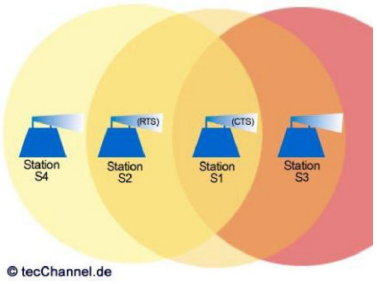


Abbildung 4.12: Hidden-Station-Problem: Beispiel [1]

Die zweite Variante, die im Bild 4.12 gezeigt wird ist, das ein Sender den nächsten erreicht, aber für den Übernächsten die Sendeleistung nicht mehr ausreicht und somit diese Sendestation nicht wahrgenommen wird. Nun entsteht hier das gleiche Problem wie eben beschrieben.

Deswegen werden im weiteren als erstes die Übertragungsmethoden ohne Berücksichtigung dieses Problem beschrieben, und dann die implementierte Lösung mit der Behandlung in Abschnitt 4.2.1 auf Seite 76.

Distributed Coordination Function Im Wesentlichen ist die Lösung der Collision Avoidance mit Hilfe von Timing-Maßnahmen realisiert worden. Hierfür wird eine Zeitspannen, die **Interframe Space** (*IFS*) genannt wird, definiert. Diese gibt an, wie lange eine sendewillige Station wartet, bis sie versucht, ihr Pakt zu senden.

Hiermit lassen sich auch bedingt Prioritäten definieren und realisieren, denn eine Station, die eine kürzere IFS hat, wird automatisch eher anfangen zu senden und andere werden das feststellen und ihre Sendung entsprechend zurückstellen (Carrier Sense). Dieses war gewünscht, da Wireless LAN auch für einen möglichen Einsatz in zeitkritischen Umgebungen geeignet sein sollte. Dieses wurde durch die sog. Point Coordination Function realisiert, welche auf Seite 77 näher beschrieben wird. Dennoch sind diese Zeitspannen ein gefährliches Gebiet, da schon kleinere Änderungen der Zeitspannen, welche im Folgenden dargestellt werden, große Auswirkungen auf den Datendurchsatz haben können.

Die grundlegende IFS-Zeit ist die **Distributed IFS**-Zeit (*DIFS*), auf der die *Distributed Coordination Function* operiert. Diese DIFS-Zeit wartet jede Station und hört sich das Medium an, bevor es eine Übertragung beginnt. Diese Verhalten nennt man *Listen Before Talk*.

Sollte das Medium aber in dieser Zeit besetzt sein, so beginnt der sog. **Backoff-Prozess**. Dieser recht aufwendige Prozess soll die Wahrscheinlichkeit einer Kollision besonders nach dem Beenden einer laufenden Übermittlung *aller dann noch wartenden* Stationen minimieren. Hierfür wird eine Pseudo-Zufallszahl generiert, die zwischen Null und einem (beliebigen, implementierungsabhängigen) Maximum liegt. Dieses Maximum wird als **Contention Window** bezeichnet.

Diese Zahl wird mit der Dauer eines Übertragungszeitschlitzfensters multipliziert und als Countdown verwandt. Die Station wartet nun so lange, bis die gerade laufende Über-

tragung beendet ist, dann wartet es die DIFS-Zeit und dann wird der Backoff-Countdown herunter gezählt. Ist das Medium immer noch frei, kann die Übertragung beginnen.

Dennoch kann dieses Verfahren das Kollisionsproblem **nicht verhindern**, sondern nur minimieren. Dieses zeigt das Beispiel 4.13, welches aus [1] entnommen wurde.

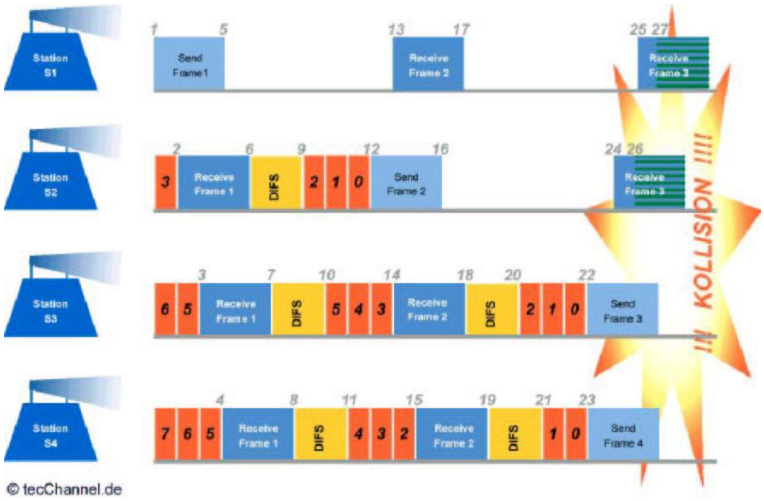


Abbildung 4.13: Kollisionsgefahr trotz ausgefeilter Timing-Algorithmen [1]

1: Station S1 beginnt die Übertragung eines Rahmens Frame 1. Alle anderen Stationen befinden sich in verschiedenen Stadien des Backoff. Sie dekrementieren also den Backoff-Zeitgeber in jedem Zeitschlitz, in dem das Medium als frei erkannt wird.

2: Nach einer (von der räumlichen Entfernung der beiden Stationen abhängig) Verzögerungszeit empfängt S2 den Rahmen. Da das Medium als belegt erkannt wird, stoppt das Herunterzählen des Backoff-Zählers.

3, 4: Nach entsprechenden Verzögerungszeiten empfangen auch S3 and S4 den Rahmen von S1 and halten ihre Backoff-Zähler an.

5: S1 beendet die Übertragung des Rahmens.

6, 7, 8: S2, S3 and S4 erkennen das Medium wieder als frei and warten eine DIFS-Zeit ab.

9, 10, 11: Nach Ablauf der DIFS-Zeit beginnen die Stationen erneut, ihre Backoff-Zähler zu dekrementieren.

12: Der Backoff-Zähler von S2 läuft ab. Daher beginnt S2 unmittelbar mit der Übertragung des Rahmens Frame 2.

13: Nach einer Verzögerungszeit empfängt S1 den Rahmen. Da sich S1 nicht im Backoff befindet, hat dies fur diese Station keine Auswirkungen.

14, 15: Sobald S3 and S4 den Rahmen empfangen, erkennen sie das Medium als belegt and stoppen das Herunterzählen des Backoff-Counters.

16: S2 beendet die Übertragung.

17: S1 erkennt das Medium wieder als frei. Da keine Informationen zur Übertragung anstehen, hat dies aber keine weiteren Auswirkungen.

18, 19: S3 and S4 erkennen das Medium wieder als frei and warten eine DIFS-Zeit ab.

20, 21: Nach Ablauf der DIFS-Zeit beginnen beide Stationen, ihren Backoff-Zähler zu dekrementieren.

22: Der Backoff-Zähler von S3 läuft ab, die Station beginnt unmittelbar mit der Übertragung des Rahmens Frame 3.

23: Gleichzeitig läuft der Backoff-Zähler von S4 ab. Auch diese Station beginnt unmittelbar mit der Übertragung eines Rahmens (Frame 4). Eine Kollision bahnt sich an.

24, 25: S2 and S1 empfangen Frame 3 zunächst störungsfrei.

26, 27: Station S2 empfängt die Überlagerung der Übertragenen Rahmen Frame 3 and Frame 4.

DIFS und SIFS Um nun zur Lösung des Hidden-Station-Problems zu kommen, welches in Abschnitt 4.2.1 auf Seite 72 beschrieben wurde, ist der erste Schritt, einen Bestätigungsmechanismus einzubauen. Hierbei wird jede empfangene Übertragung durch den Empfänger mit einem **Acknowledgement** (ACK) bestätigt. Dieses ermöglicht es auf Kosten der effektiven Datenübertragungsleistung, da ja der Overhead steigt, dass sich zum einen der Sender sicher sein kann, dass seine Übertragung fehlerfrei angekommen ist. Zum anderen ist das eine Möglichkeit, das Hidden-Station-Problem zu lösen, was im Folgenden gezeigt wird.

Die Quittierung wird nach einer kurzen Verzögerung gesendet, welche als **Short Interframe Space** (*SIFS*) bezeichnet wird und kürzer sein muss als die DIFS. Somit hat die Bestätigung eine höhere Priorität als die Sendungen anderer Stationen. Die Abbildung 4.14 zeigt dieses an einem Beispiel und wurde aus [1] entnommen.

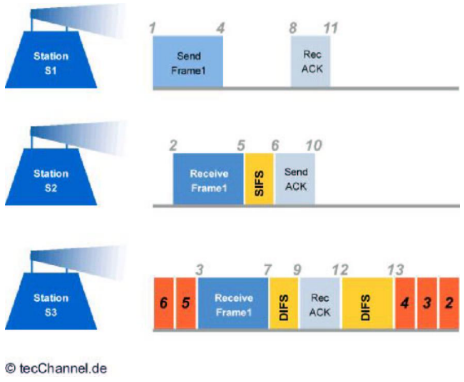


Abbildung 4.14: ACK werden bevorzugt versandt [1]

1: Station S1 beginnt die Übertragung eines Rahmens Frame 1 an S2. Zu diesem Zeitpunkt befindet sich S3 im Backoff-Prozess.

2: S2 beginnt mit dem ordnungsgemäßen Empfang von Frame 1.

3: Auch S3 empfängt den Rahmen. Das Medium wird als belegt erkannt, der Backoff-Zähler stoppt.

4: S1 beendet die Übertragung des Rahmens.

5: S2 erkennt das Medium wieder als frei. Da die Übertragung ordnungsgemäß abgeschlossen wurde, beginnt die Station ein SIFS-Warte-Intervall, um anschließend eine Quittung zu versenden.

6: Nach Ablauf von SIFS versendet S2 den ACK-Rahmen.

7: S3 erkennt das Medium wieder als frei and beginnt ein DIFS-Warte-Intervall.

8: S1 empfängt den ACK-Rahmen.

9: S3 empfängt den ACK-Rahmen und stellt fest, dass das Medium belegt ist. Daher bricht sie das noch nicht beendete DIFS-Warte-Intervall ab.

10: S2 beendet die Übertragung des ACK-Rahmens.

11: S1 beendet den Empfang des ACK-Rahmens.

12: S1 beendet den Empfang des ACK-Rahmens und beginnt ein neues DIFS-Warte-Interval 1.

13: Nach dessen vollständigem Ablauf kann der Backoff-Zähler von S1 weiter dekrementiert werden.

Während dieses Quitierungsverfahrens kann es aber auch dazu kommen, dass die Bestätigung den Sender nicht erreicht. Dieses kann (mindestens) zwei Gründe haben. Zum einen kann die Bestätigung selber durch eine Kollision (einer Hidden-Station) zerstört werden. Zum anderen spielen aber auch die Laufzeiten eine Rolle. So kann es gerade in ausgedehnten Netzen mit vielen Teilnehmern dazu kommen, dass die Summe aus SIFS und Hin- und Rückwegzeit größer ist als die DIFS.

Wenn keine Bestätigung durch den Sender empfangen wird, bereitet dieser eine erneute Sendung des Paketes vor und begibt sich in den Backoff-Zustand. Hierbei wird aber die obere Schranke des Intervalls, aus dem die Pseudo-Zufallszahl generiert wurde, verdoppelt. Dieses wird bei jedem erneuten erfolglosen Versuch der gleichen Nachricht erneut gemacht, bis die obere Grenze einen Wert *CWmax* erreicht, der von der Implementierung abhängig ist. (Der CWmax-Wert sollte mit Hilfe von Optimierungsverfahren zusammen mit den DIFS und SIFS bestimmt werden.)

Wenn das Paket einmal erfolgreich übermittelt worden ist, wird diese obere Intervallgrenze wieder zurückgesetzt.

Dieses Verfahren kann in kleineren Netzen[8] das Hidden-Station-Problem schon so weit reduzieren, dass es alleine ausreicht (bei moderaten Overhead). Bei größeren ist die Lösung aber noch nicht gefunden. Dieses bekommt man endgültig mit dem RTS-CTS-Mechanismus in Griff.

Collision Avoidance; RTS und CTS Mechanismus Das Wesen dieses Mechanismuses besteht darin, dass eine anstehende Übertragung durch den Sender angekündigt wird (**Request for Transition**, *RTS*) und der Empfänger diese bestätigt (**Clear for Transition**, *CTS*).

In der RTS wird die Identifikation des Senders zusammen mit der Länge des Paketes übersandt. Die selben Informationen, aber mit geändertem Absender wird dann vom Empfänger zurückgeschickt. Hieraus ergibt sich folgende Situation zur Lösung des Hidden-Station-Problems: Der Sender (A) sendet sein RTS an den Empfänger (B). Diese RTS wird

[8] bezogen auf die Anzahl der Teilnehmer

durch die verdeckte Station (C) natürlich nicht wahrgenommen, womit dieses ihre DIFS Zeit wartet, um eine Übertragung zu beginnen. Jetzt wird aber die Empfangsstation (B) das CTS senden, welche die verdeckte Station (C) genauso mit bekommt, wie die sendende (A). Hiermit erfährt aber die verdeckte Station (C), dass eine Übertragung im laufen ist und wird einen Countdown starten, der genau so lange wartet, wie die angekündigte Übertragung gemäß der CTS benötigen wird. Nun wird der DIFS-Countdown ebenfalls zurückgesetzt um eine Kollision mit der Quittierung zu verhindern.

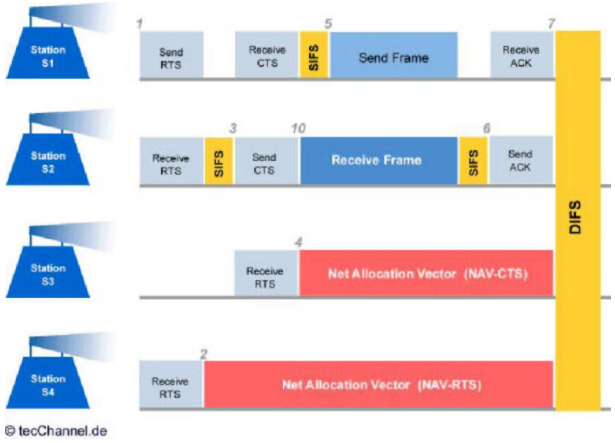


Abbildung 4.15: ACK werden bevorzugt versandt [1]

Hiermit wird die anstehende Kollision ausgeschlossen und das Hidden-Station-Problem gelöst. Dennoch erkauft man sich diese Lösung mit einem erheblich höherem Datenaufkommen, was den effektiven Datendurchsatz verringert. In der Praxis werden diese beiden Verfahren gemischt eingesetzt. Bei geringen Verkehr wird auf das RTS-CTS-Verfahren gänzlich verzichtet. Steigt das Datenaufkommen oder ist die zu übertragende Dateie sehr groß, was natürlich nicht von der MAC-Schicht entschieden wird, sondern von einer höheren veranlasst wird, wird dieser Mechanismus mit genutzt.

Point Coordination Function Mit der Point Coordination Function ist eine Möglichkeit durch die IEEE 802.11 spezifiziert worden, mit der zeitkritische Anwendungen unterstützt werden. Hiermit wird dem **Point Coordinator**, üblicherweise einem Access Point (vgl. 4.3.2 auf Seite 84), der ausschließliche, priorisierte Zugriff auf das Medium erteilt. Dieser regelt dann den Zugriff für alle anderen im Empfangsbereich liegenden Geräte. Hierfür synchronisiert der *Point Coordinator* mit Hilfe eines **Beacon**[9], welches er in einem wettbewerbsfreien Zeitraum sendet, alle anderen. In diesem Beacon ist die Dauer der **Contention Free Period** untergebracht.

Innerhalb dieser Zeitspanne nach dem Beacon fragt der Point Coordinator nun alle Station ab, ob sie einen Übertragungswunsch haben und bildet intern eine *Polling List*. Nach der Contention Free Period beginnt nun einheitlich, da durch das Beacon synchronisiert, der Übertragungszeitraum. In diesem senden nun die einzelnen Stationen anhand der Polling List des Point Coordinator.

Das IEEE 802.11 fordert diese Implementierung zwingend, auch wenn sie nicht genutzt wird. Somit ist sichergestellt, dass jedes konform betriebene Wireless LAN Gerät diese Betriebsart erkennt und sich dynamisch anpasst.

