HOCHSCHULE
MITTWEIDA
UNIVERSITY OF
APPLIED SCIENCES

MASTERARBEIT

Frau
Sara Kischnick

Aufbau von Kommunikationsnetzen in der Elektromobilität und Definition von auftretenden Zuverlässigkeitsproblemen

Mittweida, 2013

Fakultät **Elektro- und Informationstechnik**

MASTERARBEIT

Aufbau von Kommunikationsnetzen in der Elektromobilität und Definition von auftretenden Zuverlässigkeitsproblemen

Autorin:
Frau
Sara Kischnick

Studiengang:
Elektrotechnik

Seminargruppe:
ET12wA-M

Erstprüfer:
Prof. Dr. rer. nat. Peter Tittmann

Zweitprüfer:
Prof. Dr.-Ing. Alexander Lampe

Einreichung:
Mittweida, 16.12.2013

Verteidigung/Bewertung:
Mittweida, Jan. 2014

Faculty of **Electrical Engineering and Information Technology**

MASTERTHESIS

Structure of communication networks in electromobility and suitable reliability measures

author:
Ms.
Sara Kischnick

course of studies:
Electrical Engineering

seminar group:
ET12wA-M

first examiner:
Prof. Dr. rer. nat. Peter Tittmann

second examiner:
Prof. Dr.-Ing. Alexander Lampe

submission:
Mittweida, 16.12.2013

defense/ evaluation:
Mittweida, Jan. 2014

Bibliografische Angaben

Kischnick, Sara: Aufbau von Kommunikationsnetzen in der Elektromobilität und Definition von auftretenden Zuverlässigkeitsproblemen, 91 Seiten, 25 Abbildungen, Hochschule Mittweida, University of Applied Sciences, Fakultät Elektro- und Informationstechnik

Masterarbeit, 2013

Referat

Die vorliegende Arbeit beschäftigt sich mit der Struktur und den Komponenten in mobilen Kommunikationsnetzen. Es wird zudem herausgearbeitet, welche Kennzahlen notwendig sind, um die Verfügbarkeit und die Überlebensfähigkeit von mobilen Kommunikationsnetzen einschätzen und bewerten zu können. In diesem Zusammenhang wird zudem untersucht, welche Forschungsergebnisse bezüglich der Zuverlässigkeit in diesen Netzen zur Verfügung stehen. Der letzte Punkt, der in dieser Arbeit berücksichtigt wurde, ist die Anwendung der mobilen Kommunikationsnetze in der Praxis. Dabei wird auch die Kommunikation zwischen den Fahrzeugen beschrieben, die zukünftig eine wichtige Rolle dabei spielen soll, die Verkehrslage in Deutschland sicherer zu machen.

I. Inhaltsverzeichnis

II. Abbildungs- und Tabellenverzeichnis

Abbildungen

Tabellen

III. Abkürzungsverzeichnis

3GPP	Third Generation Partnership Project
AGCH	Access Grant Channel
AM	Amplitude Modulation
AMPS	Advanced Mobile Phone System
AODV	Ad hoc On-demand Distance Vector
AP	Access Point
ARQ	Automatic Repeat Request
ASK	Amplitude Shift Keying
ATM	Asynchronous Transfer Mode
AU	application unit
AuC	Authentification Centre
BSC	Base Station Controller
BSS	Base Station Subsystem
BSS	Basic Service Set
BTS	Base Transceiver Station
C2C	Car-to-Car
C2C-CC	Car 2 Car Communication Consortium
C2I	Car-to-Infrastructure
C2X	Car-to-X
CA	Carrier Aggregation
CAN	Campus/City Area Network
CN	Core Network
CRC	Cyclic Redundancy Check
CS	Coding Scheme
CSD	Circuit Switched Domain
CUP	channel unit processor
DRNC	Drift RNC
DS	Distribution System
DTAP	Direct Transfer Application Part
DVMRP	Distance Vector Multicast Routing Protocol
EDGE	Enhanced Data Rates for GSM Evolution
EIR	Equipment Identity Register

EPC	Evolved Packet Core
ESS	Extended Service Set
ETSI	European Telecommunications Standards Institute
EUTRAN	Evolved UMTS Terrestrial Radio Access Network
FM	Frequency Modulation
FSK	Frequency Shift Keying
GAN	Global Area Network
GGSN	Gateway GPRS Support Node
GMSC	Gateway MSC
GMSK	Gaussian Minimum Shift Keying
GPRS	General Packet Radio Service
GR	GPRS Register
GSM	Global System for Mobile Communications
GSN	GPRS Support Node
HLR	Home Location Register
HS	hot spot
HSCSD	High Speed Circuit Switched Data
HSDPA	High Speed Downlink Packet Access
HSPA	High Speed Packet Access
HSS	Home Subscriber Server
HTTP	Hypertext Transfer Protocol
IBSS	Independent BSS
IEEE	Institute of Electrical and Electronics Engineers
IM	Interworking Manager
IMEI	International Mobile Equipment Identity
IMS	IP Multimedia Subsystem
IMSI	International Mobile Subscriber Identity
IP	Internet-Protokoll
ISDN	Integrated Services Digital Network
ISI	Intersymbolinterferenz
ISO	International Organization for Standardization
ITU	International Telecommunication Union
LA	Location Area
LAI	Location Area Identification
LAN	Local Area Network

LLC	Logical Link Control
LTE	Long Term Evolution
MAC	Medium Access Control
MAN	Metropolitan Area Network
MANET	Mobile Ad hoc Network
MCS	Modulation and Coding Schemes
MIMO	Multiple Input Multiple Output
MME	Mobility Management Entity
MOSPF	Multicast Open Shortest Path First
MS	Mobile Station
MSC	Mobile Switching Center
MSISDN	Mobile Subscriber ISDN
MSRN	Mobile Subscriber Roaming Number
NSS	Network and Switching Subsystem
OBU	on-board unit
OFDM	Orthogonal Frequency Division Multiplexing
OMC	Operation and Maintenance Centre
OSI-Modell	Open Systems Interconnection Model
OSS	Operation (Support) Subsystem
PAN	Personal Area Network
PCRF	Policy and Charging Rules Function
PDA	Personal Digital Assistant
PDN	Packet Data Network
PDN-GW	PDN-Gateway
PIN	Personal Identity Number
PLL	Phase-Locked-Loop
PM	Phase Modulation
PSD	Packet Switched Domain
PSK	Phase Shift Keying
PUK	PIN Unblocking Key
QAM	Quadrature Amplitude Modulation
QoS	Quality-of-Service
R99	Release 99
RN	Relay Node
RNC	Radio Network Controller

RNS	Radio Network Subsystem
RSU	road-side unit
RWP	Random way point mobility model
S-GW	Serving Gateway
S1-CP	S1-Control Plane
S1-UP	S1-User Plane
SCTP	Stream Control Transmission Protocol
SDCCH	Stand-alone Dedicated Control Channel
SDH	Synchronous Digital Hierarchy
SDU	Service-Data-Unit
SGSN	Serving GPRS Support Nodes
SIM	Subscriber Identity Module
SMS	Short Message Service
SNR	signal-to-noise ratio
SONET	Synchronous Optical Network
SRNC	Serving RNC
SS7	Signalisierungssystem Nr. 7
SSID	Service Set Identifier
TCP	Transmission Control Protocol
TMSI	Temporary Mobile Subscriber Identity
UE	User Equipment
UMTS	Universal Mobile Telecommunications System
UTRA	UMTS Terrestrial Radio Access
UTRAN	UTRA Network
VANET	Vehicular Ad hoc Network
VLR	Visitor Location Register
W-CDMA	Wideband Code Division Multiple Access
WAN	Wide Area Network
WLAN	Wireless Local Area Network

1 Einleitung

In der heutigen Gesellschaft sind mobile Geräte nicht mehr wegzudenken. Jeder möchte überall erreichbar sein oder andere erreichen können, Informationen abrufen und Bilder hochladen. Diese Dienste sollen zum einen möglichst schnell zur Verfügung stehen und zum anderen soll der Zugang zum Internet lückenlos sein. Des Weiteren bieten die mobilen Kommunikationsnetze auch Möglichkeiten, die Sicherheit der Bevölkerung zu verbessern. Durch eine Kommunikation zwischen Fahrzeugen ist es beispielsweise vorstellbar, dass die Fahrzeuge selbst auf Gefahrensituationen reagieren und damit Unfälle vermeiden, die aufgrund der menschlichen Reaktionszeit anderenfalls nicht zu umgehen gewesen wären. Damit solche Szenarien in der Zukunft denkbar sind, werden an die mobilen Kommunikationsnetze hohe Anforderungen gestellt. Die Informationen müssen schnell und zuverlässig übertragen werden. Außerdem ist eine lückenlose Abdeckung notwendig.
Diese Forderungen bergen große Schwierigkeiten, denen die Mobilfunkanbieter gewachsen sein müssen. Damit Informationen schnell übertragen werden können, sind große Datenraten notwendig. Diese können durch neue Techniken, die in den letzten Jahren standardisiert wurden, erreicht werden. Neben der Datenrate sind auch eine hohe Verfügbarkeit und die Überlebensfähigkeit der Netze von großer Bedeutung.
Diese Arbeit beschäftigt sich mit den grundlegenden Techniken, die in der Mobilkommunikation verwendet werden sowie ihrer Umsetzung und Verwendung in der Praxis. Außerdem soll untersucht werden, welche Forschungsergebnisse es für die Verfügbarkeit und die Überlebensfähigkeit in mobilen Kommunikationsnetzen gibt. Das Ziel ist es, einen Überblick über den derzeitigen Stand der Forschung bezüglich der Technik und der Zuverlässigkeit in der Elektromobilität zu geben.

Als erstes werden in der Arbeit die Grundlagen der mobilen Kommunikationstechnik vorgestellt. Das zweite Kapitel beschäftigt sich mit den Übertragungstechniken und -verfahren, während das dritte Kapitel auf den Aufbau eines Kommunikationsnetzes eingeht und die Komponenten beschreibt, die verwendet werden.
Darauf aufbauend befasst sich das folgende Kapitel mit den mobilen Kommunikationsnetzen, die heutzutage verwendet werden und erläutert die Funktionsweise dieser Netze (Kapitel 4). Der nächste Teil beschäftigt sich mit der Zuverlässigkeit in den Kommunikationsnetzen (Kapitel 5) und geht dabei auf veröffentlichte Forschungsergebnisse ein. Das sechste Kapitel befasst sich schließlich mit der Anwendung und der Umsetzung von Technik und Zuverlässigkeit in der Praxis.

2 Grundlagen zur Übertragung von Signalen

2.1 Signale und ihre Ausbreitung im Medium

„Signale sind die physikalische Repräsentation von Daten.“ [33, S. 49] Die Übertragung erfolgt immer nach dem in Abbildung 2.1 (nach [23, S. 10]) gezeigten Prinzip. Durch den Sender, z. B. ein Mobilfunktelefon, werden die Daten ausgegeben. Um sie übertragen zu können, müssen sie durch den Wandler in Signale umgeformt werden.
Auf der Empfängerseite (in Abbildung 2.1 als „Senke“ bezeichnet) muss das empfangene Signal zurückgewandelt werden, damit der Empfänger die Daten auslesen kann.

Auf der Empfängerseite kann das gesendete Signal - auch Nutzsignal genannt - von Störsignalen überlagert sein. Diese können einerseits durch natürliche Störeinflüsse hervorgerufen werden wie atmosphärisches bzw. kosmisches Rauschen oder durch die vorhandenen Störungen auf den Übertragungsleitern (das Übersprechen zwischen benachbarten Leitern wird hier vernachlässigt). Bei der mobilen Kommunikation, die hier v.a. betrachtet werden soll, kommen Interferenzen mit anderen Signalen hinzu. Diese werden einerseits durch andere Übertragungen hervorgerufen, die gleichzeitig getätigt werden (z. B. mehrere Telefonate), andererseits durch das aktuell betrachtete Nutzsignal. Dieses wird im Medium Luft nicht auf einem bestimmten Weg übertragen, sondern kann sich in der Luft in alle Richtungen ausbreiten. Dabei wird es durch überall vorhandene Hindernisse, die von Landschaftsgegebenheiten über Häuser bis hin zu Straßenschildern gehen können, gebeugt, gestreut und gebrochen. Daraus ergeben sich unter Umständen unterschiedlich lange Übertragungswege bis das Nutzsignal den Empfänger erreicht. Der Anteil des Nutzsignals, der den längeren Weg zurückgelegt hat, tritt nun als Störung am Empfänger auf, da die einzelnen Nutzsignale sich zeitlich unterscheiden. Um diese Störsignale ausblenden zu können gibt es verschiedene Filtertechniken. Außerdem können auch durch Modulationsverfahren Verbesserungen im Signal-Rausch-Abstand (signal-to-noise ratio, SNR) erreicht werden.
Neben den beschriebenen Störungen, die sich auf ein Nutzsignal aufaddieren können, gibt es auch nachteilige Änderungen direkt am Nutzsignal selbst. Diese reichen von Beugung, Brechung und Streuung (siehe Kapitel 2.3.2), die Verluste hervorrufen können, über die schon beschriebenen Signalinterferenzen (siehe Kapitel 2.3.3) bis hin zu Dämpfung und Abschattung (siehe Kapitel 2.3.1).

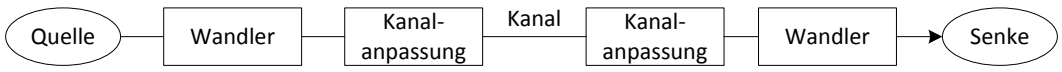


Abbildung 2.1: Grundstruktur eines jeden Übertragungssystems

2.2 Signalaufbereitung und -codierung

2.2.1 Modulationsverfahren

Die Länge einer Antenne muss in ihrer Größenordnung etwa der Wellenlänge der zu sendenden Signale entsprechen, um effizient senden zu können [33, S. 70]. Bei niederfrequenten Signalen (z. B. Musiksignale, die Frequenzen von wenigen MHz haben [33, S. 44]) werden daher mehrere hundert Meter hohe Antennen benötigt, um die Signale auf der Empfängerseite aufnehmen zu können. Des Weiteren werden große Bandbreiten benötigt, um alle Daten übertragen zu können. Wenn man diese Annahmen beispielsweise auf die Radiosender überträgt (Beispiel aus [42]), würde es nur einen Radiosender geben können, der den kompletten Niederfrequenzbereich abdeckt. Einen Ausweg bietet hier die Frequenzmodulation. Dabei werden die niederfrequenten Signale auf eine Trägerfrequenz gelegt, die aus dem hochfrequenten Bereich gewählt wird. Diese Frequenzen können auf der Empfängerseite gefiltert werden und somit, bezogen auf des beschriebene Beispiel, die verschiedenen Radiosender ausgewählt werden.
Es gibt sowohl analoge als auch digitale Modulationsarten. Im Folgenden soll sich auf die digitalen Verfahren beschränkt werden. Es gibt drei verschiedene Möglichkeiten der digitalen Modulation - Amplitudenumtastung[1] (Amplitude Shift Keying, ASK), Frequenzumtastung (Frequency Shift Keying, FSK) und Phasenumtastung (Phase Shift Keying, PSK). Die drei Modulationsarten werden in der Abbildung 2.2 (erstellt nach [17]) kurz skizziert. Der erste Abschnitt sind die zu sendenden Daten. Die drei darunterliegenden Abschnitte stellen das amplitudenmodulierte („AM“), das frequenzmodulierte („FM“) und das phasenmodulierte („PM“) Signal dar.

Bei der *Amplitudenumtastung* werden die digitalen Bits 0 und 1 in der Trägerfrequenz durch unterschiedliche Amplituden dargestellt (siehe Abbildung 2.2, zweiter Abschnitt). Durch Störungen, die das Signal beeinflussen können (siehe Kapitel 2.3.3), ist dieses Verfahren sehr fehleranfällig. Auf der Empfängerseite wird ein Referenzwert eingeführt, mit Hilfe dessen entschieden werden kann, ob der jeweilige Amplitudenwert eine 0 oder eine 1 darstellt. Durch auftretende Störungen können die Amplituden des Signals dermaßen gestört werden, dass diese Schranke nicht mehr zuverlässig eingesetzt werden kann. Anwendung findet die Amplitudenumtastung zum Beispiel in Glasfaserkabeln, bei dem je nach Codierung ein „Lichtimpuls [...] eine 1 [repräsentiert], Dunkelheit eine 0“ [33, S. 72].

Die zweite Möglichkeit der Modulation ist die *Frequenzumtastung*. Dabei werden die 0 und die 1 nicht durch verschiedene Amplituden dargestellt, sondern durch unterschiedliche Frequenzen (siehe Abbildung 2.2, dritter Abschnitt). Auf der Empfängerseite werden zum Auswerten der Frequenzen zwei Bandpässe[2] genutzt. Anhand der Signalstärke

[1] Bei digitalen Modulationsarten wird auch von „Umtastung“ gesprochen [33, S. 70].
[2] Ein Bandpass ist eine Reihenschaltung eines Hoch- und eines Tiefpasses, um eine untere und eine obere Grenzfrequenz zu erzielen. [12, S. 365-368, 34, S. 65f.]

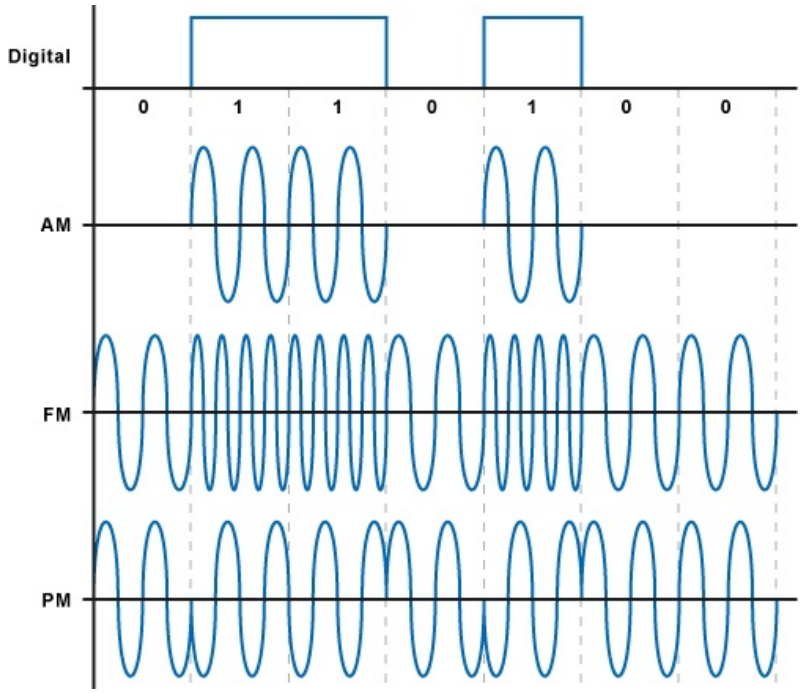


Abbildung 2.2: Digitale Modulationsarten - zu sendende Daten, amplitudenmoduliertes, frequenzmoduliertes und phasenmoduliertes Signal (von oben nach unten)

am Ausgang der Bandpässe kann entschieden werden, ob eine 1 oder eine 0 gesendet wurde. Da die Frequenz durch vorhandene Störungen weniger stark beeinflusst wird, ist die Frequenzumtastung robuster als die Amplitudenumtastung, „benötigt jedoch eine größere Bandbreite“ [33, S. 73].

Als drittes Modulationsverfahren gibt es die *Phasenumtastung*. Dabei werden die Binärwerte durch verschiedene Phasen dargestellt (siehe Abbildung 2.2, letzter Abschnitt). Bei einem zweistufigen Verfahren (2-PSK) würde das Phasensprünge von $\pi(\triangleq 180^\circ)$ ergeben und es können die einzelnen Bits 0 und 1 dargestellt werden [33, S. 73f.]. Wenn mehrstufige Verfahren verwendet werden, z. B. 4-PSK oder 8-PSK, können dementsprechend gleichzeitig 2 oder 3 Bits dargestellt werden. Dabei kann die Anzahl der angezeigten Bits nicht beliebig erhöht werden, da sich bei mehrstufigen Verfahren die Abstände zwischen den einzelnen Phasen immer weiter verkleinern. Durch die Störungen und das Rauschen, die bei der Übertragung auftreten, könnten die einzelnen Bits bei zu kleinen Abständen nicht mehr korrekt dekodiert werden. [33, S. 76] Die Phasenumtastung ist robuster als die Frequenz- und die Amplitudenumtastung, benötigt jedoch einen höheren technischen Aufwand, da die Schaltungen sowohl auf der Sender- als auch auf der Empfängerseite komplexer sind [33, S. 74] (siehe PLL- (Phase-Locked-Loop)-Schaltung, [12, S. 509–520, 26, S. 485–523]).

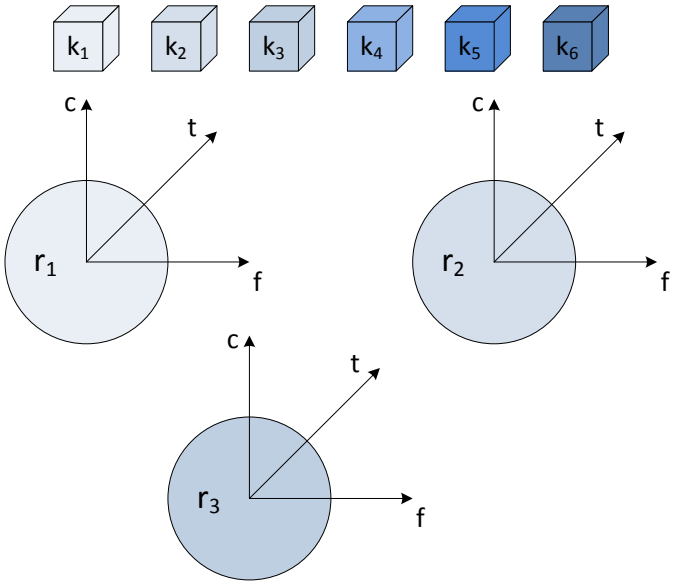


Abbildung 2.3: Raummultiplex

2.2.2 Multiplex-Verfahren

Multiplexen bedeutet, dass mehrere Geräte[3] das gleiche Medium nutzen können, ohne dass es zu Interferenzen (z. B. Kollisionen und Überschneidungen), kommt. Um dies realisieren zu können, müssen für die einzelnen Nutzer unterschiedliche Parameter definiert werden. Diese Parameter können z. B. die Zeit, der Ort oder die Frequenz sein. Nach diesen Parametern werden die verschiedenen Multiplexverfahren benannt: Raummultiplex, Frequenzmultiplex, Zeitmultiplex und Codemultiplex.

Beim *Raummultiplex* können die Nutzer die gleiche Frequenz und den gleichen Zeitraum nutzen. Um Störungen zu vermeiden, müssen die Nutzer weit genug von einander entfernt sein. Der Raum zwischen den Nutzern wird „Schutzabstand“ genannt [33, S. 63]. In der Abbildung 2.3 (gezeichnet nach [33, S. 64]) wird das Prinzip des Raummultiplex' verdeutlicht. Es wird angenommen, dass sechs Übertragungskanäle k_i vorhanden sind. An den Achsen der Diagramme sind die Größen Frequenz, Zeit und Code aufgetragen. Es ist erkennbar, dass die Kanäle k_1 bis k_3 die gleichen Frequenzen, die gleiche Zeit und den gleichen Code nutzen, jedoch verschiedene räumliche Bereiche r_1 bis r_3. Angewendet wird das Raummultiplex beispielsweise bei regionalen Rundfunksendern oder bei Mobiltelefonnutzern, die einen genügend großen Schutzabstand aufweisen. Probleme treten auf, wenn beispielsweise in einer Stadt mehrere Radiosender gleichzeitig ausstrahlen. In diesem Fall muss ein anderes Multiplexverfahren genutzt werden. [33, S. 63]

Ein weiteres Verfahren ist das *Frequenzmultiplex* (siehe Abbildung 2.4, gezeichnet

[3] Im weiteren Verlauf wird für Gerät häufig auch die Bezeichnung „Nutzer“ verwendet. Gemeint ist jedoch immer das kommunizierende Gerät.

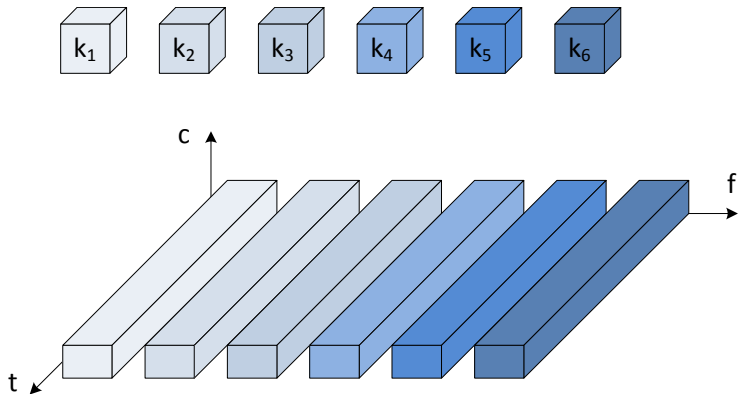


Abbildung 2.4: Frequenzmultiplex

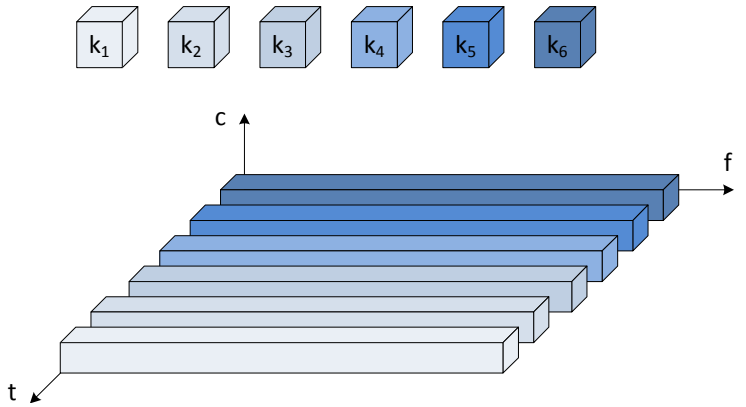


Abbildung 2.5: Zeitmultiplex

nach [33, S. 65]). Dabei wird jedem Sender ein bestimmtes Frequenzband zugewiesen. Zwischen den Bändern muss wie schon beim Raummultiplex ein ausreichend großer Schutzabstand vorliegen, damit es nicht zu einem Übersprechen kommt. Eingesetzt wird dieses Verfahren beispielsweise im Rundfunkbereich. Jeder Rundfunksender nutzt eine andere Frequenz, der Empfänger stellt die zu empfangende Frequenz ein. Es ist somit keine Synchronisation oder Absprache zwischen Sender und Empfänger nötig. [33, S. 65] Im Mobilfunkbereich hat dieses Verfahren einen entscheidenden Nachteil. Der zugewiesene Frequenzbereich würde im Mobilfunk nur kurzzeitig genutzt, die restliche Zeit würde das Frequenzband ungenutzt bleiben. Da nur eine bestimmte Menge an nutzbaren Frequenzen zur Verfügung steht, wird im Mobilfunkbereich ein kombiniertes Verfahren aus Frequenz- und Zeitmultiplex verwendet (siehe folgende Abschnitte).

Beim *Zeitmultiplex* (siehe Abbildung 2.5, gezeichnet nach [33, S. 66]) wird den Sendern zeitabhängig das gesamte Frequenzband zugeordnet. Dieses Verfahren ist viel komplexer als das Frequenz- und Raummultiplex, da alle Sender und Empfänger synchronisiert werden müssen. Allerdings können durch dieses Verfahren auch die begrenzten Ressourcen effektiver genutzt werden, da Nutzer, die nur eine geringe Datenmenge übertragen möchten, weniger Zeit zur Verfügung bekommen als andere mit einer großen Datenmenge. Das Zeitmultiplex wird daher heutzutage in fast allen digitalen Mobilfunksystemen

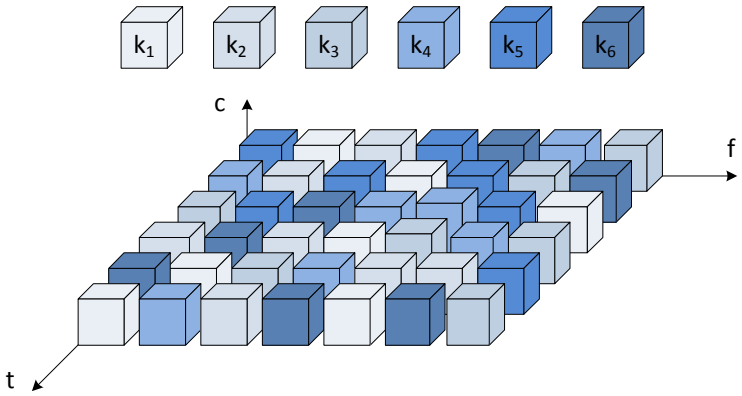


Abbildung 2.6: Frequenz- und Zeitmultiplex

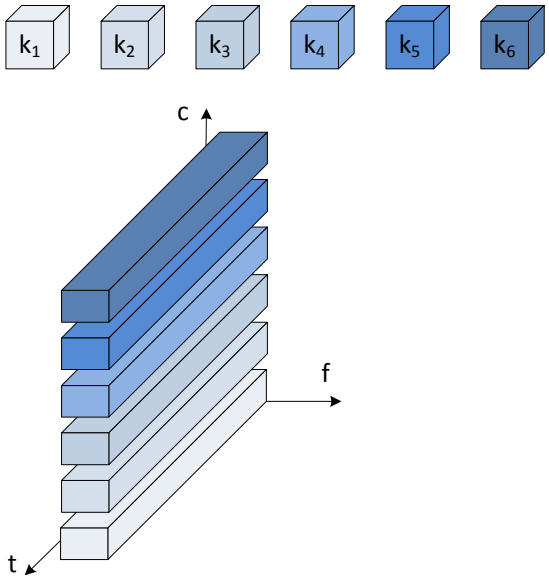


Abbildung 2.7: Codemultiplex

verwendet. [33, S. 66]

In der Praxis wird oftmals ein kombiniertes Verfahren aus *Frequenz- und Zeitmultiplex* (siehe Abbildung 2.6, gezeichnet nach [33, S. 67]) genutzt. Dabei werden den Sendern zu bestimmten Zeiten bestimmte Frequenzbänder zugewiesen. Da diese Zeitabstände sehr kurz sind, betreffen auftretende Störungen einen einzelnen Kanal nur kurzzeitig und daraus resultierende fehlerhafte Übertragungen können unter Umständen auf der Empfängerseite wiederhergestellt werden. Das Verfahren benötigt einen hohen technischen Aufwand, da Sender und Empfänger synchronisiert sein müssen und ein schneller Wechsel von Frequenzen erfolgen muss. Dennoch überwiegen hier die Vorteile, die in der Ausnutzung der vorhandenen Frequenzen liegen. Des Weiteren bietet das Verfahren einen gewissen Schutz vor Abhören, wenn nicht bekannt ist, in welcher Reihenfolge die Frequenzen genutzt werden. Ein solch kombiniertes Frequenz- und Zeitmultiplex wird heutzutage im GSM-Standard bei der Übertragung zwischen mobilen Endgeräten und den Basisstationen verwendet. [33, S.66f.]

Ein weiteres Multiplexverfahren ist das *Codemultiplex* (siehe Abbildung 2.7, gezeichnet nach [33, S. 67]). Bei diesem Verfahren nutzen die Sender alle zur gleichen Zeit die gleiche Frequenz. Alle Kanäle sind jedoch mit separaten Codes versehen. Der Empfänger entschlüsselt den Code und liest den betreffenden Kanal aus. Dabei muss der Kanal von den anderen Kanälen und Störsignalen getrennt werden. Wenn das Hintergrundrauschen, d. h. die übrigen Kanäle und die Störsignale, zu groß ist, kann der zu empfangende Kanal nicht korrekt oder vollständig empfangen werden. Dieses Multiplexverfahren ist durch die Codierung verhältnismäßig abhörsicher. Außerdem können damit mehr Kanäle gleichzeitig gesendet werden als mit den drei bereits beschriebenen Verfahren, da der Coderaum größer ist als der Zeit- oder Frequenzbereich. Nachteil des Verfahrens ist die technische Komplexität. Sender und Empfänger müssen exakt synchronisiert sein, damit der Code richtig angewendet werden kann. Außerdem müssen die Empfänger das erhaltene Signal von dem Hintergrundrauschen trennen. Durch das verbesserte Kosten-Leistungs-Verhältnis von hochintegrierten Schaltungen wird das Verfahren heutzutage immer häufiger in Mobilfunknetzen verwendet. [33, S. 68f.]

2.3 Ursachen für Signaländerungen am Empfänger

2.3.1 Dämpfung

Signale, die im freien Raum ohne Sichtbehinderungen ausgesendet werden, verlieren an Leistung, je größer die Entfernung zum Sender ist. Der Grund ist die Richtung der Ausstrahlung. Die elektromagnetischen Wellen werden kugelförmig vom Sender abgestrahlt. Mit steigendem Radius r wird die Kugeloberfläche s größer (siehe Formel 2.1). Demnach nimmt die Leistung mit größer werdendem Radius ab (siehe Formel 2.2), wobei P_0 der Leistung des Senders entspricht und P_r die Leistung abhängig vom Radius r ist. Außer von der Entfernung zum Sender hängt die Leistung des Signals auch von der Richtcharakteristik der Antennen und von der Wellenlänge ab. [33, S. 57]
Neben den Eigenschaften der Antenne sind die Beschaffenheit der Umgebung und die Atmosphäre weitere entscheidende Faktoren. Regen, Schnee, Nebel und Rauch können die Leistung eines Signals stark abschwächen. Vor allem Regen kann die Energie von elektromagnetischen Wellen gut aufnehmen. Signale können auch durch Gegenstände in der Umgebung gedämpft werden. Bei hohen Frequenzen können die Signale sogar abgeschattet, d. h. ausgelöscht werden. [33, S. 57]

$$s = 4\pi r^2 \tag{2.1}$$

$$P_r \approx \frac{P_0}{r^2} \tag{2.2}$$

2.3.2 Reflexion, Refraktion, Beugung und Streuung

Das in Kapitel 2.3.1 beschriebenen Verhalten bezog ich auf eine Umgebung ohne Hindernisse. Sobald Hindernisse auftreten, kommt es neben der Dämpfung zu weiteren Effekten der Ausbreitung – Reflexion, Brechung, Beugung und Streuung.
Bei der *Reflexion* wird das Signal von einem Hindernis zurückgeworfen. Dabei wird ein Teil des Signals absorbiert, d. h. das reflektierte Signal hat gegenüber dem einfallenden an Leistung verloren. In Städten ist es somit möglich, dass Signale ausgelöscht werden oder zu viel Energie verlieren, so dass sie nicht mehr empfangen werden können. Ein weiterer Effekt, der an Grenzflächen auftreten kann, ist die *Brechung* oder *Refraktion*. Die Ursache für die Brechung sind die unterschiedlichen Geschwindigkeiten von elektromagnetischen Wellen, die in verschiedenen Medien auftreten. In einem Medium großer Dichte ist die Ausbreitungsgeschwindigkeit kleiner als in Medien kleinerer Dichte, nur im Vakuum breiten sich die Wellen mit Lichtgeschwindigkeit aus. Anhand der Gleichung 2.3 [16] kann man erkennen, dass der Austrittswinkel β kleiner sein muss als der Einfallswinkel α, wenn die elektromagnetische Welle in ein Medium größerer Dichte eintritt ($c_1 > c_2$). Dies ist beispielsweise der Fall, wenn Nebel auftritt. In diesem Fall werden die elektromagnetischen Wellen Richtung Boden abgelenkt und können eine eventuell vorhandene Sichtverbindung (Verbindung ohne Hindernisse) nicht mehr ideal verfolgen. [33, S. 58f.]

$$\frac{\sin\alpha}{\sin\beta} = \frac{c_1}{c_2} \tag{2.3}$$

Die beschriebenen Effekte treten an Hindernissen auf, die größer sind als die Wellenlängen der Wellen. An Hindernissen, die genauso groß oder kleiner als die Wellenlänge sind, können die elektromagnetischen Wellen auch gestreut werden. Bei der *Streuung* wird die elektromagnetische Welle in mehrere kleine Wellen aufgeteilt, deren Energie kleiner ist als die der ursprünglichen Welle. Diese Wellen verlaufen in verschiedene Richtungen. Die Mobilfunkwellen haben Wellenlängen im Zentimeter- bzw. Dezimeterbereich, es können demzufolge auch Hindernisse wie Bäume oder Schilder zu Objekten werden, die Wellen streuen. Ein weiteres Phänomen ist die *Beugung*. Dabei werden die Wellen an Kanten von Hindernissen in andere Richtungen abgelenkt. Es ist somit möglich, auch Empfänger zu erreichen, die beispielsweise hinter Bergen sind. [33, S. 59]
Die beschriebenen Effekte treten alle gleichzeitig auf. Es ist somit nicht möglich, genaue Vorhersagen über die Stärke von Signalen an bestimmten Orten zu treffen und damit über die Abdeckung von Netzen. Um vorhandene Versorgungslücken zu finden, werden vor der Installation dreidimensionale Geländemodelle verwendet, die Signalstärken in hoher Auflösung berechnen können. Außerdem werden bei der Installation und während des Betriebs von Basisstationen und der Ausrichtung der Sektoren Messungen ausgeführt, die Lücken aufdecken sollen. [33, S. 58f.]