[9]engl.: Leuchtfeuer, Signal

Sicherheitskonzepte

Drahtlose Übertragungen bringen einige Sicherheitsrisiken mit sich, die höher sind als die von drahtgebundenen Übertragungsverfahren. Bei drahtgebundenen Übertragungen lässt sich der Leiter mehr oder weniger gut abschirmen, denn ein möglicher Angreifer muss Zugang zu diesem Draht bekommen, was schwierig sein kann bei autarken Installationen.

Bei Funknetzen ist dieses komplizierter und für den Angreifer einfacher, denn die verwendeten Funkwellen lassen sich nicht ohne weiteres gegen Abhörer schützen. Um jegliche Funkübertragung von einem Gebäude aus nach außen hin abzuschirmen, wäre eine funksichere Isolation aller Außenwände und Fenster nötig, was wohl in den meisten Fällen nicht gemacht wird. Daher werden diese Funkübertragungen auch außerhalb des eigentlichen Anwendungsgebietes durch andere IEEE 802.11 Geräte abzuhören sein.

Aus diesem Grund heraus wurden einige Sicherheitsmechanismen schon in der MAC-Schicht implementiert. Diese sind aber nicht mit den Sicherungsmechanismen auf den höheren Schichten des OSI-Modells zu verwechseln, welche unabhängig von denen der 2. koexistieren.

Authentifizierung Auf unterster Ebene kann die Zulassung der Teilnehmer durch eine **Electronical System ID** geregelt werden. Hierbei wird durch den Administrator eine solche ID vorgegeben, welche alle von ihm gewünschten Rechner haben. Dieses ist aber aus zwei Gründen nur sehr unzureichend. Zum einen gibt diese ID nur die allgemeine Zugangsberechtigung wieder und liefert keine eindeutigen Informationen zur Identifikation des Teilnehmers selber. Zum anderen ist es relativ einfach, sich eine gültige Electronical System ID zu verschaffen, indem man sich den Datenverkehr des Wireless LAN Netzes anhört und beobachtet, welche Teilnehmer mit welcher ID zugelassen werden, und welche nicht. Außerdem wird die Situation durch falsche Implementierungen einiger Anbieter noch verschlechtert, da diese in der Konfigurationsdatei die Option *any* für die Electronical System ID zulassen, was quasi unbeschränktem Zugang entspricht.

Eine weitere Möglichkeit, die die Sicherheit erhöhen soll, ist, dass der Administrator den Access Points eine Liste von zulässigen MAC-Adressen vorgibt, die zu akzeptieren sind. Diese eigentlich weltweit eindeutige Nummer, die jede Netzwerkkarte unterscheidbar macht, ist aber durch entsprechende Manipulationen am Treiber der Netzwerkkarte änderbar. Eine gültige MAC-Adresse lässt sich genauso wie bei der Electronical System ID bestimmen und somit ist auch dieser Mechanismus nach kurzer Zeit ausgehebelt.

Ein weiterer Nachteil dieses Verfahrens ist eher ein administratorisches, denn in größeren Netzen kommt einiges an Arbeit auf den Admin zu, um u.a. Roaming zu realisieren, denn die Konfigurationstools sind zumeist eher eingeschränkt nützlich. Außerdem benötigt der Admin auch eine sichere Übertragung für das Verwalten neuer und alter Teilnehmer und durch dieses Management steigt der Overhead erheblich.

Eine Erweiterung dieses Verfahrens wird daher von vielen Anbietern von Wireless LAN angeboten, welches sich **Remote Authentification Dail-In User Service** (RADIUS) nennt und eine zentrale Verwaltung von Benutzeridentifikationen und Passwörtern. Dieses RADIUS geht aber über die eigentliche Norm heraus. Näheres zu diesem Thema kann man auf der Internetseite [11] der IETF nachlesen.

WEP Als letztes ist in der IEEE 802.11 eine Verschlüsselung der zu übertragenden Pakete auf der MAC-Ebene **optional** implementierbar. Dieses steht für **Wired Equivalent Privacy** und wird im Abschnitt 4.4 auf Seite 88 eingehend untersucht.

4.2.2 Erweiterungen

Im weiteren werden jetzt die schon durchgeführten und geplanten Erweiterungen vorgestellt.

Schon während der Verabschiedung des IEEE 802.11 war klar, dass die Übertragungsraten nicht ausreichen würden, was dazu geführt hat, dass zwei Untergruppen der IEEE 802.11 gebildet wurden, um entsprechende Erweiterungen zu schaffen. Hierbei konzentrierte sich die IEEE 802.11b auf die Erhöhung dieser unter *Beibehaltung* des Frequenzbandes bei 2,4 Ghz. Beim IEEE 802.11a erreicht man die Erhöhung durch den Wechsel auf das, ebenfalls ja schon spezifizierte 5 GHz Band. Dieses ist in Europa teilweise durch HiperLAN der ETSI reserviert und somit nicht nutzbar.

802.11a

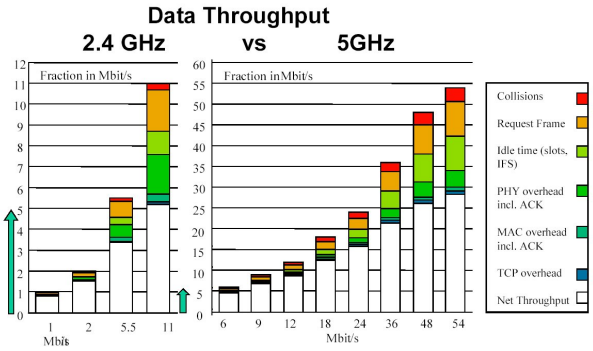


Abbildung 4.16: Vergleich der Übertagungsraten zwischen IEEE 802.11a und IEEE 802.11b nach [7]

Bei diesem Standard wird die Trägerfrequenz im 5 GHz Bereich genutzt, um eine höhere Datenübertragungsrate zu ermöglichen. Hierfür wird ein anders Modulatonsverfahren verwandt (*Orthogonal Frequency Division Multiplexing*), welches besonders mit der Varianz der Signallaufzeiten auftretenden Probleme besser umgehen kann. Die maximalen Übertragungsreichweite ist nur unwesentlich geringer, was auf die kürzere Wellenlänge zurückzuführen ist, dafür steigt die maximale

Datenübertragungsrate auf 54 MBit/s und ist somit 27 Mal so schnell wie der ursprüngliche IEEE 802.11 Standard! Diese Technik ist aber nicht kompatibel zu den Geräten im 2.4 GHz Bereich, was eine erneute Komplettinvestition bei höheren Anschaffungskosten nötig macht. Dieses und die schon beschriebene Einschränkung des Frequenzbandes auf Amerika haben daher z.Z. noch nicht zu einem Durchbruch geführt. Dennoch könnte damit gerechnet werden, besonders wegen des enorm gesteigerten Datendurchsatzes.

Einen schönen Vergleich der zu erwartenden Übertragungsraten im Vergleich zu der gleich vorgestellten Parallelentwicklung IEEE 802.11b sieht man in Abbildung 4.17 auf der nächsten Seite. Hier erkennt man auch, dass die maximale Übertragungsrate von 54

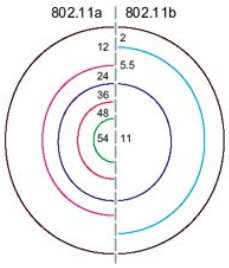


Abbildung 4.17: Vergleich der Reichweiten zwischen IEEE 802.11a und IEEE 802.11b nach [2]

MBit/s in einem Bereich von ca. 10 bis 15 Metern zur Verfügung steht. Selbst in einer Entfernung von 50 Metern, bei der die maximale Übertragungsrate von 11 MBit/s beim neuen gleich vorgestellem IEEE 802.11b auf 5 MBit/s sinkt, sinkt sie bei IEEE 802.11a gerade mal von 36 auf 24 MBit/s fällt. Somit ist diese Variante an sich die Zukunft, wird aber wohl von dem Bedarf der Kunden abhängen.

802.11b

Dieser Standard, der z.Z. der als quasi-Standard Verwendung findet spezifiziert eine maximale Übertragungsrate von 11 MBit/s. Dieses wird aber auf dem gleichen Frequenzbereich von 2.4 GHz realisiert und kann als eine direkte Weiterentwicklung des IEEE 802.11 betrachtet werden. Die IEEE 802.11 wurde 1997 verabschiedet. In dieser legte man schon mit der Spezifizierung des DSSS den Grundstock für eine leistungsfähigere Variante. Ein Jahr später wurde dann die IEEE 802.11b verabschiedet, welche durch ausschließliche Verwendung von DSSS (vgl. 4.2.1 auf Seite 69), einer dynamischen Geschwindigkeitsanpassung an die Signalqualität, einer verbesserten Signalempfindlichkeit und besonders einem Wechsel des Modulationsverfahrens realisiert. Hierbei wurde von dem urspünglichem **Binary Phase Shift Keying** zum **Quadrature Phase Shift Keying** gewechselt, was mehr Bits pro Übertragungssymbol übertragen kann. Weiterhin wird von der 11bit *Baker* Sequenz zu einer kürzeren 8 bit **Complimentary Code Keying** genannten PN-Sequenz übergegangen.

Durch die Abwärtskompatibilität zu allen nach IEEE 802.11 betriebenen Geräten und somit der, aus der Sicht der Benutzer, einfache integration in das bestehende Netz fand dieses System sehr schnell Verbreitung und wird heutzutage fast ausschließlich verwendet.

sonstige verbreitete Funknetze

Gerade in der Zeit, in der noch kein einheitlicher Standard gegeben war, aber der Bedarf in den Unternehmen immer größer wurde, dies ist besonders die Zeit zw. 1990 und 1997, wurden einige Firmen selber aktiv und erschufen ihre eigenen Lösungen samt Spezifikation. Im Folgenden seinen die verbreitetsten genannt (nach [3]).

Bluetooth Bluetooth ist eine Art Kabelersatz auf kurze Distanzen. Für Ad-hoc-Vernetzungen zwischen Laptops, Handys and PDAs ist er hervorragend geeignet. Er wird voraussichtlich die Infrarotverbindungen ablösen.

Hiper LAN Unter HiperLAN (High Performance Radio LAN) versteht man den Europäischen Standard (ETSI) im S-GHz-Band. Type 1 spezifiziert ein drahtloses Ethernet mit 24 MBit/s.

Wireless ATM Wireless ATM ist eine Variante des HiperLAN mit 20 MBit/s im 5-GHz-Band.

HiperACCESS HiperACCESS ist eine Variante des HiperLAN. Die Technik ist Wireless Local Loop mit 20 MBit/s im 5-GHz-Band.

HiperLINK HiperLINK ist eine Variante des HiperLAN. Die Technik ist eine drahtlose Punkt-zu-Punkt-Verbindung mit 155 MBit/s im 17-GHz-Band.

HomeRF Home Radio Frequency ist eine abgespeckte and dann kostengünstige IEEE-802.11-Variante fur Heimanwender. Mit dem Shared Wireless Access Protocol (SWAP) gibt es die Möglichkeit zur Sprachübertragung. Die HomeRF Working Group ist ein Industriekonsortium.

Openair Drahtloser Netzstandard vor IEEE 802.11 von der Firma Proxim.

4.2.3 Zukünftige Erweiterungen und Projekte

802.11h

Hier versucht die IEEE die Erhöhung der Datenübertragungsrate beim bestenhenden IEEE 802.11b durch Timing-Maßnahmen und Verbesserungen der technischen Komponenten zu schaffen. Angepeilt sind ca. 22 MBit/s.

802.11i

Diese Gruppe beschäftigt sich mit der Verbesserung der Sicherheit in allen IEEE 802.11 Standards, indem sie die schon bekannten und in Abschnitt 4.4 auf Seite 88 aufgezeigten Sicherheitslücken schließen und die Sicherheit darüber hinaus erhöhen wollen.

4.3 Nutzen und Verbreitung des W-LAN

In diesem Abschnitt werden die Vor- und Nachteile des Wireless LAN aus ökonomischer Sicht behandelt. Es sollen Verwendungsmöglichkeiten aufgezeigt werden, wo ein Einsatz sich lohnt und diese anhand von konkreten Anwendungen betont werden.

Hiernach wird auf die möglichen technischen Topologien eingegangen, welche mit Beispielen visualisiert werden. Dann soll mit praktischen Erfahrungsberichten der Aufbau solcher Netze beschrieben werden um schließlich mit einer Abgrenzung zu Bluetooth und UMTS einen Ausblick in die Zukunft zu geben.

4.3.1 Abgrenzung gegenüber Wired LAN, Vor- und Nachteile

Die Verbreitung von Wireless LAN hat in den letzten Jahren explosionsartig zugenommen. Diese hat besonders mit der Spezifizierung durch die IEEE zu tun, da nun auch kleiner

Unternehmen Produkte anbieten konnten, die kompatibel zu anderen waren. Des weiteren sanken dadurch die Preise und das Produkt Wireless LAN wurde massentauglich.

Die Vor- und Nachteile haben große Ähnlichkeit mit denen, beim Einsatz von Mobiltelefonen, erweitert um Dienstleistungen, die für den Computereinsatz typisch sind. So kann man folgende, wenn auch nicht alle, nennen:

Vorteile, Unternehmenssicht

Kostenersparnis

1. Zumeist sind drahtgebundene Infrastrukturen schon vorhanden, Wireless LAN ermöglicht aber das Erschließen von Gebieten, wo eine Datenverbindung nur schwer möglich oder gar unmöglich ist. (z.B. Lagerhaltung)
2. Das Verlegen von Kabeln kann erhebliche bauliche Maßnahmen nach sich ziehen, die entsprechend teuer sind.
3. Gelegentlich ist es nötig, Netzwerke nur für kurze Zeiten aufzubauen und dann wieder abzubauen, wie es beispielsweise bei Konferenzen der Fall ist. Hier eignet sich ein Funknetz.
4. Das Aufbauen eines Netzwerkes ist, bei existierenden Access Points, eine Sache von Sekunden und kann somit genutzt werden, um schnell Updates durchzuführen.

Gesteigerte Wirtschaftlichkeit und Effizienz

1. Während man mobil ist, erhält man dennoch die Verbindung zu seinem Unternehmen, was die Reaktionszeit des Einzelnen drastisch erhöhen kann.
2. Im Zuge von industriellen Fertigungsanlagen, besonders denen, die einen hohen Grad an Individualisierung realisieren müssen, kann jedes Produkt bis zur Fertigstellung mit einem mobilen Anschluss ausgestattet werden, der die einzelnen Arbeitsschritte aufzeigt.
3. Die Logistik, besonders in der Lagerverwaltung, kann durch den Einsatz eines Wireless LAN Anschlusses optimiert werden.

Verbesserter Kundenservice

1. Durch die flexible Anbindung der firmeninternen Geräte können Verkäufer sich freier bewegen und dem Kunden somit (wörtlich) entgegenkommen.
2. Ebenfalls ist eine Art Wegführung und Rundführung durch Mobile Geräte realisierbar, wie es in einigen Museen möglich ist.
3. Weiterhin kann durch ortsunabhängiges Arbeiten (vgl. auch Roaming) besser auf die Bedürfnisse der Angestellten eingegangen werden, was die Produktivität und Arbeitsfreude steigert.

Nachteil, Ausgaben Zumeist die Kosten, die für eine entsprechende Neuanschaffung zu tätigen sind. Daher wird jedes Unternehmen eine Kostenrechnung machen um zu sehen ob und wie weit sich eine solche Investition lohnt. Zumeist sind die Vorteile durch gesteigerte Produktivität aber weit höher als die Kosten für Aufbau und Betrieb der Infrastruktur.

Nachteil, Sicherheit Besondere Probleme ergeben sich, wenn sensiblere Daten übertragen werden sollen bzw. werden könnten. Auf diese Problematik wird im Abschnitt 4.4 auf Seite 88 gesondert eingegangen.

In Anbetracht dieser Vorteile lassen sich folgende Anwendungsbeispiele auflisten, welche aber nur ein kleiner Teil der möglichen Bandbreite wiedergeben.