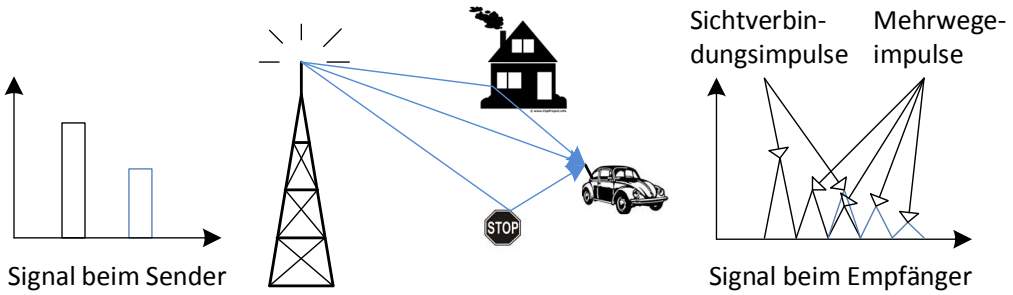


Abbildung 2.8: Mehrwegeausbreitung

2.3.3 Mehrwegeausbreitung und Signalinterferenzen

Alle die in den Kapiteln 2.3.1 und 2.3.2 beschriebenen Effekte treten gleichzeitig auf und führen damit zu der sogenannten „Mehrwegeausbreitung“, „einem der größten Probleme der drahtlosen Kommunikation“ [33, S. 60]. Durch die unterschiedlichen Phänomene wie Brechung und Beugung treten am Empfänger mehrere Signale auf, die von dem gleichen Sender abgeschickt werden. In der Abbildung 2.8 (gezeichnet nach [33, S. 61]) ist dies am Beispiel von drei verschiedenen Wegen verdeutlicht. Diese unterschiedlichen Signale können durch die endliche Ausbreitungsgeschwindigkeit Laufzeitunterschiede (auch Laufzeitdispersion genannt) aufweisen. [33, S. 60]

Die Aufteilung in mehrere Signale verbreitert den Impuls am Empfänger. Durch die unterschiedlichen Laufzeitwege mit ihren jeweiligen Dämpfungen haben die Empfangsimpulse auch unterschiedliche Stärken. Ein Großteil der Impulse am Empfänger (in der Realität weitaus mehr als drei) wird so wenig Energie haben, dass er nicht mehr detektierbar und nur noch als Hintergrundrauschen erkennbar ist. [33, S. 60f.]
Ein weiterer Effekt dieser Impulsverbreiterung ist die sogenannte „Intersymbolinterferenz“ (ISI). Die einzelnen Impulse oder eine Impulsfolge bilden ein Symbol, mehrere Symbole wiederum ein Bit. Durch die Verbreiterung eines Impulses können die zeitlich verschobenen Signale die Signale überlappen, die das nächste Symbol darstellen sollten. Bei größer werdender Symbolrate werden die Auswirkungen der ISI größer. Durch diese Interferenz wird somit die Bandbreite begrenzt, die ein Funkkanal mit Mehrwegeausbreitung nutzen kann. Der Einfluss der ISI kann bis zur Auslöschung von Signalen führen und somit zu Übertragungsfehlern. [33, S. 61]
Um die Auswirkungen der ISI zu reduzieren, sollten die Eigenschaften eines Übertragungskanals den Empfängern bekannt sein. Um diesen beispielsweise die typischen Übertragungswege „beizubringen“, kann der Sender in regelmäßigen Abständen „Trainingssequenzen“ aussenden, „deren ursprüngliches Signalmuster allen Empfängern bekannt ist“ [33, S. 61]. Anschließend kann der Empfänger einen sogenannten „Entzerrer“ („Equalizer“) programmieren, der zukünftig ankommende Signalfolgen somit anpassen und Störungen durch Mehrwegeausbreitung ausgleichen kann. [33, S. 61]

Zusätzlich zu den Interferenzen durch die Laufzeitdispersion kann es durch sich bewe-

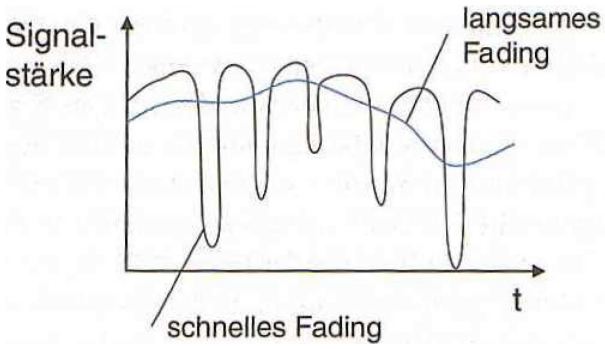


Abbildung 2.9: Kurzfristiger und längerfristiger Signalschwund ([33, S. 62])

gende Sender und Empfänger zu einer Verstärkung des Effektes kommen. Wenn sich ein Empfänger zu schnell bewegt, kann sich der Equalizer nicht schnell genug anpassen an die neuen Kanalcharakteristiken, die sich durch ändernde Übertragungswege ergeben. Es kann somit in Abhängigkeit von Geschwindigkeit und den Eigenschaften der Übertragungskanäle zu Fehlern in der Übertragung kommen. Dieses Phänomen wird als „short term fading“ bezeichnet (siehe Abbildung 2.9). [33, S. 62]
Des Weiteren kann es durch die Bewegung des Empfängers zu einem langsamer verlaufenden Effekt kommen, der auch „long term fading“ genannt wird (siehe Abbildung 2.9). Die Empfangsleistung wird bei größer werdendem Abstand zwischen Sender und Empfänger kleiner. Kompensiert werden kann der Effekt oftmals durch den Sender, der die Sendeleistung steigern kann, damit die Signalleistung am Empfänger eine bestimmte Grenze nicht unterschreitet. [33, S. 62]

3 Aufbau und Elemente globaler Kommunikationssysteme

3.1 Elemente eines Kommunikationssystems und ihre Verwendung

3.1.1 Aufbau des ISO/OSI-Referenzmodells

Das ISO/OSI-Referenzmodell ist ein „Referenzmodell für Netzwerkprotokolle als Schichtenarchitektur“ [43]. Es wurde von der „Internationalen Organisation für Normung“ (ISO) entwickelt und sollte zur „internationalen Standardisierung der verschiedenen Protokolle“ [36, S. 66] beitragen. Als Gründe für die Erstellung des Schichtenmodells werden von Tanenbaum und Wetherall [36, S. 67] folgende genannt:

1. Eine neue Schicht sollte dort erstellt werden, wo ein neuer Abstraktionsgrad benötigt wird.
2. Jede Schicht sollte eine genau definierte Funktion erfüllen.
3. Die Funktionen jeder Schicht sollten mit Blick auf die Definition international genormter Protokolle gewählt werden.
4. Die Grenzen zwischen den einzelnen Schichten sollten so gewählt werden, dass der Informationsfluss über die Schnittstellen möglichst gering ist.
5. Die Anzahl der Schichten sollte so groß sein, dass unterschiedliche Funktionen nicht in einer Schicht zusammengewürfelt werden müssen, aber so klein, dass die gesamte Architektur nicht unhandlich wird.

Aus diesen Überlegungen ergaben sich sieben Schichten (siehe Abbildung 3.1). Dabei muss beachtet werden, dass in diesem Referenzmodell das eigentliche physikalische Medium (z. B. Glasfaserkabel) nicht enthalten ist, sondern sich unterhalb der Bitübertragungsschicht befindet. Die Funktion und Bedeutung der sieben Schichten soll im Folgenden kurz erläutert werden.

Die erste Schicht, genannt *Bitübertragungsschicht* oder *Physical Layer*, dient als die eigentliche Verbindungsschicht zwischen einer Quelle und einer Senke. Hier werden die einzelnen Bits übertragen und es muss entschieden werden, ob das übertragene Bit eine 0 oder 1 darstellt. Dazu sind in dieser Schicht verschiedene technische Möglichkeiten vorhanden, z. B. Modulations- und Multiplex-Verfahren (siehe Kapitel 2.2).

Die zweite Schicht, als *Sicherungsschicht* oder *Data Link Layer* bezeichnet, erfüllt mehrere Aufgaben. Zum einen muss sie eine gesicherte Übertragung gewährleisten. Diese

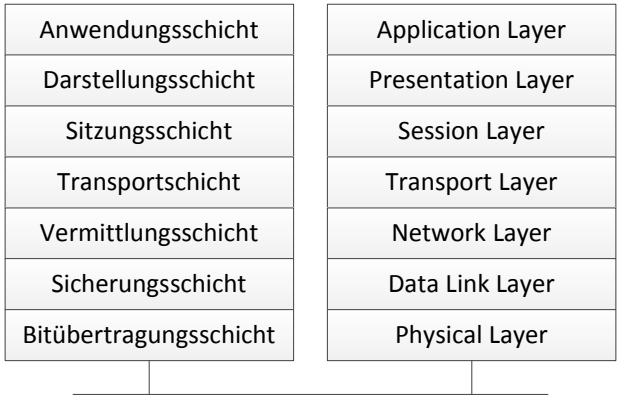


Abbildung 3.1: Schichten des ISO/OSI-Referenzmodells

wird beispielsweise per CRC- und ARQ- Codierung erreicht (siehe [23, S. 422]).
Des Weiteren erfolgt durch die Sicherungsschicht die Flusskontrolle. Dabei wird verhindert, dass eine schnelle Quelle eine langsamere Senke mit Daten „überflutet“. Bevor ein neues Datenpaket gesendet wird, muss der Empfang des vorherigen Paketes vom Empfänger quittiert werden. Eine Alternative dazu bietet das so genannte Pipelining. Dabei ist es möglich, mehrere Datenpakete ohne eine erhaltene Quittierung zu versenden. Bei einer Fehlermeldung wird das Senden ab dem fehlerhaften Paket wiederholt oder es wird nur das fehlerbehaftete Paket nochmals versendet. Das Pipeling bietet gegenüber der „normalen“ Flusskontrolle den Vorteil, dass nicht bei jedem Paket auf eine Quittung gewartet werden muss, was vor allem bei Kommunikationsnetzen mit vielen Knoten längere Zeit beanspruchen kann. [23, S. 423f.]
Die dritte Aufgabe der Sicherungsschicht ist die Überwachung des Vielfachzugriffs, der vor allem bei Broadcast eine Rolle spielt. Dabei soll verhindert werden, dass es bei einem gleichzeitigen Zugriff mehrerer Nutzer Konflikte gibt. Die Kontrolle des Vielfachzugriffs wird durch die MAC-Teilschicht übernommen, eine spezielle Zwischenschicht der Sicherungsschicht. [23, S. 423]

Die *Vermittlungsschicht* sucht über mehrere Punkt-zu-Punkt-Verbindungen einen Weg durch das Netz. Dieses Routing kann durch verschiedene Arten praktiziert werden (siehe Kapitel 4.2.5). Zum einen können vorhandene Routen in statischen Tabellen festgelegt werden. Da dieses Verfahren bei ausgefallenen Komponenten zu einer fehlerhaften oder unmöglichen Übertragung führen kann, besteht die Alternative, das die Tabellen sich automatisch aktualisieren. Eine weitere Möglichkeit ist die Festlegung der zu verwendenden Route zu Beginn einer Kommunikation, z. B. Terminalsitzung. Außerdem können Routen auch hochdynamisch gesucht werden, indem für jedes Paket der Weg neu bestimmt wird. Damit wird eine optimale Netzauslastung erreicht. Die Vermittlungsschicht ist dabei auch für die Qualität des Dienstes zuständig, z. B. für Verzögerungen, Übertragungszeiten und Jitter (Schwankungen bei den Taktzeiten). [36, S. 69]
Eine weitere Aufgabe der Vermittlungsschicht ist die Anpassung der Pakete für verschiedene Netze. Wenn ein Paket auf dem Weg zum Empfänger unterschiedliche Netze durchqueren muss, können diese verschiedene Protokolle verwenden, unterschiedliche

Paketgrößen zulassen oder eine andere Adressierung anwenden. [36, S. 69]
In Broadcast-Netzen ist das Routing-Problem einfacher gestaltet (siehe Abschnitt 4.2.5). Daher gibt es in diesen Netzen nur eine dünne Vermittlungsschicht oder sie ist gar nicht vorhanden. [36, S. 69]

Die *Transportschicht* hat die Aufgabe, eine Verbindung mit dem Kommunikationspartner aufzubauen und die Datenübertragung bis zum Ende auf Fehler zu überprüfen. Außerdem übernimmt sie die End-zu-End-Flusskontrolle. [23, S. 423]
Des Weiteren wird durch die Transportschicht das Übertragungsnetz ausgewählt, die korrekte Reihenfolge der Pakete überwacht und gegebenenfalls wiederhergestellt sowie die Datensegmentierung durchgeführt. Es können mehrere parallel laufende logische Verbindungen auf eine einzige physikalische Verbindung multiplext werden, um damit Kosten zu sparen, es kann im umgekehrten Fall auch eine Nutzverbindung (logische Verbindung) auf mehrere physikalische Verbindungen multiplext werden. Somit erhöht sich der Durchsatz. [23, S. 424]
Als weitere Aufgabe soll die Transportschicht die oberen Schichten und damit die Anwender von der Hardwaretechnik in den unteren Schichten abschirmen. [36, S. 69]

Die *Sitzungsschicht* ermöglicht Verbindungen zwischen verschiedenen Rechnern und „eröffnet, überwacht und beendet eine Sitzung“ [23, S. 424]. Sie ist für die Dialogsteuerung zuständig, legt also fest, wer aktuell Daten senden darf, sowie für die Tokenverwaltung, die verhindert, dass zwei Operationen gleichzeitig ausgeführt werden. Des Weiteren gehört die Synchronisation zu dem Aufgabenbereich der Sitzungsschicht. Dabei werden „bei langen Übertragungen Fixpunkte gesetzt [...], ab denen die Übertragung nach einem Absturz und der nachfolgenden Wiederherstellung fortgesetzt werden kann“ [36, S. 70].

Die *Darstellungsschicht* legt die Semantik und die Syntax der übertragenen Information fest. Die ausgetauschten Datenstrukturen müssen abstrakt dargestellt werden, damit ein Datenaustausch mit Computern, die unterschiedliche Strukturen verwenden, durchgeführt werden kann. Außerdem muss eine Standardcodierung verwendet werden, die die Übertragungsleitung "versteht". [36, S. 70]

Die *Anwendungsschicht* identifiziert den Kommunikationspartner und legt die Kommunikationsparameter fest. In dieser Schicht sind viele Protokolle abgespeichert, die von den meisten Benutzern verwendet werden, z. B. das HTTP-Protokoll (Hypertext-Übertragungsprotokoll), ein Datenübertragungsprotokoll im Netzwerk. „Andere Protokolle werden für Dateiübertragungen, E-Mail und Netznachrichten verwendet.“ [36, S. 70]

3.1.2 Elemente eines Kommunikationssystems

Um die Datenpakete korrekt zwischen verschiedenen Kommunikationspartner versenden zu können, sind unterschiedliche Geräte notwendig. Diese müssen in den einzelnen Schichten des ISO/OSI-Referenzmodells die Aufgaben bewältigen können, die zur Sicherstellung der Kommunikation nötig sind. Die wichtigsten dieser Geräte sollen kurz vorgestellt werden.

Repeater und *Hub* arbeiten auf der untersten Schicht, der Bitübertragungsschicht.
Der Repeater verstärkt ein empfangenes Signal und sendet es anschließend weiter ohne es auszuwerten oder zu verändern. Der Hub leitet Signale weiter, verstärkt sie jedoch nicht. Er ist in Sterntopologie aufgebaut und hat somit mehrere Ports. Im Gegensatz zum Switch (siehe nächster Abschnitt) wertet der Hub ankommende Signale nicht nach der Empfängeradresse aus, sondern sendet diese an alle Ports weiter. Wichtig dabei ist, dass alle Leitungen des Hubs mit gleicher Geschwindigkeit arbeiten. [36, S. 395]

In der nächsten Schicht, der Sicherungsschicht, arbeiten *Bridge* und *Switch*.
Die *Bridge* verbindet zwei oder mehrere Teilnetzwerke. Wenn diese gleiche Protokolle verwenden, wird in der MAC-Teilschicht (siehe Abschnitt 4.5.2) gearbeitet, anderenfalls in der LLC-Teilschicht (siehe Abschnitt 4.5.2). Es existieren wie bei Hubs mehrere Ports, im Gegensatz zu Hubs werten Bridges jedoch die Zieladressen aus und entscheiden, welcher Port für das Weitersenden benutzt werden soll. Ein weiterer Unterschied zu Hubs ist die Leitungsgeschwindigkeit, die bei Bridges nicht an jedem Port gleich sein muss. Dies setzt voraus, dass Bridges einen Zwischenspeicher benötigen, damit Pakete von einem schnelleren Netz in ein langsameres Netz gepuffert werden können.
Switches sind den Bridges sehr ähnlich. Nach Tanenbaum und Wetherall haben die Unterschiede „eher mit Marketing- als mit technischen Gründen zu tun“ [36, S. 396]. Die Bridges mussten früher weniger Teilnetze verbinden und benötigten demzufolge weniger Ports. Heutzutage ist der Begriff „Switch“ gebräuchlicher. Viele Rechner werden zudem direkt an einen Switch angeschlossen und es werden somit mehr Ports benötigt [36, S. 396].

Der *Router* entscheidet, auf welcher Ausgabeleitung ein Paket versendet wird. Dieses sogenannte Routing kann dabei auf mehrere Arten erfolgen (siehe Kapitel 4.2.5). Die Wegefindung soll dabei optimal durch den Vergleich von Entfernungen, Interferenzen und Kosten erfolgen [33, S. 537]. Router arbeiten in der Vermittlungsschicht (3. Schicht des OSI-Referenzmodells). [36, S. 397]

Das letzte typische Element eines Netzwerkes, das hier vorgestellt werden soll, ist das *Gateway*. Es kann in mehreren Schichten arbeiten. Es gibt zwei Arten Gateways – Transport- und Anwendungsgateways.
Transportgateways arbeiten im OSI-Referenzmodell in der 4. Schicht, der Transportschicht. Sie verbinden zwei Rechner, die unterschiedliche Transportprotokolle verwenden,

z. B. TCP/IP und SCTP. Das Transport-Gateway kann Daten versenden und gegebenenfalls zwischen den einzelnen Protokollen formatieren.
Anwendungsgateways arbeiten in der obersten Schicht des Referenzmodells, in der Anwendungsschicht. Sie können zusätzlich zu den Transportprotokoll-Umwandlungen auch die Inhalte und Formate der Pakete umsetzen. So besteht beispielsweise die Möglichkeit aus einem E-Mail-Format eine SMS zu generieren. [36, S. 397]

3.2 Grundlegender Aufbau eines Kommunikationsnetzes

Es gibt verschiedene Möglichkeiten Netzwerke zu klassifizieren. Dazu gehören die physikalische Reichweite, die administrative Methode, die Topologiearten und die Technologie [31, S. 14].

Die *Topologie* eines Netzes beschreibt im Wesentlichen die Art der Verkabelung zwischen den einzelnen Netzteilnehmern und -komponenten. Die drei gängigsten Varianten sind die Stern-, die Bus- und die Ringtopologie [31, S. 17]. Da die Verkabelung der Komponenten in der Mobilkommunikation eine geringere Rolle spielt, soll auf diese Klassifizierungsvariante hier nicht weiter eingegangen werden.

In der Kommunikationstechnik und für die Auswahl der mobilen Netzkomponenten spielt die *physikalische Reichweite* der Netze eine Rolle. Aus diesem Blickwinkel heraus werden die Netzwerke in sechs Bereiche eingeteilt:

- Personal Area Networks (PAN)
- Local Area Networks (LAN)
- Campus Area Networks / City Area Networks (CAN)
- Metropolitan Area Networks (MAN)
- Wide Area Networks (WAN)
- Global Area Networks (GAN)

Alle diese Netze können sowohl als drahtgebundene als auch als drahtlose Variante vorliegen. Die Reichweiten der Netze steigern sich in der Reihenfolge ihrer Erwähnung. Während das LAN in seinem Grundzustand nur 500 m Reichweite hat, erreicht man mit MAN schon bis zu 100 km [31, S. 15]. Die Reichweiten, die mit den einzelnen Netzwerken erlangt werden können, sind in der Tabelle 3.1 zusammengefasst.

Neben der Topologie und der geographischen Ausdehnung gibt es noch die Einteilung nach der *administrativen Methode*. Dabei gibt es zwei Möglichkeiten – Peer-to-Peer-Umgebung und Client-Server-Architektur. Bei der Peer-to-Peer-Umgebung sind alle Knoten gleichberechtigt und alle Nutzer sind gleichzeitig Administrator. Bei Verwendung der Client-Server-Architektur gibt es einen Server mit Administratorrechten, der die

Netzwerkbezeichnung	Reichweite	Drahtloses Kommunikationssystem
Personal Area Network	wenige Meter	Bluetooth, WLAN (IEEE 802.11)
Local Area Network	bis zu 500 m	WLAN (IEEE 802.11)
Campus Area Network	bis zu 5 km	–
Metropolitan Area Network	bis zu 100 km	WiMAX (IEEE 802.16)
Wide Area Network	unbegrenzt	Satellitensystem
Global Area Network	unbegrenzt	Satellitensystem

Tabelle 3.1: Reichweiten verschiedener Netzwerke

zentrale Verwaltung des Netzwerkes übernimmt und Sicherheitsaspekte überwacht, und mehrere Rechner, die Clienten, die Informationen und Dienste vom Server abrufen. Das zuletzt beschriebene Prinzip ist das heute gebräuchlichere, da die Zentralisierung der Verwaltung die Sicherheit eines Netzwerkes erhöht. [31, S. 15f.]

Die letzte Klassifizierungsmethode ist die Technologie. Die berücksichtigten Informationen zur Einteilung sind hier z. B. Topologie, Kabeltyp, Entfernungsbeschränkungen, Kontrollinformationen und Adressen [31, S. 19]. Für mobile Kommunikationsmittel spielt der Kabeltyp nur eine geringe bzw. gar keine Rolle, da die Übertragung über das Medium Luft erfolgt. Verbindungen und Eigenschaften der Kabel können daher vernachlässigt werden. Die Entfernungsbeschränkungen entsprechen den Reichweiten der einzelnen mobilen Netzwerkelementen, z. B. Antennen oder Access Points (siehe Kapitel 4). Die Sicherheitsaspekte und Adressen werden von den Netzwerkelementen verarbeitet (siehe Kapitel 3.1), z. B. für das Weiterleiten von Datenpaketen an bestimmte Empfänger und die Auswahl günstiger Wege (Routing, siehe 4.2.5). Die Sicherheitsaspekte spielen in der mobilen Kommunikation eine große Rolle, da das Übertragungsmedium kein Kabel ist, das durch seinen Aufbau schon einen bestimmten „Grundschutz“ hat, sondern die Informationen durch die Luft übertragen werden. Dadurch ist es recht einfach, Daten abzuhören. Da sich die vorliegende Arbeit v. a. auf die Zuverlässigkeit in mobilen Kommunikationsnetzen beziehen soll, wird hier nicht weiter auf die Sicherheitsfragen eingegangen.

4 Mobile Kommunikationssysteme

4.1 Übersicht und geschichtliche Entwicklung

In der mobilen Kommunikationstechnik gab es in den letzten Jahren einen enormen Wandel in den Standards. Die Ursache dafür liegt einerseits in den steigenden Ansprüchen der Bevölkerung, andererseits an den fortschreitenden Entwicklungen in der Technik. Die Bevölkerung verlangt mit steigender Tendenz nach mehr Möglichkeiten ortsunabhängig auf Daten zugreifen zu können. Dieser Dienst soll möglichst lückenlos und schnell zur Verfügung stehen.

Die Geschichte der Kommunikation reicht Jahrhunderte zurück. Anfangs erfolgte die Übermittlung mittels Licht. Das erste bekannte Kommunikationsmittel, dass nicht auf der Basis von Licht funktionierte, war Marconis[4] Telegraph (1895). Diese Erfindung beruhte auf der Entdeckung der elektromagnetischen Induktion durch Michael Faraday[5] und Joseph Henry[6] (1831) und der ersten Übertragung elektromagnetischer Wellen durch den Raum durch Heinrich Hertz[7] (1886). Die mit den Wellen überbrückbare Distanz konnte durch Nikola Tesla[8] erweitert werden. Ein wichtige Rolle zur Übertragung von elektromagnetischen Wellen spielt zweifelsohne auch James C. Maxwell[9], der mit den *Maxwellschen Gleichungen* 1864 die Grundlagen für elektromagnetische Felder legte. [33, S. 25f.]
Die Technik zur Funkübertragung entwickelte sich im Laufe der folgenden Jahre rasch weiter. Radiosender und Fernsehübertragung entwickelten sich. Allerdings gab es zu diesem Zeitpunkt noch keine international verbreitet festgelegten Standards. Das erste Mobilfunknetz, das es in Deutschland gab, war das sogenannte A-Netz, das noch analog funktionierte und 1958 startete. Mit diesem Netz war noch kein Handover möglich, d. h. keine Übergabe zwischen verschiedenen Basisstationen. Das Netz arbeitete mit 160 MHz. Es besaß 1971 eine „Flächendeckung von 80% und ca. 11.000 Kunden“ [33, S. 27]. Der Nachfolger des A-Netzes, ebenfalls zur 1. Generation gehörend, wurde als B-Netz bezeichnet. Es löste das A-Netz 1972 ab. Gegenüber dem Vorgänger

[4] Guglielmo Marconi (1874-1937): italienischer Radiopionier und Unternehmer. Er baute 1896 den ersten Telegraphen. Die erste Funkverbindung über den Atlantik gelang 1901. Für seine „Verdienste um die Entwicklung der drahtlosen Telegrafie“ [24] erhielt er 1909 zusammen mit Ferdinand Braun den Nobelpreis für Physik.

[5] Michael Farady (1791-1867): englischer Naturforscher und Experimentalphysiker: Er entdeckte 1821 die elektromagnetische Rotation, den Grundstein für den Elektromotor, und zehn Jahre später die elektromagnetische Induktion.

[6] Joseph Henry (1797-1878): US-amerikanischer Physiker. Er entdeckte das Phänomen der Selbstinduktion und baute die erste Spule und das erste elektromagnetische Relais (1831).

[7] Heinrich Rudolf Hertz (1857-1894): deutscher Physiker. 1886 gelang ihm die erste Übertragung elektromagnetischer Wellen.

[8] Nikola Tesla (1856-1943): Erfinder, Physiker und Elektroingenieur

[9] James Clerk Maxwell (1831-1879): schottischer Physiker. Sein größter Erfolg war die Formulierung der Maxwellschen Gleichungen, die die Grundlage für den Magnetismus und die Elektrizitätslehre bilden.

konnten nun „Anruf[e] aus dem Festnetz auf den mobilen Teilnehmer [weitergeleitet]“ [33, S. 27] werden. Dieses Mobilfunknetz war auch in Luxemburg, in den Niederlanden und in Österreich erreichbar. Ein ähnlich international angelegtes Mobilfunknetz gab es auch in den Nordländern Dänemark, Norwegen, Schweden und Finnland. Das „Nordic-Mobile-Telephone-“ (NMT-) System nutzt eine Trägerfrequenz von 450 MHz. Zusätzlich gab es in Europa verschiedene nationale Mobilfunksysteme, die alle unterschiedliche Standards verwendeten, die nicht miteinander kompatibel waren. Um einen gesamteuropäischen Standard zu erreichen, wurde 1982 die „Groupe Spèciale Mobile“ (GSM) ins Leben gerufen, die ein System entwickeln sollte, das bei 900 MHz und vollkommen digital arbeiten und ein „Roaming“ ermöglichen sollte, ergo den Wechsel zwischen verschiedenen Netzbetreibern. Während 1983 in den USA das analoge Mobilfunksystem AMPS („Advanced Mobile Phone System“) eingeführt wurde, startete in Deutschland das analoge C-Netz. Im Gegensatz zu seinen Vorgängern A-Netz und B-Netz war nun eine „Gesprächsübergabe zwischen verschiedenen ’Funkzellen’ möglich“ [33, S. 28]. Das Netz unterstütze zudem neben dem Sprachdienst auch Fax, E-Mail und Datenübertragung. Es arbeitete bis zum Jahr 2000. [33, S. 26ff.]
Seit Beginn der 90er Jahre wurde die ersten digitalen Systeme eingeführt, so 1991 der Standard DECT („Digital European Cordless Telephone“) für schnurlose Telefone, das mit Frequenzen zwischen 1880 und 1900 MHz bei einer Reichweite von 100-500 m arbeitete. Dieses System ist für Großstädte geeignet, da „Zehntausende von Teilnehmern pro Quadratkilometer“ [33, S. 28] unterstützt werden können. Es wird heute in 110 Ländern der Erde verwendet und ist inzwischen unter dem Namen „Digital Enhanced Cordless Telecommunciations“ bekannt. Im gleichen Jahr wie DECT wurde auch der GSM-Standard veröffentlicht (heute „Global System for Mobile Communication“), die zweite Generation der Mobilfunkstandards („2G“). Die damalige erste Version nutzte ca. 900 MHz und bot „internationales Roaming, automatische Teilnehmerlokalisierung, Teilnehmer- und Geräteauthentifizierung, Datenverschlüsselung auf der drahtlosen Strecke, einfache Integration in bestehende ISDN-Systeme [...] eine relativ gute Sprachqualität dank fortschrittlicher Sprachkodierer“ [33, S. 28] und einen SMS-Dienst („Short Message Service“, Kurznachrichtendienst). Der GSM-Standard hat heute einen Marktanteil von über 70% und wird in mehr als 190 Ländern genutzt. Die GSM-Architekturen sind überall miteinander kompatibel. Als die Kapazität des Frequenzbandes bei 900 MHz vor allem in Großstädten erschöpft war, wurde in Europa eine weitere Frequenz von 1800 MHz für GSM freigegeben. Auf dieser Frequenz konnten mit dem GSM-1800-Netz eine verbesserte Sprachqualität realisiert werden sowie kleinere Zellen höherer Qualität, die besonders für große Städte gut geeignet waren. [33, S. 28f.]
1998 startete das erste Satellitensystem für Mobilfunk („Iridium“), das eine Frequenz von 1,6 GHz nutzt und Sprach- und Datenkommunikation erlaubt. Im gleichen Jahr wurde der UMTS-Standard („Universal Mobile Telecommunications System“, 3. Generation des Mobilfunkstandards, „3G“) verabschiedet. 1999 folgten Standards für das lokale Funknetz - IEEE 802.11b.[10] Im Jahr 2000 wurden für GSM eine höhere Datenrate und paketorientierte Übertragung eingeführt (HSCSD - High Speed Circuit Switched Data,

[10] Der Vorgänger IEEE 802.11 arbeitet schon seit 1997.

GPRS - General Packet Radio Service). In diese Zeit fiel auch der Hype um die UMTS-Lizenzen und das „Internet auf dem Mobiltelefon“ [33, S. 30]. 2001 folgte ein neuer WLAN-Standard, der IEEE 802.11a, in den folgenden Jahren weitere Versionen wie 802.11g (2003) oder 802.11n (2009). Weiterhin wurden neue Bluetooth-Standards[11] entwickelt und der terrestrische digitale Betrieb gestartet. Das analoge Satellitenfernsehen ist seit 2012 (öffentlich-rechtlich) abgeschaltet [13]. Andere Sender folgten bzw. folgen nach und nach. [33, S. 31] Seit 2010 (Deutschland) wurde die bisher letzte Generation der Mobilfunknetze eingeführt – „Long Term Evolution“ (LTE) [19]

In den folgenden Unterkapiteln sollen die wichtigsten heute genutzten Standards kurz erläutert werden. Da die benutzte Technik und die Verfahren der einzelnen Standards sehr umfangreich sind, werden sich die Beschreibungen auf das Wesentlichste reduzieren. Der Schwerpunkt soll dabei auf den Eigenschaften der Standards liegen, die für die Zuverlässigkeit bei der Kommunikation und für zeitkritische Aspekte eine Rolle spielen. Bei der Auswahl der beschriebenen Standards wurde auch darauf geachtet, dass sie in der C2X-Kommunikation Anwendung finden.