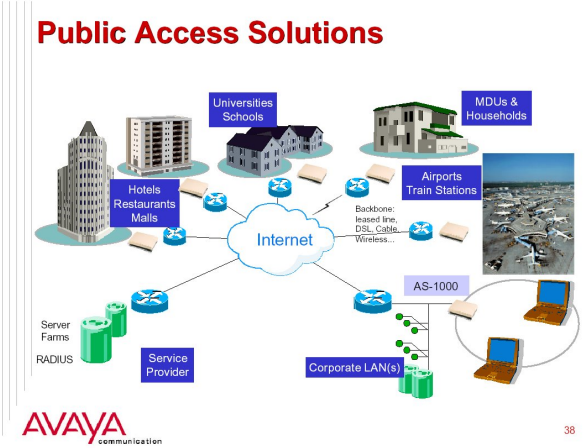


Abbildung 4.18: Beispiele für die Anwendung von Wireless LAN nach [6]

Briefing Rooms Gerade in solchen ist ein Wireless LAN sinnvoll anwendbar, da heutzutage die meisten solcher Meetings mit elektronischen Medien unterstützt werden. Somit kann der Vortragende schnell und unkompliziert seine z.B. PowerPoint-Präsentation starten, ohne das großartige Maßnahmen (nach der Vorkonfiguration des Raumes natürlich) nötig sind. Außerdem ist der Datenaustausch untereinander sehr einfach.

Lagerhaltung In Lagern kann die Effizienz von nicht automatisierten Einrichtungen gesteigert werden, indem jeder Fahrer z.B. einen Laptop erhält, auf dem seine nächsten Aufträge automatisch erscheinen und er sie sofort nach Erledigung quittieren kann. So werden Wege gespart und die Leistung gesteigert.

Flexible Arbeitsplätze Man kann das Arbeitsklima von Angestellten, die für die Arbeit am PC nicht zwingend Ortsgebunden sein müssen, erhöhen, indem man z.B. diesen erlaubt, das Büro zu verlassen um im Grünen weiter zu arbeiten, da sie ja ihren Laptop mitnehmen können.

Rundgänge und Führungen Z.B. in Museen könnten Führungen realisiert werden, indem jeder Besucher ein Headset bekommt, gekoppelt an einen tragbaren Computer, der dann je nach Lokalität eine entsprechende Beschreibung gibt. (Diese Art der Anwendung ist technisch anders und finanziell günstiger möglich)

Universitäten könnten den Studenten die Möglichkeit bieten, überall auf dem Campus Zugang zum Uni-Netz zu erhalten und somit auch außerhalb des Rechenzentrums zu arbeiten. Außerdem könnten Vorlesungen ggf. effizienter gestaltet werden.

Flughäfen Ähnliche Angebote sind auf Flughäfen, Bahnhöfen und Hotels denkbar, wo zumeist gegen eine Pauschale Nutzungsgebür Zugang zum Internet ermöglicht wird.

4.3.2 Topologie des W-LAN

Beim Wireless LAN gibt es im Wesentlichen zwei Arten des Betriebes: sog. Ad-hoc Netze und die BSS/ESS, gemanagte ausgedehnte Netze.

IBSS, **Independent Basic Service Set**, oder auch **Ad-hoc Netze** genannt, sind aufgebaute Netzwerke, die durch die Präsenz von mindesten zwei IEEE 802.11 Geräten entstehen. Hierbei versuchen beide Geräte durchgängig, einen Partner zu finden und gehen dabei nur eine voreingestellte Frequenz ab. Finden sie einen Partner, wird die Netzwerkverbindung sofort aufgebaut, was ähnlich von statten geht, wie beim Anschließen eines neuen USB-Gerätes.

Vorteil dieses Netzes ist die Spontanität, Nachteil die Tatsache, dass die Verbindung nur untereinander funktioniert. Ein Kontakt zum Internet lässt sich nur in Ausnahmefällen herstellen. Dennoch hat dieses Netz praktische Relevanz, denn in einigen Situationen ist dies völlig ausreichend. So z.B. in den Angesprochenen Meetings, zum schnellen Datenaustausch oder zum Drucken von Dokumenten im einer Office-Umgebung.

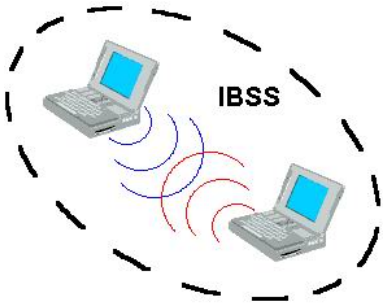


Abbildung 4.19: Ein Independent Basic Service Set besteht aus Stationen, die miteinander ohne Hilfe eines Repeaters eine Netz aufbauen [3]

BSS und ESS Dennoch werden die meisten Anwender eine andere Art der Netzverwaltung wünschen. Denn in den IBSS ist keine Roaming möglich, weiterhin sind die Zugriffsmöglichkeiten auf existierende kabelgebundene Infrastrukturen (z.B. des Unternehmens) nicht möglich, ebenso nicht der Internetzugang. Weiterhin sind Sicherungsmaßnahmen durch z.B. Firewalls nicht realisierbar.

Deshalb sind in den **Basic Service Set** (BSS) sog. Access Points eingeführt worden. Diese haben zum einen einen Zugang zum Kabelnetz des Unternehmens und werden dort als unabhängige Komponente wie ein Repeater behandelt. Hiermit ist die Möglichkeit des Netzzuganges in der Umgebung eines solchen Access Points gegeben. Dieser wiederum kann durch die unternehmensinterne Sicherungsstruktur nach außen abgesichert werden.

Werden mehrere dieser Stationen betrieben, die alle auf dem selben Kanal arbeiten, kann somit ein Roaming realisiert werden. Hierbei haben diese Access Points die gleiche Aufgabe wie die Funkzellen im Mobilfunkbereich. In diesem Fall spricht man von einem **Extended Service Set** (ESS), welches die gängig Anwendung des Wireless LAN darstellt.

Eine Abwandlung ist die Möglichkeit, innerhalb des selben Raumes bis zu drei Netzwerke und somit ESS gleichzeitig zu betreiben, die jeweils eine von den drei unabhängigen Frequenzen nutzen (vgl. 4.2.1 auf Seite 69). So lässt sich die Bandbreite innerhalb eines Raumes steigern, wobei aber zu beachten ist, das jede Wireless LAN Karte nur eine

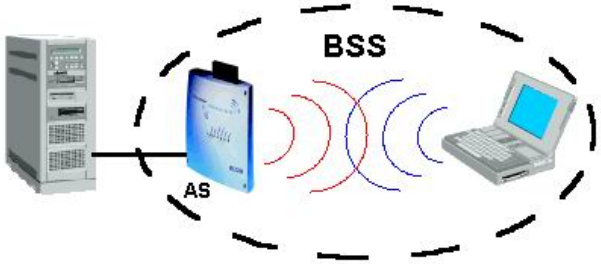


Abbildung 4.20: Ein Basic Service Set unter IEEE 802.11b besteht aus mehreren Stationen und einem Access Point, der als Repeater fungiert [3]

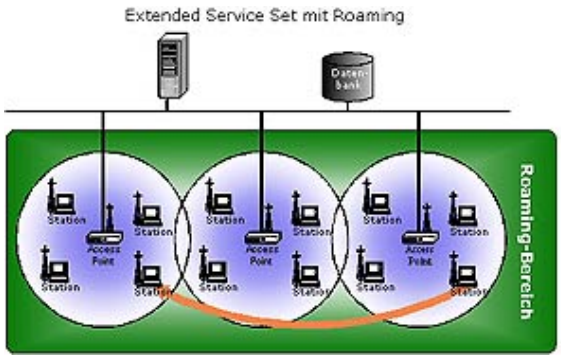


Abbildung 4.21: Mehrere sich überlappende Basic Service Sets bilden zusammen einen Extended Service Set innerhalb dessen Roaming möglich ist [2]

Frequenz gleichzeitig benutzen kann, somit könnte ein Rechner mit mehreren Karten ausgestattet werden und so seine Übertragungsrate steigern, was eigentlich nicht genutzt wird, sondern vielmehr, dass im selben Raum mehrere Team gleichzeitig arbeiten. Jedes dieser Team hätte dann eine eigene Frequenz um mit Teammitgliedern zu kommunizieren, währende die einzelnen Team für sich die volle Bandbreite von 11 MBit/s nutzen können.

Diese Anwendung ist auch sinnvoll, wenn innerhalb eines Betriebes die Lagerverwaltung einen anderen Kanal bekommt, als z.B. das Management. Dieses steigert auch die Sicherheit des Gesamtnetzes.

4.3.3 Praktische Tips für den Aufbau

In dem Artikel **Campus W-LAN Design** [14] von Dave Molta wird beschrieben, wie er einen Campus mit Hilfe der Wireless LAN Technologie so vernetzt hatte, dass alle Studenten Netzzugang bekamen. In seinem veröffentlichtem Erfahrungsbericht schreibt er über die Probleme und Gedanken, die ihm, da er nun erstmalig so etwas machte, begegneten und wie er diese gelöst hatte. Es werden auch praktische Hilfestellungen gegeben.

4.3.4 Praktische Tips für ein Unternehmen

Ein weitere Artikel, auf den hier verwiesen sei, ist im Internet mit dem Titel **Corporate Connectivity** [13]

zu finden. Hier beschreibt der Autor Stefan Ruber sehr ausführlich, welche Gesichts-

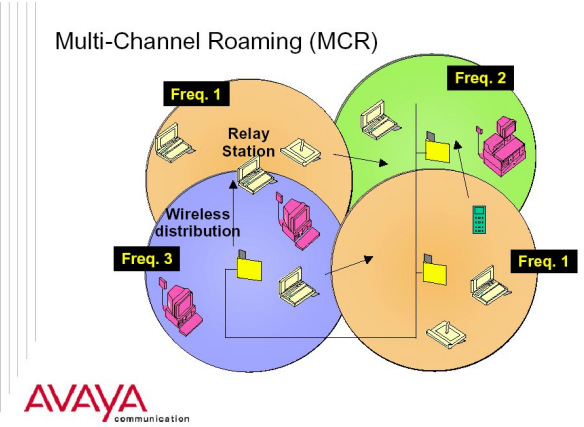


Abbildung 4.22: Office-Umgebung, in der mehrere Kanäle parallel genutzt werden [2]

punkte aus der Sicht eines Unternehmens berücksichtigt werden müssen, um sinnvoll ein Wireless LAN im Unternehmen aufzubauen. Er geht auf Problemquellen ein, liefert Lösungsvorschläge und vermittelt bewährte Verfahren. Jeder, dem die Aufgabe gestellt wird, ein solches Netz in einem Unternehmen aufzubauen, kann hier viele wissenswerte Informationen herausziehen.

4.3.5 W-LAN vs. Bluetooth

Im Wesentlichen sind diese beiden Technologien sehr verwandt. Beide ermöglichen die Verbindung von Endgeräten mit einer Basisstation. Bei Bluetooth wurde das Augenmerk aber auf einen anderen Schwerpunkt gesetzt. Hier sollten vornehmlich PDAs als Endgeräte genutzt werden, die anhand ihrer geringen Energiereserven nur über eine sehr begrenzte Reichweite verfügen. Die Vorteile von Bluetooth sind somit die gleichen, nur dass der Anwendungsbereich sehr eingeschränkt ist. Wireless LAN hingegen lässt sich flexibel einsetzen, schon alleine, weil die Möglichkeiten eines Laptops erheblich größer sind, als die PDAs.

Die verwendeten Frequenzbereiche sind zwar die selben (ebenfalls aus den selben Gründen wie bei Wireless LAN), die beiden Verfahren haben sich aber im Alltag als verträglich erwiesen. Somit macht der parallele Betrieb eines IEEE 802.11 Gerätes zusammen mit einem Bluetooth Gerät im allg. keine Probleme. Die maximale Datenübertragungsrate sinkt natürlich anhand von Kollisionen.

Im Wettbewerb werden sich diese beiden System also nicht direkt in Konkurenz stehen.

4.3.6 W-LAN vs. UMTS, ein Ausblick

Da Wireless LAN ausschließlich als lokale Erweiterung eines drahtgebundenen Netzwerkes gedacht war, ist seine Reichweite in Gebäuden nur ca. 50m. Hierbei werden Übertragungsraten von z.Z. 11 MBit/s (IEEE 802.11b) erzielt von denen nur ca. 50% effektiv zur Verfügung stehen.

„*Versucht man nun die WLAN-Technik mit dem neuen Mobilfunkstandard UMTS (Universal Mobile Telephone Standard) zu vergleichen so muss man als Erstes bemerken, dass sich hier zwei recht unterschiedliche Techniken gegenüberstehen. Während UMTS aus dem*

Standard fur Mobiltelefone GSM hervorgegangen ist, war WLAN von Anfang an nur darauf ausgerichtet, die Kabel aus lokalen Netzwerkinstallationen zu verbannen. Doch dort wo UMTS nun versucht das Internet aufs Mobiltelefon zu bringen, bekommt es Konkurrenz von Wireless LAN , das mobile Endgeräte kurzer Hand in lokale Netze integriert, die ihrerseits ohnehin längst mit dem Internet verbunden sind. Ausgehend von dieser Grundlage sieht es bei einem Vergleich dieser beiden Techniken ziemlich duster aus für UMTS, denn WLAN bietet heute schon einiges mehr als UMTS zu seiner Markteinführung Ende 2003 bieten wird. So werden die Daten bei WLAN mit heute 11 MBit/s wesentlich schneller übertragen als bei den versprochenen 2 MBit/s von UMTS, wobei in der Anfangsphase von diesen 2 MBit/s maximal 384 KBit/s zur Verfügung stehen werden. Wenn UMTS dann endlich Ende 2003 auf den Markt kommt, wird WLAN zudem voraussichtlich 54 MBit/s übertragen können. Selbst in der UMTS-eigensten Domäne, der Übertragung von Sprache bei Telefonaten, hat WLAN inzwischen aufgeholt. Denn seit breitbandige Datenverbindungen Firmennetze miteinander verbinden, wird daran gearbeitet, auch Telefonate über diese zu übertragen, um die Anzahl der verschiedenen Leitungen möglichst gering zu halten. Da Wireless LAN diese Verbindungen ja lediglich durch Funkverbindungen ersetzt, lässt sich auch dieselbe Technik der so genannten IP-Telefonie auf WLAN übertragen. Die ersten Mobiltelefone für WLAN sind mittlerweile verfügbar, und es wird sogar schon von WLAN als Technik der vierten Mobilfunkgeneration (G4) gesprochen. Ein Standard lässt hierbei bisher jedoch auf sich warten.“ [4]

Der einzige Vorteil, der bei UMTS noch angeführt werden kann, wird die bessere Flächenabdeckung sein. Da die UMTS-Sendestationen einen größere Reichweite haben und die Telefongesellschaften langfristig UMTS als Nachfolger des GSM Netzes sehen wollen, wird eine ähnliche Netzdichte zu erwarten sein. Dennoch kann man im Wireless LAN Bereich eine gewisse Eigendynamik erkennen, die mit den Anfängen des Internets zu vergleichen ist.

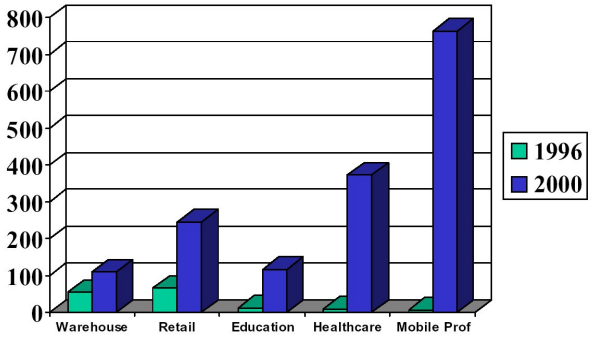


Abbildung 4.23: Anstieg der Verkaufszahlen von Wireless LAN Geräten nach [8]

So haben beispielsweise in Sydney Internet-Begeisterte 170 Access Points installiert, um ein stadtweites Funknetz zu etablieren. Diesem kann für eine Gebühr von ca. 60 EURO jeder beitreten. In Wien hat eine Firma Namens Metronet ein entsprechendes Netz installiert, das man für eine Nutzungsgebühr von monatlich ca. 20 bis 35 EURO zzgl. 30 bis 60 Cent je MB nutzen kann. Weiter hin sein noch ein Beispiel in der Altstadt von Aachen genannt, wo die Firma Accom als Pilotprojekt ein Wireless LAN Netz aufgebaut hat, welches mit 1 MBit/s für 5 Cent zu haben ist.

Auch die Bundesregierung ist mittlerweile auf die Möglichkeiten dieser Netze aufmerk-

sam geworden. So hat sie, beeindruckt von Wireless LAN Netz der *Uni Rostock*, ein Förderprogramm ins Leben gerufen, mit dem 41 deutsche Hochschulen ein solchen Netz aufbauen sollen.

„*Firmen wie Wayport and Aerozone sind darüber hinaus dabei, in ganz Nordamerika Flughäfen, Hotels and Einkaufszentren mit WLAN-Stationen zu vernetzen. Mit den US-Fluglinien Delta and United Airlines hat Aerozone bereits 15-Jahres-Verträge geschlossen. In Europa hat Aerozone gerade den Service auf dem Amsterdamer Flughafen Schiphol eröffnet. London-Heathrow, Frankfurt am Main and Charles de Gaulle in Paris sollen folgen. Bis zum Start der UMTS-Netze will Aerozone viele der an mobilen Internetzugängen interessierten Geschäftsleute längst ins eigene Netz gebracht haben. Ïn UMTS Geld zu investieren, ist schon sehr mutig,ßagte der Münchner Aerozone-Repräsentant Jürgen Vollmer.* “[4]

Über die Verbreitung des Wireless LAN sagen Analysen Voraus, dass im Jahr 2006 eine Nutzerzahl von ca. 20 Millionen erreicht werden wird. Dies wären ungefähr 1000 Mal so viele wie heute. Der künftige Umsatz wird auf 3 Milliarden EURO geschätzt und ungefähr ein Zehntel der Mobilfunkkunden wird Wireless LAN Dienste nutzen. Über die momentane Vertreitungsgeschwindigkeit gibt die Abbildung 4.23 auf der vorherigen Seite einen Eindruck.

Abschließend lässt sich sagen, dass das Aufsteigen von Wireless LAN wohl ganz auf die Kosten von UMTS gehen wird. Zwar werden die UMTS-Anbieter mit allen Mittel versuchen, ihr Produkt an den Mann zu bringen, aber solange es ihnen an einer „Killer-Applikation“ fehlt, werden besonders die Laptop-Besitzer eher zu existierenden Wireless LAN Lösungen greifen, die ihnen wesentlich mehr bringen wird. Ebenso kann davon ausgegangen werden, dass Wireless LAN keine großflächige Verbereitung erfahren wird, sondern durch einzelne Kleinunternehmer an „HotSpots“ wie Flughäfen und in Hotels ihre große Verbreitung erfahren wird. Dadurch wird es wohl bald zu einer selbstverständlichen Dienstleistung, bei der es sich wohl bald niemand mehr leisten kann, sie nicht anzubieten.

4.4 Sicherheit im Wireless LAN

Da Wireless LAN im Gegensatz zu Ethernet auf Funkwellen übertragen wird und diese sehr schwer bis zu überhaupt nicht gegen unbefugten Zugriff sicherbar sind, hat man in der IEEE 802.11 einen Sicherheitsstandard vorgesehen, der eine ähnliche Sicherheit gewährleisten soll, wie es von drahtgebundenen zu erwarten wäre. Dieser Standard ist WEP und steht für **Wired Equivalent Privacy**. Im weiteren soll dieser Mechanismus von mehreren Seiten beleuchtet werden und Schwachstellen hervorgehoben werden, die den Schluss zulassen werden, dass das Ziel mit der momentanen Implementierung und Spezifizierung nicht erreicht wurde.

4.4.1 Aufbau und Funktion von WEP

WEP nutzt einen *RC4 encryption code*, der auch als *stream cipher*. Ein solcher stream cipher arbeitet mit einem kurzen Schlüssel, der durch eine Pseudo-Zufallszahl zu einem längeren Key-stream erweitert wird. Der Sender verknüpft nun den Datenstrom mit Hilfe der XOR-Funktion mit dem Schlüsselstrom und erzeugt so den verschlüsselten Strom. Zum Dekodieren benötigt der Empfänger nun den selben Schlüssel und den gleichen Me-

chanismus zum Erstellen der Zufallszahl. Hat er den Schlüsselstrom erzeugt, kann er den Empfangenen und kodierten Strom wird mir Hilfe der XOR Funktion dekodieren und hat wieder den ursprünglichen Text.

Somit verlangt WEP nach IEEE 802.11 einen geheimen Schlüssel, den alle Stationen haben müssen. Des weiteren wird an jedes Paket eine Prüfsumme *vor der Verschlüsselung* angehängt, was Veränderungen während der Übertragung aufdecken sollen. Der Standard definiert nicht genau, wie dieser Schlüssel erzeugt und verteilt werden soll. In der Praxis wird zumeist ein einzelner Schlüssel verwandt, den sich alle Teilnehmer teilen, was den Angriff vereinfacht.

Diese Arbeitsweise des WEP-Verfahrens macht es gegenüber mehreren Angriffsmöglichkeiten verwundbar, die hier im weiteren erläutert werden. Wenn der Angreifer ein Bit im verschlüsselten Paket verändert, wird an der selben Stelle im Originaltext ebenfalls dieses Bit geändert sein. Weiterhin ist es einem Angreifer möglich, wenn er mindestens zwei verschlüsselte Pakete abfängt, die mit dem selben Schlüssel erstellt wurden, diesen mit Hilfe von statistischen Methoden herauszurechnen.

WEP hat beide Angriffsvarianten bedacht und beinhaltet deswegen entsprechende Vorsichtsmaßnahmen. Um sicherzustellen, das die Nachricht während der Übertragung nicht verändert wurde, fügt WEP einen **Integrity Check** (IC) ein. Um sicherzustellen, das zwei Nachrichten mit dem selben Schlüsselstrom verschlüsselt werden, wird ein sog. **Initialization Vector** (IV) benutzt, der zusammen mit dem kurzen für alle gleichen Schlüssel den Schlüsselstrom erzeugt. Dieser IV ist aber immer verschieden (wenn richtig implementiert) und erzeugt somit auch immer einen unterschiedlichen Schlüssel für jedes Paket. Dieser IV ist ebenfalls im Paket enthalten. *Leider sind aber beide Methoden falsch implementiert worden, was die Sicherheit erheblich reduziert.*

Der IC ist eine CRC-32 Kontrollsummenwert, welcher mit dem Inhalt zusammen verschlüsselt wird. Die CRC-32 Funktion ist aber linear, was es ermöglicht, die Änderung im Inhalt des Datenpaketes zu errechnen und den CRC-Wert entsprechend zu ändern. Somit ist es einem Angreifer möglich, beliebige Bits innerhalb einer Nachricht zu verändern und die Checksumme so zu ändern, dass das Paket als gültig anerkannt wird.

Der IV in einem WEP Paket ist ein 24-bit Feld, welches in Klartextteil versandt wird und den Zweiten Teil des Schlüssels darstellt. Wenn man nun davon ausgeht, dass dieser IV bei jedem neuen Paket verändert wird ergibt sich somit unter den Annahmen:

1. jedes versandt Pakte hat eine Größe von 1500 byte,

2. Übertragungsrate ist konstant bei 11 MBit/s,

dass nach

$$\frac{1.500 * 8}{(11 * 10^6)} * 2^{24} \approx 18.000sec \approx 5h$$

die maximale Anzahl der verschiedenen Schlüssel erreicht ist und diese sich nun wiederholen. Wenn man von der Praxis ausgeht, in der ein Großteil der Pakte kleiner ist als 1.500 byte und somit wird die Zeit sogar eher kürzer sein. Somit ist es möglich, dass eine Angreifer nach relativ kurzer Zeit mindesten zwei Pakete mit der selben IV angreifen kann. Mit diesen kann er mit Hilfe von statistischen Angriffsmethoden den Schlüssel errechnen. Hat er diesen erst einmal, kann er jede andere Nachricht, die den gleichen Schlüssel verwendet, was alle des gleichen Netzes sein sollten.

Noch schlimmer ist die Tatsache, das gerade heutzutage noch viele Angebote auf dem Markt sind, die bei gleichem Hersteller in jeder Installation **den selben geheimen Schlüssel** vorgeben und die meisten zumeist unerfahrenen Admins diesen nicht ändern oder noch schlimmer nicht ändern können.

4.4.2 Angriffsmöglichkeiten

Passiver Angriff um den Datenverkehr zu entschlüsseln Die erste Angriffsmöglichkeit beruht genau auf der eben dargestellen Beobachtung. Hierbei wird der Angreifer alle Datenpakete passive abfangen, indem er sie einfach speichert. Irgendwann wird er mindesten eine Kollision von gleichen IVs haben und kann mit einem Angriff beginnen.

Wenn zwei Pakte mit gleichem Schlüssel XORt werden ergibt das das gleiche Ergebnis, als wenn die Originaltexte miteinander XORt werden. Weiterhin beinhaltet Netzverkehr von sich aus schon eine große Redundanz und ist zumeist vorhersehbar. Diese Redundanz kann auch verwandt werden um den möglichen Lösungsraum einzuschränken. So lässt sich der exakte Inhalt leichter bestimmen.

Wenn ein solcher statistischer Angriff nicht erfolgreich ist, können die Chancen durch weitere Kollisionen erhöht werden. Im Allgemeinen steigt die Erfolgsrate mit der Anzahl der Kollisionen sehr schnell. Sobald ein Text entschlüsselt worden ist, sind es auch alle anderen!

Man kann diesen Angriff auch noch erweitern, indem man dafür sorgt, dass eine bekannte Nachricht über das Internet an den Access Point gesendet wird um von diesem über Wireless LAN ausgestrahlt zu werden. Wenn der an Angreifer nun seine eigene Nachricht verschlüsselt abfängt kann er mit dem ihm ja bekanntem Inhalt die Nachricht entschlüsseln und so den verwendeten Schlüssel herausfinden.

Aktiver Angriff um Pakete zu ändern Die nächste mögliche Angriffsform ist eine direkte Konsequenz des oberen Sachverhaltes. Wenn man davon ausgeht, dass ein Angreifer den exakten Inhalt eines Paketes kennt, kann er sein Wissen nutzen, um eine andere Nachricht zu erstellen, die sicher akzeptiert wird. Hierfür muss er eine neue Nachricht erstellen, den CRC-32 Wert errechnen und bei einer abgefangenen Nachricht die Inhalte entsprechend ändern. Hierbei wird der Tatsache Rechnung getragen, dass $RC(XxorXxorY) = RC(Y)$ ist. Somit wird die veränderte Nachricht auf jeden Fall akzeptiert, auch wenn der Schlüssel nicht bekannt ist. So könnte man z.B. in einer TelNet Umgebung die Befehle abändern.

Aktiver Angriff von beiden Seiten Dieser eben beschriebene Angriff kann noch erweitert werden, um sich das Entschlüsseln der Nachrichten quasi vom Feind machen zu lassen. Hierfür versucht der Angreifen nicht den Inhalt selber zu erraten, sondern nur den Kopf der Nachricht. Dieser Teil ist meistens leicht zu erraten, denn zumeist muss nur die Zieladresse bekannt sein. Mit diesem Wissen kann der Angreifer diese Adresse durch eine ersetzen, die er kontrolliert. Da die meisten Access Points Internetzugang haben (was ja der Grundgedanke war), würde die von ihm abgefangene und veränderte Nachricht vom Access Point entschlüsselt werden und dann *unverschlüsselt* durch das Internet an die Zieladresse befördert. Diese würde schon durch alle Gateways und Routers gehen. Kann er weiterhin noch den Teil des TCP Kopfes entschlüsseln, so könnte der den Port

nachträglich auf 80 stellen, was die Nachricht dann auch durch die meisten Firewalls hindurch zu einem beliebigen Rechner transportiert.

Tabellengestützte Attacken Da die Anzahl möglicher IVs begrenzt ist

$$24bit \Rightarrow 2^{24} \approx 16.777.216,$$

kann ein Angreifer eine Tabelle von IVs samt erkannten Schlüssel erstellen, die über die Zeit mit Hilfe der oben genannten Techniken gefüllt wird. Diese Tabelle wird eine Größe von ca. 15 GB haben, wenn sie komplett ist und es ermöglichen, das sofort alle Pakete entschlüsselt werden.

Dieses ist um so bedeutsamer, wenn man sich nochmal das am Anfang gesagte vor Augen führt, dass sich die Schlüssel im allg. nämlich nicht ändern!

Der Schlüssel unter WEP Nach der ursprünglichen Definition hat der Schlüssel nach IEEE 802.11 eine Länge von 64 bit, von denen aber nur 40 bit real auf den Schlüssel verfallen und die anderen 24 bit der IV ist. Es gibt eine inoffizielle Erweiterung des Schlüssel auf 128 bit wo der Schlüssel dann eine reale Länge von 104 bit hat.

Bei der 64 bit Variante wird zumeist aus einem Passwort, welches der Admin eingibt, vier Schlüssel erstellt. Während der Übertragung wird dann festgelegt, welcher verwendet wird. Hierbei ist auf die gleichen Sicherheitsmerkmale wie bei der Erstellung von Passwörtern zu achten, da sonst das Passwort selber einfacher angreifbar ist. Hierbei werden zumeist zwei Varianten verwandt. Zum einen eine Angriff mit Hilfe eines Wörterbuches und zum anderen „Brutal-Force“.

Benötigtes Material und geschätzte Zeiten Das schlimme für den Angegriffenen und das Angenehmen für den Angreifer ist, dass die benötigte Hard- und Software einfach zu beschaffen ist. Auf der Hardwareseite benötigt der Angreifer nur eine Wireless LAN Karte samt Rechner. Dieser sollte natürlich beruhend auf den rechenintensiven Angriffsalgorithmen möglichst schnell sein. Die Software ist einfach zu beschaffen, es gibt einige Programme, die das Mitspeichern kostenlos und einfach ermöglichen, versierte Programmierer können auch einfach den Treiber so abändern, dass dieser das selber tut.

Das Passwort selber anzugreifen kann sehr Zeitaufwendig sein, denn um einen 40 bit Schlüssel direkt anzugreifen werden ungefähr 210 Tage bei 60.000 Versuchen/sec benötigt. Da der Angriff aber relativ einfach zu parallelisieren ist, kann dieser Angriff mit ca. 100 Rechnern in akzeptabler Zeit durchgeführt werden. Hierbei sei erwähnt, dass es bessere Angriffsmöglichkeiten gibt, z.B. mit Hilfe eines Wörterbuches. 104 bits Schlüssel sind nicht in brauchbarer Zeit entschlüsselbar, denn es würde ca. 10^{19} Jahre benötigen.

Wie schon beschrieben ist es verhältnismäßig einfacher, die Pakete selber anzugreifen. Denn nach schon 4823 Paketen besteht eine 50% Wahrscheinlichkeit auf eine Kollision von IVs. Noch bessere Chancen hat man, wenn man dafür sorgt, das e-mails an einen selbst gesendet werden.

Zusammenfassend lässt sich sagen, das die doch erheblichen Schwächen der Implementierung auf Unwissenheit der kryptographischen Grundlagen zusammen mit Missverständnissen bei der Umsetzung zurückzuführen sind.

4.4.3 Juristische Einordnung auf Grund einer Fallstudie

Gemäß einer Studie der Uni Bonn zum Thema *Tatsächliche und rechtliche Risiken drahtloser Computernetzwerke* [5] wird besonders auf die rechtliche Situation beim Einsatz von Wireless LAN eingegangen und einige Gedanken aufgeworfen, die nicht ohne weiteres von der Hand zu weisen sind; der kryptographische Gesichtspunkt wurde eben besprochen und wird jetzt außer Acht gelassen. Auf diese Studie wird im Folgenden nun eingegangen.

Vorab muss erwähnt werden, dass Angriffe auf Wireless LAN Netze eine enorme Dunkelziffer haben dürften, da ein solcher Angriff zumeist gar nicht festgestellt werden kann. Dieses ist bei drahtgebundenen Netzen anders, da hier der Zugang meist nicht unbemerkt bleibt. In der Computerkriminalität machen Angriffe auf Funknetze nur ca. 1% (offiziell) aus.

Ausschlaggebend für die rechtliche Bewertung ist der

§ 202a StGB, Ausspähen von Daten :

1. Wer unbefugt Daten, die nicht für ihn bestimmt und die gegen unberechtigten Zugang **besonders gesichert** sind, sich oder einem anderen verschafft, wird mit Freiheitsstrafe bis zu drei Jahren oder mit Geldstrafe bestraft.

2. Daten im Sinne des Absatzes 1 sind nur solche, die elektronisch, magnetisch oder sonst nicht unmittelbar wahrnehmbar gespeichert sind oder **übermittelt werden**.