4.2 Global System for Mobile Communications (GSM)

4.2.1 Übertragung und Standardisierung

GSM arbeitet leitungsvermittelt. Dabei wird zum Gesprächsbeginn eine direkte Leitung zwischen den Gesprächsteilnehmern reserviert. Die Leitung wird dabei von einer Vermittlungsstelle („Switching Center“) mit einer so genannten Vermittlungsmatrix („Switching Matrix“), die einen „beliebigen Eingang mit einem beliebigen Ausgang verbinden kann“ [32, S. 2], geschaltet. Seit der Abschaffung der analogen Übertragung werden in den Vermittlungsstellen auch die Umwandlungen zwischen digitalen und analogen Daten vorgenommen. Bei „Integrated Services Digital Network“- (ISDN-) Anschlüssen wird die Digitalisierung schon in den Endgeräten vorgenommen, die Übertragung findet hier rein digital statt. [32, S. 2]
Mit der Einführung der GSM-Mobilfunktechnik mussten teilweise Neuerungen gegenüber der Festnetztechnik vorgenommen werden. Die Hardware und „die unteren Softwareschichten, die für das Schalten der Verbindungsmatrix und die Signalisierung zuständig sind“ [32, S. 3], konnten weiterverwendet werden. Gegenüber dem Festnetz, bei dem der Standort eines Empfängers fest ist, können die Teilnehmer im Mobilfunk ihren Standort frei wählen. Es ist auch nicht sicher, ob sich ein Teilnehmer immer im gleichen geographischen Gebiet aufhält, d. h, ob er immer über die gleiche Vermittlungsstelle erreichbar ist. Aus diesen Gründen musste im Mobilfunkstandard die Gesprächsvermittlung neu definiert werden. Dazu wurde in den Vermittlungsstellen eine „Mobilitätsmanagementkomponente“ [32, S. 3] integriert, der die Standorte aller erreichbaren Teilnehmer bekannt

[11] Bluetooth wurde 1999 eingeführt bei einer Frequenz von 2,4 GHz und einer Übertrgaungsrate von 11 Mbit/s.

sind. Im Mobilfunk ist es zudem möglich, dass während des Gesprächs ein neuer Standort gewählt wird. Daher musste auch die Gesprächsverwaltung neu gestaltet werden, da es möglich ist, dass die Leitung während eines Gesprächs gewechselt wird. [32, S. 3f.] Neben den Sprachübertragungen, die leitungsorientiert versendet werden, gibt es zusätzlich noch andere Nutzdaten, die gesendet werden müssen. Diese Daten werden paketorientiert übertragen. Daher gab es über lange Zeit im GSM zwei parallele Netzwerke. Da das gleichzeitige Betreiben von zwei Netzwerken ineffizient und somit auch teuer ist, wurden in der Vermittlungsstelle die Verbindungsmatrizen für die leitungsvermittelte Übertragung durch ein so genanntes „Media Gateway“ ersetzt. Durch dieses Gateway werden Sprachverbindungen nicht mehr als leitungsvermittelte Daten übertragen, sondern virtuell als IP-Pakete. Das leitungsvermittelnde Netzwerk kann somit eingespart werden. [32, S. 4]

Damit alle Kommunikationsteilnehmer problemlos und unabhängig vom Standort Zugriff auf das mobile Netz haben, mussten sich die Netzbetreiber auf Standards einigen. Einer der wichtigsten ist das Signalisierungssystem 7 (SS-7) für die Gesprächsvermittlung. Durch das Roaming ist es auch möglich, international zu telefonieren und mobil Daten zu senden. Ein weiterer Vorteil der Standardisierung ist die Reduzierung der Entwicklungskosten, da Geräte weltweit verkauft werden können, ohne dass kostspielige Spezifikationen vorgenommen werden müssen. Zuständig für die Standardisierung von GSM-Netzen ist das „European Telecommunication Standard Institute“ (ETSI), das neben GSM weitere europäische Telekommunikationsstandards festgelegt hat. Zusammen mit anderen weltweit verteilten Standardisierungsgremien wurde durch ETSI das „3rd Generation Partnership Project“ (3GPP) gegründet. Dieses Gremium ist seitdem zuständig für die Standardisierung von GSM, UMTS und LTE und aller Systeme, die zu den drei genannten dazu gezählt werden (z.B. GPRS (siehe Kapitel 4.2.4)). [32, S. 4f.]

4.2.2 Architektur von GSM

Über die Strukturierung der GSM-Architektur gibt es in der Literatur widersprüchliche Angaben. Diese Meinungsunterschiede betreffen dabei nicht nur die Bezeichnung der einzelnen Subsysteme, sondern auch deren Anzahl, die sich zwischen drei und vier bewegt. Die folgenden drei Subsysteme werden jedoch in den meisten Quellen genannt [21, 32, 33, 40]:

- Base Station Subsystem (BSS, Feststationssystem)
- Network and Switching Subsystem (NSS, Mobilvermittlungssystem)
- Operation (Support) Subsystem (OSS, Betriebs- und Wartungssystem)

In manchen Quellen findet man als erstes Subsystem noch die „Mobile Station“ (MS, Mobilstation) [21, 40]. Bei Sauter [32, S. 13] hingegen wird das dritte Subsystem nicht als OSS sondern als „Intelligent Network Subsystem“ (IN) bezeichnet. Bei Schiller

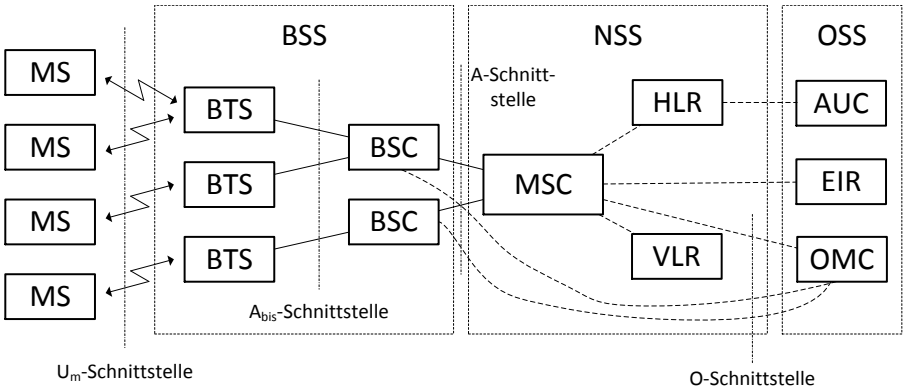


Abbildung 4.1: GSM-Architektur (gezeichnet nach [33, S. 134])

[33, S. 134] wurden die Mobilstationen und die BSS zum „Funk-Feststationssystem“ (Radio Subsystem, RSS) zusammengefasst. Im weiteren Verlauf werden die letzte Schicht als Operation Subsystem bezeichnet und die Mobilstationen als eigenständige Subsysteme betrachtet.

Mobile Station (MS)

Mobilstationen bestehen aus einer Hardware- und einer Software-Komponente. Die Hardwarekomponente, das mobile Gerät, ist der nutzerunabhängige Teil der MS, der nutzerspezifische Teil, die Softwarekomponente, ist das sogenannte Subscriber Identity Module (SIM). [33, S. 135]
Das SIM ist der nutzerspezifische Speicher, auf dem alle Daten gesichert werden, die zur Identifizierung des Nutzers benötigt werden. So kann man dort den Typ, die Seriennummer, registrierte Dienste, die persönliche Kennung (Personal Identity Number, PIN), den PIN-Entsperrschlüssel (PIN Unblocking Key, PUK), Authentifizierungsschlüssel K_i und die internationale Teilnehmerkennung (International Mobile Subscriber Identity, IMSI) finden. Um die SIM entsperren und damit das Endgerät nutzbar machen zu können, muss die PIN eingegeben werden, nach dreimaligem falschem Eingeben wird das Endgerät gesperrt und kann nur die durch Eingabe der PUK wiederum freigegeben werden. Wenn auch diese zehnmal falsch angegeben wurde, ist die SIM-Karte endgültig gesperrt und unbrauchbar. [33, S. 135f.]
Der nutzerunabhängige Teil der MS, die Hardwarekomponente, kann „unabhängig von einem Nutzer [...] über eine eindeutige Gerätekennung, die *International Mobile Equipment Identity (IMEI)*, identifiziert werden“ [33, S. 135]. Die IMEI dient auch der Diebstahlsicherung, da sie für jede Mobilstation eindeutig ist. Es gibt im Geräteidentifikationsregister (Equipment Identity Register, EIR, siehe Operation Subsystem) eine schwarze Liste, in der alle als gestohlen gemeldet oder anderweitig gesperrte IMEI gespeichert sind. Über diese Geräte ist somit keine Kommunikation mehr möglich. [33, S. 138]
Sobald sich eine MS in einem GSM-Netz anmeldet, werden verschieden weitere Daten gespeichert. Zum einen sind das der Schlüssel zur Datenverschlüsselung K_C, anderer-

seits Daten zur Aufenthaltsbestimmung. Zu letzterem gehören eine temporäre Teilnehmerkennung (Temporary Mobile Subscriber Identity, TMSI) und die Kennung für den Aufenthaltsort (Location Area Identification, LAI). [33, S. 136]
Neben den Komponenten, die zur Funktion und zur Erkennung der Mobilstationen zwingend notwendig sind, kann eine MS auch über weitere optionale Hardwarekomponenten verfügen. So besitzt sie Schnittstellen zur Kommunikation mit dem Anwender (z. B. Anzeige, Lautsprecher, Mikrofon, programmierbare Tasten) und kann auch eine Bluetooth-Schnittstelle haben. Des Weiteren sind auch zusätzliche Softwarekomponenten wie Kalender, Notizbücher, Browser oder Spiele möglich, die häufig in sogenannten Applikationen (Apps) zur Verfügung gestellt werden. [33, S. 136f.]

Network and Switching Subsystem (NSS)

Die GSM-Architektur, siehe Abbildung 4.1, besteht aus verschiedenen Geräten, die alle unterschiedliche Aufgaben erfüllen. Der zentrale Knoten der Architektur ist das *Network and Switching Subsystem (NSS)*. Es ist die Verbindungsstelle des drahtlosen Netzes mit den herkömmlichen öffentlichen Netzen. Außerdem führt sie die Übergabe von Verbindungen zwischen mehreren Base Station Subsystemes (BSS) durch und überwacht sie. Des Weiteren unterstützt sie die Ortung von Teilnehmern weltweit und die Abrechnung und das Roaming von Teilnehmern zwischen den Netzbetreibern unterschiedlicher Länder. [33, S. 137] Das NSS setzt sich aus drei Komponenten zusammen: die Dienstvermittlungsstelle (Mobile Services Switching Centre, MSC), dem Heimatregister (Home Location Register, HLR) und dem Besucherregister (Visitor Location Register, VLR).
Die Hauptaufgabe einer MSC ist der Verbindungsaufbau zu anderen MSC und zu den Base Station Controllern (BSC, siehe Base Station Subsystem), die sich in ihrem Zuständigkeitsbereich befinden. Außerdem verarbeitet sie die Signalisierungsdaten für den Verbindungsauf- und -abbau und die Verbindungsübergabe zwischen verschiedenen BSCs. Zur Verarbeitung dieser Signalisierung wird das Signalisierungssystem Nr. 7 (SS7) verwendet. Dieses System unterstützt neben den „normalen“ Diensten auch die Rufnummernportabilität, 0800/0900-Nummern und die automatische Anrufweiterleitung. Spezielle MSCs, sogenannte Gateway MSCs (GMSC), können durch zusätzliche Funktionen auch Verbindungen zu herkömmlichen Festnetzen, z. B. ISDN, aufbauen. [33, S. 137]
Die zweite Komponente im NSS ist das Heimatregister. Das HLR dient als Speicher für alle Teilnehmerdaten, von denen sowohl statische als auch dynamische Daten gespeichert werden. Zu den statischen Daten gehören die Rufnummer (Mobile Subscriber ISDN, MSISDN), vom Nutzer freigegebene Dienste wie Anrufweiterleitung und Roaming-Einstellungen sowie eine internationale Kennung (International Mobile Subscriber Identity, IMSI), die für die „eindeutige[..] Identifizierung des Netzteilnehmers und [für das] Routing innerhalb des Mobilfunknetzes“ [14] genutzt wird. Die dynamischen Daten sind der aktuelle Aufenthaltsort (Location Area, LA) und die aktuelle physikalische Telefonnummer MSRN (Mobile Subscriber Roaming Number) sowie die derzeitige Adresse

von VLR und MSC. Die physikalische Telefonnummer dient der Aufenthaltsbestimmung des Teilnehmers und setzt sich aus der Kennziffer des Landes und des Netzes sowie der zugehörigen MSC zusammen. Sobald eine MS ihren Aufenthaltsort verändert und somit die Zelle wechselt, werden die LA-Daten im HLR aktualisiert. Somit kann das HLR die Abrechnung von Roaming-Gebühren und von Verbindungsnachweisen unterstützen. Gleichzeitig ist es jedoch auch notwendig, dass die Datenbanken der HLR Echtzeitanforderungen unterstützen. [33, S. 137f.]
Die dritte Komponente ist das Besucherregister, eine hochdynamische Datenbank, in der die Daten der Mobilstationen gespeichert werden, die sich derzeit in der Location Area einer MSC aufhalten. Dazu gehören wiederum IMSI, MSISDN und außerdem die HLR-Adresse. Wenn eine Mobilstation in das LA einer MSC eintritt, werden die Daten aus der HLR in die VLR kopiert. Dadurch wird erreicht, dass die HLR-Daten nicht ständig aktualisiert werden müssen (nicht bei jedem Zellwechsel) und dass die Signalisierung von Teilnehmerdaten nicht fortwährend über teilweise weite Entfernungen erfolgen muss. [33, S. 138]

Base Station Subsystem (BSS)

Das Base Station Subsystem vereint alle Funktionen, die für eine permanente Funkverbindung zu einer Mobilstation nötig sind. In der BSS erfolgt die „Kodierung [bzw.] Dekodierung der Sprachdaten und die Anpassung der Datenraten vom [bzw.] zum drahtlosen Netz“ [33, S. 135]. Dies ist notwendig, da die GSM-Funkschnittstelle mit 13 kbit/s arbeitet, die Daten zwischen der Base Transceiver Station (BTS) und dem Base Station Controller (BSC) jedoch mit 64 kbit/s übertragen werden. [33, S. 135] Ein BSS besteht aus mehreren Komponenten. Es gibt mindestens eine Feststationssteuerung (Base Station Controller, BSC) und mehrere Sende- und Empfangsstationen (Base Transceiver Station, BTS).
Der Base Station Controller steuert mehrere Base Transceiver Stations. Zu seinen Aufgaben gehört die Reservierung von Funkfrequenzen und die Verbindungsübergabe zwischen verschiedenen BTSs innerhalb einer BSS (siehe Kapitel 4.2.3). Des Weiteren ist es für den Rundruf (Paging) einer MS zuständig und multiplext die Funkkanäle auf Festnetzverbindungen an der Schnittstelle zwischen dem BSC und dem MSC (A-Schnittstelle). [33, S. 135]
Die BTSs enthalten „alle funktechnischen Einrichtungen wie Antennen, Signalverarbeitung und Verstärker für die Übertragung“ [33, S. 135]. Jede BTS bildet eine Zelle, deren Größe von der Umgebungsbeschaffenheit und von der Anzahl der Mobilstationen in der Zelle abhängig ist. Die maximale Reichweite einer BTS beträgt 70 km im Durchmesser, diese wird allerdings durch die Sendeleistung der Mobilstationen eingeschränkt, sodass auf dem Land mehr als 15 km im Radius selten anzutreffen sind. In der Stadt kann die Reichweite sogar nur wenige 100 m betragen. Diese kleinen Zellen entstehen, weil eine BTS nur eine bestimmte Anzahl an Mobilstationen abdecken kann und diese Kapazitätsgrenze in Großstädten oder dicht besiedelten Gegenden schnell erreicht wird. Die Zellen

können in separate Sektoren unterteilt werden, wenn die BTS mehrere sektorisierende Antennen hat. [32, S. 29] Die Funkverbindung zwischen der BTS und der MS ist die sogenannte U_m-Schnittstelle, die alle nötigen Mechanismen für die drahtlose Übertragung beinhaltet. [33, S. 135]

Operation (Support) Subsystem (OSS)

Im Betriebs- und Wartungssystem (OSS) sind alle Funktionen für einen zuverlässigen Betrieb und die Wartung von GSM vereint. OSS kommuniziert mit allen anderen Subsystemen über das Signalisierungssystem Nr. 7 (SS7). [33, S. 138] Das OSS besteht aus drei verschiedenen Komponenten: die Betriebs- und Wartungszentrale (Operation and Maintenance Centre, OMC), die Authentifizierungszentrale (Authentification Centre, AuC) und das Geräteidentifikationsregister (Equipment Identity Register, EIR).
Das OMC ist für die Überwachung und Steuerung der Netzkomponenten zuständig. Die Kommunikation mit diesen erfolgt über die sogenannte O-Schnittstelle (siehe Abbildung 4.1). Zu den Aufgaben des OMCs gehören die Überwachung des Verkehrs, die Erstellung von Statusberichten einzelner Komponenten und das Sicherheitsmanagement. Des Weiteren ist das OMC zuständig für die Verwaltung der Teilnehmer sowie die Abrechnung und die Rechnungsstellung. [33, S. 138]
Das AuC ist für den Schutz der Identität eines Teilnehmers und der Datenübertragung verantwortlich. Diese Komponente ist im Mobilfunk notwendig, da die Luftschnittstelle und die mobilen Teilnehmer in Mobilfunknetzen einfacher angreifbar sind als in Festnetzen, weil beispielsweise geschirmte Kabel fehlen. Im AuC sind alle Schlüssel gespeichert und es erstellt alle Parameter, die für eine Teilnehmeridentifizierung im HLR notwendig sind. Da die zwei Komponenten HLR und AuC eng zusammenarbeiten, kann sich das AuC auch in einem speziell geschützten Teil des HLR befinden. [33, S. 138]
Der letzte Bestandteil des OSS ist das Geräteidentifikationsregister (EIR). Hier sind Datenbanken gespeichert, die alle IMEIs des jeweiligen Netzbetreibers enthalten. Wie schon erwähnt wurde, liegt hier auch eine schwarze Liste mit allen als gestohlen gemeldet oder anderweitig gesperrten Gerätekennungen. Es ist somit vollkommen aussichtslos, ein gestohlenes Gerät in dem Netz des gleichen Anbieters nutzen zu wollen. Leider werden die Daten zwischen den verschiedenen Netzbetreibern nicht immer verglichen, daher ist es möglich, gestohlene Geräte widerrechtlich in dem Netz eines anderen Betreibers zu nutzen. Neben der schwarzen Liste gibt es noch eine weiße und eine graue Liste; während die erstere alle gültigen IMEIs enthält, beinhaltet die graue Liste alle Geräte, die Fehlfunktionen aufweisen. [33, S. 138f.]

Luftschnittstelle

„Die interessanteste Schnittstelle in GSM aus Sicht der drahtlosen Kommunikation ist U_m“ [33, S. 129] (siehe Abbildung 4.1), der Übertragungsweg zwischen einer BTS und einem Mobilfunkteilnehmer, der auch als *Luftschnittstelle* oder Air Interface bezeichnet

wird. Da eine BTS im Bedarfsfall mit mehreren Teilnehmern gleichzeitig kommunizieren muss, wird das kombinierte Frequenz- und Zeitmultiplex verwendet (siehe Kapitel 2.2.2). Zur kabellosen Übertragung von Daten werden auf der Trägerfrequenz „Frames“ mit einer Zeitdauer von 4,615 ms versendet. Jeder dieser Frames hat acht „Zeitschlitze“ („Timeslots“). Daraus ergibt sich ein auch als „Burst“ bezeichnetes Zeitintervall für einen Timeslot von 577 μ. Wenn einem Endgerät z. B. der Timeslot Nr. 2 zugeordnet wird, „darf es in jedem Frame in diesem Timeslot senden und empfangen“ [32, S. 30], in der restlichen Zeit des Frames muss es warten. Anhand dieser Überlegung kann die Gesamtkapazität einer BTS abgeschätzt werden. Dazu verwendet Sauter folgendes Beispiel [32, S. 30f.]:
Die Zelle einer BTS wurde in drei unabhängige Sektoren unterteilt[12]. Jeder dieser Sektoren kann zwei Frequenzen nutzen. Da auf jeder Frequenz acht Timeslots zur Verfügung stehen, hat jeder Sektor somit $2 * 8 = 16$ Timeslots. Von diesen 16 Timeslots werden zwei für Signalisierungsaufgaben genutzt und vier oder mehr für die paketorientierte Übertragung GPRS (siehe Kapitel 4.2.4), somit bleiben für die Sprachübertragung pro Sektor zehn Timeslots erhalten, pro BTS sind es damit 30 Timeslots. Folglich können bis zu 30 Teilnehmer gleichzeitig kommunizieren. In der Praxis versorgt eine BTS allerdings weitaus mehr Mobilfunknutzer, da diese nicht alle zur gleichen Zeit auf die BTS zugreifen. „Mobilfunknetzbetreiber gehen davon aus, dass im Durchschnitt ein Teilnehmer pro Stunde 1 min telefoniert“ [32, S. 31]. Somit kann rechnerisch davon ausgegangen werden, dass die tatsächliche Anzahl der BTS-Nutzer 60-mal so groß ist, wie die Zahl, die zu einem bestimmten Zeitpunkt auf die BTS zugreift. In dem beschriebenen Beispiel würde die BTS somit in etwa 1800 Teilnehmer versorgen. [32, S. 30f.] Aus diesen Annahmen kann man wiederum die Anzahl der Basisstationen berechnen, die für einen Mobilfunknetzbetreiber notwendig sind, um alle Kunden zu versorgen. Diese Berechnungen sind allerdings nur grobe Näherungen, da nicht jede Basisstation die gleichen Kapazitäten abdeckt [32, S. 31].
Jeder Burst ist wiederum in verschiedene Bereiche eingeteilt. Diese Abschnitte enthalten die Nutzdaten, die Signalisierungsdaten und auch bestimmte Bereiche die zur Synchronisation und für Zeitpuffer genutzt werden, da sich einerseits in die Übertragung Fehler einschleichen können durch z. B. Reflexion, Absorption und Mehrwegeausbreitung (siehe Kapitel 2.3) und andererseits Teilnehmer mehr oder weniger weit von der BTS entfernt sein können und die Übertragungswege und somit auch -zeiten dadurch unterschiedlich lang sein können. [32, S. 31] Da sich diese Arbeit hauptsächlich mit der Zuverlässigkeit von Mobilkommunikation beschäftigt, soll hier nicht näher auf den Aufbau der Timeslots eingegangen werden.

[12] Dieses Vorgehen wird in der Praxis häufig angewendet, da auf diese Weise die zur Verfügung stehenden Frequenzen besser ausgenutzt werden können und die Effizienz einer BTS gesteigert wird [32, S. 29f.].

4.2.3 Verbindungsauf- und -abbau und Verbindungsübergabe (Handover)

Für den Aufbau einer Verbindung gibt es zwei verschiedene Szenarien. Zum einen kann eine Verbindung zustande kommen, wenn ein Teilnehmer ein Gespräch beginnen oder eine SMS verschicken möchte. Für den Aufbau der Verbindung schickt das Endgerät eine Channel-Request-Nachricht an den BSC. Der Base Station Controller überprüft, ob ein Signalisierungskanal (SDCCH) frei ist und antwortet nach dessen Aktivierung auf dem Access Grant Channel (AGCH) mit einer Immediate Assignment-Nachricht. Diese enthält unter anderem die Nummer des zugeteilten Signalisierungskanals. [32, S. 37] „Über die [...] aufgebaute Verbindung können nun DTAP-Nachrichten transparent zur MSC weitergeleitet werden“ [32, S. 37]. Das genannte Direct Transfer Application Part (DTAP) ist ein Übertragungsprotokoll, über das Nachrichten zwischen der MSC und den Endgeräten ausgetauscht werden [32, S. 11].
Die zweite Möglichkeit, eine Verbindung aufzubauen, ist die Anfrage von außen. In diesem Fall erhält der BSC bei einem ankommenden Gespräch oder einer SMS eine Paging-Nachricht von der mobilen Vermittlungsstelle. In dieser Nachricht enthalten sind die IMSI, die TMSI und die Location Area, in der sich der Empfänger gerade aufhält (siehe Kapitel 4.2.2). Die BSC hat eine Location Area-Datenbank, über die sie herausfinden kann, welche Zellen sich in der gewünschten Location Area befinden. Die Paging-Nachricht kann nun an diese Zellen weitergeleitet werden. Wenn das Endgerät die Nachricht erhalten hat, antwortet es dem BSC mit der Channel-Request-Nachricht, die schon bei dem ersten Verfahren zum Verbindungsaufbau verwendet wurde. Im Weiteren läuft der Verbindungsaufbau nun wie bei dem ersten Szenario ab. [32, S. 37f.]

Neben dem Aufbau einer Verbindung muss der Base Station Controller auch das Handover verwalten, die Übergabe der Verbindung von einer Zelle zur nächsten ohne Unterbrechung der Verbindung. Es gibt mehrere Arten der Verbindungsübergabe (siehe Abbildung 4.2). Je nachdem, welche Zellengrenzen überschritten werden, muss ein Wechsel der BTS, der BSC oder gar der MSC erfolgen. Außerdem kann bei bestimmten Bedingungen auch innerhalb einer Zelle eine Übergabe erfolgen. Daraus ergeben sich vier verschiedene Übergabearten [33, S. 153f.]:

- Intrazellenübergabe: Diese Übergabe beinhaltet nur den Wechsel der Trägerfrequenz. Wenn innerhalb einer Zelle schmalbandige Störungen auftreten, die nur einige Frequenzen betreffen, kann die BSC die Verbindung auf eine andere Trägerfrequenz legen.
- Interzellenübergabe (Intra-BSC-Übergabe): Diese Übergabeart ist „vermutlich das typischste Beispiel“ [33, S. 153]. Wenn eine MS eine Zellengrenze überschreitet, jedoch in der LA einer BSC verbleibt, führt die BSC einen Wechsel zwischen den BTSs durch.
- Inter-BSC-Übergabe (Intra-MSC-Übergabe): Wenn bei einem Zellenwechsel der Bereich einer BSC verlassen wird, übernimmt die MSC den Handover-Vorgang.

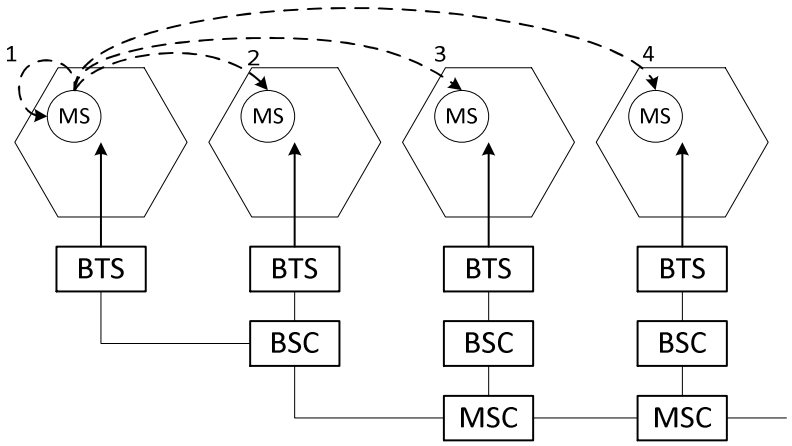


Abbildung 4.2: Übergabearten in GSM (nach [33, S. 153])

- Inter-MSC-Übergabe: Das vierte Handover-Verfahren ist der Wechsel zwischen Zellen, die von unterschiedlichen MSCs verwaltet werden. In diesem Fall steuern die MSCs das Handover gemeinsam.

Da die Interzellenübergabe nach Schiller [33, S. 153] die häufigste Übergabeart ist, soll das Handover an dieser Art erläutert werden.
Während des Bestehens der Verbindung werden fortwährend die Signalqualitäten und -stärken von Uplink und Downlink gemessen. Die Downlink-Qualität wird dabei vom Endgerät an die BSC übermittelt, die Qualität des Uplinks von der BTS. Um ein Endgerät an die Zellen mit der besten Signalqualität übergeben zu können, ist es zudem erforderlich, die Signalqualität der benachbarten Zellen zu ermitteln. Dazu sendet die BSC dem Endgerät die Frequenzen der benachbarten Zellen und das Endgerät kann die Qualitäten in Sendepausen messen und an die BSC überliefern. Wenn beispielsweise die Signalqualität in der aktuellen Zelle zu schlecht wird, weil sich der Teilnehmer zu weit von der BTS entfernt, übergibt die BSC das Endgerät an eine neue Zelle. Dazu wird als erstes in der neuen Zelle ein Traffic-Kanal aktiviert und anschließend ein Handover-Kommando an das Endgerät gesendet. In diesem Kommando sind unter anderem die neue Frequenz und die Timeslot-Nummer des zugewiesenen Kanals enthalten. Nun kann das Endgerät seine Sende- und Empfangsfrequenz ändern und sich mit der neuen BTS synchronisieren. Es „sendet in vier aufeinander folgenden Bursts des Timeslots eine Handover Access-Nachricht“ [32, S. 39] und im fünften Burst eine sogenannte SABM-Nachricht. Anschließend sendet die BTS eine Establish Indication-Nachricht an das BSC und eine sogenannte UA-Nachricht zum Endgerät, wenn sie das Handover korrekt erkannt hat. Nun kann die BSC die Verbindung in die neue Zelle schalten. Zum Schluss muss die BSC den Traffic-Kanal in der alten Zelle abbauen und das MSC über den erfolgreich ausgeführten Handover informieren. Diese Nachricht hat auf den weiteren Verlauf der Verbindungen keinen Einfluss. [32, S. 40]
Bei einer Inter-BSC-Übergabe sendet das alte BSC die Nachricht, dass ein Zellenwechsel erforderlich ist, an das MSC. Dieses fragt bei dem neuen BSC an, ob ein Timeslot zur Verfügung steht. [33, S. 155] Die weiteren Schritte erfolgen wie bei der Interzellenübergabe,

allerdings erfolgt die Kommunikation über die MSC.

4.2.4 Besonderheiten von GPRS und EDGE

Der direkte Nachfolger von GSM ist *High Speed Circuit Switched Data (HSCSD)*. Dieses Verfahren ist noch leitungsorientiert, konnte jedoch im Gegensatz zu GSM mehrere Kanäle zusammenfassen und somit höhere Datenraten erzielen. Eine MS fordert dazu nicht nur einen Zeitschlitz an, sondern mehrere. Dieses Verfahren bietet den Vorteil, dass es asymmetrisch sein kann, d. h., dass beispielsweise die Download-Richtung mehr Zeitschlitze hat als die Upload-Richtung. „Dies würde [...] dem typischen Verhalten eines Nutzers des Internets entsprechen, bei dem im Allgemeinen mehr Daten heruntergeladen als in das Netz geschickt werden“ [33, S. 159].
Neben den Vorteilen der Effizienz hat HSCSD jedoch immer noch die Nachteile der verbindungsorientierten Übertragung. Wenn eine MS einen Zeitschlitz belegt hat, kann dieser nicht mehr von anderen MSs genutzt werden. Bei der Nutzung von Webdiensten werden allerdings nicht fortwährend Informationen übertragen, sondern nur abschnittweise große Datenmengen, für die kurzzeitig mehrere Zeitschlitze benötigt werden. Die übrige Zeit bleiben die Timeslots ungenutzt. Ein zweites Problem bildet die Signalisierung der Zeitschlitze. Da nur die MS und die BSC „wissen“, dass mehrere Kanäle zusammengehören und diese auch zusammensetzen können, werden die Timeslots separat betrachtet. Dadurch wird die Wahrscheinlichkeit größer, dass „bei einer Übergabe eine oder mehrere Verbindungen blockiert werden und damit die Bandbreite sinkt“ [33, S. 160]. In Anbetracht dieser Vor- und Nachteile bezeichnet Schiller HSCSD als „Interimslösung“, die viele Betreiber überspringen und stattdessen die nächste Ausbaustufe, General Packet Radio Service (GPRS), anwenden würden [33, S. 160f.].

Der entscheidendste Unterschied zu GSM und HSCSD ist, dass die Übertragung in GPRS paketorientiert erfolgt. GPRS nutzt die gleichen Zeitschlitze wie die ursprünglichen Verfahren, der Zugriff erfolgt jedoch nicht über eine zeitlich begrenzte Reservierung einer einzelnen MS, sondern wird bedarfsgerecht verwaltet, d. h. mehrere Nutzer können gleichzeitig auf einen Timeslot zugreifen. So wird laut Schiller oft „zumindest ein Zeitschlitz pro Zelle für GPRS [von den Betreibern] reserviert, um [...] in Hochlastzeiten zumindest noch eine minimale Datenrate zu gewährleisten“ [33, S. 161], dieser Timeslot muss allerdings auf alle Nutzer aufgeteilt werden. Um diese bedarfs- und paketorientierte Übertragung umsetzen zu können, sind allerdings nicht nur eine neue Software für die Netzwerkkomponenten nötig, sondern auch zusätzliche Hardwarekomponenten. [33, S. 161]
Diese neuen Elemente, die „im Prinzip Router sind“ [33, S. 164], werden als *GPRS Support Nodes (GSN)* bezeichnet und es werden zwei verschiedene benötigt. Zum einen gibt es den *Gateway GPRS Support Node (GGSN)*, der die „Verbindung[..] zwischen dem GPRS-Netz und externen *paketorientierten Netzen (Packet Data Network, PDN)*“ [33, S. 164] darstellt. Der Knoten ist zuständig für die Wandlung von Adressdaten und

Kodier-schema	Anzahl Schlitze							
	1	2	3	4	5	6	7	8
CS-1	9,05	18,2	27,15	36,2	45,25	54,3	63,35	72,4
CS-2	13,4	26,8	40,2	53,6	67	80,4	93,8	107,2
CS-3	15,6	31,2	46,8	62,4	78	93,6	109,2	124,8
CS-4	21,4	42,8	64,2	85,6	107	128,4	149,8	171,2

Tabelle 4.1: GPRS Nutzdatenraten in kbit/s (nach [33, S. 162])

für die Tunnelung von Nutzdaten zu einem Teilnehmer unter Zuhilfenahme der Datenkapselung. Außerdem enthält er die Wegwahldaten für GPRS-Teilnehmer und überträgt Daten über ein internes Verbindungsnetz an die *Serving GPRS Support Nodes (SGSN)*. Dieser SGSN unterstützt eine MS via BSS. Er „erfragt Teilnehmeradressen vom *GPRS Register (GR)*, verfolgt den Aufenthaltsort von Teilnehmern, ist die für die Sammlung von Abrechnungsdaten verantwortlich (z.B. durch Zählen der Bytes) und führt verschiedene Sicherheitsdienste (z. B. Zugangskontrolle) durch“ [33, S. 165]. Das GR, das Bestandteil des HLR sein kann, ist für die Speicherung aller GPRS-relevanten Daten zuständig.

GPRS kann in Abhängigkeit der Anzahl der zur Verfügung stehenden Zeitschlitze Datenraten bis zu 171,2 kbit/s (acht Zeitschlitze, keine Fehlerkorrektur) erreichen [33, S. 162]. Neben der Anzahl der Timeslots spielt auch die Kodierung der Daten eine Rolle[13]. Schiller fasst unter Berücksichtigung von Kodierschema und Zeitschlitzen die Übertragungsgeschwindigkeiten in Tabelle 4.1 zusammen [33, S. 162]. Die Zahl der Zeitschlitze, die zur Verfügung steht, ist stark abhängig von der Anzahl der MSs, die gerade mit der BTS kommunizieren, da GPRS lediglich freie Timeslots verwenden kann.
Neben der Anzahl Zeitschlitze und der Kodierung hängt die Datenrate auch von der Geräteklasse der MS ab. Nur wenige Geräte können gleichzeitig senden und empfangen. Des Weiteren können sie nur eine bestimmte Anzahl von Zeitschlitzen nutzen. Nach Schiller haben heutige Geräte eine „Empfangsdatenrate von maximal 53,6 kbit/s und eine Sendedatenrate von 13,4 kbit/s“ [33, S. 162], da sie der Geräteklasse 8 angehören und CS-2 nutzen (siehe Tabelle 4.2).

[13] Es existieren für GPRS vier *Coding Schemes (Kodierungsverfahren)* [32, S. 84f.]. Sie unterscheiden sich in der Anzahl der zulässigen Fehler, die übertragen werden dürfen. Bei der Kodierung der Daten entstehen Fehlerkorrekturbits, die jedoch aufgrund einer Bitbegrenzung nicht alle übertragen werden. Da dem Empfänger bekannt ist, welche Fehlerkorrekturbits nicht gesendet wurden, kann er die entsprechenden Bits „punktieren“ und als Fehler betrachten. Da bei steigender CS-Klasse mehr Fehlerkorrekturbits nicht übertragen werden, dürfen gleichzeitig auch weniger Fehler auftreten. Vor allem CS-1 und CS-2 werden in allen Netzwerken verwendet, CS-3 und CS-4 sind hingegen weniger weit verbreitet. Da bei einer Verwendung von CS-3 und CS-4 für eine Steigerung der Übertragungsgeschwindigkeit ein Ausbau des Transportnetzwerkes erforderlich wäre, nutzen viele Betreiber EDGE, das eine Geschwindigkeitssteigerung durch ein anderes Modulationsverfahren erreicht (siehe nachfolgende Abschnitte). [32, S. 84f.]