Im Wesentlichen lässt sich zusammenfassen, dass das gezielte Ausspähen von Daten, die sich auf dem Rechner des „Opfers“ befinden voll unter den Schutz des oben genannten Paragraphen fallen. Als besonders geschützt gelten die Daten dann schon, wenn die wie schon beschriebenen Mangel haben Schutzmechanismen nutzen, die Wireless LAN zur Verfügung stellt.

Das Problem liegt aber, gem. [5], bei der sich in der Übertragung befindlichen Daten, welche nicht die Ursprüngliche Datei darstellen, sondern ein für den freien Zugriff extra erzeugten Duplikat. Zum einen kann bei künftigen Anwendern und der momentanen Sachlage davon ausgegangen werden, das mit WEP verschlüsselte Daten nicht ausreichend geschützt sind (bei dogmatischer Auslegung des § 202a). Des weiteren ist es das Wesen der Übertragung, dass alle sie mitbekommen. Daher kann bei bei entsprechender Auslegung des Gesetzes dazu führen, dass das Ausspähen der Daten **nicht strafbar ist**.

Dennoch wird darauf hingewiesen, dass das Richterrecht, also die Auslegung vor Gericht wohl auch diesen Tatbestand unter Strafe stellen wird. Es wird aber dennoch geraten, bei sicherheitsrelevanten Übertragungen gänzlich auf Wireless LAN als Übertragungsmedium zu verzichten.

4.4.4 Ausblick in die Zukunft

Abschließend lässt sich sagen, das die Sicherheitsmechanismen des Wireless LAN nicht ausgereift sind. Dieses wurde aber erkannt und wird voraussichtlich bis Ende 2003 abgestellt sein (vgl. 802.11i). Um heutzutage diese Netze sicher betreiben zu können muss zwingend dafür gesorgt werden, dass entweder das Protokoll reimplementiert wird, wie es einige Anbieter für eigene Lösungen getan haben oder es wird auf alternative Sicherungsmaßnahmen der oberen OSI-Schichten, wie SSL.

4.5 Fazit

Die Vorteile des Wireless LAN sind eindeutig. Dieses lässt sich verhältnismäßig einfach in schon existierende Netzwerke einbinden und ermöglichen diese erweitertes Einsatzspektrum. Somit wird auf einfachste Weise das Problem der Unerreichbarkeit einiger Ortschaften, zumeist mobiler Natur, gelöst. Die so gewonnene Mobilität ist ohne weiteres ein Verkaufsargument, welches das Produkt Wireless LAN für eine kleinere aber zumeist kaufstarke Gruppe von Anwendern und Managern interessant macht, da die Produktivität gesteigert wird oder eine nahezu durchgängige Anbindung im Unternehmen und während der Dienstreise auf Flughäfen, in Bahnhöfen und Hotels ermöglicht wird.

Auch der Aufbau ist verhältnismäßig kostengünstig, wenn auch von der Reichweite und der Roamingeigenschaft eher beschränkt. Wireless LAN Funknetze werden voraussichtlich eine Eigendynamik entwickeln und diese Technologie stark verbreiten. Nicht zuletzt kann auch in der nahen Zukunft mit vernetzten Haushalten gerechnet werden, was sich schon auf den diesjährigen Messen abzeichnete.

Die Sicherheitslücken, die aufgezeigt worden sind, können je nach Einsatzgebiet von unbedeutend bis erheblich reichen. Dennoch kann auch heute schon ein erheblich höherer Sicherheitsstandard realisiert werden, indem man auf Sicherungsmechanismen der höheren OSI-Schichten zurückgreift. Das löst zwar nicht die juristischen Probleme (zumindest in Deutschland), aber diese werden wohl in den nächsten Jahren nachgebessert. Auf der anderen Seite sollten Verfahren wie SSL bis hin zu PGP[10] ein sehr hohes Maß an Sicherheit realisieren lassen.

[10]hier als Beispiele

Literaturverzeichnis

[1] Artikel **Standard für drahtlose Netze**;Axel Sikora; *www.tecchannel.de*

[2] Online Veröffentlichung **Zukunft des Wireless LAN**; von ?; *www.zdnet.de*

[3] Online Veröffentlichung **Wireless LAN - WLAN**; Patrick Schnabel; *www.e-online.de*

[4] Online Veröffentlichung **Wireless LAN vs. UMTS**; Lucas Adler; *www.wswg.org*

[5] Veröffentlichung **Tatsächliche und rechtliche Risiken drahtloser Computernetzwerke**; M. Dornseif, K. Schumann, C. Klein; *www.uni-bonn.de* oder *md.hudora.de/publications/200204-dud-wlan/200204-dud-wlan.pdf*

[6] Aus Vortrag **Mobile Broadband Connection**; Steinkühler GmBH und Co KG, *www.steinkuehler.de*

[7] Aus Vortrag **Avaya Wireless 5 GHz FAQs**; Steinkühler GmBH und Co KG, *www.steinkuehler.de*

[8] Aus Vortrag **WaveLAN eine Einführung**; Steinkühler Netzwerk Systeme, *www.steinkuehler.de*

[9] Wireless LAN association; *www.wlana.com*

[10] Wireless Ethernet Compatibility Alliance; *www.wi-fi.com*

[11] The Internet Engineering Task Force; *www.ietf.org*

[12] Institute of Electrical and Electronics Engineers; *www.ieee.org*

[13] Corporate Connectivity; *www.zdnet.de*

[14] Campus W-LAN Design; *www.networkcomputing.com*

Kapitel 5

The Future of Mobile Services

Stephan Kiene

Inhaltsverzeichnis

5.1 Einleitung

Diese Arbeit beschäftigt sich mit den „Mobile Services“ der Zukunft. Das Hauptaugenmerk liegt hier auf den Konzepten und Ideen der neuen Technologien, dabei aber weniger auf den technischen Aspekten, da es den Rahmen dieser Arbeit verlassen würde. Es soll zum einen geschildert werden, über welche Kapazitäten die neuen Dienste verfügen, was uns in den nächsten Jahren an Neuerungen erwarten wird, aber auch wo die Grenzen dieser Technologien liegen: Da sind zum einen die Endgeräte, die auf kleinstem Raum Hard- und Software unterbringen und zum anderen eine Netzarchitektur, die es dem Nutzer ermöglichen muss die von ihm gewünschte Datenmenge an jedem beliebigen Ort der Welt zur Verfügung zu stellen. Als erstes werden hier die Übertragungsstandards mit ihren Leistungsdaten vorgestellt und dabei auf Schwachstellen hingewiesen. Dann auf das Angebot von „primitiven“Diensten (mit primitiv sind hier die Dienste wie MMS, SMS,

etc. gemeint, da der eigentlich neue Aspekt der 3. Generation in der sinnvollen Verknüpfung verschiedener Übertragungsarten [GPRS, IrDA, Bluetooth] liegt und diese dann zu einem neuen komplexeren Dienst zusammenschliesst) kurz aufzeigen. Danach wird an einem kleinen Beispiel gezeigt, was Mobilität des Nutzers für das momentan zur Verfügung stehende Netz bedeutet. Danach wird die Leistungsfähigkeit der Software unter die Lupe genommen und dort Grenzen aufgezeigt. Nachdem dann die Grundlagen für die Interaktion der einzelnen Dienste und Netzarchitekturen besprochen wurden, werden sinnvolle Konzepte zur Kooperation dieser Dienste vorgestellt, die die Industrie bis jetzt entwickelt und erfolgreich auf den Markt gebracht hat. Als letztes werden die angebotenen Dienste und ihre Qualität für den potentiellen Markt diskutiert. Abschliessend folgt ein Fazit über den Nutzen und die Erweiterbarkeit der sogenannten 3. Generation. Um hier zu einem vernünftigen Ergebnis zu kommen ist es notwendig alle Einzelfaktoren vom Netz über die Hard- und Software bis hin zu den einzelnen Diensten und ihren Anwendungen zu gehen. Vorher werden grundlegende Begriffe definiert. Da es sich um eine relativ neue Technologie handelt, wurde hier nicht mit Fachausdrücken und Anglizismen gespart.

5.2 Mobile Dienste

Das Thema Mobile Dienste wird hier in seine zu diskutierenden Faktoren unterteilt: Netz, Dienste, Roaming, Betriebssystem, Endgeräte und sicher auch die sinnvolle Verknüpfung all dieser Faktoren zu einem Produkt der 3. Generation. Nach einer Vorstellung der Ideen und Daten folgen dann die Schwachstellen dieser Faktoren.

5.2.1 Entwicklung der Übertragungsstandards

Es gibt weltweit 662 Mio. GSM-Benutzer. Das sind etwa 70 Prozent des „total digital Wireless markets“ (Stand Ende Februar 2002). Für die Versteigerung der UMTS-Lizenzen gingen alleine in Deutschland 100 Milliarden DMark über den Tisch. An diesen Zahlen kann man erkennen, wie wichtig dieser Bereich für die Wirtschaft ist und was für Erwartungen an UMTS gestellt werden.

Die Anfänge der mobilen Telefonie

Die Mobile Telekommunikation begann in Deutschland 1958 mit dem damaligen A-Netz der Bundespost und wurde dann 1973 vom B-Netz abgelöst. Diese Netze arbeiten mit einem Frequenzbereich von 150-450 MHz. Die nächste Stufe bildete 1981 das analoge zelluläre System des C-Netzes.

GSM 1 und 2

Zu Beginn der 80er begann man mit der Planung eines Paneuropäischen Mobilfunktnetzes. Daran hauptsächlich beteiligt war die GSM (Groupe Speciale Mobile später Global System for Mobile Communication) eine deutsch-französische Kooperation. Der wesentliche Vorteil an GSM liegt heute an den Roamingmöglichkeiten, da dies in 170 Länder möglich ist. Das daraus resultierende Netz CSD (Circuit Switched Data) ist mit einer Geschwindigkeit von 9,6 kbit/sec der heutzutage etwas veraltete Standard der GSM Datenübertragung [0].

GSM 2,5

Da es nach Expertenmeinung noch bis zu 4 Jahren dauern kann bis UMTS oder vergleichbare Formate (W-LAN) für den breiten Markt nutzbar sind, aber trotzdem die Nachfrage nach schnellerer Datenübertragung besteht, wurden die „alten“ GSM-Netze erweitert zu HSCSD HighSpeed CSD (Empfangen: 4 Kanäle à 14,4kBit/sec = 57,6 kBit/sec; Senden: 1 Kanal mit 13,4KBit/sec). Der Geschwindigkeitsvorteil gegenüber CSD begründet sich demnach hauptsächlich im Prinzip der Kanalbündelung. Desweiteren soll der Kanal der SMS-Übertragung vom GSM Steuerkanal gelöst werden, um diesen zu entlasten. Eine letzte Neuerung dieses Netzes ist das sogenannte EDGE (Enhanced Data Rates for GSM Evolution). Hiermit ist eine verändertes Modulationsverfahren gemeint, dass in Verbindung mit GPRS eine Datenübertragungsgeschwindigkeit von etwa 384 kbits/sec ermöglicht. Dies und beispielsweise Dienste wie CAMEL (Customized Application for Mobile Enhanced Logic, eine Dienst zur unkomplizierten Einbindung von weiteren Diensten des Betreibers) werden zu der sogenannten Generation 2,5 (oder auch GSM Phase 2+) gerechnet [0].

GSM 3

Die 3. Generation dieses GSM-Netzes stützt sich auf UMTS (Universal Mobile Telecommunications System). Dies ist ein paketorientierter Übertragungsstandard. Dieses wird durch die Erweiterung GPRS (General Packet Radio Service - laut Spezifikation: 8 Kanal à 21,4kBit/sec = 171,2kBit/sec) ermöglicht. Über UMTS lassen sich Sprache, Bild und Text-Daten mit einer Geschwindigkeit bis zu zwei Megabit pro Sekunde etwa auf ein Handy übermittelt. UMTS ist bis zu 30 Mal schneller als ISDN und bis zu 200 Mal schneller als heutige WAP-Handys. Die Datenübertragungsgeschwindigkeit ist bei UMTS/GPRS abhängig von der Zelle in der man sich bewegt und von der Geschwindigkeit in der man dieses tut: Makrozelle 144 kBits/sec bei max 500km/h; Mikrozelle 184 kBits/sec bei max 120km/h; Pikozelle 2Mbit bei quasi-stationärem Endgerät. Die alten Protokolle rechneten nach online Zeit ab. Da das dem Kunden nicht sehr entgegenkommt, wurde dies mit GPRS geändert. Hierdurch wird ermöglicht, dass mehreren Nutzer einen Kanal gleichzeitig nutzen oder sogar ein einzelner Nutzer mehrere Kanäle nutzt. Ebenso wird sicher gestellt, dass die Ressourcen nicht ungenutzt bleiben. Auf der anderen Seite wird hier aber auch eines der Hauptprobleme sichtbar: Das System hat jetzt bereits wenig Ressourcenbuffer. Hier wird nun nach Datenvolumen bezahlt. Somit ist ständige Empfangsbereitschaft (Always-On Architektur) nicht mit weiteren Kosten verbunden. Die Verbindung wird aufrecht gehalten, ohne die Datenkanäle zu beanspruchen. Dies mindert die „unnütze“ Ressourcenverschwenudung und erhöht somit den Grad der Ausschöpfung.

5.2.2 Probleme und Leistungsgrenzen der einzelnen Systeme

So kontinuierlich sich die einzelnen Netze auch weiter entwickeln und sich mehr den Ansprüchen des Kunden anpassen sind bei vielen Ideen bereits Schwachstellen festzustellen. Aber anstatt die Kunden darauf aufmerksam zu machen oder ihre Idee zu überdenken, werden die Kunden mit phantastisch klingenden Zahlen geblendet, die die theoretische Kapazität darstellen. Leider wird oft verschwiegen, unter welchen meist obskuren Bedingungen diese nur erreicht werden können (siehe. Beispiel UMTS).

HSCSD

Der Hauptnachteil von HSCSD ist, dass alle Kanäle belegt bleiben solange das HSCSD-Handy önlineïst. Auch dann, wenn der Nutzer am anderen Ende gerade keine Daten überträgt. Zum Beispiel, weil er eine E-Mail oder eine Webseite liest. Diesen Nachteil gleicht GPRS aus, indem es jedem Mobilfunkgerät dynamisch einen Anteil an den insgesamt im Netz vorhandenen Kapazitätsreserven zuteilt. GPRS arbeitet dabei paketorientiert.Es werden keine festen Datenkanäle reserviert [2].

GPRS

Eindeutiger Nachteil von GPRS: Sind keine Kapazitätsreserven vorhanden, weil das Netz schon stark ausgelastet ist oder weil viele Nutzer gleichzeitig Daten übertragen wollen, sinkt die GPRS-Übertragungsrate in Richtung Null. Einer der wesentlichen Nachteile von GPRS liegt darin, dass der Netzbetreiber notgedrungen auch der Internet-Provider ist und somit ist man an dessen Dienste und Tarife gebunden [2].

UMTS

Die Datenrate von 2 MBit/s ist nur der Idealwert. Auch in technischer Hinsicht scheint UMTS nicht alles halten zu können, was die Ausrüster vollmundig versprochen haben. Die maximale Datenrate von 2 MBit/s lässt sich nämlich nur im Idealfall erreichen: Wenn das Netz voll hochgerüstet ist, sich der Nutzer nicht vom Fleck bewegt und am besten noch mit seinem Wunder-Handy allein auf weiter Flur steht. Ein Problem ist, dass die Übertragungsrate mit zunehmender Geschwindigkeit, größerer Entfernung und steigender Netzauslastung stark sinkt. Der schnelle Internetzugang und die Videotelephonie wird auf der Autobahn oder im InterCityExpress wohl vorerst eine Wunschvorstellung bleiben. Bei Tempo 120 auf der Autobahn beträgt die Übertragungsrate nur noch ein Fünftel der maximalen Rate und bei Tempo 300 im ICE nur noch ein Vierzehntel, wobei dort praktisch jedoch kein stabiler Empfang mehr möglich ist. Selbst Telefonie per UMTS wird mit heutiger Technik im ICE kaum möglich sein. Zum einen wegen der hohen Geschwindigkeit und vor allem wegen der unzureichenden Netzabdeckung. Wenn in der Werbeaussage für Ende 2003 eine Funkversorgung für 40 Prozent der Bevölkerung angepriesen wird, muss das noch lange nicht bedeuten, dass auch 40 Prozent der Fläche Deutschlands abgedeckt sind, denn diese 40 Prozent der Bevölkerung bewohnen gerade einmal 7 Prozent der Fläche Deutschlands: Versorgt sind dann nur die dicht besiedelten Ballungsgebiete. Im ländlichen Raum (das heißt 93 Prozent Deutschlands) könnte UMTS noch viele Jahre auf sich warten lassen. Die Lizenzbestimmungen schreiben bis 2003 eine Versorgung von 25 Prozent und bis 2005 eine Versorgung von 50 Prozent der Bevölkerung vor. Deshalb wird erwartet, dass es Multiband-Handys geben wird, die sowohl im UMTS-Netz als auch in den GSM-Netzen arbeiten können. Durch solch einen dualen Betrieb von GSM und UMTS ist es den Mobilfunkanbieter möglich, ihre Netze langsam aufzubauen und den Kunden bereits UMTS-Geräte von Beginn an zur Verfügung zu stellen, auch wenn das Netz noch nicht flächendeckend ausgebaut ist. Die Betreiber werden, um die Investitionen in das neue Netz nicht ausufern zu lassen, ihre UMTS-Netze mit einer weitaus geringeren Datenrate als die theoretisch erreichbaren 2 MBit/s starten. Aber sicherlich ist es nur eine Frage der Zeit, bis das Netz einmal nahezu flächendeckend zur Verfügung steht und auch im Zug oder

auf der Autobahn ein schnellerer Datenanschluss realisiert werden kann. Bis dahin sollten die Erwartungen an UMTS jedoch gedämpft werden [1], [2].

5.2.3 Dienste der zweiten und dritten Generation

Nun folgt ein kleiner Überblick über die Systeme, Netze und Dienste. Er dient hauptsächlich dazu, dem Leser die vielen Abkürzungen etwas mit Inhalt zu füllen, die relevant für das Verständnis der neuen Dienste sind. Das eigentliche Augenmerk liegt hierbei nicht auf den Dienste im einzelnen, sondern - wie später in den Beispielen ersichtlich - auf der sinnvollen Kooperation der einzelnen Dienste und ihre Kooperation zu anderen Medien wie beispielsweise dem Internet, W-LAN, Bluetooth.

GPRS

Mit GPRS(General Packet Radio Service) werden die Daten erstmals paketweise mit einer Geschwindigkeit von bis zu 115 Kilobit pro Sekunde (kbps) übertragen. Die Funkleitung wird jeweils nur dann genutzt, wenn ein Datenpaket verschickt wird. Die übrige Zeit steht sie für andere Dienste, etwa für Telefongespräche zur Verfügung. Die Mobilfunkbetreiber wollen bei GPRS nicht mehr nach Zeit, sondern nach übertragener Datenmenge abrechnen [2].

i-mode

Mobilfunkstandard des grössten japanischen Mobilfunk- und Internetanbieters NTT DoCoMo mit mehr als 30 Millionen Abonnenten. i-mode ist das bislang weltweit einzige Netzwerk, das seinen Kunden schon heute einen kontinuierlichen Zugang zum Internet bietet. Die Technologie ermöglicht e-Mail, den Austausch von Bildern und Videos sowie die Darstellung spezieller Websites auf dem Handy-Display. Die Geschwindigkeit der Datenübertragung erfolgt mit 9600 bps. Dies ist kein sehr hoher Wert wird aber durch die Seiten im sogenannten Leichtgewicht (ca. 1,2k) relativiert. Genaueres an einem Beispiel folgt weiter unten.

MMS

Der „Multimedia Messaging Standard“ermöglicht den Versand von Texten, Melodien und Bildern. Dabei ist die Nachrichtenlänge, Gestaltung und Dateigrösse nicht begrenzt - im Gegensatz zur herkömmlichen SMS. Auch Videosequenzen sind mit MMS möglich. [5]

SMS

Short Message Service. Kurznachrichtendienst bei Mobiltelefonen, bei dem bis zu maximal 160 Zeichen übertragen werden können.

WAP

Das Wireless Application Protokoll (etwa: Protokoll für drahtlose Anwendungen) bringt speziell aufbereitete kurze Texte und einfache Grafiken aus dem Internet auf das Handy-Display. Damit kann der Nutzer beispielsweise seinen Kontostand per Mobiltelefon abfragen oder Kinokarten reservieren [2].

EMS

Enhanced Media Service ist die Weiterentwicklung des SMS-Protokolls, das wie SMS über den Signaling-Kanal der Mobilfunknetze läuft, der ansonsten dazu dient, Lokalisierungs- und Verbindungsdaten auszutauschen. EMS ist ein offener, vom ETSI festgelegter Standard. Der Unterschied zu SMS liegt darin, dass der Text bei EMS formatiert werden kann, und dass EMS einfarbige Grafiken unterstützt, die allerdings ohne Grauabstufungen auskommen müssen - dafür können sie animiert werden. Ausserdem stehen akustische Signale nach dem bereits eingesetzten, von der IrDA festgelegten iMelody-Standard zur Verfügung, auch benutzerdefinierte Geräusche sind möglich. Eine EMS-Nachricht kann also in Absätze unterteilt, rechts-oder linksbündig angeordnet und mit einer fettgeschriebenen Überschrift versehen werden, ausserdem lassen sich einfache Grafiken und Piktogramme integrieren und Worte oder Bilder mit akustischen Hinweis-Tönen verbinden [4].

5.2.4 Die Geräte und Ihre Netzeinbettung

Im vorherigen Abschnitt wurden ein paar technische Eckdaten der Netze und die möglichen Dienste der 2,5. und 3. Generation von GSM Geräten vorgestellt. Dies war hauptlastig im Bereich der Telekommunikation angesiedelt. Der Kern der neuen Leistungsfähigkeit liegt aber in der Kommunikation und Kooperation von Mobilfunkgeräten und anderen technischen Geräten wie Notebooks, Subnotebooks, Organizer, PDAs und anderen ähnlich gearteten Geräten. Die folgenden Beispiele sollen etwas illustrieren wie diese Kooperation aussehen kann und wie sie unseren Alltag erleichtern oder zumindest verändern kann. Um die Kooperation und vor allem Kommunikation dieser einzelnen Systeme gewährleisten können, müssen die Schnittstellen definiert werden und die Betriebssysteme der Endgeräte müssen dafür ausgelegt sein. Ich werde hier kurz auf die Möglichkeiten von Roamingverfahren eingehen, um die Verstrickung und Interdependenzen zwischen den einzelnen Netzen und deren Betreibern zu verdeutlichen. Nachfolgend werden einige Endgeräte und die darauf laufenden Betriebssysteme am Beispiel von Symbian OS vorgestellt. Hierdurch soll die Vielseitigkeit und Offenheit der einzelnen System gezeigt werden.

Die Endgeräte

Zuerst wird eine, den meisten Handys zu grundliegenden, Architektur vorgestellt, um zu verdeutlichen, dass es sich hierbei nicht mehr um simple Elektronik sondern um hoch integrierte Systeme handelt. Auch hier liegt eine der Leistungsgrenzen, da das Maximum an Integration (zu einem bezahlbaren Preis) bei vielen Endgeräten schon nahezu erreicht ist. Die Hardware lässt sich zumeist in drei Hauptschichten einteilen: Den Prozessorkern (hierbei handelt es sich meist um einen ARM Architektur RISC Prozessor der über einen V4 Befehlssatz oder höher verfügt) samt MMU (Memory Managing Unit), den SoC (System-on-Chip: dieser beinhaltet alle für das System relevanten Peripheriegeräte wie Timer, Interrupt-Controller, IO-Controller und andere Ports) und den PCB der alle Operationen mit den externen Schnittstellen regelt: Display, Tastatur, AudioCodec-Memory, etc. Das Betriebssystem des Handys befindet sich in einem ROM, das während der Bootphase meist mittels Flashverfahren in das RAM geladen wird. Die MMU verfügt über Second-Level-Cache in einer Seitengrösse von etwa 4KB und First-Level-Cache in einer Seitengrössen von 1MB. Näher wird hier nicht auf die konkrete Hardware eingegangen, da es sonst zu

technisch würde. Stattdessen werden hier ein paar der Features der neuen Geräte präsentiert um den Sinn oder Unsinn der oben genannten Dienste zu veranschaulichen. Da sich die momentanen Top-Geräte der einzelnen Hersteller nicht sehr stark vom Leistungsumfang unterscheiden, werden hier stellvertretend für diese Geräte Klasse die Möglichkeiten des Ericsson R380 vorstellen. Noch in diesem Jahr werden (hauptsächlich in Japan) diverse java-fähige Endgeräte auf dem Markt eingeführt. Alle diese Geräte mit ihren Features und Werten hier zu anzusprechen würde den Rahmen dieser Arbeit verlassen. Bei Interesse können alle diese Neuerungen auf: http://www.javamobiles.com nachgelesen werden. Neben den heutzutage schon üblichen Features wie Kontakte, Spiele, Organizer Telefonbuch, SMS-Nachrichten, Start-up/Shutdown Animationen, Kurzwahl und den üblichen Anrufoptionen wie wir sie von ISDN her kennen kommen jetzt noch Möglichkeiten wie Handschriftenerkennung, Notepad, Bildschirmschoner, Sprachgesteuerte Anrufannahme, Sprachsteuerung, Weltuhr, SMS Cell Broadcast, Synchronisation mit dem PC, Unified Messaging , E-Mail-SMS, Fax über SMS und WTLS-Sicherheitsklasse 1 und 2. Die Eingabe geschieht über Touch-Screen oder On-Screen Tastatur. Der Nokia Communicator bietet sogar eine QWERTY-Tastatur, die man findet wenn man das Handy aufklappt und den Communicator nutzen möchte. Das Ericsson hat die Abmessungen 130x50x26mm und ein Gewicht von 164g. Das Nokia 7650 wiegt knapp 10g weniger wartet aber zusätzlich mit einer Digitalkamera auf und bietet eine grafische Auflösung von 640x480 Bildpunkten (VGA) in Farbe oder ein Standarddisplay: 176x208 Bildpunkte ebenfalls in Farbe untergebracht auf 35x41mm.
Diese Angaben sollten noch einmal deutlich machen, wie hochkomplex diese Geräte sind und welche Erweiterungen nötig sind, um die oben aufgeführten Dienste auch nutzen zu können.

Roaming

Zum Thema Kooperation und Kommunikation soll kurz auf die drei Möglichkeiten des Roaming hinwiesen werden, um die Verflechtung der einzelnen Netze zu schildern. Unter Roaming kann man das Wechseln des eigentlichen Betreibernetzes (Providing Network) in ein anderes Netz (Foreign Network) verstehen. Das Wechseln einer der oben angesprochenen Funkzellen nennt man hingegen Handover. Das Netz in dem man sich zu einem Zeitpunkt befindet, nennt man Serving Network. Es kommt öfter dazu, dass gewisse Dienste vom Serving Network (insofern es nicht das Providing Network ist) nicht erbracht werden können (beispielsweise die Erreichbarkeit in anderen Ländern). Nun gibt es mehrere Wege der Dienstbereitstellung. Einmal die Möglichkeit des Pipelinenings - hier werden die Daten im Prinzip durch das Serving Network zum Providing Network durchgetunnelt. Desweiteren besteht die Möglichkeit der Einrichtung eines Proxy-Servers der Requests und ankommende Daten zwischen dem Benutzer und dem Anbieter vermittelt. Die letzte und wohl auch komplizierteste Möglichkeit ist das direkte Versenden des Dienstes (vgl. Abbildung). Jeder Dienst verfügt über mobile und immobile Bestandteile. Die Mobilen können durch das Netz transferiert werden und so dem Serving Network zur Verfügung gestellt werden, während die immobilen über die ersten zwei Arten nutzbar sind. Dies ist der schematische Aufbau eines solchen Dienstes. Die Service-Components können sehr viel seitig sein und ein Dienst kann auch über eine grössere Anzahl von ihnen verfügen. Der genau Aufbau variiert zwischen den Diensten sehr stark. Die einzigen zwei Komponenten die aber für den Versand des Dienstes zwingend nötig sind, ist zum einen die Service

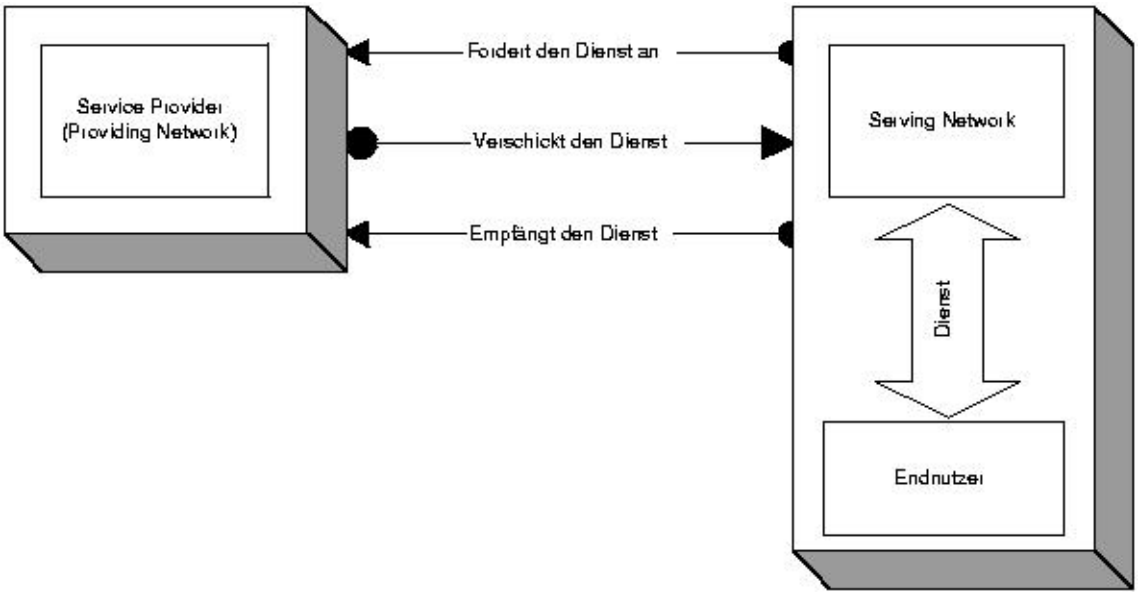


Abbildung 5.1: *Protokoll zur Übermittlung*

Factory und zum anderen der Service Handle. Die Factory hat den Auftrag die einzelnen Komponenten zu erzeugen und dann zusammenzusetzen. Die Aufgabe des Service Handles liegt in der Kommunikation und Koordinierung der einzelnen Komponenten untereinander und auch mit der Aussenwelt -sprich dem Benutzer. Die Service UI bildet mit dem Frontend die Schnittstelle mit dem Terminal, von dem der Dienst aufgerufen wird, und mit dem Backend die Verbindung zum Server. Damit ist klar, dass sowohl die Komponenten und die UI zu den mobilen Teilen des Service gehören müssen. Die Aufgabe des Mobility Controllers liegt darin, die fertiggestellten und zusammengesetzten Komponenten dem Serving Network zukommen zu lassen. Da für einige Dienste grössere Datenbanken (eines der Beispiele für Immobile Service-Anteile) oder Informationen von anderen Diensten benötigt werden, benötigt der Dienst den ISCA (Inter-Service-Communication-Agent).
Hieran soll deutlich werden, wie stark auf diesem Markt die Kooperation und Offenheit von Netzen und Systemen ist. Dies ist auch nötig, wenn man bedenkt, dass es weltweit 170 Länder gibt, die man mit GSM Geräten ohne Schwierigkeiten erreichen kann. Hier liegt aber auch ein Problem begründet: Neuerungen brauchen Zeit bis sie sich durchsetzen können, da zum einen Standards festgelegt werden müssen (und das zwischen relativ vielen Partnern) und die Infrastrutkur muss in einem grossen Raum dafür ausgelegt werden [13].