Klasse	Empfangsschlitze	Sendeschlitze	Maximale Anzahl an Schlitzen
1	1	1	2
2	2	1	3
3	2	2	3
5	2	2	4
8	4	1	5
10	4	2	5
12	4	4	5

Tabelle 4.2: Beispiele für GPRS-Geräteklassen [33, S. 163]

Es ist mit GPRS möglich, einzustellen, ob bei der Übertragung mehr Wert auf die „*Rangfolge des Dienstes* (service precedence: hoch, normal, niedrig), die *Zuverlässigkeitsklasse* (reliability class) und *Verzögerungsklasse* (delay class) [...] [oder] die *Nutzdatenrate* (user data throughput)“ [33, S. 163] Wert gelegt werden soll. Dazu kann ein Nutzer von GPRS ein sogenannten *QoS- (Quality-of-Service-, Dienstgüte-) Profil* anlegen. GPRS muss die zur Verfügung stehenden Ressourcen dermaßen verwalten, dass die Nutzeranforderungen bestmöglich erfüllt werden. [33, S. 163]
Es existieren für GPRS drei Zuverlässigkeitsklassen, die die Wahrscheinlichkeiten für Fehlerübertragungen beinhalten. Diese fehlerhaften Übertragungen teilen sich in vier Arten – Service-Data-Unit- (SDU-) Verlust, SDU-Duplizierung, SDU-Vertauschung und SDU-Veränderung bei Übergabe an eine höhere Schicht [33, S. 163]. Die jeweiligen Wahrscheinlichkeiten für die einzelnen Zuverlässigkeitsklassen sind in Tabelle 4.3 [33, S. 163] angegeben. Aus den Werten ist ersichtlich, dass die Klasse 3 mit den größten Werten verwendet werden kann, wenn Anwendungen ein sehr gutes Verfahren zur Fehlerbeseitigung haben oder fehlertolerant sind. Die Klasse 1 hingegen, die nur eine sehr kleine Fehlerübertragung hat, ist für Anwendungen geeignet, die intolerant für Fehler sind.
Verzögerungsklassen sind in GPRS vier definiert. Die Werte setzen sich aus der „Zugriffsverzögerung auf das Medium [...] der Kodierung zum Fehlerschutz [und] der Übertragungsverzögerung über den festen und drahtlosen Teil des Netzes“ [33, S. 163] zusammen. Die Verzögerungszeiten nach Schiller sind in Tabelle 4.4 [33, S. 164] aufgezeigt. Es ist deutlich zu erkennen, dass die Zeiten „um Größenordnungen über denen von herkömmlichen Festnetzen“ [33, S. 164] liegen. Bei Anwendungen und Sicherheitsfragen muss diese Eigenschaft beachtet werden, damit das Netz auf die Anforderungen der Nutzer (z.B. Nutzung „interaktiver Internetanwendungen“ [33, S. 164]) angepasst werden kann.

Die nächste Ausbaustufe nach GPRS ist *Enhanced Data Rates for GSM Evolution (EDGE)*. Mittels des 8-PSK-Verfahrens (siehe Kapitel 2.2.1) soll die Übertragungsgeschwindigkeit weiter gesteigert werden. Bei 8-PSK werden gleichzeitig drei Bit übertragen. Außerdem wurden für EDGE neue Kodierungsverfahren festgelegt (siehe Tabelle 4.5).

Zuverlässigkeitsklasse	SDU-Verlust	SDU-Duplikate	Vertauschte SDUs	Veränderte SDUs
1	10^{-9}	10^{-9}	10^{-9}	10^{-9}
2	10^{-4}	10^{-5}	10^{-5}	10^{-6}
3	10^{-2}	10^{-5}	10^{-5}	10^{-2}

Tabelle 4.3: Zuverlässigkeitsklassen in GPRS nach ETSI (1998c) [33, S. 163]

Verzögerungsklasse	SDU-Größe 128 byte		SDU-Größe 1024 byte	
	Durchschnitt	95%-Quantil	Durchschnitt	95%-Quantil
1	$< 0,5s$	$< 1,5s$	$< 2s$	$< 7s$
2	$< 5s$	$< 25s$	$< 15s$	$< 75s$
3	$< 50s$	$< 250s$	$< 75s$	$< 375s$
4	nicht spezifiziert			

Tabelle 4.4: Verzögerungsklassen in GPRS nach ETSI (1998c) [33, S. 164]

Damit können Übertragungsgeschwindigkeiten bis zu knapp 60 kbit/s erreicht werden.
Des Weiteren bildet auch die Anzahl der für EDGE vorhandenen *Modulation and Coding Schemes (MCS)* einen Vorteil. Durch die neun MCS-Klassen besteht die Möglichkeit, exakt auf die aktuelle Übertragungsqualität zu reagieren und das entsprechende Modulations- und Kodierungsschema zu verwenden. Netzwerk und Endgerät informieren sich „ständig gegenseitig über die Signalqualität beim Empfang der vorhergehenden Datenpakete“ [32, S. 86], somit ist es möglich, schnell auf sich verändernde Bedingungen zu reagieren. Durch dieses Verfahren kann einerseits die Fehlerrate gesenkt werden, andererseits können bei sehr guten Übertragungsbedingungen auch MCS-8 und MCS-9 verwendet werden, die hohe Übertragungsgeschwindigkeiten bieten. [32, S. 86]
EDGE bietet auch bei der Fehlerkorrektur einen gegenüber GPRS erweiterten Standard. Wie bei GPRS (siehe Fußnote 13) schon erwähnt, werden nicht alle Fehlerkorrekturbits, die berechnet wurden, auch gesendet. EDGE bietet ein Verfahren namens *Incremental Redundancy*, das bei einem Übertragungsfehler nicht das komplette Paket neu überträgt, sondern nur die nicht gesendeten Fehlerkorrekturbits. Somit stehen am Empfänger mehr Bits zur Verfügung und der Fehler kann mit einer höheren Wahrscheinlichkeit korrigiert werden. [32, S. 87]
Ein zweites Verfahren für die Korrektur einer fehlerhaften Übertragung ist die sogenannte *Re-Segmentation*. Dabei wird das Paket, das zuvor mit einer hohen MCS-Klasse gesendet wurde, in zwei Pakete aufgeteilt, die eine kleine MCS-Klasse haben. [32, S. 87]
Die dritte Möglichkeit, einen Übertragungsfehler zu korrigieren, besteht in dem soge-

	Modulation	Geschwindigkeit pro Timeslot
MCS-1	GMSK	8,8 kbit/s
MCS-2	GMSK	11,2 kbit/s
MCS-3	GMSK	14,8 kbit/s
MCS-4	GMSK	17,6 kbit/s
MCS-5	8-PSK	22,4 kbit/s
MCS-6	8-PSK	29,6 kbit/s
MCS-7	8-PSK	44,8 kbit/s
MCS-8	8-PSK	54,4 kbit/s
MCS-9	8-PSK	59,2 kbit/s

Tabelle 4.5: Übertragungsgeschwindigkeiten und Modulationsverfahren von EDGE [32, S. 85]

nannten *Interleaving*. Bei diesem Verfahren, das schon bei GPRS existierte, für EDGE jedoch erweitert wurde, werden die Bits gemischt, um „punktuelle Übertragungsfehler über den Block zu streuen“[14] [32, S. 87]. Bei EDGE wird ein Datenblock im Gegensatz zu GPRS nicht über vier Bursts versendet, sondern nur über zwei. Daher müssen bei einer Fehlerübertragung nur zwei Bursts wiederholt werden. [32, S. 87]
Eine weitere Entwicklung von der 3GPP-Organisation, um die Datenübertragungsgeschwindigkeit zu steigern, ist *Envolved-EDGE*. Dabei sollte einerseits die Latenzzeit verringert werden, andererseits durch das Endgerät eine Bündelung von zwei Frequenzkanälen erfolgen. Dadurch wäre es möglich, bei einer Übertragung Timeslots auf zwei Kanälen zu verwenden und somit die Übertragungsgeschwindigkeit zu steigern. Dieses Verfahren hat allerdings den Nachteil, dass es aus Netzwerksicht keine nennenswerte Geschwindigkeitssteigerung erzielt, da die Anzahl der Frequenzkanäle pro Basisstation begrenzt ist. Nach Sauter wäre es „somit fraglich, ob Netzbetreiber in Zukunft diese Verfahren verwenden werden“ [32, S. 124].

4.2.5 Routing

Wenn ein Paket verschickt wird, stehen meist mehrere Wege offen, die zum Empfänger führen. Das Verfahren, die Wege möglichst effektiv zu wählen, wird als *Routing* bezeichnet. Der Router arbeitet auf der Vermittlungsschicht (siehe Kapitel 3.1.1). Um den günstigsten Weg zu finden und die zur Verfügung stehenden Ressourcen möglichst

[14] Beim Interleaving werden die kodierten Bits auf der Senderseite „gemischt“, d. h., auf einen ganzen Block (vier Bursts) verteilt. Auf der Empfängerseite wird dieser Vorgang rückgängig gemacht. Dies hat den Vorteil, dass bei der Übertragung auftretende Fehler, die sich meist über mehrere Bits ziehen, anschließend durch das „Entmischen“ der Bits auf den ganzen Block verteilt werden (siehe auch [32, S. 48f.]).

effektiv zu verteilen, stehen verschiedene Routing-Algorithmen zur Verfügung. Zwischen einem Sender und dem Empfänger existieren in den meisten Fällen mehrere Router, die jeweils neu entscheiden, welchen Weg ein Paket nehmen soll. Dabei wird der neue Weg nicht unter Berücksichtigung von Quelladresse und dem bisherigen Weg gesucht, sondern nur anhand der Zieladresse gewählt. Diese Eigenschaft wird als „Quellenunabhängigkeit“ bezeichnet [22, S. 390].
Zur Ermittlung der Wege existieren Routing-Tabellen. Diese Tabellen sollten immer aktuell sein, damit ein optimaler Weg gefunden werden kann. Riggert teilt den Begriff „optimal“ in folgende vier Merkmale [30, S. 25]:

- gleichmäßige Lastverteilung
- hoher Datendurchsatz
- gebührenminimal
- Sicherheit

Je nach Art ihrer Erstellung können Routing-Tabellen statisch (auch nicht adaptiv) oder dynamisch (auch als adaptiv bezeichnet) sein. Die statischen Tabellen werden manuell eingegeben und bleiben anschließend erhalten. Diese Art der Wegewahl bietet einige Vorteile, da „keine aufwendigen Aktualisierungen von Paketleitwegen notwendig“ [30, S. 27] sind und sich die Fehlersuche einfach gestaltet, weil die Wege der Pakete bekannt sind. Statische Routing-Verfahren haben allerdings den Nachteil, dass sie sich nicht der Netzwerktopologie anpassen, wenn beispielsweise Netzknoten ausfallen oder neue hinzukommen. Daher ist ein Einsatz nur sinnvoll, wenn der Weg immer bekannt ist (beispielsweise, wenn nur ein Nachbarrouter existiert) [36, S. 421].
Bei den adaptiven Routing-Verfahren aktualisieren sich die Routing-Tabellen in Abhängigkeit der jeweiligen Netztopologie und reagieren somit beispielsweise auf Ausfälle von Netzwerkkomponenten. Die Tabellen enthalten die möglichen Endknoten und den jeweils nächsten Router, den das Paket anwählen muss. Dabei ist es auch möglich, dass bei unterschiedlichen Endknoten die jeweils nächsten Router die gleichen sind, sich die Einträge somit gleichen. In diesen Fällen können die Tabellen verkürzt werden mittels des *Vorgabe-Routings (Default Routing)*, das in der Routing-Tabelle mit der geringsten Priorität abgespeichert wird. Wenn der Router in der Routing-Tabelle keinen Eintrag für die Zieladresse eines Paketes findet, schickt er das Paket an die Default-Adresse weiter. [22, S. 394] Es gibt für das adaptive Routing mehrere Verfahren, die meist den kürzesten Weg suchen [32, S. 557], die hier nicht alle näher beleuchtet werden sollen (siehe auch Tanenbaum und Wetherall (2012) und Meinel und Sack (2004) [22, S. 391-405, 36, S. 420-438]).

Broadcast- und Multicast-Routing

Beim Broadcasting wird ein Paket an alle Empfänger gesendet. Für das Routing existieren mehrere Methoden.

Am häufigsten eingesetzt wird ein Verfahren, bei dem die Quelle ein Paket an alle Ziele schickt, obwohl dieses Verfahren sehr langsam ist und viel Bandbreite benötigt. Außerdem muss die Quelle alle Ziele gespeichert haben. Allerdings sind die Anforderungen an das Netzwerk nicht sonderlich hoch. [36, S. 439]
Eine Alternative stellt das *Multidestination-Routing* dar. Dabei wird ein Paket an eine bestimmte Anzahl von Zielen versendet. Der Router, der das Paket mit diesen Zielen empfängt, verschickt das Paket über alle Ausgänge, die mindestens eines der Ziele abdecken. Dabei werden jedoch Kopien der Pakete versendet, die nur noch die Ziele im Header enthalten, die durch den jeweiligen Router-Ausgang erreicht werden. Wenn dieser Vorgang ausreichend oft wiederholt wird (nach einer bestimmten Anzahl von Wegabschnitten), haben die Pakete nur noch ein Ziel im Header stehen und können wie „ganz normale[..] Paket[e]“ [36, S. 439] behandelt werden. Der Vorteil dieses Verfahrens liegt in der Auslastung der Bandbreite. Die Pakete werden zwar durch die Router wie einzeln adressierte Pakete behandelt, „bei mehreren Paketen auf dem gleichen Pfad [muss jedoch nur] eines davon den vollen Fahrpreis [...] zahlen [...], während die übrigen umsonst mitreisen“ [36, S. 439]. Allerdings muss die Quelle auch bei dieser Methode die Liste der Ziele gespeichert haben und die Router haben weiterhin den gleichen Aufwand wie bei einzelnen Paketen.
Eine dritte Methode für Broadcast-Routing stellt das *Fluten* dar. Dabei wird ein Paket über jeden Ausgang versendet, außer dem, über den es den Router erreicht hat. Um eine unendliche Anzahl von Paketen zu vermeiden, wird dabei einerseits ein Teilstreckenzähler eingesetzt, der bei jedem Router dekrementiert wird bis er null erreicht. Dann wird das Paket gelöscht. Andererseits merken sich die Router, welche Pakete sie schon verschickt haben, damit Pakete, die sie noch einmal erreichen, nicht zum zweiten Mal versendet werden. Allerdings wird auch bei diesem Verfahren unter Umständen sehr viel Bandbreite und Rechenzeit benötigt. [36, S. 439]
Ein weiteres Verfahren für Broadcast-Routing ist das *Reverse Path Forwarding* (Weiterleitung auf dem umgekehrten Pfad). Wenn ein Router ein Paket erhält, überprüft er, ob der Eingangspfad der gleiche Pfad ist, auf dem normalerweise Pakete versendet werden, die an den Absender des gerade erhaltenen Paketes adressiert sind. Wenn dies der Fall ist, wird das erhaltene Paket über alle Kanäle verschickt mit Ausnahme des Ausgangs, über den das Paket erhalten wurde. Anderenfalls wird das Paket verworfen. Das Verfahren ist zum einen effizient und zum anderen leicht zu implementieren. Es basiert im Wesentlichen auf dem Fluten-Algorithmus, jedoch muss ein Router sich nicht merken, welches Paket ihn schon einmal erreicht hat, sondern nur wissen, welche Wege üblicherweise verwendet werden. [36, S. 439f.]
Ein drittes Verfahren basiert auf der Berechnung von Spannbäumen eines Netzwerkes. Wenn ein Router den Spannbaum seines Netzes kennt, kann er ankommende Pakete nur über die Kanäle weiterleiten (außer an denjenigen Ausgang, an dem das Paket angekommen ist), die zu einem Spannbaum gehören. Dieses Verfahren „nutzt die Bandbreite hervorragend, da nur die absolute Mindestanzahl an Paketen erzeugt wird“ [36, S. 441]. Dieses Verfahren ist nur in Abhängigkeit von dem verwendeten Routing-Algorithmus anwendbar, da nicht jeder Algorithmus die Spannbäume von Netzwerken

berechnet. [36, S. 441]

Ähnlich dem Broadcasting, jedoch nur auf eine ausgewählte Menge der Ziele bezogen, ist das *Multicasting*. Dieser Fall wird verwendet, wenn nur ein bestimmter Teil der Netzteilnehmer Informationen beziehen möchte oder darf, der zahlenmäßig zwar sehr groß ist, jedoch nur einem kleinen Anteil des Netzes entspricht, beispielsweise bei Multiplayer-Spielen, Internet-Fernsehen oder Telefonkonferenzen. Voraussetzung für das Multicast-Routing ist, dass jeder Router Multicast-Gruppen erstellen, bearbeiten und löschen kann. Es gibt dafür verschiedene Methoden, auf die hier aus thematischen Gründen nicht näher eingegangen werden soll. Multicast-Routing-Methoden bauen auf dem Broadcast-Routing auf. Besonders wenn eine Multicast-Gruppe sehr dicht ist, d. h., die „Mitglieder“ über das ganze Netz verteilt sind, bietet sich Broadcast-Routing an, da somit alle Teilnehmer des Netzes erreicht werden. Allerdings werden auch in diesem Fall die Pakete an Teilnehmer gesendet, die nicht in der Multicast-Gruppe sind, und dadurch werden Ressourcen unnötig verbraucht. Dieses Problem kann mittels eines Verfahrens von Deering und Cheriton [7, S. 99f.] gelöst werden. Sie benutzen den Broadcast-Spannbaum und löschen aus ihm alle Verbindungen, die nicht zu Teilnehmern der Multicast-Gruppe führen. Für das Entfernen der nicht benötigten Verbindungen gibt es mehrere Verfahren, Tanenbaum und Wetherall nennen dafür beispielsweise das *Multicast Open Shortes Path First (MOSPF-) Protokoll* und das *Distance Vector Multicast Routing Protocol (DVMRP)* [36, S. 442f.]. Der Nachteil dieses Verfahrens liegt im Speicheraufwand, da jeder Router für jede Gruppe einen Spannbaum erstellt. Nach Tanenbaum und Wetherall wären das mn Bäume, wenn ein Netz n Gruppen mit durchschnittlich m Knoten hat [36, S. 443]. Einen Ausweg bietet der *Core-Based Tree* (wurzelbasierter Baum). Tanenbaum und Wetherall [36, S. 443] beschreiben dabei das Verfahren von Ballardie et al. [3]. Für alle Router einer Gruppe wird ein einziger Spannbaum berechnet. Ein Knoten wird als sogenannter *Core* oder *Rendevouz-Punkt* festgelegt, der die Wurzel des Baumes darstellt. Jeder Router sendet ein Paket an den Core, die dabei zurückgelegten Wege bilden den Baum. Wenn ein Knoten ein Paket an die Gruppe senden möchte, wird das Paket an die Wurzel des zugehörigen Spannbaumes versendet. Das Paket muss den Core jedoch nicht zwangsläufig erreichen. Sobald es einen Knoten des Spannbaumes erreicht, wird es auf allen Wegen, die zum Spannbaum gehören, weiter verschickt. [36, S. 443]

Routing für mobile Hosts

Für Endgeräte, die ständig ihren Standort wechseln können, beispielsweise Laptops, ergibt sich ein Problem für die Routing-Verfahren. Die Geräte müssen für eine Übertragung neu geortet werden, wenn sie ihren Standort verlassen haben. Die Geräte bekommen einen ständigen *Heimatstandort* (home location) zugewiesen, der sich nie ändert [36, S. 445] Außerdem erhält jeder der sogenannten Hosts (Geräte) eine Heimatadresse, die zur Ermittlung des Heimatstandortes genutzt wird. Wenn das mobile Gerät seinen Standort verändert, teilt es einem Host an seinem Heimatstandort, dem *Heimatagenten* (home agent) seine aktuelle Adresse mit, die *Care-of-Adresse* genannt

wird. [36, S 446] Wenn der mobile Host ein Paket an seine Heimatadresse gesendet bekommt, wird es durch den Heimatagenten abgefangen, „weil der mobile Host nicht zu Hause ist“. [36, S. 447]. Der Heimatagent *verkapselt* das Paket und vergibt einen neuen Header mit der derzeitigen Care-of-Adresse des mobilen Hosts. Dieses Verfahren wird als *Tunneling* bezeichnet. [36, S. 447] Der mobile Host entpackt das Paket und erhält somit das ursprüngliche, an ihn versendete Paket. Die Antwort wird direkt an den ursprünglichen Sender adressiert, der somit die aktuelle Care-of-Adresse erhält. Die nachfolgende Kommunikation kann somit über den direkten Weg erfolgen und benötigt den Umweg über den Heimatagenten nicht mehr. Wenn der mobile Host seinen Standort erneut wechselt und eine neue Care-of-Adresse erhält, muss für die Pakete wiederum der Weg über den Heimatagenten gewählt werden. Die Gesamtroute (Weg über Heimatagenten an mobilen Host und Direktverbindung zu der Care-of-Adresse) wird als Triangle-Routing bezeichnet, „da sie einen Umweg darstellt, falls der entfernte Standort weit vom Heimatstandort entfernt ist“ [36, S. 447].
Das Verfahren bietet gegenüber dem herkömmlichen Routing den Vorteil, dass bei einem Standortwechsel eines mobilen Gerätes nicht alle Routing-Tabellen neu berechnet werden müssen. Bei einer Topologieveränderung, wie sie bei einem Standortwechsel verursacht wird, muss bei den bisher vorgestellten Verfahren eine Aktualisierung der Routing-Tabellen erfolgen.

Routing in Ad-hoc-Netzen

Gegenüber dem Routing bei mobilen Hosts weist das Routing in Ad-hoc-Netzwerken ein weiteres Problem auf. Alle Mobilgeräte, die sich in einem Ad-hoc-Netz befinden, fungieren gleichzeitig als Nutzer und als Router. Daher ist es möglich, dass Router genau wie die mobilen Hosts ihren Standort wechseln können, wegfallen oder neu hinzukommen. Aus diesen Gründen ist es erforderlich, dass Routen in einem Ad-hoc-Netzwerk bei Bedarf berechnet werden. Ein Vorausberechnen der Routen ist nicht effizient, da sich die berechneten Wege schon geändert haben können, wenn sie benötigt werden. Nach Tanenbaum und Wetherall existieren „viele, viele Routing-Algorithmen“ [36, S. 448], bei denen jedoch nicht geklärt ist, welches der Verfahren am geeignetsten ist [36, S. 448].
Tanenbaum und Wetherall beschreiben einen Ad-hoc-Routing-Algorithmus, welcher von ihnen als einer der bekanntesten angesehen wird [36, S. 448], den *Ad hoc On-demand Distance Vector (AODV, Ad-hoc-Distanzvektor nach Bedarf)* von Perkins und Royer [27]. Der Algorithmus ermittelt nach Bedarf einen Pfad vom Sender- zum Empfängerknoten. In jedem Knoten des Netzwerkes ist eine Distanzvektortabelle gespeichert, die nach dem jeweils gewünschten Ziel indiziert ist und denjenigen Nachbarknoten enthält, an den das Paket zunächst gesendet werden muss. Ist bei einem Sendewunsch kein Eintrag zu dem gewünschten Zielknoten vorhanden, verschickt der Sender ein sogenanntes „Route Request“-Paket mittels Broadcast. Jeder nachfolgende Knoten wird die Nachricht weiter versenden bis der Zielknoten erreicht wird. Dieser erstellt eine „Route Reply“-Nachricht und versendet sie auf dem umgekehrten Pfad, auf dem er die Request-

Nachricht erhalten hat. Der Sender weiß mit Erhalt der Reply-Nachricht den günstigsten Weg zu dem Zielknoten. Gleichzeitig speichern auch alle Zwischenknoten auf dem ermittelten Pfad den Weg zu dem Zielknoten in ihrer Distanzvektortabelle ab. Bei der Ermittlung des besten Weges können viele Broadcast-Sendungen entstehen. Um die Anzahl zu reduzieren, bekommen die Nachrichten eine Lebenszeit („Time to Live“), die durch einen Zähler realisiert wird, der dekrementiert wird durch jeden Knoten, den er passiert. Wenn der Sender innerhalb einer fest gelegten Zeit keine Antwort erhält, versendet er die Request-Nachricht nochmals mit erhöhter Lebenszeit. Dieser Prozess wird solange wiederholt, bis der Sender eine Antwort durch den Empfänger erhält. [36, S. 449f.]
Berücksichtigt werden muss bei diesem Verfahren, dass die einzelnen Knoten nicht erkennen, wenn ein anderer Knoten ausfällt und Wege daher nicht mehr verfügbar sind. Um dieses Problem zu vermeiden, senden alle Knoten in regelmäßigen Abständen „Hello“-Nachrichten eines Nachbarknotens an alle ihre Nachbarn mittels Broadcast. Wenn ein Knoten mehrere „Hello“-Nachrichten nicht erhalten hat, weiß er, dass sich dieser Knoten nicht mehr in seinem Funkbereich aufhält oder ausgefallen ist. Er löscht jetzt alle Zielknoten in seiner Distanzvektortabelle, die über diesen ehemaligen Nachbarknoten erreicht wurden. Außerdem informiert er alle sonstigen Nachbarknoten über den Ausfall des anderen Knotens, die nun ihrerseits alle Einträge in ihren Tabelle löschen, die diesen Knoten verwendeten und wiederum ihre eigenen Nachbarn informieren. Durch diesen Domino-Effekt werden die Einträge, die den ehemaligen Knoten verwenden, in allen Distanzvektortabellen gelöscht. [36, S. 450f.]
Einen weiteren Vorteil, den AODV bietet, ist das Verwenden der vorhandenen Routen zum Erstellen neuer Pfade. Wenn ein Sender den Pfad zu einem Zielknoten ermitteln möchte und die Request-Nachricht erreicht einen Knoten, der einen Eintrag für dieses Ziel hat, teilt dieser Knoten dem Sender den vorhandenen Pfad mit. Das Broadcast kann somit an dieser Stelle in diese Richtung abgebrochen werden. [36, S. 451]
Neben dem beschriebenen AODV-Verfahren erwähnen Tanenbaum und Wetherall weitere Ad-hoc-Routing-Verfahren, auf die hier jedoch nicht weiter eingegangen werden soll. Welches dieser Routing-Verfahren sich in der Praxis bewähren wird, hängt von der Art des Ad-hoc-Netzes ab, das sich in der Praxis durchsetzen wird. [36, S. 451]

4.3 Universal Mobile Telecommunications System (UMTS)

4.3.1 Übersicht und Standardisierung

UMTS ist eine Weiterentwicklung von GSM und GPRS. Es „vereinigt die Eigenschaften eines leitungsvermittelnden Sprachnetzwerkes mit denen eines paketvermittelnden Datennetzwerkes“ [32, S. 127] und liefert damit viele neue Möglichkeiten gegenüber früheren Technologien. Viele Eigenschaften wurden von GSM und GPRS übernommen.

Daher werden die folgenden Kapitel sich nicht in aller Ausführlichkeit mit UTMS beschäftigen, sondern sich auf Neuerungen und Verbesserungen gegenüber dem Vorgänger beziehen. Des Weiteren spielt in der Car-to-X-Kommunikation v. a. LTE eine entscheidende Rolle, welches wiederum eine Weiterentwicklung von UMTS ist.
Die Standardisierung von UMTS wird durch die 3GPP-Organisation vorangetrieben. Alle Neuerungen werden als sogenannte *Releases* (Versionen) etwa alle 18 Monate veröffentlicht. [32, S. 128] Die erste Version, das „Release 99“ (R99), spezifiziert die erste Stufe von UMTS. Dazu gehört u. a. das Zugangsnetzwerk *UMTS Terrestrial Radio Access Network (UTRAN)* (siehe Abschnitt 4.3.2) mit erneuerter Identifizierung (anstatt des Frequenz- und Zeitmultiplex wird *Wideband Code Division Multiple Access (W-CDMA)* verwendet) und vergrößerter Bandbreite auf der Luftschnittstelle. Als Frequenzband wurde der Bereich von 1920-1980 MHz (Uplink) und 2110-2170 MHz (Downlink) genutzt. [32, S. 129]
Das „Konzept der Basisstationen und übergeordneten Controllern“ [32, S. 129] wurde hingegen von GSM übernommen. Allerdings wurden neue Bezeichnungen für die BTS und BSC gewählt – Node B und Radio Network Controller (RNC). Die Mobilstation (Mobile Station, MS) wurde in User Equipment (UE) umbenannt.
Die Kernnetze von GSM und GPRS wurden bis auf wenige Softwareänderungen und die Schnittstellen zum Zugangsnetzwerk nicht verändert. Daher soll hier nicht weiter darauf eingegangen werden (für nähere Informationen [32, S. 129f.]).
In Release 4 wurde „als wichtigste Neuerung“ [32, S. 130] das Kernnetz, das seit GPRS teilweise leitungsvermittelt, teilweise paketvermittelt arbeitet, auf vollständige IP-Übertragungen umgestellt. Diese Umstellung erfolgte auf Wunsch vieler Netzbetreiber, die sich durch das gemeinsame IP-Protokoll „deutliche Kostenvorteile“ versprechen. [32, S. 130f.] Mit dem Release 5 wurde durch die Einführung des *IP Multimedia Subsystem (IMS)* ein „weiterer Schritt in Richtung All-IP-Netzwerk“ [32, S. 132] getan. Das IMS hat sich trotz jahrelanger Standardisierung in der Praxis noch nicht durchgesetzt, da es in der Umsetzung eine Reihe von Problemen liefert. Dazu gehören u. a. die Komplexität des Systems, das für „Videotelephonie, Austausch von Bildern [...], Instant Messaging, Presence, etc.“ [32, S. 133] entwickelt wurde, und die Überführung von laufenden Telefonaten in das GSM-Netz, wenn die UMTS-Abdeckung endet. Release 5 standardisierte zudem das Übertragunsgverfahren *High Speed Downlink Packet Access (HSDPA)*, das gegenüber R99 mit maximal 384 kbit/s eine theoretische Übertragungsgeschwindigkeit von 14 Mbit/s liefert. Die praktischen Datenraten liegen bei etwa 5-7 Mbit/s im optimalen Fall bzw. 800 kbit/s bei vielen Nutzern und schlechteren Empfangsbedingungen. [32, S. 134]
Die folgenden Releases erweiterten die Standards für UMTS. Dabei wurde v. a. auf Nutzerfreundlichkeit (steigende Datenraten, Reaktionsgeschwindigkeiten, etc.) und auf Kostensenkung für die Anbieter und damit für die Nutzer geachtet. HSDPA wurde um einen weiteren Standard erweitert. Beide Standards werden in Kombination als *High Speed Packet Access (HSPA)* bezeichnet. [32, S.135-139] Seit Release 2008 wird auch LTE spezifiziert, mit Release 10 erschien LTE-Advanced [18, 41] (siehe auch Abschnitt 4.4). Derzeit wird an Release 12 gearbeitet, dass im Juni 2014 veröffentlicht

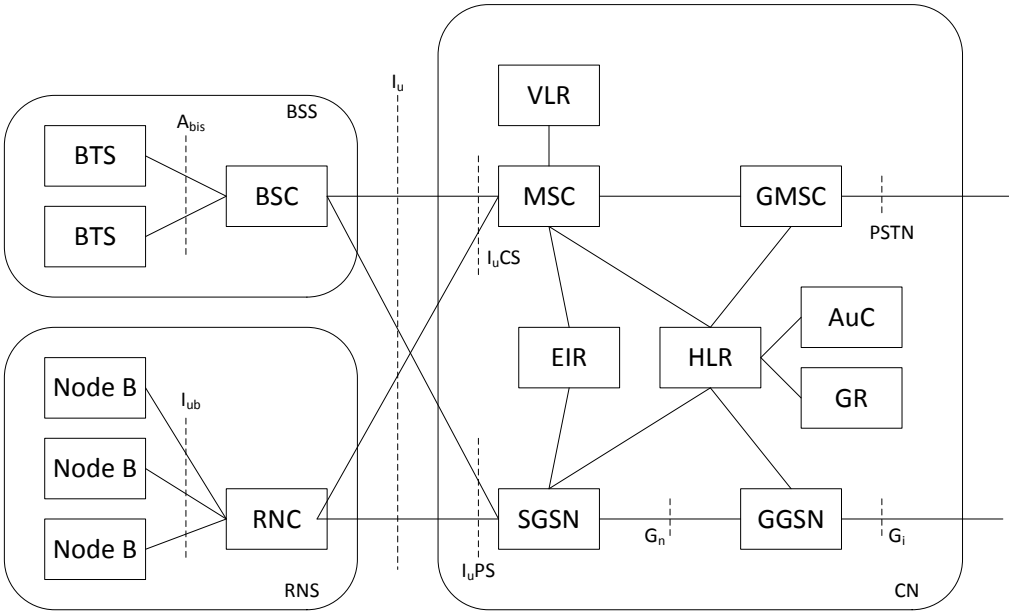


Abbildung 4.3: UMTS-Kernnetz zusammen mit einem 3G-RNS und einem 2G-BSS (nach [33, S. 193])

werden soll [29].

4.3.2 Architektur

Das UMTS-Netz besteht wie schon GSM und GPRS aus mehreren Komponenten, die zu eigenständigen, kleinen Netzwerken zusammengefasst werden können. Diese UMTS-eigenen Netzwerke sind zum einen das UTRAN (UMTS Terrestrial Radio Access Network), das mehrere *Radio Network Subsystems* (RNS, Funknetzsubsysteme) beinhaltet, und zum anderen das *Core Network* (CN, Kernnetz). Der Aufbau eines UMTS-Netzes mit dem CN und dem RNS ist in Abbildung 4.3 dargestellt. Zudem sieht man in der Darstellung, dass auch Verbindungen zu GSM/GPRS-Netzelementen (BSS) möglich sind.
Die Aufgaben und Funktionen der einzelnen Komponenten ähneln denen in der GSM/GPRS-Architektur. Es soll im Folgenden v. a. auf die Neuerungen und Unterschiede zwischen GSM/GPRS und UMTS eingegangen werden.