Betriebssysteme: am Beispiel Symbian OS

Neben den oben beschriebenen Leistungsbegrenzungen durch das Netzwerk oder besser gesagt durch die knappen Buffer, die die Netzarchitektur offen lässt ist ein weiterer begrenzender Faktor das Endgerät und der mangelnde Platz für Hard- und Software. Das bedeutet, dass das Endgerät und somit die dem Nutzer zur Verfügung stehenden Dienste auch durch die Software definiert werden. Ich möchte hier einen kleinen Einblick in die Leistungsfähigkeit eines solchen Betriebssystem geben, dass auf den neuesten Handys der Klassen: Nokia 9210 Communicator, Ericsson R380 Smartphone und Nokia 7650 zur An-

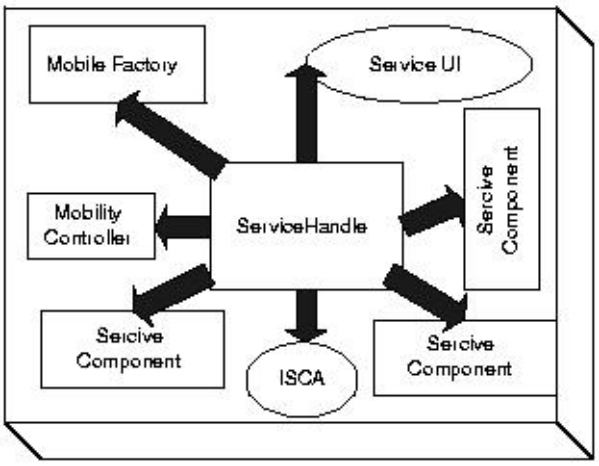


Abbildung 5.2: *Aufbau eines mobilen Dienstes*

wendung kommt: Symbian OS.

Symbian OS ist ein robustes Multi-Tasking Betriebssystem, ausgerichtet für drahtlose Umgebungen und den problemlosen Einsatz von Mobiltelefonen. Das problemlose Ablaufen der einzelnen Prozesse wird durch eine virtuelle Maschine (VM) gewährleistet. Ein weiterer Vorteil dieser VM ist das erleichterte Bearbeiten oder Bug-Fixen des Codes, da die Umgebung klar definiert ist. Durch die VM ist es auch möglich, dass mehrere Applikationen auf dem gleichen Code arbeiten, was wiederum Speicher spart. Die VM wird ermöglicht durch die MMU, die die Daten im virtuellen Adressraum hin- und herschieben kann. Eine Anwendung besteht aus einem einzelnen Prozess. In diesem Prozess wird dafür gesorgt, dass die Daten geschützt bleiben und die einzelnen Threads ungestört laufen können. Wird eine Applikation durch eine polymorphe DLL gestartet werden ihr Speicherseiten zugeordnet, während die Daten über die Threads in den Second-Level gelegt werden. Wird der Thread gewechselt, weist der Kernel die MMU an, die Thread Daten an einen vordefinierten Ort abzulegen. Die Ausführung des nächstens Thread wird eingeleitet. Somit besteht kein Unterschied zu einem normalen Single User Multi-Tasking Betriebssystem.

Desweiteren ist es bereits darauf ausgelegt, mit einem begrenzten Speicher zu operieren. Symbian OS ist IP-basierend und für die Datenkommunikation ausgelegt. Sämtliche Industriestandards werden eingehalten, was ein problemloses Arbeiten ermöglichen soll. In Version 6 sind dies folgende:

Netzwerk: (TCP/IP, PPP, TSL, SSL, IPSec, FTP)
Schnittstellen: (Bluetooth, IrDA, Obex)
Sicherheit: (DES, RSA, DSA, DH)
Nachrichtenprogramme: (POP3, IMAP4, SMTP, SMS, BIO)
Browser: (HTML, HTTPS, WAP, WML)
Telefonie: (GSM, GPRS, fax)
Multimedia: (WAV, AU, WVE, JPEG, BMP, MBM, GIF)

Auch hier kann man an der offen Plattform diese Systems erkennen, dass es für das Integrieren von Diensten Dritter offensteht. Bei Software Dritter wird eine Datei angelegt, die angibt welche Bibliotheken von dieser Software benötigt werden. Dies wird dann über ein Zertifizierungssystem geprüft. Dies soll sicherstellen, dass die Software problemlos

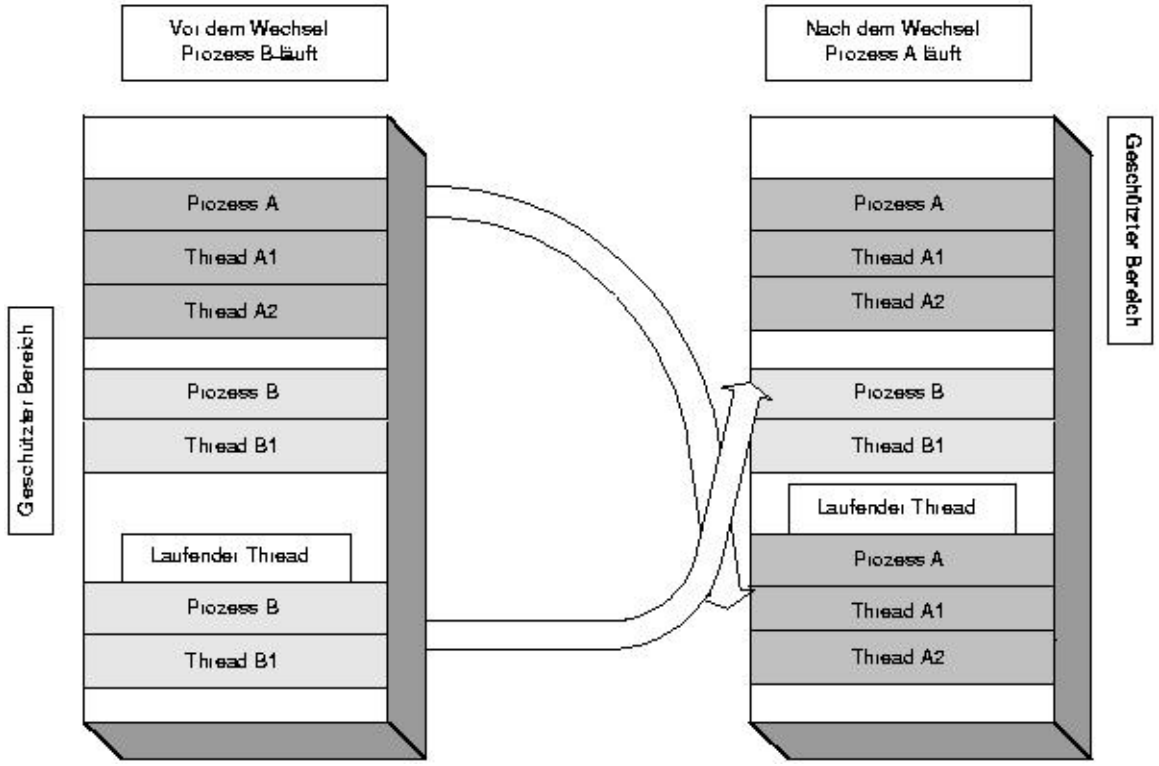


Abbildung 5.3: *Ablauf des Prozesswechsel*

mit den bereits vorhandenen Komponenten kooperiert. Eine der Schwachstellen ist bei jeder Software sicher immer der geringe Speicher, der zur Verfügung steht. Um es den Programmieren hier zu erleichtern, liefert Symbian mehrere Frameworks für geringe Speichernutzung - beispielsweise das client-server-Framework für Auslagerung von Ressourcen - und ein gutes Stack-Managment mit. Weitere Software kann mit C++ geschrieben werden und dann mittels Windows und einem Tool auf die Systeme des Handys übertragen werden. Symbian liefert eine virtuelle Maschine für Java mit, die es ermöglicht Java auf dem Handy laufen zu lassen. Ein weiteres Verfahren um Platz zu sparen stellen die polymorphen DLLs dar. Sie ähneln dabei kleinen Fabriken, die Instanzen von Programmen, auf die Schnittstellen anpassen. Durch diesen skelettartigen Aufbau durch die Frameworks bleibt es jedem Hersteller sehr weit offen wie er sein Gerät nachher konkret gestalten will.

5.2.5 Beispiel-Anwendungen

Wie bereits angeführt, liegt der grosse Nutzen von GPRS/UMTS und den hochmobilen Endgeräten nicht nur in der ortungebungebundenen Einsatzfähigkeit und der ortgebundenen Dienstleistung von LBS (Location-based-Services, der wohl bekannteste, aber unscheinbarste LBS ist wohl der von Genion angebotene Homezone-Tarif: hier wird die Position des Nutzers mit seiner Homezone abgeglichen und somit kontrolliert ob er sich im günstigen Tarifgebiet befindet oder nicht), sondern in der Kooperationen all dieser Dienste. Was diese Dienste im einzelnen sein können, möchte ich kurz am Beispiel von IMOCOS (http://imocos.com) aufzeigen. IMOCOS bietet eine offene Software-Plattform an, Diese Offenheit ist, wie sicher schon ersichtlich geworden, keine Besonderheit, sondern eine

Notwendigkeit, um die so oft angesprochene Kooperation, Individualität oder das nahezu weltweite Roaming zu gewährleisten. Diese Plattform enthält folgende Komponenten: Eine Mobile Search Engine, einen Online Browser, einen Mobile Homepage Editor, ein Online Bookmark System, ein Technical Center, Mobile Optimzed Content und das eben genannte LBS. Hinter dem Optimized Content verbirgt sich ein Standard den IMOCOS zusammen mit dem Fraunhofer Institut entwickelt hat. Es handelt sich hierbei um den MOC-Standard (Mobile Optimized Content Standard). Dieser soll es ermöglichen, dass die Unmenge von Daten und Diensten, die derzeit angeboten wird, auf den einzelnen Geräten, die diese nutzen wollen, nicht zu unnötigen Fehler oder gar Abstürzen führt. Denn viel Anbieter achten bei ihren Diensten nicht auf die Eignung dieser Dienste bei den einzelnen Endgeräten. Faktoren wie Speicher, Darstellungsmöglichkeit, Browserkapazität oder die Eignung von Benutzerschnittstellen werden aber auch oft von den Herstellen der Mobilfunkgerät-Hersteller nicht den auf dem Markt erhältlichen Diensten angepasst. So ist auch nachzuvollziehen, dass es momentan (laut Angaben IMOCOS) zu einer Fehlerrate von etwa 50 Prozent bei WAP basierenden Inhalten kommt. Der MOC-Standard soll zum einen eine Art Kompatibilitätstest anbieten und zum anderen den Dienstanbietern ein Portal zur Verfügung stellen, dass die Kompatibilität sicherstellen soll. Wie man an den oben an den Bestandteil der Plattform bereits erkennen konnte, weist beispielsweise der mobile Homepage Editor, auf einen hochmobilen Benutzer hin. Die Leistungsfähigkeit einer solchen Plattform lässt also in etwa eine grobe Analyse der potentiellen Nutzer solcher Systeme zu. Hierauf werde ich in einem späteren Absatz noch genauer eingehen.

Ein Blick in die Zukunft

Der Terminplaner wurde gestern abend via Computer auf den aktuellen Stand gebracht. Persönliche Notizen wurden automatisch mit den Einträgen verknüpft. Das Handy sorgt via GPS und Bluetooth für korrekte Lokalisierung. Die Uhrzeit wird Synchronisiert. Laut Handy noch 10 Minuten bis der Zug planmässig in München eintrifft. Das Servicenetz der Bahn zeigt keine Verspätung an. Das Wetter ist laut PDA heiter bis wolkig. Eine SMS trifft ein. Wollen wir uns nachher im Englischen Garten treffen? Mmmh, wie komme ich dahin? Ein kurzer Crosslink zum PDA und eine detaillierte Karte zeigt den kürzesten Weg zwischen der aktuellen Position und dem gewünschten Ziel. Kaum am Hauptbahnhof angekommen, meldet sich das Handy: Kino Premiere eines neuen Films, der dem Profil aus dem Internet entspricht. Sollen Karten bestellt werden? Ein letzter Blick auf den Terminplan: Ok Einkaufen, Englischer Garten, Essen, Deutsches Museum, BMW-Museum,... wie kriege ich das am besten unter einen Hut? Schnell wird der digitale Einkaufszettel geprüft und die nötigen Lokalitäten bestimmt. Ein Klick und das Handy schlägt die optimales Route vor. Mmmh tja alles etwas unübersichtlich, vielleicht doch lieber Darstellung via PDA. OK, bei den Entfernungen vielleicht doch einen Leihwagen. Schnell ist der nächste Anbieter ausfindig gemacht. Kurz durch das verfügbare Angebot gestöbert: Ah genau den nehme ich. Wieder schlägt das Handy in Kooperation mit dem PDA den kürzesten Weg vor. Bezahlt wird mit dem Handy, via Bankverbindung, da die EC Karte zuhause liegengeblieben ist. Weiter wird mit dem PDA und dem Handy als GPS-Peiler durch die Stadt navigiert. Das Handy schlägt Alarm: Vollsperrung des nördlichen Rings, der PDA korrigiert den Kurs und stellt den Zeitverlust fest, um den Terminplan um zu stellen, denn noch weiter können wir das mit dem Mittag sicher nicht hinauszögern. Unterwegs ein paar Bilder mit der Digitalkamera gemacht und dann mit Infrarot übers Handy und

als MMS als digitale Postkarte an die Freunde verschickt... So oder so ähnlich könnten uns UMTS, Bluetooth, IrDA, W-Lan-Geräte in der Zukunft unseren Alltag erleichtern. Zwar ist diese starke Verknüpfung der einzelnen Dienste und Endgeräte noch nicht Stand der Dinge, aber bereits möglich. Ein paar Beispiele, die bereits Realität sind folgen im Weiteren.

M-Commerce

Bald wird auch das Bezahlen via Handy möglich sein, denn HP, Lucent, Sun, Oracle und Siemens haben die Gesellschaft Paycircle gegründet. Das Ziel dieser Unternehmung liegt darin es den Kunden ohne die Installation weiterer Software auf den verschiedensten Zahlungssystemen einfach via Handy zu bezahlen. Wieder einmal handelt es sich hierbei um eine offene Schnittstelle. Somit können Anwendungen unabhängig vom Endgerät und vom Zahlungssystem einfach auf diese Schnittstelle aufgesattelt werden [12]

Beispiel: i-mode

I-Mode wurde im Februar 1999 in Japan eingeführt und zählt mittlerweile zu den weltgrössten Anbietern von Diensten wie mobilen Internetzugängen. Die Seiten von i-mode lassen sich in zwei Kategorien aufteilen: Die offiziellen und die inoffiziellen - letztere lassen sich nicht direkt über das Menü ansteuern. Im Juni 200 wurden etwa 12.000 Seiten für i-mode angeboten; davon 500 offizielle. Von den inoffiziellen sind einige umsonst, manche andere müssen aber monatlich bezahlt werden (100-300 Yen). Ein Grossteil all dieser Seiten ist jedoch auf japanisch verfasst. i-mode arbeitet auch mit GPRS, was bedeutet, dass man nach Datenvolumen und nicht nach Onlinezeit bezahlt. Neben den oben schon kurz angeführten Möglichkeiten wie Buchen von Flügen oder Konzertkarten, Banktransaktionen, E-Mail und News sind auch Ansätze von LBS Diensten vorhanden: ortsbezogene Stadtkarten oder Fahrpläne für Züge oder ähnliches [10].

Beispiel: City-Companion/DCSMA

Zur diesjährigen Cebit beabsichtigte die Daimler Chrysler Services Mobile Application GmbH (DCSMA) ihren Informations- und Navigationsdienst City-Companion auf fünf weitere Städte auszuweiten: Hamburg, München, Frankfurt am Main, Köln und Dresden. Zuvor war dieser Dienst nur für Berlin ausgelegt. „Nach der Planungsphase und nachdem sich unser System seit der IFA in Berlin bestens bewährt hat, gehen wir jetzt mit Hochdruck daran, City-Companion für Städte in ganz Europa fit zu machen“, so Dr. Afsabeh Haddadi, Projektleiterin für den City Companion. Mit diesem Dienst können Termine im Internet geplant werden und unterwegs kann der Nutzer dann diese Informationen mit einem mobilen WAP-fähigen Endgerät abrufen. Der Benutzer ist dann unterwegs in der Lage, mit einem Klick direkt über seinen Organizer beispielsweise ein geeignetes Parkhaus, die optimale Route oder einen anderen gesuchten Ort zu finden oder Anbindungen mit anderen Verkehrsmittel abzufragen. Der City-Companion-Service bietet Benutzer auch zu andere nicht verkehrsbezogene Informationen zur unmittelbaren Umgebung wie Restaurants und Hotels. Für die Zukunft sind auch andere Orte wie beispielsweise Apotheken, Tankstellen oder ein Geldautomat hiermit zu finden. Das Leistungsspektrum dieses Dienstes gliedert sich in drei Teilbereiche auf: City Highlights, City Navigation, City Services. Die City Highlights bieten aktuelle Infos zum städtischen Geschehen: Durch Wählen aus

den Kategorien „Kino, Kultur, Hotels, Nightlife, Essen und Trinken, Sehenswürdigkeiten, Sport/ Wellness und Events“hat man die Möglichkeit, je nach Vorliebe weiter zu spezifizieren. So lassen sich z. B. ganz einfach ein Chinarestaurant oder ein Fitness-Center in der Nähe finden.
City Navigation zeigt den Weg durch die Metropole: Startpunkt und gewünschtes Ziel werden eingegeben und der Dienst liefert den Anfahrtsweg und den entsprechenden Kartenausschnitt. Außerdem kann man Parkmöglichkeiten abrufen und jede beliebige Adresse innerhalb der Stadt kann auf einer digitalen Karte angezeigt werden.
Im Bereich City Services kann man medizinische Dienste wie z.B. Krankenhäuser, Services rund ums Auto, wie z.B. Pannendienste oder Nützliches wie beispielsweise die zentrale Zugauskunft finden. Mittels cross-funktionalen Methoden kann man die Suche nach einem Parkplatz an einem dieser Orte erleichtern.
Neben diesen ortsbezogenen Diensten ist das Gerät auch noch mehr als nur ein persönlicher Tagesplaner. Terminpläne die im Internet erstellt worden, können via WAP eingelesen werden. Dank der oben angeführten Fähigkeit der Cross-Funktionen lassen sich die Daten so sinnvoll verknüpfen. Beispielsweise kann man sich auch an jeden Termin per SMS erinnern lassen. Ebenso kann man die Adresse der eingeplanten „Highlights“ auch an Freunde oder Bekannte weiterschicken. Leider lassen sich Termine nur über WAP einlesen nicht aber ändern [11].

Beispiel: Viag Interkom

Als letztes hier ein Beispiel, dass nicht zu den High-End-Geräten greift, sondern auch dem Kunden mit seinem „stinknormalen“ Handy die Vorteile von LBS zur Verfügung stellt. Unter der Nummer 3463 können Viag Interkomkunden komplett ohne WAP oder ähnliches über eine einfache SMS mit dem gewünschten Suchbegriff anfragen und bekommen via SMS dann die nächstgelegene Adresse zurück. Gefällt diese Auswahl nicht, können weitere angefordert werden.

Schwachstellen von LBS

Location-based-Services machen nur dann Sinn, wenn der angebotene Dienst und der momentane Standort zusammenpassen. Dies ist leider nicht immer der Fall. Zwar steht den Betreibern die Information über die momentane Funkzelle zur Verfügung und durch Anpeilen über mehrere Antennen erhöht sich die Genauigkeit, kann in der Stadt (Mikrozellen) aber nur eine Genauigkeit von 100-500m erzielt werden. Das mag für einen Grossteil von LBS ausreichen, nur stellt man sich jetzt beispielsweise einen LBS vor, der Informationen über Kultur und Geschichte von Bauwerken zur Verfügung stellt, dann ist man mit 100-500m zum Beispiel in Rom nachher genauso klug wie vorher. Da die Makrozellen auf dem Lande noch grössere Dimensionen annehmen, kann dort teilweise eine exakte Routenplanung oder das Herbeirufen von Rettungskräften nicht mehr als sicher angenommen werden (hier können Abweichungen von bis zu 10km auftreten!). Um solche Dienste trotzdem gewährleisten zu können, müssen Laufzeiten ausgewertet werden oder gar GPS-Dienste herangezogen werden. Um dies alles leisten zu können, steigen aber auch die Kosten. Zum einen muss die Netzarchitektur verändert werden (wobei die Kosten irgendwie wieder auf die Kunden zurückfallen) und zum anderen müssen die Endgeräte ebenfalls dafür ausgelegt sein. Desweiteren zieht das eine stärkere Stromversorgung mit

sich, die dadurch auch ein Grösser- und Schwererwerden des Handys mit sich bringt. Eine andere Möglichkeit wäre noch eine zusätzliche Positionsbestimmung über Bluetooth , etwa über die Koordinaten eines Access-Points, mit dem das zu ortende Bluetooth-Gerät eines Nutzers kommuniziert. Aber auch hier ist die Infrastruktur noch nicht ausreichend. Selbst wenn man das Problem der exakten Positionierung überwinden würde, würde sich hier ein neues Problem ergeben. Nicht nur man selber könnte jeder Zeit seinen Standort bestimmen, sondern auch die einzelnen Betreibergesellschaften wüssten darüber bescheid. Da das Thema Datensicherheit auch noch nicht (und wahrscheinlich nie) geklärt ist, hätte man eine ständige Überwachung zu befürchten.

5.2.6 Potentielle Nutzer

Um die Grösse des Marktest in etwa abschätzen zu können, kann man folgende Prognose als Richtwert betrachten: Ende 2004 wird es mehr als eine Milliarde mobile internetfähige Geräte weltweit geben [9].

Der deutsche Markt

Stellvertretend für Europa soll hier der deutsche Markt etwas genauer betrachtet werden. Der UMTS-Hype ist zwar in Deutschland vorbei, aber das Marktpotenzial bleibt. So lautet das Ergebnis einer Studie von Mercer Management Consulting. Es wird davon ausgegangen, dass die Zahl der Handynutzer um 19 Prozent oder neun Millionen auf 57 Millionen steigen wird. Der Umsatz der Mobilfunkbranche wird laut dieser Studie auch steigen: Ein Plus von fast 21 Prozent. Dass lässt die Erwartungen steigen. Aber es wird noch Jahre dauern bis UMTS mit seiner vollen Leistung an den Markt geht. Nach wie vor aber verspricht der Breitbandmobilfunk langfristig ein enormes Potenzial mit nachhaltigen Auswirkungen auf Wirtschafts- und Sozialleben, so die Studie. Zum ersten Mal übertrafen die Mobilfunkverträge in Deutschland die Festnetzanschlüsse. Wann in Deutschland UMTS sich etabliert haben wird, steht auch nach Expertenmeinung noch nicht fest: Erst hiess es 2003/2004. Dieses wurde anschließend schnell auf 2005/2006 korrigiert. Heute spricht die Branche vom Jahr 2008. Auch im Jahr 2005 gilt Sprache mit 60 Prozent Anteil als Umsatzbringer Nummer eins. Deshalb werden unter Druck stehende Netzbetreiber versucht sein, sich durch massive Preissenkungen und explizite Annäherung an Festnetzpreise ein Stück vom 25-Milliarden-Euro-Kuchen des Festnetzsprachverkehrs zu sichern. Nur ein rascher UMTS-Markterfolg könne verhindern, dass übermäßig Sprachvolumen in mobile Netze abwandert und die Festnetzbetreiber unter Zugzwang geraten. „Etabliert sich UMTS erst sehr spät, könnte das Festnetz einer der größten Verlierer der dritten Mobilfunkgeneration sein", betont Gauer von Mercer Management Consulting. Ein Schwerpunkt der neuen mobilen Portale werden die Inhalte sein. Hiermit versuchen dann die Anbieter ihre Kunden zu binden. Durch nur mangelhafte Verfügbarkeit der Endgeräte wird noch das Wachstum des Marktes gehemmt. Mercer Managment Consulting geht davon aus, dass UMTS ein eine Vielzahl von unterschiedlichen Endgerätetypen hervorbringen wird , deren Hersteller nicht nur aus dem Mobilfunkmarkt, sondern auch dem PC-, Spielkonsolen- und Unterhaltungselektronikmarkt kommen werden. Gleichwohl werde die Polarisierung des Endgerätemarktes in Lowtech- und Hightech-Endgeräte weiter zunehmen. „Bei UMTS wird sich der Kunde daran gewöhnen, mehrere mobile Endgeräte zu nutzen - je nach Kommunikationssituation", konstatiert Gauer.

Schenkt man diesen Prognosen Glauben, dann kommt man zu dem Ergebnis, dass UMTS - wenn es einmal den Markt erobert hat; ähnlich wie heute bereits das Handy - in das soziale Leben integrieren und überall Anwendung finden. Es wird neben dem Wirtschaftssektor auch Einzug in die Unterhaltungswelt finden. Da UMTS bisweilen noch nicht im grossen Rahmen nutzbar ist, lässt sich nur schwer etwas über den vermeintlichen Benutzerkreis sagen. Aufgrund des hohen Preises und den momentan angebotenen Diensten liegt der Schwerpunkt der Benutzer aber noch auf Seiten von Handelsreisenden oder Geschäftsleuten, die viele Termine unter einen Hut zu bringen haben. Dies wird sich sicher aber in den nächsten Jahren auf dem deutschen Markt ändern [8].

Der japanische Markt

Wie oben bereits im Abschnitt über i-mode angeführt, ist dieses Produkt von NTT DoCoMo in Japan das meistgenutzte.

Platz	Anbieter	Nutzer in Mio.
1.	i-mode	32.63
2.	EZweb	9.97
3.	J-Sky	10.35
4.	DDI Pocket	2.41
Gesamt:		55.36

Daher sind die Nutzerprofile sicher auf den gesamten Mobile-Service-Sektor übertragbar. Der Kern der i-mode Nutzer in Japan liegt bei den jüngeren Kunden zwischen 24 und 35 Jahren. Die Benutzer, die die Dienste am stärksten in Anspruch nehmen sind Frauen in den späten 20ern. Im November 2000. Wurde die Benutzerzahl von i-mode auf 14.9 million geschätzt. Wenn man darauf achtet, dass zu diesem Zeitpunkt i-mode erst 1 1/2 Jahre auf dem Markt war, muss man von einer beachtlichen Zahl sprechen. Da i-mode in Japan der vorherrschende Standard ist, hier ein kleiner Vergleich der beliebtesten Endgeräte (vgl. Abschnitt Endgeräte für UMTS, diese sind für den europäischen Markt aussagekräftig aber nicht für den i-mode Einsatz).

Platz	Gerät	Hersteller	Höhe	Breite	Dicke	Gewicht	Gesprächszeit	Standy-By zeit
1	P208	Panasonic	ähnlich 209	ähnlich 209	ähnlich 209	k.A	k.A	k.A
2	P209iS	Panasonic	92mm	47mm	25mm	84g	135 min	380 Stunden
3	P209i	Panasonic	123mm	39mm	15mm	60g	135min	350Stunden
4	N502i	NEC	93mm	48mm	22mm	98g	120min	420Stunden
5	P502i	Panasonic	130mm	43mm	16mm	69g	125min	300Stunden
6	N821i	NEC	93mm	48mm	24mm	105g	120min	380Stunden
7	F209i	Fujitsu	125mm	40mm	15mm	63g	135min	450Stunden
8	N209i	NEC	90mm	46mm	19mm	86g	120min	500Stunden
9	D502i	Mitsubishi	132mm	43mm	20mm	84g	130min	350Stunden
10	SO502i	Sony	122mm	42mm	17mm	73g	120min	210Stunden

Wie man an diesen Zahlen sehen kann, ist für den Endnutzer der Faktor Hersteller (in diesem Fall Panasonic) am entscheidensten. Faktoren wie Grösse oder Akkuleistung stehen hierbei im im Hintergrund. Ein wesentlicher Punkt, der mit dem Hersteller zusammenhängt könnte hierbei die Menüführung oder die Zusammenstellung de Features sein. Genaueres lässt sich leider nicht sagen, da sich Informationen über diese Endgeräte leider nur auf japanischen Seiten finden lassen und diese Geräte auf dem europäischen oder amerikanischen Markt nicht auftauchen. Die oben aufgeführten Geräte sind allesamt - abgesehen von Position 1- Geräte der neusten Generation. Diese Konstellation aus i-mode, den angeführten Endgeräten und dem stärksten Benutzer-Klientel lässt darauf

schliessen, dass in Japan der Schwerpunkt auf dem Unterhaltungssektor nicht aber im wirtschaftlichen Bereich liegt.

5.3 Schluss

Abschliessend sollen hier die wichtigsten Punkte noch einmal wiederholt und zusammenhängend dargestellt werden, um eine Übersicht zu geben. Danach folgt ein Fazit, das eine Prognose über die Entwicklung und eventuelle Marktverschiebung.

5.3.1 Zusammenfassung

Das alte GSM Netz mit der CSD Software wird verbessert zu HSCSD. WAP wird durch GPRS günstiger. SMS wird durch EMS oder MMS erweitert. Aus einfachen kleinen Programmen, die das Funktionieren des Handys sicher stellten, werden Multitasking Betriebssysteme. Aus IrDA-Schnittstellen werden Bluetooth- oder W-Lan Adapter. Einfache Dienste werden zu sinnvollen komplexen Diensten im Stil der LBS zusammengeschlossen. PDA und Handy verschmelzen immer mehr. Durch verbesserte Roamingverfahren und durch mehr Kompatibilität wird die Welt immer kleiner und UMTS hält langsam aber sicher den Einzug. Wichtig für diese weltweite Kommunikation und Kooperation der einzelnen Netze ist die Offenheit der einzelnen Schnittstellen und die Sicherstellung der Kompatibilität. Leider ist diese Kompatibilität nicht überall geben (siehe Markt Europa/Japan). Das ist einer der Gründe der es verhindern wird, dass UMTS zu einem Weltstandard wird. Dennoch kann man feststellen, dass der Gedanke aller dieser Netze in die gleiche Richtung geht: Die Dienste werden immer komplexer und lauffähig auf den unterschiedlichsten Geräten.

5.3.2 Fazit

Noch sind die Endgeräte für UMTS zu teuer und das Netz noch nicht ausreichend ausgebaut. Für den Normalbürger scheint dieses Netz noch nicht ausgereift, aber dies wird sich in den nächsten 3 bis 4 Jahren sicher ändern. Die Möglichkeiten von UMTS sind zwar - wie in einigen Absätzen gezeigt - in einigen Bereichen schon fast am Limit. Es ist also davon auszugehen, dass das Netz sicher noch einmal überarbeitet wird, um mehr Kapazität zu zusichern. Aber gerade die Kapazität raubenden Dienste wie beispielsweise Videostreaming sind sicher für Geräte wie Handy und PDA ungeeignet aufgrund der kleinen Darstellflächen und der geringen Akkuleistung. Wie in einem der oberen Absätze aber bereits prognostiziert wurde, wird UMTS nicht nur für Handys und PDA eine Rolle spielen, sondern auch im Bereich der Multimediaunterhaltung seinen Platz finden. Zum Beispiel in Spielkonsolen oder Fernsehgeräten. Hier wiederum würde gerade dieser Dienst einen Sinn machen. Was den Alltag sicher stark beeinflussen wird, sind die angeführten LBS und das interaktive Zusammenspiel mit Terminkalendern oder Persönlichkeitsprofilen. Es wird die Zeit kommen, wo nahezu jeder mit einem solchen Gerät unterwegs sein wird und sämtliche Navigation, Orientierung und Zahlung mit einem solchen Gerät tätigen wird. Wann das sein wird steht sicher noch in den Sternen und ist in Europa von der Realisierung der UMTS Netze abhängig. Jede Medaille hat aber immer ihre zwei Seiten: Mit einem solchen Gerät bei sich und einer Telefonnummer oder IP ist man jeder Zeit

und allerorts lokalisierbar. Sogar ein Verhaltensprofil und der Freundes- und/oder Interessenkreis ist nun kein Geheimnis mehr. Für Leute mit entsprechender Hardware und dem nötigen Know-How liegt dann alles offen. Big Brother lässt grüssen.

Literaturverzeichnis

[0] http://www.gsmworld.com

[1] http://www.teltarif.de/i/umts-utopie.html

[2] http://www.syfex.com/pages/techno/glossar/glossar.htm

[3] http://www.tele-fon.de

[4] http://www.mobileEms.com

[5] http://www.mobileMMS.com

[6] http://www.symbian.com

[7] http://www.imocos.com

[8] http://news.zdnet.de/story/

[9] Ovum, Durlacher Report, Boston Consulting Group

[10] http://imodelinks.com/desktop/faq.html

[11] http://houns54.clearlake.ibm.com/solutions/industrial/indpub.nsf/detailcontacts/

[12] http://www.funkschau.de/heftarchiv/pdf/2002/fs0302/fs0203007.pdf

[13] http://www.emorphia.com/downloads/ServiceMobility.pdf