UMTS Terrestrial Radio Access Network (UTRAN)

Jedes UTRAN besteht aus mehreren RNSs. Diese wiederum enthalten jeweils einen RNC und mehrere Node Bs, die durch den RNC gesteuert werden. Daneben hat der RNC als „Herz des UMTS-Radionetzwerkes“ [32, S. 171] eine Fülle anderer Aufgaben [33, S. 190/191]:

- *Zugangssteuerung*: Der RNC überwacht das aktuelle Datenaufkommen in einer Zelle und entscheidet darüber, ob weitere Verbindungen aufgebaut werden können.
- *Staukontrolle*: Die Funkressourcen, die von mehreren UEs genutzt werden, müssen entsprechend der geforderten Dienstgüte durch den RNC verteilt werden.
- *Verschlüsselung/Entschlüsselung*: Das RNC ist für die Verschlüsselung von Daten aus dem Festnetz zuständig. Im umgekehrten Fall müssen die Daten entschlüsselt werden, bevor sie aus dem Funknetz in das Festnetz weitergeleitet werden können.
- *ATM-Switching*[15]*, Multiplex, Protokollumsetzung*: Der RNC muss Verbindungen zwischen RNCs, Node Bs und dem CN, die auf ATM basieren, weiterleiten und Datenströme multiplexen. Eine weitere Aufgabe ist die Umsetzung von Protokollen.
- *Funkressourcenüberwachung*: Zu den Aufgaben des RNC gehört auch die Überwachung der Funkressourcen der Zellen, deren Node Bs er verwaltet. Dazu zählt u. a. die Erfassung von Interferenzen und Auslastungen.
- *Funkträgerverwaltung*: Der RNC ist für die Verwaltung (Aufbau, Halten, Beenden) von Funkverbindungen zu den UEs zuständig.
- *Codeauswahl*: Der UE nutzt für die Übertragung verschiedene Codes, die durch den RNC ausgewählt werden müssen und die auch während der Übertragung gewechselt werden können.
- *Leistungsregelung*: Es gibt für die Vermeidung von z. B. Interferenzen zwei Regelkreise. Die schnelle zelleninterne Regelung erfolgt durch den Node B. Der RNC steuert den sogenannten äußeren Regelkreis, der für langsame Änderungen zuständig ist. In die Entscheidungen des RNCs fließen zum einen die Interferenzwerte anderer eigener Zellen mit ein, zum anderen auch die Messwerte anderer RNCs.
- *Verbindungsübergabesteuerung und RNS-Verlagerung*: Bei einem Abschwächen der Verbindung in einer Zelle wird eine Übergabe des UEs durch das RNC eingeleitet. Wenn sich das UE dabei in den Bereich eines neuen RNC hinein bewegt, wird die RNC-Übergabe als RNS-Verlagerung bezeichnet.
- *Allgemeine Verwaltung*: Neben den genannten Steuerungs- und Überwachungsvorgängen ist das RNC auch für das Sammeln von Informationen, z. B. Auslastung einer Zelle, Verkehrsaufkommen und Fehlerzustände, verantwortlich. Diese Daten sind notwendig für die Netzbetreiber, damit ein Netz effizient genutzt werden kann.

Das zweite Element in einem RNS ist der Node B, der im GSM/GPRS-Netz als BTS bezeichnet wurde. Der Node B verwaltet eine oder mehrere Antennen, die die Zellen aufspannen. Zu den Aufgaben des Node Bs gehören die Überwachung der Verbindungsqualitäten und Signalstärken, die an den RNC gesendet werden, und die Leistungsregelung im inneren Regelkreis. Diese Regelung ist notwendig, damit Interferenzen („Nah-Fern-

[15] *Asynchronous Transfer Mode (ATM)* ist ein Kommunikationsprotokoll, das für die Übertragung von Sprache, Daten und Videosignalen genutzt werden kann. Es entwickelte sich in den 1990er Jahren und konnte sich in einigen Anwendungsbereichen durchsetzen, obwohl es Probleme bezüglich Timing, Technologie und Implementierung aufwies [36, S. 296] (für nähere Informationen siehe Tanenbaum und Wetherall [36, S. 296f.]).

Effekte“ [33, S. 192]) weitestgehend vermieden werden können. Des Weiteren ist der Node B für die Verbindungsübergabe zwischen zwei „eigenen“ Zellen zuständig („weiche Übergabe“, siehe Abschnitt 4.3.3). [33, S. 192]

Das dritte und letzte Element, das in einem RNS arbeiten kann, ist das UE. Es stellt das „Gegenstück zu mehreren Komponenten innerhalb der Architektur“ [33, S. 192] dar.
Es misst analog zum Node B die Signalqualität und nimmt dementsprechend am inneren Regelkreis zur Anpassung der Sendeleistung teil. Außerdem führt es in diesem Rahmen die Modulation und die Ratenanpassung durch. Das UE kooperiert weiterhin mit dem RNC bei der Verbindungsübergabe und der Auswahl der günstigsten Zelle. Er muss die Daten „ver- und entschlüsseln und an der Belegung der Funkressourcen teilnehmen“ [33, S. 192]. Als drittes fungiert er als Gegenstück zum CN. In dieser Funktion unterstützt er die Auswahl eines Trägers und implementiert die Funktionen für die Mobilitätsverwaltung. Des Weiteren muss er Dienste vom Netz abfragen. [33, S. 192]
Neben den Aufgaben, die ein Endgerät bezüglich der UMTS-Netzarchitektur erfüllen muss, erwarten auch die Nutzer eine große Auswahl an Funktionen und Programmen (Kalender, Planer, Spiele, Kamera und ähnliche Applikationen (Apps)) sowie eine möglichst lange Laufzeit. [33, S. 192]

Die RNC verwalten heute im Normalfall „mehrere hundert Node Bs“ [32, S. 171], die an das RNC mittels IP-basierten, breitbandigen DSL, Glasfaser oder Ethernet Microwave angeschlossen sind, deren „Datenraten [...] sich im Bereich von hundert Megabit pro Sekunde oder mehr“ [32, S. 171] bewegen.

Core Network (CN)

Das CN benutzt prinzipiell die gleichen Komponenten, die auch die GSM- bzw. GPRS-Infrastruktur verwendete. Das Kernnetz besteht aus zwei grundsätzlichen Komponenten: der *Circuit Switched Domain* (CSD, leitungsvermittelter Bereich) und der *Packet Switched Domain* (PSD, paketvermittelter Bereich). Zur CSD gehören das Mobile Switching Centre (MSC), das Gateway MSC (GMSC) und das Visitor Location Register (VLR). Zum PSD zählen der Serving GPRS Support Node (SGSN) und der Gateway GPRS Support Node (GGSN). Das Equipment Identity Register (EIR) und das Home Location Register (HLR) mit dem Authentification Centre (AuC) und dem GPRS Register (GR) werden von beiden Domänen genutzt. [33, S. 192f.]
Die Verbindung zwischen dem CN und dem RNS bzw. auf GSM/GPRS-Ebene dem Base Station Subsystem (BSS) wird über die I_u-Schnittstelle realisiert. Diese Schnittstelle wurde zweigeteilt in I_u(cs) für die Verbindung zu dem leitungsvermittelten Bereich und in I_u(ps) zu dem packetvermittelten Bereich. Die paketvermittelten Dienste werden über die I_u(ps)-Schnittstelle abgewickelt. An den SGSN, der auf der CN-Seite mit der I_u(ps)-Schnittstelle verbunden ist, können auch die Base Station Controller aus der GSM/GPRS-Architektur angeschlossen werden. Der SGSN muss daher sowohl GPRS-Protokolle als auch UMTS-Protokolle verwalten können. In der Datenverarbeitung gibt es bei UMTS

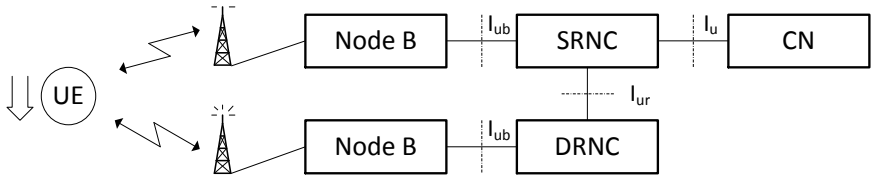


Abbildung 4.4: Serving RNC und Drift RNC [33, S. 197]

eine Neuerung. Während der SGSN bei GSM/GPRS die Daten, die vom GGSN kamen, noch verarbeiten und an die richtige Zelle weiterleiten musste, sendet er bei UMTS die Pakete nur direkt an den RNC weiter, der die Verarbeitung übernimmt. [32, S.173f.]

4.3.3 Verbindungsübergabe (Handover)

In UMTS-Netzen existieren zwei Arten der Verbindungsübergabe: *Soft Handover* (Weiche Übergabe) und *Hard Handover* (Harte Übergabe). Die harten Übergaben sind Handover, bei denen der Timeslot gewechselt werden muss. „Sobald die Trägerfrequenz gewechselt werden muss, also ein so genannter *Inter Frequency Handover* durchgeführt wird, ist dies immer eine harte Übergabe“ [33, S. 195]. Des Weiteren wird es als harte Übergabe bezeichnet, wenn ein Wechsel zwischen verschiedenen Systemen durchgeführt wird (beispielsweise UMTS nach GSM). Im Gegensatz zu der harten Übergabe werden bei einer weichen Übergabe die Datenströme nicht kurzzeitig unterbrochen. Der Übergang zu einer neuen Zelle erfolgt fließend und langsam. Ein UE kann durch die Protokolle, die in UMTS verwendet werden, mit mehreren Antennen gleichzeitig kommunizieren. Diese Antennen können zudem zu unterschiedlichen Node Bs gehören. Die Datenströme, die zum UE gelangen sollen, werden durch den RNC aufgeteilt und durch den UE wieder zusammengesetzt. In der umgekehrten Richtung sendet der UE die Daten, die von mehreren Antennen empfangen werden und durch den RNC wieder kombiniert werden müssen. [33, S. 196]

Diese Eigenschaft, mit mehreren Antennen gleichzeitig kommunizieren zu können, wird als *Makrodiversität* bezeichnet. Die Übertragungen sind weniger anfällig gegenüber schnellem Schwund und Wegeausbreitung sowie gegen Abschattung. Wenn der Pfad zu einer der Antennen blockiert sein sollte, bricht die Verbindung nicht wie normalerweise zusammen, sondern kann durch die anderen Antennen aufrechterhalten werden. Während der weichen Übergabe korrespondiert das UE mit allen teilnehmenden Antennen. Dabei nimmt der UE von diesen Antennen Anweisungen entgegen. Es ist somit möglich, dass die Sendeleistung an die Antenne angepasst wird, die dem UE am nächsten ist. Wenn der UE sich bezüglich seiner Sendeleistung nach einer weiter entfernten Antenne richten würde, wäre die Sendeleistung für die am nächsten gelegene Antenne unter Umständen zu groß und würde Interferenzen in dieser Zelle hervorrufen. Durch dieses Verfahren können die Interferenzen auf ein Minimum reduziert werden. [33, S. 196]

Die weiche Übergabe wird nicht durch den CN unterstützt. Damit sie dennoch auch für UEs möglich ist, die zwischen Zellen wechseln, die zu unterschiedlichen RNCs gehören, muss einer der betreffenden RNCs die Übergabe steuern. Der RNC, der für die „alte“ Zelle zuständig ist, arbeitet als *Serving RNC (SRNC)*, der RNC der „neuen“ Zelle wird als *Drift RNC (DRNC)* bezeichnet (siehe Abbildung 4.4, gezeichnet nach Schiller [33, S. 197]). Der SRNC empfängt die Daten für die Node Bs durch den CN und leitet diese an „seinen“ Node B und an den DRNC weiter (splitting). Wenn der UE Daten sendet, werden diese vom SRNC zusammengesetzt und anschließend an den CN weitergeleitet. Somit ist die Makrodiversität für den CN unsichtbar. Wenn der UE sich aus dem Bereich des aktuellen SRNC herausbewegt, muss eine SRNC-Verlagerung durchgeführt werden, damit nicht zwei RNCs (SRNC und DRNC) Ressourcen zur Verfügung stellen müssen. Diese Übergabe erfolgt mit Unterstützung des CN und ist somit eine harte Übergabe. [33, S. 196f.]

Wie bei den GSM-Netzen schon beschrieben, gibt es auch für UMTS-Netze mehrere Arten der Übergabe. Die Einteilung erfolgte nach den Elementen, die jeweils die Übergabe steuern. In der Abbildung 4.5 sind die unterschiedlichen Übergabeszenarios dargestellt. Schiller teilt die Übergaben in fünf Kategorien ein [33, S. 197f.]:

- Intra-Node B, Intra-RNC: UE_1 wechselt die Zelle, bleibt jedoch im Bereich von Node B_1. Die Kombination und Teilung der Ströme wird durch den Node B_1 gesteuert. Diese Übergabe wird als weichere Übergabe (Softer Handover) bezeichnet.
- Inter-Node B, Intra-RNC: UE_2 bewegt sich von einer Zelle von Node B_1 in eine Zelle von Node B_2. Die Datentrennung und -kombination wird durch RNC_1 gesteuert.
- Inter-RNC: UE_3 bewegt sich aus dem Bereich von Node B_2 in den Bereich des Node B_3, wobei RNC_1 die Funktion des SRNC übernimmt und RNC_2 als DRNC fungiert. Wie schon beschrieben, ist dies eine weiche Übergabe. Wenn der UE_3 den Bereich des RNC_1 endgültig verlassen hat, wird eine SRNC-Verlagerung durchgeführt. Diese wird durch den 3G MSC_1 gesteuert und ist somit eine harte Übergabe.
- Inter-MSC: Auch bei dieser Übergabeart verlässt der UE_3 den Bereich des Node B_2 und wechselt in den Bereich des Node B_3. Dabei besteht die Möglichkeit, dass der RNC_2 die Verbindung übernimmt. In diesem Fall wäre es eine harte Übergabe.
- Intersystem: Die letzte von Schiller genannte Übergabeart ist der Wechsel von einem UMTS-Netz in ein GSM-Netz. Dies ist eine harte Verbindungsübergabe und „wesentlich für die Akzeptanz die Benutzbarkeit des gesamten UMTS-Systems auf Grund der Versorgungslücken in der Fläche“ [33, S. 198].

Der Hard Handover funktioniert ähnlich dem GSM-Handover und soll daher hier nicht näher betrachtet werden. Der Soft Handover bietet gegenüber dem Hard Handover einige Vorteile. Durch den langsamen Übergang wird die Wahrscheinlichkeit für eine Unterbrechung der Verbindung sehr gering. Dadurch erhöht sich die Verbindungsqualität für die Nutzer. Des Weiteren wird der Handover-Vorgang schon eingeleitet, wenn die

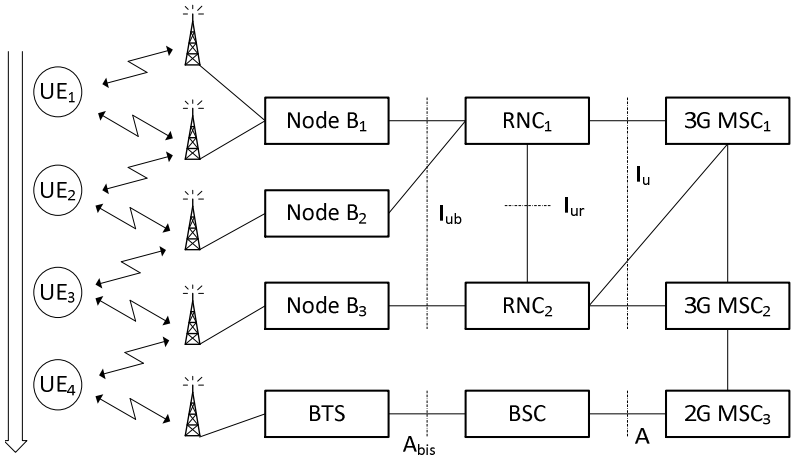


Abbildung 4.5: Übersicht über verschiedene Typen von Verbindungsübergaben [33, S. 198]

Signalqualität, die fortwährend durch das Netzwerk kontrolliert wird, noch „akzeptabel“ [32, S. 185] ist. Daher ist bei einer plötzlichen Verschlechterung der Signale in der alten Zelle ein Abbruch der Verbindung sehr unwahrscheinlich. [32, S. 184f.]
Ein weiterer Vorteil des Soft Handovers ist die verringerte Sendeleistung des Endgerätes. Sauter erklärt diesen Umstand an folgendem Beispiel [32, S. 185f.]: Ein Endgerät kommuniziert über eine gute Verbindung mit einer Zelle 1 und hat in Folge des Handover-Vorgangs auch eine gute Verbindung zu Zelle 2. Wenn die Verbindung zu Zelle 1 durch Gebäude zeitweise verschlechtert wird, müsste das Endgerät seine Sendeleistung erhöhen um weiterhin mit Zelle 1 kommunizieren zu können. In diesem Fall ist dies jedoch nicht notwendig, weil die Daten von der Zelle 2 weiterhin gut empfangen werden können. Dabei wird die Verbindung zu der Zelle 1 jedoch nicht abgebaut.
Die Pakete werden dabei immer noch sowohl an Zelle 1 als auch an Zelle 2 gesendet. Der RNC entscheidet im Anschluss, welches der Pakete verworfen wird (im Beispiel von Sauter wird das von Zelle 1 empfangene Paket verworfen, da die Signalqualität schlechter war). [32, S. 186]
Für das Netzwerk hat der Soft Handover ebenfalls Vorteile. Wie schon einmal erwähnt, passt sich das Endgerät immer der Zelle mit der besten Signalqualität an (beispielsweise der am nächsten gelegenen). Dadurch minimieren sich die Interferenzen, die in den Zellen auftreten. [32, S. 187]

Allerdings bietet der Soft Handover nicht nur Vorteile.
In Downlink-Richtung werden die Pakete an alle Node Bs gesendet, mit deren Antennen das Endgerät kommuniziert. Das Endgerät empfängt die Datenpakete demzufolge mehrmals. Diese Daten werden als „zwei separate Datenströme auf der physikalischen Schicht behandelt“ [32, S. 186]. Damit erhöht sich der Rechenaufwand, da die Daten mehrmals dekodiert werden müssen. [32, S. 186]
Durch das mehrmalige Versenden der Daten an die Node Bs und in umgekehrter Richtung das Versenden eines Datenpaketes von mehreren Node Bs an den RNC werden mehr Ressourcen genutzt als es bei der Kommunikation mit einem Node B der Fall

gewesen wäre. Um diesen Ressourcenverbrauch möglichst gering zu halten, versuchen die Netzwerkbetreiber keine Stellen zu schaffen, an denen ein Endgerät mit mehr als drei Node Bs gut kommunizieren kann. [32, S. 187] Theoretisch könnte ein Endgerät jedoch mit bis zu sechs Node Bs gleichzeitig kommunizieren [32, S. 184].
Der Ressourcenverbrauch und die Komplexität der Kommunikation wird noch größer, wenn mehrere RNCs an dem Handover beteiligt sind, die dann als SRNC und DRNC fungieren (siehe oben) [32, S. 187].

Um die nicht flächendeckenden UMTS-Netze in der Praxis nutzbar zu machen, müssen die Verbindungen zwischen UMTS und GSM wechseln können. Dieser Wechsel muss auch in laufenden Verbindungen möglich sein. Dieser Handover wird als Intersystem Handover bezeichnet. Es gibt in UMTS mehrere Arten des Intersystem Handovers. Sauter nennt zwei verschiedene Verfahren: Blind Intersystem Handover und gesteuerter Intersystem Handover [32, S. 190f.].
Beim Blind Intersystem Handover existiert neben der aktiven UMTS-Zelle nur eine GSM-Nachbarzelle. Wenn sich das Endgerät aus der UMTS-Zelle herausbewegt, wird durch den RNC der Handover eingeleitet. Dieser Handover wird „blind“ genannt, weil für die GSM-Zelle zum Zeitpunkt der Übergabe keine Messdaten zur Verfügung stehen. [32, S. 190] Der Vorteil dieses Verfahrens ist nach Sauter eine einfache Implementierung im Netzwerk und in den Endgeräten [32, S. 190] Es existieren allerdings auch eine Reihe von Nachteilen [32, S. 190]:

- Das Netzwerk hat keine Informationen, ob die GSM-Zelle mit dem Endgerät kommunizieren kann.
- Es besteht keine Synchronisation zwischen Endgerät und GSM-Zelle. Daher wird die Zeit verlängert, die das Endgerät benötigt um mit der Zelle Kontakt aufzunehmen. Dies kann z. B. zu kurzen Unterbrechungen in aktiven Telefonaten führen.
- Wenn mehrere GSM-Nachbarzellen vorhanden sind, weiß der RNC nicht, an welche der Zellen das Endgerät übergeben werden muss.

Eine Alternative zum Blind Intersystem Handover stellt der gesteuerte Intersystem Handover dar, der „heute in der Praxis verwendet wird“ [32, S. 191]. Das Endgerät misst die Signalqualität nicht nur in benachbarten UMTS-Zellen, sondern auch in den GSM-Nachbarzellen. Die Messwerte werden an den RNC gesendet. Dieser kann zusammen mit dem MSC den Handover in die GSM-Nachbarzellen durchführen. [32, S. 191]

4.4 Long Term Evolution (LTE)

4.4.1 Übersicht und Standardisierung

Die nächste Entwicklungsstufe nach GSM/GPRS und UMTS war *Long Term Evolution (LTE)*. Die 3GPP-Organisation entwickelte die Technologie, weil das UTMS-System

„seine Grenzen in ähnlicher Weise [erreichte], wie GSM und GPRS etwa ein Jahrzehnt früher“ [32, S. 229].
Bei UMTS wurde eine Bandbreite von 5 MHz genutzt. Für eine nochmalige Erhöhung der Übertragungsgeschwindigkeit müssen die Übertragungsschritte verkürzt werden. Diese Kürzung der Abstände zwischen den Schritten würde jedoch das sogenannte „Multipath Fading“ verstärken. Dieses Fading beschreibt den Einfluss der Datenkopien, die durch die Mehrwegeausbreitung verzögert am Empfänger ankommen, nachfolgende Signale überlappen und möglicherweise nicht mehr detektierbar machen (siehe Abschnitt 2.3.3). Mit der gänzlich neu entwickelten LTE-Technologie sollte das Multipath Fading umgangen werden. Bei LTE wird nicht mehr nur ein Signal sehr schnell übertragen, sondern mehrere langsame Signale parallel. Dieses Verfahren wird als *Orthogonal Frequency Division Multiplexing (OFDM)* bezeichnet. Die Datenraten, die mit diesem Verfahren erreicht werden, sind bei gleicher Bandbreite mit denen von UMTS identisch, der „Multipath-Effekt ist jedoch aufgrund der sehr langen Übertragungsschritte jedes Datenstroms deutlich geringer“ [32, S. 229]. Je nach benötigter Geschwindigkeit und Datenrate und in Abhängigkeit der zur Verfügung stehenden Bandbreite wird die Anzahl der parallelen Datenströme vergrößert bzw. verkleinert. Aufgrund dieser Tatsache konnten Bandbreiten zwischen 1,25 MHz und 20 MHz definiert werden, die durch die LTE-Endgeräte unterstützt werden müssen. Die Wahl des Frequenzbandes ist abhängig vom „Frequenzband und dem Umfang des Spektrums eines Netzbetreibers“ [32, S. 230]. Bei Verwendung einer Bandbreite von 20 MHz und „sehr guten Übertragungsbedingungen“ [32, S. 230] können in der Praxis Datenraten von etwa 100 Mbit/s erreicht werden. [32, S. 230]
Neben der Bedingung, dass die Endgeräte alle Bandbreiten unterstützen müssen, gibt es eine weitere Anforderung an die LTE-Endgeräte. Sie müssen Multiple Input Multiple Output (MIMO) Übertragungen unterstützen können, d. h., dass „mehrere unabhängige Datenströme gleichzeitig über den gleichen Kanal übertragen werden“ [32, S. 230].

Neben der Veränderung der Signalübertragung gibt es eine weitere wichtige Neuerung gegenüber der UMTS-Architektur. Während bei UMTS die Sprach- und SMS-Dienste noch leitungsvermittelt waren, werden bei LTE mit Ausnahme der SMS-Dienste alle Daten paketvermittelt übertragen. Durch diese Fokussierung auf das paketvermittelnde Internet-Protokoll (IP) wird das „Design und die Implementierung der LTE-Luftschnittstelle, des Radionetzwerkes und [...] des Kernnetzes“ [32, S. 230] vereinfacht.
Das IP-Protokoll bringt neben der Designvereinfachung noch einige weitere Vorteile für LTE. Der LTE-Standard definiert nur das IP-Protokoll und lässt die darunter liegenden Schichten offen. Somit ist die darunter liegende Infrastruktur transparent und leicht austauschbar. Des Weiteren wurde die Anzahl der Netzkomponenten verringert, damit die Verzögerungszeiten kleiner werden. Um die Zeit für den Verbindungsaufbau mit dem Netzwerk weiter zu verringern, wurde zusätzlich die Signalisierung für das Herstellen und Halten der Verbindungen vereinfacht. Der Nachteil an der Verwendung des IP-Protokolls ist der Aufbau einer Sprachverbindung. Es wird an der Entwicklung von „Voice over LTE“ (VoLTE) gearbeitet, allerdings verhinderte die Komplexität der Technologie bisher eine Verbreitung. Diese Komplexität wird u. a. durch die Schwierigkeit, eine laufende

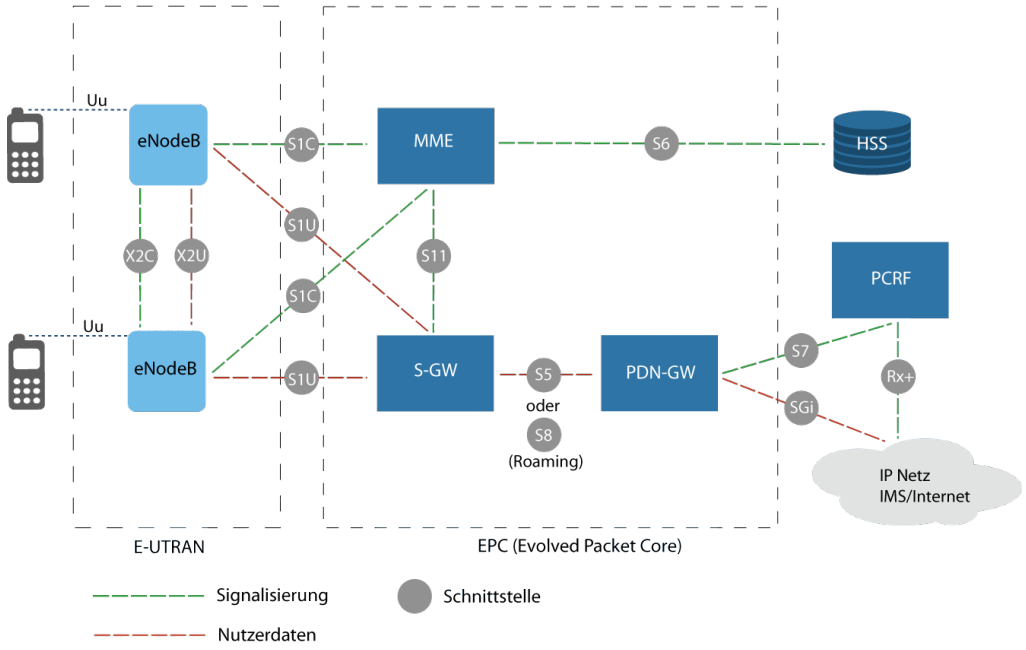


Abbildung 4.6: Aufbau des EPC mit Schnittstellen [10, S. 23]

Sprachverbindung in das GSM/GPRS- bzw. UMTS-Netz zu übertragen, hervorgerufen. Alternativ wird das so genannte „CS-Fallback“ genutzt, bei dem die LTE-Endgeräte vor dem Aufbau einer Sprachverbindung in das GSM/GPRS- oder UMTS-Netz wechseln müssen. Daraus resultierend ergibt sich eine weitere Anforderung an die Endgeräte: Sie müssen neben den LTE-Protokollen auch die Protokolle für GSM, GPRS, EDGE, UMTS und HSPA kennen. [32, S. 230f.]

4.4.2 Architektur

Die Architektur von LTE besteht aus zwei grundsätzlichen Komponenten, dem Radionetz und dem Kernnetz, und ist somit der Architektur von GSM und UMTS ähnlich. Allerdings wurde die Anzahl der Elemente verringert, damit die Effizienz der Netzwerke gesteigert wird. Gleichzeitig können damit die Kosten gesenkt und die Latenzzeiten verringert werden. [32, S. 231] Das Radionetz, auch als *Evolved UMTS Terrestrial Radio Access Network (EUTRAN)* bezeichnet, besteht aus dem Endgerät (User Equipment, UE) und der Basisstation (eNodeB). Das Kernnetz wird in LTE auch *Evolved Packet Core (EPC)* genannt. [10, S. 13]

Evolved UMTS Terrestrial Radio Access Network (EUTRAN)

Das EUTRAN ist eine Weiterentwicklung des UTRAN. Das zentrale Element bildet der eNodeB. Um eine hohe Datenrate auf der Funkschnittstelle zu gewährleisten, wurde gegenüber UMTS die Technologie verändert. Gutt nennt diesbezüglich eine „hohe Modulation, eine komplexe Antennentechnologie und neue Verfahren zur Kodierung“

[10, S. 13].
Als Modulationsverfahren in Downlink-Richtung wird die 64-QAM (Quadrature Amplitude Modulation)[16] verwendet, in Uplink-Richtung die „langsamere aber dafür robustere 16-QAM-Modulation“ [32, S. 232]. Diese Modulationsarten müssen von allen LTE-Endgeräten unterstützt werden. Des Weiteren müssen alle UEs MIMO-Übertragungen unterstützen. Mit diesem Verfahren ist es möglich, dass mehrere Datenströme auf dem gleichen Kanal zwischen mehreren Antennen im Endgerät und mehreren Antennen im eNodeB übertragen werden. Die Endgeräte sind in der Lage, die verschiedenen Datenströme zu unterscheiden, wenn sie unterschiedliche Pfade genommen haben. Diese verschiedenen Wege werden z. B. durch Reflektionen an Objekten und durch die räumliche Trennung der Antennen verursacht. Die Anzahl der Datenströme, die parallel auf einem Kanal gesendet werden können, ist abhängig von der Anzahl der Sende- und Empfangsantennen des eNodeBs und des Endgerätes. Nach Sauter haben die meisten LTE-Endgeräte heutzutage zwei Antennen, allerdings wurden zukunftsweisend auch vier Antennen spezifiziert, die in der Praxis jedoch v. a. bei kleinen Endgeräten, z. B. Smartphones, schwer umsetzbar sind. Des Weiteren sind mehr als zwei Antennen schlecht realisierbar, weil LTE, GSM und UMTS mehrere unterschiedliche Frequenzbänder nutzen, die alle durch die Endgeräte unterstützt werden müssen und dadurch die Komplexität der Antennen vergrößern. Die Anzahl der Antennen, die in einem Endgerät bzw. in den Basisstationen enthalten sind, werden durch Angaben wie 2×2 MIMO oder 4×4 MIMO angegeben. [32, S. 232f.]
Theoretisch können heute maximale Übertragungsraten von 100 bis 150 Mbit/s bei einer Bandbreite von 20 MHz erzielt werden. Praktisch sind diese jedoch nur selten erreichbar, da Interferenzen und die mit zunehmendem Abstand vom Sender sich verringernde Signalleistung die Übertragungsrate verringern. [32, S. 233]

Das Endgerät kommuniziert über die sogenannte Uu-Schnittstelle mit dem eNodeB. Die Basisstationen bestehen aus drei Komponenten – den Antennen, dem Radiomodul und dem Digitalmodul. Das Radiomodul ist für die Modulation bzw. die Demodulation der Datenströme und für die Verstärkung der Signale zuständig. Das Digitalmodul übernimmt die Signalverarbeitung und dient als „digitales Interface zum Kernnetz“ [32, S. 235].
In den GSM/GPRS- und UMTS-Netzen waren die Basisstationen für die Luftschnittstellen verantwortlich, während der RNC bzw. der BSC für die Verarbeitung der Informationen und das Management der Ressourcen zuständig waren. In LTE wurde jedoch darauf geachtet, dass die Anzahl der Elemente verringert wurde, damit die Latenzzeiten herabgesetzt werden konnten. Daher wurden die meisten Funktionen des RNC in den eNodeB integriert. Der eNodeB fungiert somit als [32, S. 235]:

- Air Interface
- User Management und Aufteilung der Ressourcen auf dem Air Interface
- Sicherstellung der QoS-Attribute (z. B. maximale Verzögerungszeit, minimale Band-

[16] Bei der QAM werden die Phasenumtastung (PSK) und die Amplitudenumtastung (ASK) miteinander kombiniert. Es können bei der 64-QAM-Modulation sechs Bit gleichzeitig übertragen werden.

breite)
- Mobilitätsmanagement
- Interferenzmanagement (Reduzierung der eigenen Sendeleistung zur Störungsminimierung bei benachbarten eNodeB)

Die Kommunikation zwischen den eNodeBs und dem Kernnetz erfolgt über das sogenannte „S1 Interface“, das als kabelgebundene, als optische oder als Mikrowellenverbindung realisiert sein kann. Die Übertragungsraten dieser Verbindung liegen im Bereich von mehreren Hundert Mbit/s bis hin in den Gbit/s-Bereich. Mit diesen Geschwindigkeiten können nach Sauter Basisstationen mit drei oder mehr Sektoren versorgt werden. Des Weiteren werden über die S1-Schnittstelle auch Daten an GSM- und UMTS-Basisstationen gesendet. Die Bandbreite, die für das S1 Interface benötigt wird, ist demzufolge wesentlich größer als die für die LTE-Sektoren. [32, S. 236]
Die S1-Schnittstelle besteht zwar aus einer physikalischen Verbindung, teilt sich funktionell jedoch in zwei verschiedene Abschnitte. Der Teil, der für die Übertragung der Nutzdaten zuständig ist, wird als S1-User Plane (S1-UP)-Protokollstack bezeichnet. Die andere Hälfte, der S1-Control Plane (S1-CP)-Protokollstack, überträgt Signalisierungsdaten. Diese Signalisierungen entstehen durch zwei Teilfunktionen, für die der S1-CP zuständig ist. Zum einen kommuniziert die Basisstation über den S1-CP mit dem Kernnetz, wenn sie sich beispielsweise beim Netzwerk anmelden will. Des Weiteren werden über den S1-CP benutzerspezifische Signalisierungen versendet. Diese Signalisierung beinhaltet den Aufbau der logischen Verbindung mit dem Kernnetz für den Nutzer und ihre Aufrechterhaltung. Außerdem wird S1-CP genutzt, um die logische Verbindung gegebenenfalls an andere Basisstationen zu übergeben. [32, S. 236f.]

Wie schon erwähnt, wurde für die LTE-Netze die Anzahl der Komponenten reduziert. Dadurch entfallen auch RNC (UMTS) bzw. BSC (GSM/GPRS), die in den Vorgängernetzen von LTE die Basisstationen überwachten und steuerten. Diese Kontrolle beinhaltete auch die Kommunikation mit mehreren Basisstationen. Um die „Absprache“ zwischen den eNodeBs dennoch zu gewährleisten, wurde das X2 Interface spezifiziert. Diese Schnittstelle zwischen den eNodeBs ist u. a. für den Handover notwendig. Wenn die Basisstationen die Nachbarzellen kennen, kann der Handover direkt über die X2-Schnittstelle erfolgen. Anderenfalls muss die Kommunikation über das S1 Interface und das Kernnetz abgewickelt werden. Alternativ besteht auch die Möglichkeit, dass mithilfe der Endgeräte die Nachbarzellen dynamisch ermittelt werden. [32, S. 237] Eine weitere Funktion des X2 Interfaces ist die Minimierung der Interferenzen zwischen Nachbarzellen, da die eNodeBs durch die Kommunikation über die Schnittstelle die Sendeleistung der Antennen anpassen können [32, S. 237].

Evolved Packet Core (EPC)

Das Kernnetz besteht aus der *Mobility Management Entity (MME)*, dem *Serving-Gateway (S-GW)*, dem *Packet Data Network Gateway (PDN-Gateway, PDN-GW)* und dem *Home*

Subscriber Server (HSS). Das HSS ist keine Neuentwicklung für LTE, sondern das umbenannte Home Location Register (HLR), das bei GSM/GPRS und UMTS verwendet wurde. Das HLR und das HSS sind üblicherweise identisch, damit ein „nahtlose[r] Übergang zwischen den unterschiedlichen Radionetzen“ [32, S. 240] möglich ist.
Die MME verwaltet die Signalisierungsdaten, die zwischen den eNodeBs und dem EPC ausgetauscht werden. In der Praxis werden typischerweise mehrere MMEs verwendet, da in großen Netzen die Kapazitäten der MMEs schnell ausgeschöpft sind. Des Weiteren ist bei einem Defekt einer MME durch die redundanten Elemente die Funktionalität des Netzes weiterhin gewährleistet. Sauter beschreibt die Aufgaben der MME folgendermaßen [32, S. 238f.]:

- Authentifizierung: Wenn ein Endgerät nach dem Einschalten eine Verbindung zum LTE-Netzwerk aufbaut, kommuniziert der eNodeB mit der MME, damit das Endgerät identifiziert werden kann. Die Informationen, die zur Authentifizierung notwendig sind, erhält die MME vom HSS.
- NAS Mobility Management: Wenn ein Endgerät längere Zeit (üblich sind Zeiten zwischen 20 und 30 s [32, S. 238]) inaktiv ist, wird die Verbindung über die Luftschnittstelle deaktiviert. Somit können einerseits Ressourcen in der Luftschnittstelle freigegeben werden, andererseits wird im Endgerät Energie gespart. Das UE entscheidet in dieser Phase selber über Zellwechsel. Wenn Daten für das Endgerät ankommen, versendet die MME eine Paging-Nachricht (siehe Abschnitt 4.2.3) an die eNodeBs, die sich in der Tracking Area[17] befinden, in der sich das UE zuletzt aufhielt.
- Handover-Unterstützung: Wenn die X2-Schnittstelle zwischen zwei eNodeBs nicht existiert, erfolgt die Kommunikation bei einem Handover-Vorgang über die MME.
- Interworking mit anderen Radionetzwerken: Wenn ein Endgerät, das sich an der Grenze eines LTE-Netzes befindet, nach GSM oder UMTS wechselt, übernimmt die MME die Steuerung.

Die zweite Komponente im LTE-Kernnetz ist das Serving-Gateway, das für die “Weiterleitung von Nutzerdaten [...] zwischen den eNodeBs und dem PDN-Gateway verantwortlich“ [32, S. 239] ist. Das PDN-Gateway ist die Anbindung zum Internet oder an Firmennetzwerke. Des Weiteren ist das PDN-Gateway für die Vergabe von IP-Adressen an die Endgeräte verantwortlich. Nachdem das Endgerät durch die MME authentifiziert wurde, fordert die MME über die S5-Schnittstelle vom PDN-Gateway eine IP-Adresse an. [32, S. 240]

[17] In LTE werden mehrere Zellen zu einem Tracking Area zusammengefasst (siehe auch Location Area bei GSM).

4.4.3 Verbindungsübergabe (Handover)

Wenn ein Endgerät eine bestimmte Signalstärke unterschreitet, sucht es nach Nachbarzellen und meldet die Messdaten an den eNodeB. Dieser entscheidet anhand der Messungen, ob eine Übergabe an einen anderen eNodeB erfolgen soll. Dieser Handover-Vorgang kann auf zwei verschiedene Arten erfolgen. Die effizientere Variante liegt vor, wenn die zwei eNodeBs über die X2-Schnittstelle direkt miteinander kommunizieren können. Wenn keine direkte Verbindung vorliegt, erfolgt die Kommunikation über die MME und das S1 Interface. [32, S. 265]

Wie schon beschrieben, liegt die effizientere Handover-Variante vor, wenn die eNodeBs über das X2 Interface kommunizieren können. Daher soll sich bei der Beschreibung des Handovers auf diese Variante bezogen werden.
Wenn der aktuelle eNodeB das Endgerät an einen anderen eNodeB übergeben möchte, sendet er eine *Handover Request*-Nachricht an die neue Basisstation, die alle notwendigen Informationen über das UE und die aktuelle Verbindung enthält. Nachdem der neue eNodeB überprüft hat, ob eine Verbindung aufgrund seiner zur Verfügung stehenden Bandbreite möglich ist, reserviert er die benötigten Ressourcen und bestätigt die Anfrage mit einem *Handover Request Acknowledge*, das alle Informationen über die neue Zelle enthält, die das Endgerät benötigt. Im Anschluss versendet der alte eNodeB ein Handover-Kommando an das Endgerät. Es werden nun keine Daten mehr übertragen zwischen dem Endgerät und dem eNodeB. Ankommende Datenpakete für das Endgerät werden über die X2-Schnittstelle an den neuen eNodeB weitergeleitet, der sie nach einem erfolgreichen Handover an das UE weiterleitet. Das Endgerät meldet sich mit einer *Random Access*-Nachricht bei dem neuen eNodeB an, der darauf mit einer *Random Access Response*-Nachricht antwortet. Aus Sicht des Endgerätes ist der Handover-Prozess somit beendet. [32, S. 266f.]
Die Daten, die über das X2 Interface weitergeleitet wurden, können nun durch den neuen eNodeB an das Endgerät versendet werden. Allerdings besteht für das Endgerät noch keine Verbindung über die S1-Schnittstelle zum Kernnetz. Dazu sendet der neue eNodeB eine *Path Switch Request*-Nachricht an die MME, das wiederum die Anmeldung beim S-GW übernimmt. Nach der Bestätigung durch das Serving-Gateway antwortet das MME der Basisstation mit einer *Path Switch Request Acknowledge*-Nachricht. Zum Abschluss des Handover-Vorgangs wird dem alten eNodeB durch die neue Basisstation mitgeteilt, dass der Handover erfolgreich war. [32, S. 267f.]

Ein Aspekt, der für die Anwendung der LTE-Netze in der Praxis sehr wichtig ist, ist die Zusammenarbeit mit GSM/GPRS- und UMTS-Netzen. Da bei einem Verlust des LTE-Netzes in den Randgebieten der LTE-Abdeckung eine Suche nach anderen Netzen unter Umständen sehr lang dauern kann, wird der Übergang zu GSM/GPRS- und UMTS-Netzen durch das Netzwerk unterstützt. Dafür sind drei Prozeduren standardisiert, die hier nur genannt werden sollen[18] [32, S. 272]:

[18] Für nähere Informationen siehe [32, S. 272ff.]

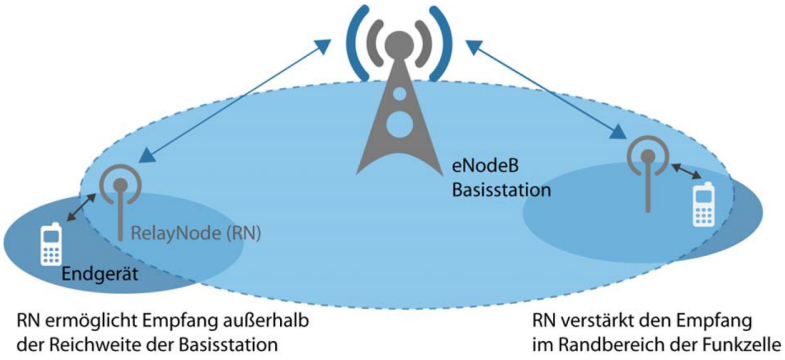


Abbildung 4.7: Signalverstärkung im Randbereich der Zelle durch Relay Nodes [10, S. 33]

- Cell Reselection von LTE nach UMTS oder GSM
- RRC Release mit Redirect von LTE zu UMTS oder GSM
- Inter-RAT (Radio Access Technology) Handover von LTE nach UMTS oder GSM

4.4.4 LTE-Advanced

Eine Weiterentwicklung des LTE-Standards, die sich in der Praxis noch nicht etabliert hat, ist LTE-Advanced. Diese Technologie soll auf eine möglichst kosteneffiziente Weise höhere Datenraten ermöglichen und die Anforderungen der ITU erfüllen, d. h. eine vollständige Anpassung an 4G-Netze umsetzen. [41] Bei der Umsetzung wurde darauf geachtet, dass eine vollständige Kompatibilität mit LTE möglich ist [10, S. 33].
Um die Anforderungen umsetzen zu können, wurden neue Funktionalitäten für LTE-Advanced integriert: Carrier Aggregation (CA), die erweiterte Verwendung von mehreren Antennen und die Unterstützung von Relay Nodes (RNs, Relay Stationen) [41]. So wurde beispielsweise für die Uplink-Richtung 2×4 MIMO spezifiziert und für den Downlink 8×8 MIMO-Antennentechnik [10, S. 33]. Die RNs wurden standardisiert, um die Abdeckung v. a. auch in den Randgebieten zu verbessern. Sie werden durch die Endgeräte als Basisstationen erkannt. Die Kommunikation zwischen den Relay Nodes und den eNodeBs kann z. B. über „optische Freiraumübertragung“ [10, S. 34] realisiert werden. Das Prinzip der Zusammenarbeit zwischen eNodeBs und den RNs ist in Abbildung 4.7 dargestellt. Carrier Aggregation bedeutet, dass einem Nutzer nicht nur ein Frequenzblock zugewiesen wird, sondern mehrere. Dadurch erhöht sich die Bandbreite. [41]
Die Bandbreiten, die für LTE-Advanced standardisiert wurden, betragen bis zu 100 MHz. Mit dieser Bandbreite sollen Datenraten bis zu 1 Gbit/s erreicht werden. [10, S. 34]

4.5 Wireless Local Area Network (WLAN)

4.5.1 Standardisierung

Das Wireless Local Area Network (WLAN) wurde durch das IEEE unter der Bezeichnung 802.11 standardisiert. WLAN basiert im Wesentlichen auf den Standards der LAN-Struktur (802.X der IEEE-Standards). Änderungen betreffen nur die untersten drei Schichten des OSI-Modells. Auf der Schicht 3 werden IP-Pakete versendet. Die Schicht 2 wurde von den LAN-Standards übernommen und um einige Management-Operationen erweitert, die für die drahtlose Übertragung notwendig waren. Die erste Schicht (Physical Layer) des OSI-Modells musste hingegen vollständig neu spezifiziert werden, da keine Kabel für die Übertragung verwendet werden sondern Funkwellen. [32, S. 297]
Für WLAN wurden in den letzten Jahren mehrere Standards entwickelt. Der „Durchbruch für WLAN erfolgte mit dem 802.11b Standard, mit dem Datenraten von 1-11 Mbit/s möglich sind“ [32, S. 298] und der im 2,4 GHz ISM[19]-Band sendet. Der Nachfolger des 802.11b Standards war 802.11g, das in dem gleichen Frequenzband sendet, aber Datenraten von 54 Mbit/s erreichen kann. Bei der Standardisierung wurde darauf geachtet, dass Rückwärtskompatibilität erhalten bleibt. Der Nachfolger 802.11a setzte sich in der Praxis nicht durch, weil er im 5 GHz Bereich sendet und eine Rückwärtskompatibilität durch die unterschiedlichen Frequenzbänder nur schwer durchführbar war. Der nach Sauter heute gängigste Standard ist 802.11n, der sowohl das Frequenzband von 2,4 GHz als auch das von 5 GHz unterstützt. Die Frequenzauswahl ist notwendig, da das WLAN-Netz in den letzten Jahren sehr dicht geworden ist und Frequenzen bei 2,4 GHz meist nicht zur Verfügung stehen. Theoretisch sind mit diesem Standard 600 Mbit/s möglich. Die nächste Stufe wird 802.11ac sein, das Geschwindigkeiten bis zu 6,9 Gbit/s erreichen kann. [32, S. 298f.] Neben den genannten Standards existieren weitere, die jeweils „zusätzliche optionale Wireless LAN-Funktionalitäten“ [32, S. 299] spezifizieren, auf die hier nicht näher eingegangen werden soll.

4.5.2 Architektur

WLAN kann in zwei verschiedenen Varianten betrieben werden. Zum einen gibt es den Ad hoc-Modus für eine direkte Kommunikation der Endgeräte miteinander, zum anderen existiert der Infrastructure BSS Mode, bei dem durch den Access Point (AP) eine zentrale Verwaltung vorhanden ist.
Im Ad hoc Mode (auch als Independent BSS (IBSS) bezeichnet) ist jede Station gleichberechtigt. Alle Daten werden von allen Teilnehmern empfangen, jedoch nicht verarbeitet, wenn die Ziel-IP nicht die eigene ist. Die Teilnehmer in einem Ad hoc-Netz müssen

[19] Das Industrial, Scientific and Medical (ISM)-Band ist ein „öffentliches Frequenzband“ [32, S. 298], das „in den meisten Ländern lizenzfrei versendet werden darf“ [32, S. 298]. Die Sendeleistung muss auf 100 mW beschränkt sein. Auf diesem Band senden neben WLAN auch andere Funksysteme, z. B. Bluetooth. [32, S. 298]

sich zu Beginn auf einige Konfigurationsparameter und auf die IP-Adressen „einigen“. Da dies nach Sauter eine „komplizierte Konfiguration“ [32, S. 300] ist, wird der Ad hoc-Modus in der Praxis nur selten verwendet. [32, S. 300] Im Weiteren wird daher auf den Infrastructure BSS Mode Bezug genommen (siehe Abbildung 4.8, gezeichnet nach [33, S. 249]).

Der zentrale Punkt in einem WLAN ist der Access Point (AP). Er bildet die Schnittstelle zwischen dem drahtlosen und dem drahtgebundenen Netzwerk und dient als Vermittler der Pakete, die zwischen den Endgeräten versendet werden. Bei einer Übertragung eines Datenpaketes wird dies zuerst an den Access Point versendet, der es anschließend weiter an die Zieladresse verschickt. Dieses Verfahren bietet nach Sauter zwei Vorteile. Zum einen müssen die Eigenschaften der Endgeräte nicht untereinander bekannt sein. Nur der Access Point benötigt alle Informationen. Zum anderen ist es mit dem Access Point als Vermittler möglich, dass drahtlose Endgeräte, die sich nicht im gegenseitigen Sendebereich befinden, dennoch Daten austauschen können. Der Nachteil des Verfahrens ist das doppelte Versenden einer Nachricht über die Luftschnittstelle. Die Bandbreite eines Basic Service Sets[20] wird somit halbiert. Um dieses Problem zu umgehen, wurde im Standard 802.11e eine optionale Direktübertragung zwischen zwei Endgeräten spezifiziert. Diese Erweiterung wird heute jedoch erst selten genutzt. [32, S. 300f.]
Wenn es aufgrund der begrenzten Reichweite eines Access Points notwendig ist, mehrere APs zu verwenden, wird dieses System als Extended Service Set (ESS) bezeichnet. Wenn sich ein Endgerät aus dem Bereich eines APs herausbewegt und von einem anderen besser versorgt werden kann, „meldet sich die Netzwerkkarte [...] beim neuen AP an“ [32, S. 302]. Die Teilnehmerinformationen werden im Anschluss zwischen den APs über das sogenannte Distribution System (DS, Ethernet-Verbindung zwischen den APs) ausgetauscht. [32, S. 302]

Für die Konfiguration eines Access Points sind zwei Parameter von großer Bedeutung: die Service Set Identifier (SSID) und die zu nutzende Frequenz. Die SSID, die gleichzeitig den Netzwerknamen darstellt, „identifiziert einen Access Point eindeutig“ [32, S. 303]. Es ist somit möglich, dass mehrere Access Points, die zu unterschiedlichen Netzwerken gehören, parallel am gleichen Ort verwendet werden. Die Frequenzwahl sollte sorgfältig bedacht werden. Die zur Verfügung stehen Frequenzbänder sind begrenzt und es kann bei örtlich parallelen WLANs schnell zu Überschneidungen kommen. Außerdem gibt es zwischen den Kontinenten teilweise Unterschiede, welche Kanäle genutzt werden können. Bei der Konfiguration der Endgeräte durchsuchen diese alle Frequenzen nach vorhanden APs. Im Anschluss kann durch den Nutzer die gewünschte SSID gewählt werden. [32, S. 304]

Der IEEE 802.11 Standard passt sich an die IEEE 802.x Standards an. Dadurch ist

[20] Ein Basic Service Set (BSS) sind alle Stationen, die mit der gleichen Frequenz übertragen. Der Begriff des BBS „umfasst auch den geographischen Bereich, in dem sich die Teilnehmer des BBS aufhalten können“ [32, S. 300].

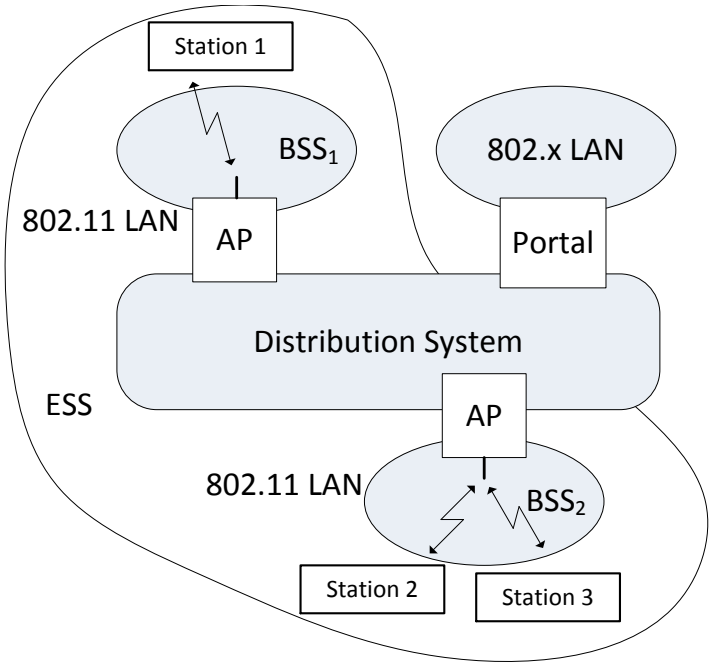


Abbildung 4.8: Architektur eines infrastrukturbasierten IEEE 802.11-Netzwerkes (gezeichnet nach [33, S. 249])

es möglich, dass die Anwendungen, die mit den oberen Schichten des OSI-Modells kommunizieren, nicht „mitbekommen“, ob sie mit einem drahtlosen Netz oder mit dem Ethernet kommunizieren. Ein Teil der Sicherungsschicht (Logical Link Control, LLC) verdeckt die unteren Schichten und macht sie für Anwendungen unsichtbar. Die unteren WLAN-spezifischen Schichten sind die Bitübertragungsschicht (PHY) und die Medienzugriffssteuerung MAC (Media Access Control). [33, S. 251]
Die MAC-Schicht ist für die Steuerung des Medienzugriffs, die Fragmentierung von Nutzdaten und die Verschlüsselung zuständig [33, S. 251]. Die Bitübertragungsschicht kann in drei Varianten ausgeführt werden. Bei zwei der Möglichkeiten werden die Daten über Funkwellen übertragen, bei der dritten wird infrarotes Licht verwendet.
Da in dieser Arbeit die Zuverlässigkeit in mobilen Netzen in den oberen Schichten des OSI-Modells betrachtet werden soll, wird hier nicht näher auf die Schichten des IEEE 802.11 Standards eingegangen.

5 Zuverlässigkeit mobiler Kommunikationssysteme

5.1 Grundlagen der Zuverlässigkeitsanalyse

In diesem Kapitel sollen die Grundlagen der Zuverlässigkeit von Kommunikationsnetzen erläutert werden. Um diese erklären zu können, müssen zuvor einige grundlegende Definitionen bezüglich der Zuverlässigkeit bekannt sein.
Es wird in den folgenden Ausführungen stets angenommen, dass ein Netzwerk als ungerichteter Graph G angesehen werden kann, d. h. dass eine Kommunikation zwischen zwei adjazenten[21] Knoten in beide Richtungen erfolgen kann. Diese Annahme liegt in der Praxis nicht zwangsläufig vor. Zwei Knoten sind in Kommunikationsnetzen zueinander adjazent, wenn sie sich in dem jeweils anderen Sendebereich befinden, d. h. über eine Verbindung direkt miteinander kommunizieren können. Es kann allerdings auch vorkommen, dass eine Kommunikation zwischen zwei adjazenten Knoten nur in eine Richtung erfolgen kann, wenn die Sendeleistung des einen Knoten beispielsweise größer ist als die des anderen [36, S. 449]. Des Weiteren gilt die Voraussetzung, dass die Knoten und Kanten des Netzwerkes stochastisch unabhängig ausfallen. Als Ausfallursachen können „gezielte Angriffe, [...] fehlerhaftes Betreiben oder zufällige (unvorhersehbare) Ereignisse“ [38, S. 13] auftreten. Im Weiteren werden nur die zufälligen Ausfallursachen berücksichtigt, zu denen beispielsweise der „Ausfall elektronischer Bauelemente in Servern oder Übertragungseinrichtungen, Schäden durch Blitzschlag [sowie] Leitungsunterbrechungen durch Erdarbeiten oder Brände“ [38, S. 13] gehören. Die Ausfälle, die durch die Betreiber oder durch mutwillige Zerstörung hervorgerufen werden, sind selten und sollen hier nicht weiter berücksichtigt werden.
Die Zuverlässigkeit in einem Netzwerk kann beispielsweise umschrieben werden mit der Erreichbarkeit in dem Netzwerk und der vorhandenen Redundanz. Die *Erreichbarkeit* wird durch die *Zusammenhangswahrscheinlichkeit* $R(G)$ definiert. $R(G)$ beschreibt die Wahrscheinlichkeit, dass ein ungerichteter Graph zusammenhängend[22] ist. Ausgehend von dieser Definition kann man verschiedene Zusammenhangswahrscheinlichkeiten definieren: die paarweise, die K- und die totale Zusammenhangswahrscheinlichkeit. Die *paarweise Zusammenhangswahrscheinlichkeit* $R_{st}(G)$ (2-terminal reliability) ist eine Verallgemeinerung der totalen Zusammenhangswahrscheinlichkeit [38, S. 18]. Mit ihr ist es möglich, zu prüfen, ob zwischen zwei Knoten s und t mindestens ein Weg besteht. Diese Zusammenhangswahrscheinlichkeit verwenden Kharbash und Wang für die Berechnung der Zuverlässigkeit in Ad-Hoc-Netzen (siehe Abschnitt 5.2.2) als Grundlage. Die *K-Zusammenhangswahrscheinlichkeit* $R(G,K)$ (K-terminal reliability) ist

[21] Zwei Knoten sind zueinander adjazent, wenn sie durch eine gemeinsame Kante verbunden sind [39, S. 13].
[22] In zusammenhängenden Graphen existiert „zwischen je zwei Knoten u und v [...] ein Weg“ [39, S. 15].

eine Erweiterung der paarweisen Zusammenhangswahrscheinlichkeit. Sie beschreibt die Wahrscheinlichkeit, dass „zwischen allen Knoten aus $K \subseteq V$ eine Kommunikation über funktionstüchtige Wege von G möglich ist“ [38, S. 18].
Neben der Erreichbarkeit spielt die *Redundanz* eine entscheidende Rolle in der Zuverlässigkeit von Netzwerken. Redundanz in Netzwerken bedeutet, dass auch nach dem Ausfall einzelner Komponenten die Funktionalität des Netzes erhalten bleibt. Nach Beichelt und Franken existieren drei verschiedene Redundanzarten [4, S. 22f.]: *unbelastete* (kalte), *erleichterte* (warme) und *belastete* (heiße) Redundanz. Bei der unbelasteten Redundanz „sind die Elemente keinerlei Beanspruchung ausgesetzt“ [4, S. 22]. Durch diese fehlende Beanspruchung können sie nicht versagen, bis ihr zugehöriges arbeitendes Element ausfällt. Die belastete Redundanz ist das Gegenteil zu der kalten Redundanz. Sowohl das „Originalelement“ als auch das redundante Element arbeiten und werden gleichermaßen beansprucht. Für die Redundanz bedeutet dieses Verhalten, dass das redundante Element nicht als ein Ersatzelement, sondern als gleichwertig zum Originalelement angesehen wird. Das Netzwerk ist funktionstüchtig, solange eine Mindestanzahl der redundanten Elemente intakt ist. [4, S. 22f.] Die erleichterte Redundanz ist eine Kombination aus belasteter und unbelasteter Redundanz. Durch eine geringe Beanspruchung der redundanten Elemente sind Ausfälle der redundanten Elemente möglich, die Ausfallwahrscheinlichkeit ist jedoch geringer als bei den „Originalelementen“. [4, S. 22].

Neben den beschriebenen Zuverlässigkeitskenngrößen Zusammenhangswahrscheinlichkeit und Redundanz existieren weitere Kenngrößen, die die Zuverlässigkeit in Netzwerken genauer definieren. Eine wichtiger Aspekt für die Berechnung der Zuverlässigkeit in Netzwerken ist die „Bedeutung der Knoten und Kanten [...] für die Funktionsfähigkeit des Systems“ [38, S. 19], die durch sogenannte *Importanzmaße* beschrieben werden kann. Wenn ein Netz als zuverlässig definiert ist, kann die Zusammenhangswahrscheinlichkeit $R(G)$ als Maß für die Zuverlässigkeit genutzt werden, sofern das Netz zusammenhängend ist. Die Bedeutung einer Kante e kann nun über die *Birnbaum-Importanz* ermittelt werden [38, S. 19]:

$$I_B(G,e) = \frac{\delta R(G)}{\delta p_e} = R(G_e) - R(G_{-e}) \tag{5.1}$$

G_e und G_{-e} sind Graphen, die aus G gebildet wurden. Für den zuerst genannten Graphen wird die Kante e kontrahiert, d. h. die Kante e wird entfernt und die dazu inzidenten[23] Knoten werden verschmolzen. Um den Graphen G_{-e} zu bilden, wird die Kante e aus G entfernt.
Wenn die Importanzbewertung die Verfügbarkeit der Kante e berücksichtigen soll, wird die *modifizierte Birnbaum-Importanz* verwendet [38, S. 19]:

$$\begin{aligned} I_0(G,e) &= R(G) - R(G_{-e}) = p_e I_B(G,e) \\ I_1(G,e) &= R(G_e) - R(G) = (1-p_e) I_B(G) \end{aligned} \tag{5.2}$$

Ein Beispiel einer für die Funktionalität des Netzes wichtigen Komponente ist das Home

[23] Eine Kante e ist zu einem Knoten v inzident, wenn v ein Endknoten von e ist [39, S. 12].

Location Register (siehe Abschnitt 4.2.2). Wenn diese Komponente ausfällt, sind die Positionen von Mobilstationen teilweise nicht mehr bekannt und die betroffenen Mobilstationen könnten nicht mehr erreichbar sein. Der Einfluss eines HLR-Fehlers auf ein UMTS-Netz wurde unter anderem von Dharmaraja et al. untersucht [8] (siehe Abschnitt 5.2.1).

Weitere Kenngrößen, die zur Einschätzung der Zuverlässigkeit und der Redundanz genutzt werden, sind die Knoten- und die Kantenzusammenhangszahlen.
Ein Graph hat eine *Knotenzusammenhangszahl* $\kappa(G)$, wenn er „k-zusammenhängend, jedoch nicht $(k+1)$-zusammenhängend ist“ [39, S. 91]. Ein Graph wird als *k-zusammenhängend* bezeichnet, wenn er mehr als k Knoten besitzt und wenn G_{-U} für alle $U \subseteq V(G)$ mit $|U| \leq k-1$ zusammenhängend ist, wenn G V Knoten besitzt. U ist eine *trennende Knotenmenge*, wenn G zusammenhängend, G_{-U} jedoch nicht zusammenhängend ist. Die Knotenzusammenhangszahl gibt in einem Netzwerk an, wie viele Knoten mindestens ausfallen müssen, damit alle Verbindungswege unterbrochen werden. [38, S. 15] „Die Knotenzusammenhangszahl findet daher für die Einschätzung der Sicherheit und Fehlerredundanz Anwendung“ [38, S. 15].
Das Gegenstück zur Knotenzusammenhangszahl ist die Kantenzusammenhangszahl. Sie gibt an, wie viele Kanten in einem zusammenhängenden Graphen G mindestens entfernt werden müssen, damit G nicht zusammenhängend ist (Schnittmenge). Ein Graph G wird als *k-fach kantenzusammenhängend* bezeichnet, wenn „G_{-F} für alle $F \subseteq E$ mit $|F| \leq k-1$ zusammenhängend ist“ [38, S. 15]. Die Kantenzusammenhangszahl $\lambda(G)$ ist gleich k, wenn G k-fach zusammenhängend ist, jedoch nicht $(k+1)$-fach zusammenhängend. Aus diesen Überlegungen kann man ableiten, dass die Kantenzusammenhangszahl eine geeignete Kenngröße für die „Übertragungssicherheit“ [38, S. 15] darstellt, da sie die Anzahl der Kanten angibt, die mindestens unterbrochen werden müssen, damit alle Wege zwischen zwei Knoten s und t unterbrochen sind. Des Weiteren ist die Kantenzusammenhangszahl „die Grundlage für die Algorithmen zur Berechnung von Schranken und Näherungswerten für Zuverlässigkeitskenngrößen in Kommunikationsnetzen“ [38, S. 15].

5.2 Erreichbarkeit in mobilen Kommunikationsnetzen

Bei der Kommunikation in mobilen Netzen spielt die Suche nach den Teilnehmern und deren Erreichbarkeit eine große Rolle. Um für andere Teilnehmer ansprechbar zu sein, wird jede Mobilstation im VLR registriert, sobald sie den Bereich einer BTS betritt (siehe Abschnitt 4.2.2). Dadurch ist die Lokalisierung einer Mobilstation im Idealfall leicht durchführbar. Die Lokalisierbarkeit kann jedoch durch Ausfälle der Hard- und Softwarekomponenten in der Infrastruktur verhindert werden. Für die Berechnung von zuverlässigen Kommunikationsnetz-Infrastrukturen gibt es mathematische Modelle, die die Ausfallwahrscheinlichkeiten von Komponenten berechnen und die Anzahl der Anfragen, die aufgrund dieser Ausfälle verloren gehen.
In der Car-to-I- (C2I-) Infrastruktur (siehe Abschnitt 6.2) spielen vor allem UMTS und LTE

sowie Ad-Hoc-Netze eine große Rolle. Daher soll anhand der UMTS-Infrastruktur die Berechnung der Ausfallwahrscheinlichkeit erläutert werden. Das Verfahren von Dharmaraja et al. basiert auf stochastischen Modellen, wie z.B. Markov-Ketten und dem Semi-Markov-Prozess [8, S. 132]. Um die Zuverlässigkeit von UMTS-Netzen zu berechnen, nutzen Dharmaraja et al. die hierarchische Struktur des Mobilfunknetzes und ermitteln zuerst die Zuverlässigkeit der einzelnen Schichten und anschließend die Zuverlässigkeit des Gesamtnetzes. Das UMTS-Netz wird in drei Schichten („level“) unterteilt. Das unterste enthält die Node Bs, die zweite die RNCs und das drittel Level die CSs (siehe Abschnitt 4.3).
Neben der UMTS-Architektur soll weiterhin auf die Ad-Hoc-Netze eingegangen werden, die v. a. in der Kommunikation zwischen Fahrzeugen (Car-to-Car-, C2C-Kommunikation, siehe Abschnitt 6.2) eine wichtige Rolle spielen. Kharbash und Wang beschäftigen sich mit der Zuverlässigkeit in Ad-Hoc-Netzen. Sie beziehen sich dabei auf das Vorhandensein von Wegen zwischen zwei Knoten und definieren die sonstigen Knoten in dem Netzwerk als Übermittler (siehe Abschnitt 5.2.2). [15]

5.2.1 Zuverlässigkeit in UMTS-Netzen

Die Zuverlässigkeit der Node Bs ist abhängig von den Hard- und Softwarebestandteilen des Elementes sowie der Stromversorgung und der Verbindung zu den RNCs, die entweder kabelgebunden oder als Richtfunk realisiert sein kann. Zu den zu berücksichtigenden Hardwarekomponenten wiederum gehören der *channel unit processor (CUP)*, die Antennen, der Sender und der Empfänger. Dharmaraja et al. bezeichnen den CUP als die „kritischste Komponente“ [8, S. 133], da alle Verarbeitungsfunktionen für die Kanäle hier durchgeführt werden und ein Defekt zu einem kompletten Ausfall des Node Bs führen würde [8, S. 133]. Um die Funktionalität der zu dem Node B gehörenden Zelle bei einem Ausfall des Node Bs nicht zu gefährden, wird durch die Netzbetreiber ein redundanter CUP eingesetzt, der aktiviert wird, wenn der erste CUP ausfällt. Der Einfluss der Antennen, Sender und Empfänger soll bei der Berechnung nach Dharmaraja et al. nicht weiter berücksichtigt werden, da ein Defekt durch Naturkatastrophen und mutwillige Zerstörungen sehr selten vorkommt [8, S. 133].
Zu den Softwareaufgaben des Node Bs gehört die Verarbeitung der Kanäle, z. B. die Kodierung, die Verschlüsselung und die Kanalerkennung. Wenn ein Fehler bei diesen Funktionen auftritt und dieser nicht beseitigt wird, sind die Kanäle nicht verfügbar für die Mobilstationen und der Node B ist im sogenannten *down state*. [8, S. 133] Die dritte Komponente, die für die Berechnung der Zuverlässigkeit berücksichtigt werden muss, ist die Spannungsversorgung. Die Node Bs verfügen über Stützbatterien, die die Funktionalität des Node Bs aufrecht erhalten, wenn die Spannungsversorgung unterbrochen wird. Die vierte und nach Dharmaraja et al. wichtigste Komponente ist die Schnittstelle zum RNC, die durch „Interferenzen, Rauschen und Ausblendung“ [8, S. 134] gestört werden kann. Deshalb werden besonders zuverlässige kabelgebundene oder kabellose Richtfunkverbindungen verwendet, um die Verbindung von Node Bs und RNC zu realisieren

[8, S. 134].
Für die Berechnung der Zuverlässigkeit der Node Bs wird durch Dharmaraja et al. angenommen, dass jeder RNC M Node Bs verwaltet. Die Zuverlässigkeit des $m-ten$ Node Bs berechnet sich in diesem Fall wie in Gleichung 5.3 angegeben. Der Zustandsraum S enthält alle Zustände, die ein Node B annehmen kann. Für die Berechnung muss der Zustand D (down state) allerdings abgezogen werden, da der Node B in diesem Fall keinerlei Funktionalität hat. $\pi_k(t)$ beschreibt die zeitabhängigen Wahrscheinlichkeiten der Zustände in S.

$$R_{NB\_m}(t) = \sum_{k \in S-\{D\}} \pi_k(t) \tag{5.3}$$

Aus der Gleichung 5.3 lässt sich die Zuverlässigkeit für alle Node Bs berechnen, die durch einen RNC verwaltet werden. Dabei muss hier allerdings berücksichtigt werden, dass Dharmaraja et al. für alle Node Bs die gleichen Ausfallwahrscheinlichkeiten annehmen. Daraus kann man ableiten, dass gilt: $R_{NB\_m}(t) = R_{NB}(t)\ \forall m$. Weiterhin muss als Bedingung erfüllt sein, dass mindestens ein Node B aktiv sein muss, damit Mobilstationen, die sich im Gebiet des betreffenden RNC befinden, das Netz erreichen können. Unter diesen Voraussetzungen lässt sich die Zuverlässigkeit für alle Node Bs des RNCs wie in Gleichung 5.4 berechnen. R_{NB} kann ersetzt werden durch die Gleichung 5.3 und m gibt die Anzahl der aktiven Node Bs an. [8, S. 134]

$$R_{GNB}(t) = \sum_{m=1}^{M} \binom{M}{m} [R_{NB}(t)]^m [1 - R_{NB}(t)]^{M-m} \tag{5.4}$$

Die zweite Schicht der UMTS-Infrastruktur beinhaltet die RNCs. Diese Komponenten setzen sich aus mehreren Hardwarebestandteilen sowie der Softwareeinheit, der Stromversorgung und der Schnittstelle zum CN zusammen [8, S. 134]. Für die Berechnung der Zuverlässigkeit wird wiederum angenommen, dass N RNCs von einem CN verwaltet werden. Die Zuverlässigkeit des $m-ten$ $(m \in \{1,2,...,N\})$ RNCs wird nach der Gleichung 5.5 berechnet. Die λ-Werte geben die Fehlerraten der einzelnen Komponenten der RNCs an, wobei λ_{eth}, λ_{comp} und λ_{adp} die Hardwarebestandteile beschreiben, λ_{pow} die Stromversorgung, λ_{link} die Verbindung zum CN und λ_{sft} die Softwareeinheit (siehe auch [8, S. 134]).

$$\begin{aligned} R_{RNC\_m}(t) = R_{GNB}(t) \times (2e^{-\lambda_{eth}t} - e^{-2\lambda_{eth}t}) \\ \times e^{-(\lambda_{comp}+\lambda_{adp})t} \times (2e^{-\lambda_{pow}t} - e^{-2\lambda_{pow}t}) \times e^{-(\lambda_{link}+\lambda_{sft})t} \end{aligned} \tag{5.5}$$

Aus der Formel für ein einzelnes RNC lässt sich wiederum diejenige für alle RNCs berechnen, die durch ein CN verwaltet werden (siehe Formel 5.6). Auch hier gilt die Bedingung, die schon für die Formel 5.4 angegeben wurde: Es muss mindestens ein RNC aktiv sein. Im Unterschied zur Gleichung für die Node Bs sind hier keine identischen Fehlerraten und Ausfallwahrscheinlichkeiten erforderlich. Die Menge S ist definiert als $S = \{1,2,...,N\}$

und $\mathbb{A}$ als $\mathbb{A}(S) = P(S) - \{\emptyset\}$, wobei $P(S)$ die Potenzmenge[24] von S ist.

$$R_{GRNC}(t) = \sum_{S_j \in \mathbb{A}(S)} \prod_{m \in S_j} R_{RNC\_m}(t) \prod_{k \notin S_j} [1 - R_{RNC\_k}(t)] \tag{5.6}$$

In der dritten Schicht der UMTS-Netzarchitektur befinden sich die CNs. Diese bestehen aus der CSD und der PSD[25] (siehe Abschnitt 4.3.2). Außerdem sind an die CNs mehrere Datenbanken angeschlossen, die das HLR und das VLR enthalten (siehe Abschnitt 4.3.2). [8, S. 135] Für einen Fehler der Datenbanken wird eine exponentielle Verteilungsfunktion angenommen [8, S. 135]. Unter der Bedingung, dass in der dritten Schicht K CNs vorhanden sind, ergibt sich für die Zuverlässigkeit eines $m-ten$ $(m \in \{1,2,...,K\})$ CN die Zuverlässigkeitsfunktion nach Gleichung 5.7. λ_{DB} bezeichnet hier die Fehlerrate der Datenbanken.

$$R_{CN}(t) = R_{GRNC}(t) \times [1 - \{1 - R_{PS}(t)\}\{1 - R_{CS}(t)\}] \times e^{-\lambda_{DB}t} \tag{5.7}$$

Um die unbekannten Größen $R_{PS}(t)$ und $R_{CS}(t)$ berechnen zu können, werden wiederum die einzelnen Bestandteile der beiden Komponenten PSD und CSD betrachtet. Die CSD besteht nach Dharmaraja et al. aus den *Interworking Managers (IM)* (λ_{IM}), die als Gateway zu netz-externen Elementen dienen, dem *Kernserver* (λ_{ser}), der die Ruffunktionen durchführt, dem *High Speed LAN* (λ_{LAN}), das IMs und Kernserver verbindet, und einer zweifachen *Spannungsversorgung* (λ_{pow}) [8, S. 135]. Mit diesen Komponenten und den jeweils in Klammern angegebenen Bezeichnungen für die Fehlerraten ergibt sich die Zuverlässigkeit für die CSD wie in Gleichung 5.8 angegeben.

$$R_{CS}(t) = e^{-(\lambda_{ser}+\lambda_{IM}+\lambda_{LAN})t}(2e^{-\lambda_{pow}t} - e^{-2\lambda_{pow}t}) \tag{5.8}$$

Die PSD beinhaltet die Komponenten SGSN (λ_{SGSN}) und GGSN (λ_{GGSN}) (siehe Abschnitt 4.3.2). Mit der Annahme einer exponentiellen Fehlerverteilung ergibt sich die Zuverlässigkeit wie in Gleichung 5.9 angegeben. [8, S. 135]

$$R_{PS}(t) = e^{-(\lambda_{SGSN}+\lambda_{GGSN})t} \tag{5.9}$$

Mit den Gleichungen 5.7, 5.8 und 5.9 kann nun die Zuverlässigkeit von mehreren CNs ermittelt werden. Auch hier hat abermals die Annahme Gültigkeit, dass mindestens einer der CNs aktiv sein muss. Jedoch müssen nicht alle CNs identisch sein, d. h. gleiche Eigenschaften aufweisen. Mit der Gleichung 5.10 kann die Zuverlässigkeit von K CNs berechnet werden. Die Menge S ist definiert als $S = \{1,2,...,K\}$ und $\mathbb{A}(S)$ als $\mathbb{A}(S) = P(S) - \{\emptyset\}$, wobei $P(S)$ als Potenzmenge von S definiert ist. [8, S. 135]

$$R_{GCN}(t) = \sum_{S_j \in \mathbb{A}(S)} \prod_{m \in S_j} R_{CN\_m}(t) \prod_{k \notin S_j} [1 - R_{CN\_k}(t)] \tag{5.10}$$

[24] Die Potenzmenge $P(S)$ einer Menge S ist die Menge aller Teilmengen der Menge S.

[25] Bei Dharmaraja et al. [8, S. 135] werden diese Elemente als CS und PS bezeichnet. Im weiteren Verlauf werden die Bezeichnungen von Abschnitt 4.3.2 weiterverwendet.

Nach der Berechnung der Zuverlässigkeit der einzelnen Schichten der UMTS-Architektur kann nun die Zuverlässigkeit des gesamten Netzwerkes berechnet werden. Diese ergibt sich nach Gleichung 5.11. Die Zuverlässigkeitsgröße R_{arch} beschreibt die Zuverlässigkeit der internen Verbindungen des Netzwerkes. Um diese zu berechnen, beziehen sich Dharmaraja et al. auf drei Topologien: Stern, Ring und Synchronous Optical Network -(SONET-)[26] Ring. [8, S. 135] Die Zuverlässigkeiten für die jeweiligen Topologien sind nach Gleichung 5.12 definiert.

$$R_{net}(t) = R_{arch}(t) \times R_{GCN}(t) \tag{5.11}$$

$$R_{arch}(t) = \begin{cases} (R_{link}(t))^K, & Ring \\ (R_{link}(t))^{K-1}, & Star \\ (1-\{1-R_{link}(t)\}^2)^K, & SONET \end{cases} \tag{5.12}$$

Bei der Ringtopologie ist die Anzahl der Verbindungen genauso groß wie die Zahl der CNs. Der „Defekt irgendeiner Verbindung führt zum Ausfall des gesamten Netzwerkes“ [8, S. 135]. Bei der Sterntopologie ist die Anzahl der Verbindungen um eins kleiner als die Anzahl der CNs, wenn alle CNs in einem Sternpunkt, der durch ein CN gebildet wird, verbunden sind. In zweifachen SONET-Ringen ist die Anzahl der Verbindungen genauso groß wie die Zahl der CNs. Jedoch gibt es eine zusätzliche Verbindung, die als „Back-up-Verbindung“ dient. [8, S. 135] Mit K als die Anzahl der CNs und $R_{link}(t)$ als die Zuverlässigkeit einer einzelnen Verbindung ergibt sich somit die Gleichung 5.12.

Unter Berücksichtigung der unterschiedlichen Ausfallursachen der Node Bs untersuchten Dharmaraja et al. den Einfluss von Node B-Fehlern. Dafür wurden für die Node Bs in Abhängigkeit von den ausgefallenen Elementen verschiedene Zustände eingeführt:
Wenn der Node B intakt ist, befindet er sich im Zustand U. Bei Ausfall der Spannungsversorgung in U_P. Wenn die Stützbatterien nicht geladen sind oder ausfallen, geht der Node B in den down state D über. Im Falle eines Softwarefehlers wechselt der Node B mit einer Wahrscheinlichkeit c in den Zustand U_S, wenn der Softwarefehler durch implementierte Fehlerkorrektur-Software und Wiederherstellungstechniken beseitigt werden kann. Anderenfalls geht der Node B in den Zustand D über. Die dritte Komponente, die ausfallen kann, ist der CUP. In diesem Fall wechselt der Node B in den Zustand U_H, wenn der redundante CUP eingesetzt wird. Anderenfalls erreicht er den Zustand D. Neben den genannten Zuständen kann der Node B auch in alle kombinierten Zustände fallen, wenn mehrere Komponenten ausfallen, es jedoch möglich war, sie zu ersetzen. Die Übergänge zwischen den Zuständen und die dazugehörigen Wahrscheinlichkeiten sind in der Abbildung 5.1 dargestellt. [8, S. 134] Wenn sich der Node B in den Zuständen U oder U_P befindet, ist er „voll funktionsfähig“ [8, S. 136], im Zustand D vollkommen aus und in den sonstigen Zuständen teilweise funktionsfähig.

[26] SONET ist ein von Bellcore entwickelter Standard für optische Übertragungssysteme auf der Bitübertragungsschicht. Nahezu identisch ist die *Synchronous Digital Hierarchy (SDH)*, die durch die Internationale Fernmeldeunion (International Telecommunication Union, ITU) standardisiert wurde. [36, S. 193]

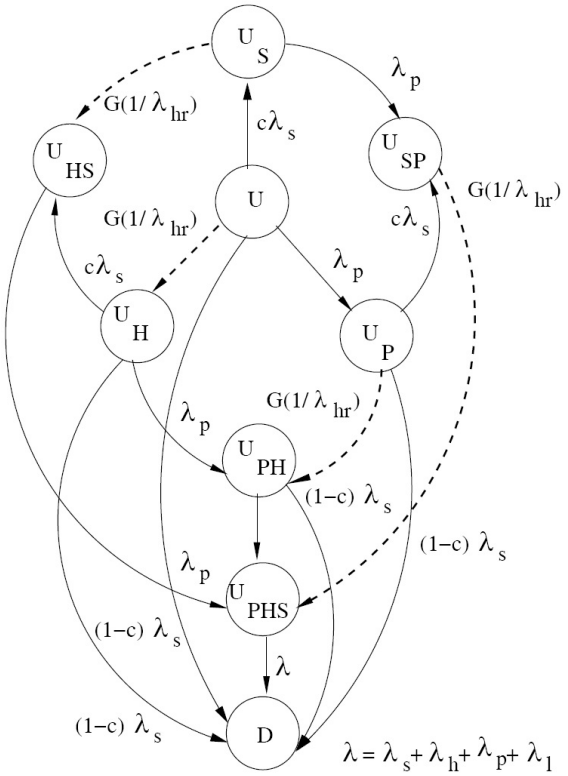


Abbildung 5.1: Zustandsübergangsdiagramm für die Zuverlässigkeit von Node Bs [8, S. 134]

Mit diesen Betrachtungen kann der Einfluss eines Fehlers eines Node Bs berechnet werden (siehe Gleichung 5.13). Hierbei müssen der down state und die Zustände, in denen eine teilweise Funktionstüchtigkeit vorliegt, betrachtet werden.

$$\begin{aligned} I_{nodeB\_i}(t) = [&\{P_b^H(t) \times (\pi_{U_H}(t) + \pi_{U_{PH}}(t))\} \\ &+\{1 \times \pi_D(t)\} + \{P_b^S(t) \times (\pi_{U_S}(t) + \pi_{U_{SP}}(t))\} \\ &+\{P_b^S(t)P_b^H(t) \times (\pi_{U_{HS}(t)} + \pi_{U_{PHS}(t)})\}] \times 100\% \end{aligned} \tag{5.13}$$

Die π-Werte sind die Wahrscheinlichkeiten der einzelnen Zustände, die ein Node B annehmen kann. P_b^H und P_b^S sind die Wahrscheinlichkeiten, dass Anrufe aufgrund von Soft- bzw. Hardwarefehlern geblockt werden.

Um die Wahrscheinlichkeit für geblockte Anrufe, die auf einen Hardwarefehler zurückzuführen sind, berechnen zu können, muss das Verhalten des Node B in einem Fehlerfall bekannt sein. Wenn das OMC (siehe Abschnitt 4.2.2), das fortwährend die Leistung des Netzwerkes überwacht, einen Fehler eines CUPs feststellt, leitet es den Umschaltprozess zu dem redundanten CUP ein. Das Umschalten erfolgt mit einer Wahrscheinlichkeit p. Im Fehlerfall wird die Umschaltfunktion zurückgesetzt und der Prozess anschließend wiederholt. Während des Umschaltprozesses und der im Anschluss erforderlichen Synchronisation des neuen CUP steht nur ein Teil der Kanäle des Node Bs für Nicht-Echtzeit-Anrufe zur Verfügung, wenn diese Anrufe „verzögerungstolerant sind und keine festen QoS-Grenzen besitzen“ [8, S. 136] (QoS siehe Abschnitt 4.2.4). Verzögerungs-

empfindliche Echtzeit-Anfragen gehen verloren. Wenn das Umschalten nicht erfolgreich war und während der Fehlerdetektion durch das OMC, werden alle Anfragen von Mobilstationen abgelehnt.
Die Wahrscheinlichkeit, dass Anfragen an einen Node B mit n Kanälen geblockt werden, sei $p_b(n)$. Die Anrufe werden geblockt, wenn alle Kanäle belegt sind. Dann berechnet sich $p_b(n)$ als

$$p_b(n) = \frac{\left(\frac{\lambda}{\mu}\right)^n / n!}{\sum_{j=0}^{n} \left(\frac{\lambda}{\mu}\right)^j / j!} \tag{5.14}$$

Die Variablen λ und μ sind die *inter-arrival time* , d. h. die Zeit zwischen den Anfragen, und die *channel holding time (CHT)*, d. h. die Zeit, die ein Anruf einen Kanal belegt [2].
Die Wahrscheinlichkeit für geblockte Anfragen, die durch einen Hardwarefehler hervorgerufen wurden, ergibt sich als

$$P_b^H(t) = \sum_{i \in I} P_b(i)\pi_i(t) \tag{5.15}$$

Dabei ist $P_b(i)$ die Wahrscheinlichkeit für geblockte Anfragen, wenn sich die Hardware des Node Bs in einem bestimmten Zustand i befindet. Wenn ein CUP ausfällt, der Wechsel zu dem redundanten CUP misslingt oder der redundante CUP ebenfalls ausfällt, ist $P_b(i) = 1$, da der Node B nicht mehr erreichbar ist für Mobilstationen. Während des CUP-Wechsels und der Synchronisation des redundanten CUPs im Falle eines erfolgreichen CUP-Umschaltens berechnet sich die Wahrscheinlichkeit als $P_b(i) = r_n \times p_b(x) + (1 - r_n) \times 1$. $\pi_i(t)$ ist die Wahrscheinlichkeit für das Eintreten eines Zustandes i zum Zeitpunkt t. r_n ist der Anteil der Nicht-Echtzeit-Anrufe, $1 - r_n$ der Anteil der Echtzeit-Anfragen. Letztere werden in der Gleichung mit Eins multipliziert, da in den zugehörigen Zuständen nur verzögerungsunempfindliche Nicht-Echtzeit-Anrufe angenommen werden. [8, S. 137]

Neben den Hardwarefehlern haben auch Softwarefehler einen Einfluss auf die Abblockwahrscheinlichkeit von Anrufen. Für die Berechnung der Wahrscheinlichkeit, dass Anrufe geblockt werden aufgrund von Softwarefehlern, ist es wiederum notwendig zu wissen, wie der Node B in einem Fehlerfall reagiert.
Dharmaraja et al. nehmen an, dass ein Node B n Kanäle hat, einer von diesen arbeitet als Steuerungskanal, d. h. auf diesem werden nur Kontrollinformationen übertragen. Demzufolge ist die Anzahl der Sprachkanäle $(n-1)$. Die Wahrscheinlichkeit, dass auf dem Steuerungskanal eine Fehler auftritt sei q, die für die Sprachkanäle $(1-q)$. Ein Fehler eines Steuerungskanals wird mit einer Wahrscheinlichkeit c_1 durch eine automatische Wiederherstellungssoftware behoben oder es wird ein Sprachkanal zu dem neuen Steuerungskanal mit Wahrscheinlichkeit c_2. Wenn der Steuerungskanal weder wiederhergestellt noch ersetzt werden kann, wechselt der Node B mit Wahrscheinlichkeit $1-(c_1+c_2)$ in den down state D. Die Wahrscheinlichkeit, dass ein Fehler in einem Sprachkanal behoben werden kann ist p, mit $(1-p)$ kann die Funktionalität des Kanals nicht wiederhergestellt werden. Für die Berechnung der Wahrscheinlichkeit, dass Anrufe

geblockt werden, wird durch Dharmaraja et al. angenommen, dass *i* Kanäle intakt, jedoch beschäftigt sind zum Zeitpunkt der Anfrage.
Mit diesen Annahmen kann die Wahrscheinlichkeit, dass Anrufe aufgrund von Softwarefehlern abgelehnt werden, berechnet werden mit:

$$P_b^S = \sum_{i=1}^{n} \pi_i(t) P_b(i) + \sum_{i=1}^{n} \pi_{F_i}(t) P_b(F_i) + \sum_{i=2}^{n} \pi_{S_i}(t) P_b(S_i) + \pi_D(t) P_b(D) \tag{5.16}$$

Es sei $P_b(i) = p_b(i)$, wenn der Node B i $(1 \leq i \leq n)$ fehlerfreie Kanäle besitzt. $p_b(i)$ ist die Wahrscheinlichkeit, dass alle „*i* Kanäle beschäftigt sind“ [8] und kann von der Gleichung 5.14 übertragen werden. Die Zustände F_i $(1 \leq i \leq n)$ erreicht der Node B, wenn ein Steuerungskanal ausfällt, die Zustände S_i $(2 \leq i \leq n)$ bei Ausfall eines Sprachkanals. Der Node B wechselt in den Zustand D, wenn in einem Fehlerfall der Steuerungskanal nicht wiederhergestellt oder kein Sprachkanal als neuer Steuerungskanal genutzt werden kann. Die π-Werte sind die zeitabhängigen Wahrscheinlichkeiten für die einzelnen Zustände. [8, S.137f.]
Mithilfe der Gleichung 5.13 kann nun die Funktion für alle Node Bs berechnet werden, die durch ein RNC verwaltet werden (siehe Gleichung 5.17). M ist die Anzahl der Node Bs eines RNCs.

$$I_{nodeB}(t) = \frac{1}{M} \sum_{i=1}^{M} I_{nodeB\_i}(t) \tag{5.17}$$

Ausgehend von diesen Erkenntnissen formulierten Dharmaraja et al. einen Ausdruck für den Einfluss der Fehler eines RNCs, wobei $Rel_{RNC}(t)$ die Zuverlässigkeit des RNCs ist:

$$I_{RNC}(t) = I_{nodeB}(t) + ([1 - Rel_{RNC}(t)] - I_{nodeB}(t)[1 - Rel_{RNC}(t)]) \tag{5.18}$$

Als drittes untersuchten Dharmaraja et al. den Einfluss eines HLR-Fehlers auf die Zuverlässigkeit eines Netzwerkes. Bei einem Ausfall des HLR können eingehende Anrufe nicht an die angesprochene Mobilstation weitergeleitet werden, da die aktuelle Position nicht aus der Datenbank ausgelesen werden kann. Um zu gewährleisten, dass nur ein Minimum an Anfragen verloren geht, wird ein Sicherungs-HLR eingesetzt. Dieses Backup ist allerdings „veraltet“ [8, S. 138]. Die Positionen der Mobilstationen, die während der Ausfallzeit des HLR ihren Ort gewechselt haben, sind nicht eingetragen. Eingehende Anfragen für diese Mobilstationen können demzufolge nicht weitergeleitet werden. Die Aktualisierung der HLR-Einträge kann auf zwei Arten erfolgen: Zum einen wird eine Aktualisierung der eigenen Position durch eine Mobilstation hervorgerufen, wenn sie in den Bereich eines neuen Node B eintritt. Die Voraussetzungen für dieses Positionsupdate sind allerdings ein fehlerfrei arbeitender Node B und ein fehlerloses RNC. Wenn das HLR innerhalb eines Zeitintervalls s keine Meldung von einer Mobilstation bekommt, sendet sie eine automatische Positionsabfrage an die betreffende Mobilstation. Demzufolge ist die maximale Zeit bis zu einem Update der Positionen die Zeit s. [8, S. 138] Dharmaraja et al. sind der Meinung, dass die Effizienz der Positionsupdate-Prozedur umgekehrt

proportional zu der Durchschnittsanzahl von Anrufen ist, die verloren gehen, weil eine ineffiziente Positionsupdate-Prozedur die Verzögerung bei Datenbankenabfragen ansteigen lässt [8, S. 138].

Dharmaraja et al. beschäftigten sich mit der Frage, wie viele Anrufe während der Fehler- und Wiederherstellungsphase des HLR verloren gehen. Sie nahmen an, dass die Anzahl der verlorenen Anrufe X_{tot} die Summe der verlorenen Anfragen während der zwei Einzelphasen ist, d. h. $X_{tot} = X_F + X_R$ mit X_F als die Anzahl der verlorenen Anrufe während der Fehlerphase und X_R als die Anzahl während der Wiederherstellungsphase. Die durchschnittliche Anzahl berechnet sich dementsprechend als $E(X_{tot}) = E(X_F) + E(X_R)$. [8, S. 138]
Zur Berechnung von X_{tot} werden zunächst die einzelnen Anzahlen definiert. Die Wahrscheinlichkeit, dass während einer Fehlerperiode der Länge $U = u\ k$ Anrufe die Mobilstation erreichen, ergibt sich nach Dharmaraja et al. wie in Gleichung 5.19. Dazu wurde von ihnen angenommen, dass die Anrufe, die eine HLR erreichen, sich gemäß einer Poisson-Verteilung mit dem Parameter λ_a verhalten.

$$P(X_F = k|U = u) = \frac{(\lambda_a u)^k}{k!} e^{-\lambda_a u}, k = 0, 1, ... \tag{5.19}$$

Daraus folgt

$$\begin{aligned} P(X_F = k) &= \int_0^\infty Pr(X_F = k|U = u) f_U(u) du \\ &= \int_0^\infty \frac{(\lambda_a u)^k}{k!} e^{-\lambda_a u} \lambda_r e^{-\lambda_r u} du = \frac{\lambda_a^k \lambda_r}{(\lambda_a + \lambda_r)^{k+1}} \end{aligned} \tag{5.20}$$

Die Variable λ_r ist der Parameter der exponentiell verteilten Fehlerzeit des HLRs. $f_U(u)$ ist die Wahrscheinlichkeitsdichtefunktion der Fehlerzeit. Die Anzahl der verlorenen Anrufe während einer Fehlerperiode ergibt sich somit als

$$E(X_F) = \sum_{k=0}^{\infty} \frac{k \lambda_a^k \lambda_r}{(\lambda_a + \lambda_r)^{k+1}} = \frac{\lambda_a}{\lambda_r},\ \lambda_r > 0 \tag{5.21}$$

Neben der Anzahl der verlorenen Anrufe während der Fehlerperiode wird nun die Anzahl während der Wiederherstellungsperiode benötigt. Die durchschnittliche Anzahl wird durch Dharmaraja et al. definiert als $E(X_R) = \lambda_a E(L)$ [8, S. 139], wenn L die Länge der Wiederherstellungsperiode und als $min(s, t_c, t_u)$ definiert ist. t_c ist das Zeitintervall des Moments, bei dem das HLR erneut arbeitet und ein Anruf ankommt. t_u ist definiert als das Zeitintervall des Moments, bei dem das HLR arbeitet und es ein Positionsupdate erreicht. Die Wahrscheinlichkeitsdichtefunktion der $f_l(t)$ der Zufallsvariable L ist nach Dharmaraja et al. definiert nach Gleichung 5.22 mit $\delta(\cdot)$ als Dirac-Funktion.

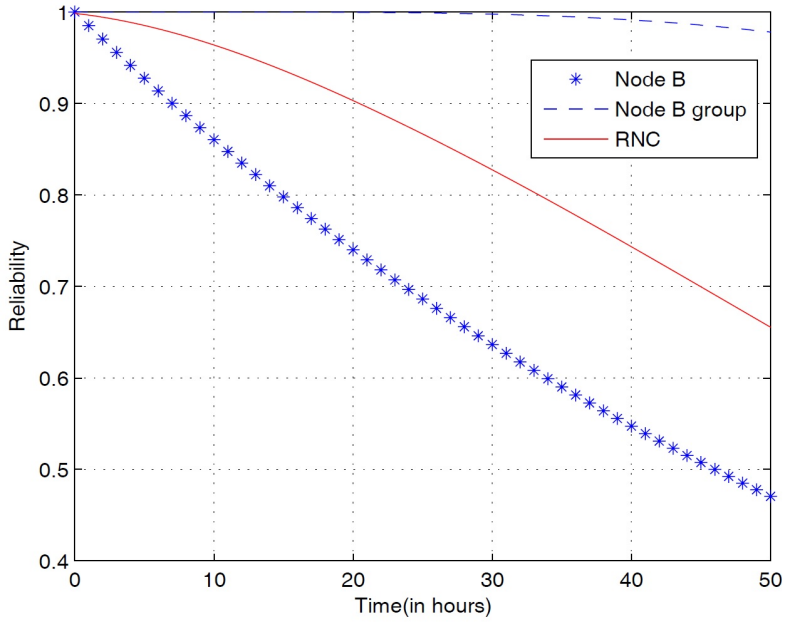


Abbildung 5.2: Zuverlässigkeit eines Node Bs, einer Gruppe mehrerer Node Bs und eines RNCs [8, S.139]

$$f_l(t) = \begin{cases} (p\lambda_u + \lambda_c)e^{-(p\lambda_u+\lambda_c)t}, & 0 \leq t < s \\ e^{-(p\lambda_u+\lambda_c)t}\delta(t-s), & t = s \\ 0, & otherwise \end{cases} \tag{5.22}$$

Die durchschnittliche Länge der Wiederherstellungsperiode ist definiert nach Gleichung 5.23. p ist die Wahrscheinlichkeit, dass eine Mobilstation bei einem Positionswechsel in den Bereich eines fehlerfreien Node Bs oder eines fehlerlosen RNCs wechselt und ist somit definiert als $p = \frac{1}{N} \cdot \frac{1}{M}$, mit M als der Anzahl Node Bs im Bereich eines RNC und N als Anzahl der RNC, die von einem MSC verwaltet werden. T_u ist der Zeitabstand zwischen den Positionsupdate-Abfragen und exponentiell verteilt mit dem Parameter $p\lambda_u$.

$$E(L) = \frac{1 - e^{-(\lambda_c + p\lambda_u)s}}{\lambda_c + p\lambda_u} \tag{5.23}$$

Somit ergibt sich für die durchschnittliche Gesamtzahl an verlorenen Anrufen während einer Fehler- und einer Wiederherstellungsperiode

$$E_{tot} = E(X_F) + E(X_R) = \frac{\lambda_a}{\lambda_r} + \lambda_a E(L) \tag{5.24}$$

Durch die Anwendung der durch Dharmaraja et al. vorgestellten Funktionen konnten die Einflüsse der einzelnen Komponenten auf die Netzzuverlässigkeit ermittelt werden.
Zum einen konnte durch die Simulation nachgewiesen werden, dass die Zuverlässigkeit bei mehreren Node Bs größer ist als diejenige für einen einzelnen Node B, da die Mobilstationen bei dem Ausfall eines Node Bs in den Bereich des Nachbarknoten wechseln

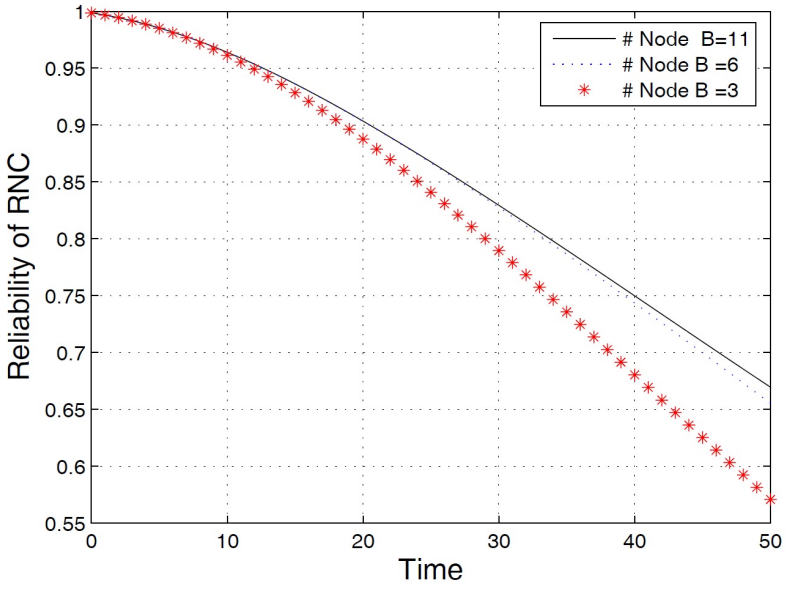
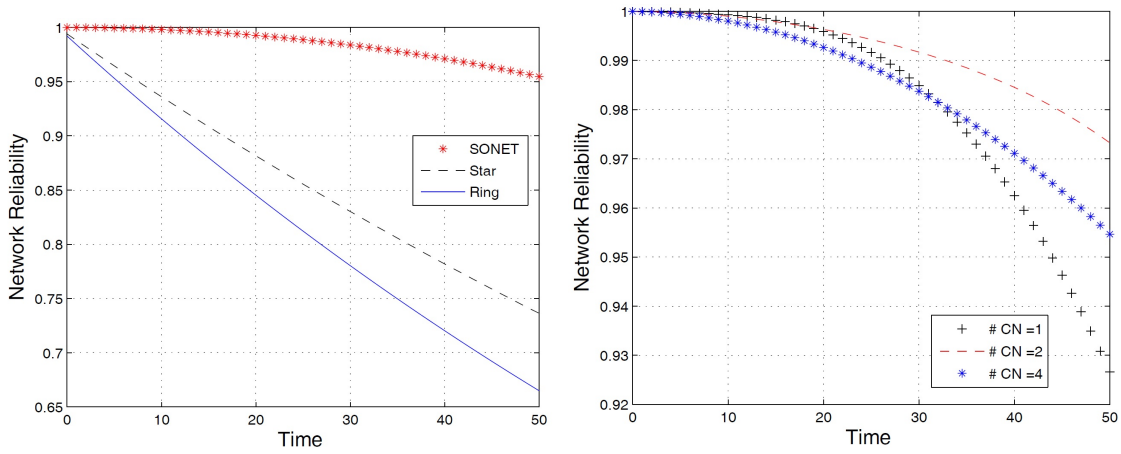


Abbildung 5.3: Einfluss der Anzahl der Node Bs auf die Zuverlässigkeit eines RNCs [8, S.140]

können (siehe Abbildung 5.2). In diesem Zusammenhang konnte durch Dharmaraja et al. auch nachgewiesen werden, dass die Zuverlässigkeit mit steigender Anzahl der Node Bs zunimmt. Die Zuverlässigkeit ist allerdings auf ein Maximum begrenzt, d. h. dass mit steigender Anzahl der Node Bs die Zuverlässigkeit immer weniger zunimmt und nicht unendlich erhöht werden kann (siehe Abbildung 5.3). Der dritte Aspekt, der im Zusammenhang mit der Netzwerkzuverlässigkeit untersucht wurde, ist die Netztopologie. Die betrachteten Topologien sind die schon beschriebenen Ring-, Stern- und SONET-Architekturen. Dharmaraja et al. erkannten, dass die Ringtopologie die geringste Zuverlässigkeit erreicht, die SONET-Struktur die größte. Bei der Ringarchitektur ist die Zuverlässigkeit von jeder einzelnen Verbindung und jedem CN abhängig, bei der Sterntopologie hingegen nur von der Zuverlässigkeit des zentralen CN. Daher liegt die Zuverlässigkeit der Sterntopologie zwischen denen der Ring- und SONET-Architekturen (siehe Abbildung 5.2.1). Die zweifachen SONET-Ringe erlauben Selbstheilung, daher ist ihre Zuverlässigkeit am größten.
In Bezug auf die Netzwerkzuverlässigkeit untersuchten Dharmaraja et al. neben der Topologie auch den Einfluss der Anzahl der CNs. Sie stellten fest, dass sich die Zuverlässigkeit des Netzwerkes verbessert, wenn die Anzahl der CNs von eins auf zwei erhöht wird. Bei einer weiteren Vergrößerung des Netzes auf vier CNs verringert sich die Zuverlässigkeit allerdings wieder (siehe Abbildung 5.2.1). Aus diesen Erkenntnissen schließen Dharmaraja et al., dass sich die Netzwerkzuverlässigkeit bei einer Vergrößerung der Anzahl der CNs nicht vergrößert, sondern dass sich nur die Unterhaltungskosten erhöhen.

Neben der Zuverlässigkeit eines UMTS-Netzes betrachteten Dharmaraja et al. auch die Überlebensfähigkeit der Netzwerke. Sie stellten fest, dass der Anteil der Mobilstationen, die durch RNC-Fehler beeinflusst wurden, größer ist als derjenige, der durch Node B-



(a) Netzwerkzuverlässigkeit für drei Netzwerkarchitekturen (b) Einfluss der CN auf die Netzwerkzuverlässigkeit

Abbildung 5.4: Einfluss auf die Netzwerkzuverlässigkeit [8, S. 141]

Fehler beeinträchtigt war. Wenn ein RNC ausfällt, beeinflusst er alle Mobilstationen, die sich in den Bereichen der Node Bs befinden, die durch den RNC verwaltet werden. Wenn ein Node B ausfällt, beeinträchtigt dies hingegen nur die Mobilstationen in dem jeweiligen Einzugsgebiet. Daher ist die Anzahl der durch das RNC beeinflussten Mobilstationen größer. [8, S. 140]
Als letztes untersuchten Dharmaraja et al. den Einfluss eines HLR-Ausfalls auf die Überlebensfähigkeit eines Netzes. Der HLR-Fehler ist exponentiell verteilt mit dem Parameter λ_r. Dharmaraja et al. erkannten, dass die durchschnittliche Anzahl verloren gegangener Anrufe kleiner wird, wenn λ_r größer wird, da mit steigendem λ_r die mittlere Ausfallzeit des HLRs sinkt.

Mit ihrer Forschung konnten Dharmaraja et al. nachweisen, dass eine dichtere Infrastruktur für ein UMTS-Netz über eine bestimmte Anzahl von Komponenten hinaus keine Verbesserung der Zuverlässigkeit erzielt. Mit diesen Erkenntnissen leisten sie einen Beitrag, der es Netzbetreibern ermöglicht, die Netze bezüglich Zuverlässigkeit, Überlebensfähigkeit und Fehlertoleranz effizienter zu gestalten.

5.2.2 Zuverlässigkeit in Ad-Hoc-Netzen

Die UMTS-Netze spielen neben LTE in der Car-to-Infrastructure- (C2I-) Kommunikation eine entscheidende Rolle. In der Kommunikation zwischen den Fahrzeugen, d. h. in der Car-2-Car- (C2C-) Kommunikation, sind die Ad-Hoc-Netze verbreitet. Wie bei den UMTS-Netzen gab es auch für die Zuverlässigkeit in Ad-Hoc-Netzen in den letzten Jahren mehrere Forschungsarbeiten. Eine davon beschäftigt sich mit der *Two-Terminal Reliability* in mobilen Ad-Hoc-Netzen [15]. Kharbash und Wang entwickelten ein Simulationsmodell, das es ermöglicht, durch die Veränderung verschiedener Parameter die Eigenschaften eines Ad-Hoc-Netzwerkes zu ermitteln, die die Zuverlässigkeit in diesen

Netzen beeinflussen. Nach Kharbash und Wang ist das typische Problem zur Ermittlung der Zuverlässigkeit in einem Netzwerk, die „Wahrscheinlichkeit zu berechnen, dass ein bestimmtes Set an Knoten über eine gegebene Zeit miteinander kommunizieren kann“ [15]. Bei der Two-Terminal Reliability werden zwei Knoten aus einer Menge von Knoten als Quelle (source, s) und als Ziel (auch als Senke bezeichnet, destination, d) definiert. Alle anderen Knoten dienen als Vermittler, um einen Kommunikationspfad zwischen der Quelle und der Senke aufzubauen. Snow et al. bezeichnen jede Komponente in einem kabellosen Netzwerk, sowohl Knoten als auch die Verbindung zwischen den Knoten, als eine „potenzielle Fehlerstelle“ [35] und die Zuverlässigkeit sei abhängig von dem Grad der Redundanz [35] und von der Zuverlässigkeit der einzelnen Komponenten [15].
Um die Zuverlässigkeit eines Ad-Hoc-Netzes zu simulieren, verwenden Kharbash und Wang zwei Verfahren von Rai et al. [28] und Aggarwal et al. [1]. Der Algorithmus von Rai et al. berechnet die minimale Schnittmenge, das sind „Verbindungen, deren Entfernen die Routen zwischen Quelle und Senke [in einem gerichteten Netzwerk] trennt“ [15]. Aggarwal et al. führten die Überlegung ein, dass ein Ausfall eines Knotens einen Fehler aller Verbindungen herbeiführt, die inzident zu dem betreffenden Knoten sind [15]. Daher kann man jeden Kommunikationszweig als eine Serie von aufeinanderfolgenden Knoten mit jeweils zwischengeschalteten Verbindungen betrachten. Diese Überlegung kann benutzt werden, um den Einfluss fehlerbehafteter Knoten zu berücksichtigen. Dazu wird zuerst der Ausdruck für die Zuverlässigkeit des Kommunikationspfades aufgeschrieben, ohne die Ausfallwahrscheinlichkeiten der Knoten zu berücksichtigen. Danach wird jeder Ausdruck durch die Funktion der Knoten und Verbindungen ersetzt. [15]
Für die Ermittlung der Zuverlässigkeit werden die Ideen von Rai et al., die nur für Netze mit feststehenden und fehlerfreien Knoten und mit unersetzbaren Verbindungen mit bekannter Wahrscheinlichkeit implementiert waren, verwendet und auf mobile Ad-Hoc-Netze angewendet. Die Fehlerraten und die Mobilität der Knoten werden durch die Integration des Verfahrens von Aggarwal et al. berücksichtigt. [15]

Für die Berechnung wird von Kharbash und Wang ein Ad-Hoc-Netz G angenommen mit N Knoten und E Kanten. Die Quelle wird mit s bezeichnet und das Ziel als d. Die Wahrscheinlichkeit, dass ein Knoten ausfällt, sei $1-p_n$, die der Verbindungen sei $1-p_e$. p_e ist abhängig von den Knoten, daher kann p_e als $p_e = Pr(e\ exists \mid n_i\ and\ n_j\ are\ operating)$ ausgedrückt werden, wenn die Kante e die Knoten n_i und n_j verbindet. Jede Kante kann zwei Zustände annehmen, sie ist entweder ausgefallen (Zustand ist Null) oder arbeitet (Zustand ist Eins). Es kann ein Zustandsvektor $\boldsymbol{S(t)}$ definiert werden, der alle Zustände der Kanten enthält. Damit ergibt sich $\boldsymbol{S(t)} = [S_1(t), S_2(t), ..., S_E(t)]$. Die Wahrscheinlichkeit für einen Zustand S(t) lässt sich berechnen mit Gleichung 5.25. [15]

$$Pr(S(t)) = \prod_{e=1}^{E} p_e^{S_e(t)} (1-p_e)^{1-S_e(t)} \tag{5.25}$$

Die Zustände werden durch eine Strukturformel $\phi_{s,d}$ überwacht. Durch diese Strukturformel wird kontrolliert, ob mindestens ein Pfad existiert, der den Knoten s mit dem Knoten d verbindet. Wenn dieser Pfad vorhanden ist, wird $\phi_{s,d}(S(t)) = 1$, anderenfalls 0. Unter diesen Annahmen ergibt sich die folgende Zuverlässigkeitsfunktion:

$$Rel_{s,d}[G(t)] = \sum_{all\ S(t)} \phi_{s,d}(S(t)) Pr(S(t)) \tag{5.26}$$

Der Algorithmus von Kharbash und Wang beinhaltet drei grundsätzliche Schritte: die Erstellung einer Verbindungsmatrix, das Ersetzen der jeweils zu betrachtenden Verbindungen und die Ableitung des Ausdrucks für die Zuverlässigkeit.
Die Verbindungsmatrix $M[c](t)$ ist eine $m \times m$-Matrix, wenn m der Anzahl der Knoten in einem Ad-Hoc-Netz entspricht. Die Einträge $x_{i,j}$ werden Null, wenn zwischen den Knoten i und j keine Verbindung besteht. Dieser Fall tritt auf, wenn sich zwei Knoten nicht in der Reichweite des jeweils anderen Knoten befinden oder wenn das ausgetauschte Funksignal am Empfänger einen festgelegten Grenzwert unterschreitet. Anderenfalls wird in das Matrixelement der Wert $x_{i,j}$ eingetragen, wobei der Index i immer kleiner ist als der Index j [15]. Die Matrix ist demzufolge symmetrisch aufgebaut. In der Abbildung 5.5 ist ein Beispiel-Ad-Hoc-Netz mit fünf Knoten dargestellt [15]. Die zugehörige Verbindungsmatrix ist in der Gleichung 5.27 dargestellt. Um die Verbindungsmatrix für die nächsten Schritte verwenden zu können, muss sie zuvor modifiziert werden. Dazu wird die Einheitsmatrix $\boldsymbol{I}$ aufaddiert. Außerdem müssen die erste Spalte und die letzte Zeile gelöscht werden. Ausgehend von der Matrix aus Gleichung 5.27 ergibt sich somit eine modifizierte Matrix wie in Gleichung 5.28.

$$M[c](t) = \begin{pmatrix} 0 & x_{1,2} & x_{1,3} & 0 & 0 \\ x_{1,2} & 0 & 0 & x_{2,4} & 0 \\ x_{1,3} & 0 & 0 & x_{3,4} & x_{3,5} \\ 0 & x_{2,4} & x_{3,4} & 0 & x_{4,5} \\ 0 & 0 & x_{3,5} & x_{4,5} & 0 \end{pmatrix} \tag{5.27}$$

$$M[c]_{mod}(t) = \begin{pmatrix} x_{1,2} & x_{1,3} & 0 & 0 \\ 1 & 0 & x_{2,4} & 0 \\ 0 & 1 & x_{3,4} & x_{3,5} \\ x_{2,4} & x_{3,4} & 1 & x_{4,5} \end{pmatrix} \tag{5.28}$$

Der nächste Schritt des Algorithmus' ist die Substitution der einzelnen Einträge. Es gibt zwei Arten: *0-Sub* und *1-Sub*. Die 0-Sub-Operation ersetzt den aktuell betrachteten Eintrag durch eine Null. Diese Operation entfernt „die Funkverbindung aus dem kabellosen Netzwerk“ [15]. Wenn in der Gleichung 5.28 beispielsweise das Element $x_{1,2}$ verwendet wird, ergibt sich die Matrix in Gleichung 5.29. Bei der 1-Sub-Operation wird das derzeit interessierende Element auf Eins gesetzt. Dieses Verfahren simuliert den nächsten Schritt in Richtung Zielknoten in einem Netzwerk, der über eine aktive Verbindung durchgeführt wird. Bei dieser Operation wird als Erstes das Element, das be-

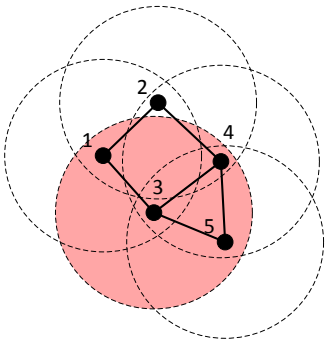


Abbildung 5.5: Ad-Hoc-Netzwerk mit fünf Knoten (nach [15])

trachtet wird, durch Eins ersetzt. Anschließend wird in der Spalte, in dem dieses Element eingetragen ist, nach einer weiteren Eins gesucht. Die Zeile mit dieser Eins wird nun auf die erste Zeile aufaddiert. Zum Schluss werden die Zeile mit der Eins und die Spalte mit dem betrachteten Element gelöscht. Die Matrix hat nun eine Größe von $m-1 \times m-1$, wenn m die ursprüngliche Zeilen- und Spaltenzahl der Matrix war. Angewendet auf das Element $x_{1,2}$ in der Matrix aus 5.28 ergibt sich die Matrix in 5.30. [15]

$$A_{new}[x] = \begin{pmatrix} 0 & x_{1,3} & 0 & 0 \\ 1 & 0 & x_{2,4} & 0 \\ 0 & 1 & x_{3,4} & x_{3,5} \\ x_{2,4} & x_{3,4} & 1 & x_{4,5} \end{pmatrix} \tag{5.29}$$

$$A_{new}[x] = \begin{pmatrix} x_{1,3} & x_{2,4} & 0 \\ 1 & x_{3,4} & x_{3,5} \\ x_{3,4} & 1 & x_{4,5} \end{pmatrix} \tag{5.30}$$

Mit diesen Grundlagen kann der Ausdruck für die Zuverlässigkeit abgeleitet werden. Dieser Schritt teilt sich in mehrere kleinere Teilschritte, die Kharbash und Wang folgendermaßen beschreiben [15]:

1. Die Zeit auf $t = 0$ setzen.
2. Die Verbindungsmatrix aufstellen und wie beschrieben modifizieren.

    a) Eine Variable *Level* wird auf $l = 0$ gesetzt.
    b) Die Spalten, in denen in der ersten Zeile der Matrix Einträge ungleich Null sind, werden nach dem Vorhandensein von Einsen in den weiteren Zeilen untersucht. Anschließend werden die Spalten sortiert: die Spalten ohne eine Eins werden an den Anfang gesetzt, die anderen Spalten folgen Index-sortiert.
    c) Auf die Elemente der ersten Zeile der modifizierten Matrix werden die logischen Ausdrücke 1, 01, 001, ... angewendet, d. h. im ersten Schritt wird in der Matrix aus 5.28 $x_{1,2} = 1$ gesetzt, im zweiten Schritt wird $x_{1,2} = 0$ und $x_{1,3} = 1$.

   d) Entsprechend der in 2c gesetzten logischen Ausdrücke werden die 0-Sub- und die 1-Sub-Operationen ausgeführt. Wenn bei einer 1-Sub-Operation keine Eins in der Spalte vorhanden sein sollte, ergibt die 1-Sub-Operation eine (1).
   e) Der Schritt 2d muss für alle Elemente in der ersten Zeile der Matrix wiederholt werden, die nicht Null sind.
3. Die Variable l wird inkrementiert und die Schritte 2a bis 2e werden wiederholt, bis das Ergebnis der Substitutionen eine (1), eine (0) oder eine einzelne Variable liefert.
4. Durch das Verfahren erhält man einen Baum mit den einzelnen Schritten, jeder Pfad in diesem Baum symbolisiert einen Weg von dem Startknoten zum Ziel. Durch die Summation aller Pfade, die als Ergebnis keine (0) haben, erhält man die Two-Terminal-Zuverlässigkeit $Rel_{s,d}(t)$ für den Zeitpunkt t=0.
5. Die Zeit t wird inkrementiert. Wenn das Ende der Simulation erreicht wurde, wird der letzte Schritt übersprungen.
6. Nach Erhöhung der Zeit t befinden sich die Knoten (außer s und d, die als feststehend angenommen werden) an anderen Orten. Die Schritte ab 2 werden unter den neuen Voraussetzungen wiederholt.

Für die Matrix aus 5.28 ergibt sich nach dem Durchführen der Schritte 2 bis 4 der in Abbildung 5.6 dargestellte Baum. Die Abbildung ist von Kharbash und Wang übernommen. In dem letzten Rechteck unten rechts müsste anstatt $\overline{X_{2,4}}X_{1,3}\ \overline{X_{4,5}}X_{2,4}$ eingetragen sein, weil die vorhandenen Variablen durch die Schritte verarbeitet werden müssen und eine Variable mit der Bezeichnung $X_{1,3}$ nicht mehr existiert. Die Zuverlässigkeit für dieses Beispielnetzwerk ergibt sich als:

$$\begin{aligned} Rel_{1,5}[G(t_0)] = {} & X_{1,2}X_{1,3}X_{3,5} + X_{1,2}X_{1,3}\overline{X}_{3,5}X_{2,4}X_{4,5} \\ & + X_{1,2}X_{1,3}\overline{X}_{3,5}\overline{X}_{2,4}X_{3,4}X_{4,5} + X_{1,2}\overline{X}_{1,3}X_{2,4}X_{4,5} \\ & + X_{1,2}\overline{X}_{1,3}X_{2,4}\overline{X}_{4,5}X_{3,4}X_{3,5} + \overline{X}_{1,2}X_{1,3}X_{3,5} \\ & + \overline{X}_{1,2}X_{1,3}\overline{X}_{3,5}X_{3,4}X_{4,5} \end{aligned} \tag{5.31}$$

Das beschriebene Verfahren wird für die Simulation der Zuverlässigkeit von Ad-Hoc-Netzen verwendet. In dieser Simulation untersuchten Kharbash und Wang zum einen den Einfluss der Fehlerraten der Knoten auf die Netzwerkleistung und zum anderen die Wirkung der Knotenbewegungen. Für die Simulation wurde ein $600m \times 600m$ großes Areal angenommen mit jeweils 6, 11 und 27 Knoten. Die Quelle s und das Ziel d sind in diesen Knotenzahlen enthalten. Die Knoten haben alle die gleiche Reichweite von 250 m und können sich, mit Ausnahme von s und d, frei bewegen. Für die Bewegung der Knoten wurde von Kharbash und Wang ein Modell von Broch et al. [5] genutzt, das *Random way point mobility model (RWP)*, das es ermöglicht, dass sich die Knoten auf einen Startbefehl hin eine zufällige Position in dem Areal suchen und sich zu dieser mit einer bekannten Durchschnittsgeschwindigkeit bewegen. Anschließend verharren sie an

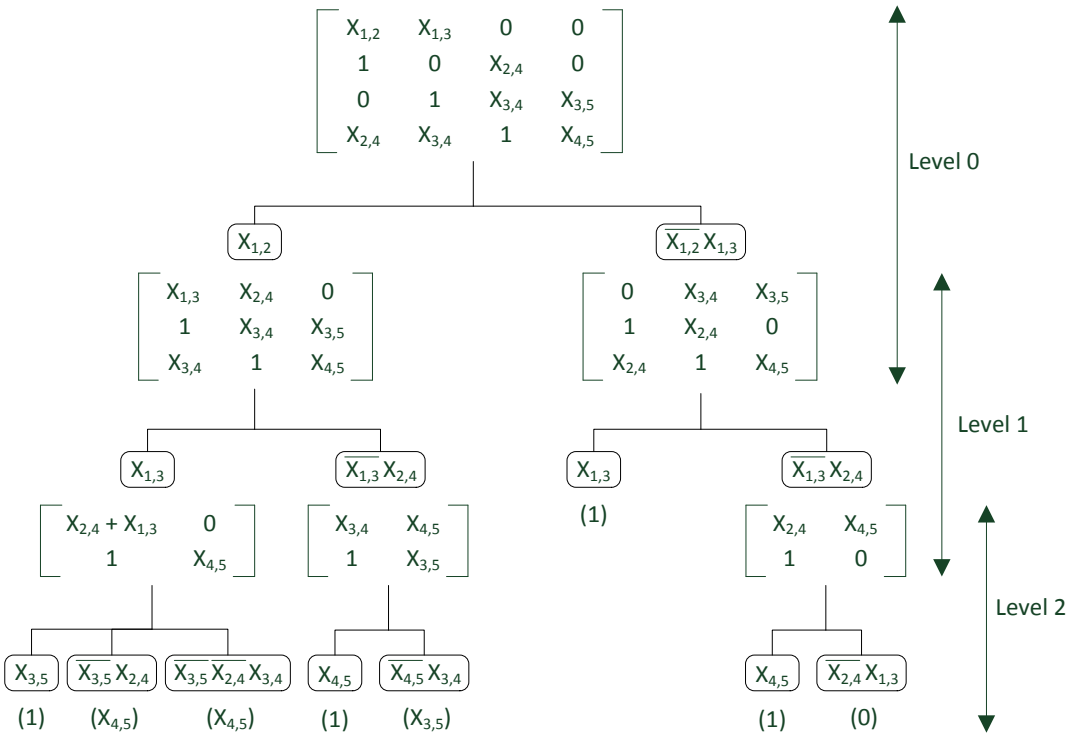


Abbildung 5.6: Berechnung der Two-Terminal-Zuverlässigkeit anhand einer Beispielmatrix (nach [15])

diesem Ort für eine vorgegebene Zeit. Wenn ein Knoten den Rand des Gebietes erreicht, entfernt er sich in dem gleichen Winkel von dem Rand weg, in dem er aufgetroffen ist. Dabei bleibt auch die Geschwindigkeit konstant. Für die Simulation variierten Kharbash und Wang die Geschwindigkeit, mit der sich die Knoten bewegen, und die Pausenzeiten bis zum nächsten Positionswechsel. [15]

Als Erstes wurde der Einfluss der Fehlerraten der Knoten auf die Zuverlässigkeit des Netzes untersucht. Jeder Ausfall eines Knotens bewirkt auch einen Ausfall der inzidenten Kanten, d. h. der Verbindungen zu anderen Knoten. Kharbash und Wang stellten fest, dass bei einer steigenden Fehlerrate das Netzwerk mit Kontrollnachrichten „überflutet“ [15] wird und die Anzahl der verloren gegangenen Pakete „dramatisch ansteigt“ [15]. Gleichzeitig nimmt der Einfluss der Fehlerraten jedoch ab, wenn mehr Knoten in dem vorgegebenen Gebiet vorhanden sind.

Neben den Auswirkungen der Fehlerraten auf die Zuverlässigkeit eines Netzwerkes wurde auch der Einfluss der Knotenbewegungen untersucht. Kharbash und Wang erkannten, dass die Zuverlässigkeit des Netzwerkes eine „bessere Stabilität“ [15] erreicht, wenn sie eine kleinere Geschwindigkeit und größere Pausenzeiten einstellten. Durch die daraus resultierende langsamere Änderung der Netztopologie sind die Wege von s nach d längere Zeit vorhanden, d. h. die Routen für die Pakete sind stabiler. Bei einer steigenden mittleren Geschwindigkeit brechen mehr Verbindungen zusammen und es gibt folglich weniger Routen von s nach d. Des Weiteren ergab die Auswertung der Versuche, dass

mit steigender Simulationszeit die Zuverlässigkeit des Netzwerkes abnimmt. Der Grund dafür ist das Bewegungsmodell von Broch et al. Das Modell konzentriert die Knoten in der Mitte des Simulationsgebietes. Demzufolge stehen weniger Wege zwischen s und d zur Verfügung. [15]
Die Simulation von Kharbash und Wang hat ergeben, dass neben der Zuverlässigkeit der Einzelkomponenten auch die Anzahl der Knoten, d. h. die Redundanz des Netzwerkes, eine bedeutende Rolle für die Zuverlässigkeit des Gesamtnetzwerkes spielt. Weiterhin ist die Geschwindigkeit der Knotenbewegungen entscheidend (siehe auch Abschnitt 6.1) und die Verteilung der Knoten.

6 Anwendung mobiler Kommunikationssysteme

6.1 Umsetzung in der Praxis

In diesem Abschnitt soll ein Überblick gegeben werden, inwieweit und wie die bisher beschriebenen Techniken in der Praxis umgesetzt wurden. Die Recherche der praktischen Umsetzung gestaltete sich teilweise recht schwierig, da Firmen und Institutionen nicht immer ihr Daten veröffentlichen. Außerdem sind diese teilweise unvollständig oder v. a. für Marketingzwecke ausgelegt. Dies bedeutet, dass sie nur einen geringen Anteil an technischen Details enthalten.
Ein Beispiel für eine solche Marketing-Veröffentlichung bietet die Motorola Mobility Germany GmbH. Das Unternehmen nutzt LTE als neue Technologie für die öffentlichen Sicherheitsbereiche, z. B. Polizei, Rettungsdienste und Feuerwehren. Für die Verwendung der LTE-Technologie ist die Verfügbarkeit und die Überlebensfähigkeit des Netzes von Bedeutung. In der Kommunikationsarchitektur finden daher Systeme Verwendung, die durch ein „Self-Organizing Network“ [37, S. 10] selbstheilend sind und sich in der Umgebung eines Netzwerkfehlers selbstständig wiederherstellen können. Des Weiteren werden Gateways genutzt, die dynamisch die Daten an die besten Netzwerke senden. [37] Auf die genaue technische Umsetzung bezüglich der Überlebensfähigkeit (beispielsweise durch Redundanz) und die Verfügbarkeit (z. B. die Dichte der Netze) wird jedoch nicht eingegangen.

Die meisten großen Netzbetreiber unterstützen heute die LTE-Technologie. Diese Mobilfunktechnologie ermöglicht es den Anbietern, die steigenden Anforderungen der Kunden bezüglich Datenraten zu erfüllen. Um die Technik an die Bedürfnisse anzupassen, sind durch die Netzbetreiber folgende Fragen zu stellen [20, S. 2]:

- Wie müssen die Bedürfnisse der Nutzer bezüglich Datenraten und Latenzzeiten berücksichtigt werden?
- Wie kann die Umsetzung kosteneffizient erfolgen und wie können die Ersparnisse an die Kunden weitergegeben werden?
- Wie müssen die Daten übertragen werden und welche Architektur sollte verwendet werden?

Nokia Siemens Networks (heute Nokia Solutions and Networks) beschreibt die einzelnen Probleme, die für die Planung eines LTE-Netzwerkes notwendig sind. Diese betreffen sowohl die notwendigen Kapazitäten (Datenraten) und Latenzzeiten als auch die benutzerdefinierten Einstellungen (QoS) und die Sicherstellung des Services. Allerdings finden sich auch hier keine Daten über die derzeitige Umsetzung in der von Nokia betriebenen

Netzstruktur.

In den Kommunikationsnetzen können mit bestimmten Wahrscheinlichkeiten Komponenten ausfallen. Wenn diese Defekte Elemente betreffen, die zentrale Aufgaben haben, kann es zum Ausfall größerer Bereiche der Netze kommen, der viele Nutzer betreffen kann. Um im Falle eines Fehlers die Funktionalität des Netzwerks zu erhalten, werden in der Praxis teilweise redundante Elemente verwendet. Eine zentral gelegene Komponente ist beispielsweise das Home Location Register (GSM/GPRS, UMTS) bzw. der Home Subscriber Server (LTE). Für diese Komponente existiert eine Sicherungsdatenbank [8, S. 138, 11]. Eine weiteres Element, dass redundant ausgeführt wird, ist der Core Unit Processor (CUP), die Steuerungseinheit in den NodeBs (siehe 5.2.1).
Der Ausfall von Komponenten hat auch Einfluss auf die Pfade, die Datenpakete in den Netzen verwenden können, um die Zielorte zu erreichen. Durch die dynamischen Routing-Algorithmen kann auf ausgefallene Elemente reagiert werden und die zur Verfügung stehenden Pfade können neu berechnet werden.

6.2 Kommunikation zwischen Fahrzeugen und der Umgebung

6.2.1 Bedeutung der Car-to-X-Kommunikation

In den letzten Jahren nahm die Bedeutung von Fahrzeugen für den Menschen stark zu. Die „Anzahl der Fahrzeuge nimmt schneller zu als die Anzahl der Straßen“ [9, S. 2]. Folglich nimmt die Zahl der auftretenden Stausituationen zu und es können auch mehr Unfälle entstehen. Mithilfe der Kommunikation zwischen den Fahrzeugen soll die Verkehrssicherheit und die Effizienz erhöht werden. Außerdem wird eine Erhöhung des Reisekomforts und eine Reduzierung der Umweltverschmutzung erhofft.
Um diese Ziele zu erreichen und die Forschung effektiv voranzutreiben, wurde das Car 2 Car Communication Consortium (C2C-CC) gegründet, eine Organisation, die aus europäischen Automobilherstellern gebildet und von Zuliefererbetrieben, Forschungseinrichtungen und anderen Partnern unterstützt wird [25].

Die Kommunikation zwischen den Fahrzeugen, die auch als Car-to-Car- (C2C-) Kommunikation bezeichnet wird, kann nach Graf in drei Szenarien die Sicherheit der Insassen erhöhen [9, S. 3]:

- Forward Collision Warning: Die Kommunikation zwischen den Fahrzeugen soll frühzeitig auf Hindernisse aufmerksam machen und somit Unfälle vermeiden.
- Pre-Cash Sensing/Warning: Bei einem nicht vermeidbarem Unfall sollen die Sicherheitsvorkehrungen für einen optimalen Schutz getroffen werden (z. B. Airbag, Sicherheitsgurtstraffer).

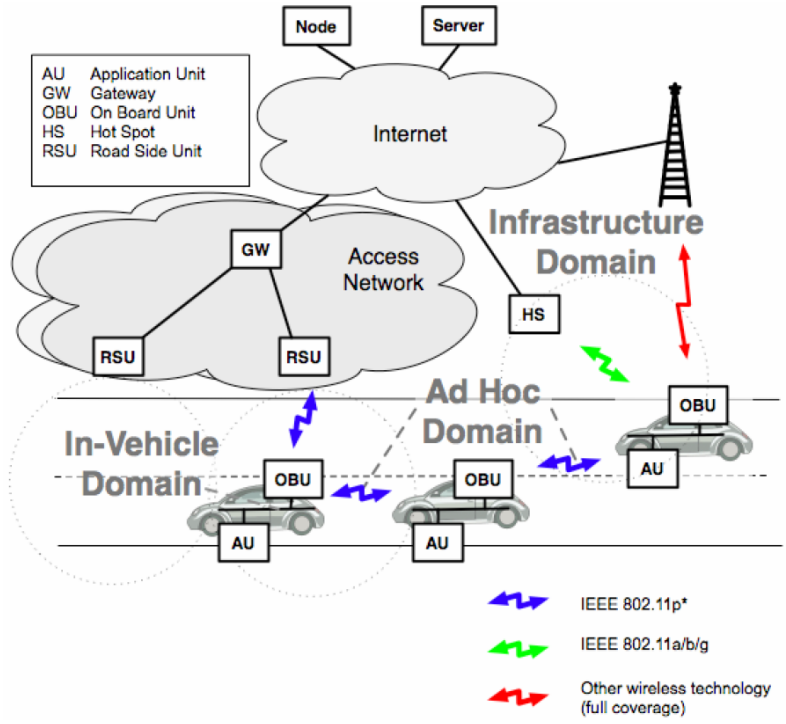


Abbildung 6.1: Architektur der C2X-Kommunikation [6, S. 27]

- Hazard Location Warning: Der Datenaustausch zwischen den Fahrzeugen soll ein rechtzeitiges Erkennen von Gefahrensituationen (z. B. Glatteis, Schlaglöcher) möglich machen.

Für die Verbesserung der Verkehrseffizienz kommunizieren die Fahrzeuge nicht nur miteinander, sondern auch mit der Infrastruktur (Car-to-Infrastructure-, C2I-Kommunikation), in der beispielsweise Daten zur Verkehrsauslastung gesammelt werden. Mit diesen Informationen können die Fahrzeuge die günstigste Route zu einem bestimmten Ziel berechnen. Außerdem bietet es Möglichkeiten, die Geschwindigkeiten für grüne Wellen zu berechnen oder das „Einfädeln“ in den laufenden Verkehr zu erleichtern. [9, S. 4].
Wenn über die Kommunikation zwischen den Fahrzeugen und mit der Infrastruktur gesprochen wird, wird dies als Car-to-X- (C2X-) Kommunikation bezeichnet.

6.2.2 Systemarchitektur

Wie in der Abbildung 6.1 dargestellt, existieren in der C2X-Architektur drei verschiedene Domänen: In-Vehicle Domain, Ad Hoc Domain und Infrastructure Domain.
Die *in-vehicle domain* besteht aus der on-board unit (OBU) und einer oder mehreren application units (AU). Die AUs sind Geräte, die die Ressourcen der OBU nutzen und einzelne oder mehrere Applikationen ausführen. Für die AUs gibt es zwei Ausführungen. Einerseits können sie fest in das Fahrzeug integriert und mit der OBU verdrahtet sein, andererseits können AUs auch mobile Geräte, beispielsweise Laptops, darstellen. In

diesem Fall kann die Verbindung zum OBU kabelgebunden oder per Funk, z. B. über Bluetooth, realisiert sein. [6, S. 25]
Die *ad hoc domain* (auch als Vehicular Ad hoc Network (VANET) bezeichnet) besteht aus den Fahrzeugen mit den OBUs und stationären Einheiten, die an den Straßenrändern positioniert sind, den road-side units (RSU). Die OBUs, die mit mindestens einer kabellosen Kommunikationsschnittstelle ausgestattet sind, kommunizieren miteinander und bilden somit ein Mobile Ad hoc Network (MANET). Für die Kommunikation zwischen zwei OBUs, die direkt oder über eine Multi-hop-Verbindung möglich ist, ist keine zentrale Verwaltung der Verbindungen notwendig. Die RSUs, die Daten in das MANET einspeisen, dienen zur Optimierung der Verkehrssicherheit. Sie können außerdem Daten senden, empfangen und weiterleiten und damit die Reichweite des Ad hoc-Netzes vergrößern. Des Weiteren kann eine RSU über eine Verbindung zum Internet verfügen. Wenn sich ein OBU im Bereich einer solchen RSU befindet, können die mit dem OBU verbundenen AUs alle Dienste im Internet nutzen. Neben der Kommunikation mit den RSUs kann ein OBU auch mit Internetservern oder hot spots (HS) Daten austauschen. [6, S. 25]
RSUs und HSs sind Komponenten der infrastructure domain. Die RSUs, die einen Zugriff auf das Internet ermöglichen, bauen typischerweise eine sichere Verbindung auf. Die HSs hingegen, die privat oder öffentlich sein können, werden normalerweise weniger überprüft. Wenn weder HSs noch RSUs mit Internetzugriff erreichbar sind, kann auch eine Kommunikation über die Funknetze GSM/GPRS, UMTS oder LTE erfolgen, sofern der OBU diese Netze unterstützt. [6, S. 26]

6.3 Offene Fragestellungen

Die Masterarbeit befasst sich mit den technischen Grundlagen in der Mobilkommunikation und ihrer Umsetzung in der Praxis. Des Weiteren wurde recherchiert, welche Maßnahmen durch Netzbetreiber ergriffen werden, um die mobilen Kommunikationsnetze möglichst zuverlässig zu gestalten. Damit die Netzbetreiber geeignete Maße haben, damit sie die Netze effizient planen können, wurden mathematische Modelle erstellt. Diese Modelle betrachten verschiedene Eigenschaften der mobilen Kommunikationsnetze.
Die wichtigste Grundlage für die Berechnung der Zuverlässigkeit ist die Topologie des Netzes. Diese Eigenschaft gibt beispielsweise Auskunft über zentrale Knoten und redundante Wege. Des Weiteren ist die Struktur des Netzes notwendig, um den Zusammenhang des Kommunikationsnetzes berechnen zu können. Die Zusammenhangswahrscheinlichkeit ist ein Maß für die Zuverlässigkeit eines Netzwerkes.
Mit der Kenngröße Zusammenhangswahrscheinlichkeit kann zudem die Erreichbarkeit in mobilen Kommunikationsnetzen betrachtet werden. Erreichbarkeit bedeutet, dass jeder Knoten durch jeden anderen Knoten erreichbar ist. Diese Eigenschaft ist wichtig für Routing-Algorithmen, die die Wege zwischen Knoten berechnen.
Ein weiterer wichtiger Aspekt für die Einschätzung der Zuverlässigkeit in einem Netzwerk ist die Schwachstellenanalyse. Um die Schwachstellen in einem Kommunikationsnetz zu ermitteln, müssen beispielsweise die zentralen Knoten in einem Netzwerk betrachtet

werden (Zentralitätsmaße). Wenn ein zentraler Knoten, z. B. der Home Subscriber Server (siehe Abschnitt 4.4.4), ausfällt, kann dies zum Ausfall eines Großteil des Netzes oder gar des ganzen Netzwerkes führen, da die Positionen der Endgeräte nicht mehr bekannt sind und diese daher nicht mehr oder nur noch schwer erreichbar sind.

Für die „alten“ Kommunikationsnetze GSM/GPRS und UMTS gibt es mehrere veröffentlichte Forschungsarbeiten (siehe z. B. Zuverlässigkeit in UMTS-Netzen von Dharmaraja et al., Kapitel 5.2.1). Weniger untersucht ist hingegen die Zuverlässigkeit in LTE-Netzen. Da diese in der C2I-Kommunikation eine bedeutende Rolle spielen und spielen werden, ist die Untersuchung dieser mobilen Kommunikationsnetze von zentraler Bedeutung in der Elektromobilität.
Um die Modelle mit den beschriebenen Kenngrößen berechnen und untersuchen zu können, ist es notwendig, dass Daten über die Kommunikationsnetze zur Verfügung stehen. Diese müssen durch die Netzwerkbetreiber geliefert werden. Für die Ermittlung der beschriebenen Kenngrößen sind die Eigenschaften der Netze notwendig, beispielsweise Ausfallwahrscheinlichkeiten der Komponenten, Redundanzen und Auslastungen. Bei der Erstellung der Modelle für die LTE-Netze ergibt sich daraus ein neues Problem: Die LTE-Technologie wird in der Praxis erst wenige Jahre eingesetzt und ist noch nicht weit verbreitet. Es stehen demzufolge nur wenige Daten über z. B. die Ausfallwahrscheinlichkeiten der Komponenten zur Verfügung.

7 Schlussfolgerung

Im Rahmen der Masterarbeit sollten die technischen Grundlagen für die Mobilkommunikation sowie ihre Umsetzung in der Praxis herausgearbeitet werden. Des Weiteren wurde recherchiert, wie zuverlässig mobile Kommunikationsnetze sind und welche Maßnahmen durch Netzbetreiber ergriffen werden, um ihre Netze zuverlässig zu gestalten. In diesem Zusammenhang sollte zudem untersucht werden, welche Forschungsergebnisse zu Zuverlässigkeit und Überlebensfähigkeit von Funknetzen veröffentlicht wurden. Es sollte zudem der Bezug zur Elektromobilität, besonders zur Kommunikation zwischen Fahrzeugen und der Umgebung hergestellt werden.

Zu den „alten“ Mobilkommunikationsnetzen wie GSM/GPRS und UMTS wurden mehrere Forschungsergebnisse vorgestellt. Zu diesen Netzen gibt es viele Untersuchungen bezüglich ihrer Zuverlässigkeit, wobei teilweise auch Auswirkungen auf die Überlebensfähigkeit beschrieben und nachgewiesen wurden. Weniger erforscht hinsichtlich der Zuverlässigkeit, der Verfügbarkeit und der Überlebensfähigkeit sind hingegen die LTE-Netze.
Die verfügbaren Untersuchungen setzen häufig voraus, dass die Parameter der Komponenten der mobilen Netze exponentiell verteilt sind. Bei der Übertragung der Forschungsergebnisse auf die Kommunikation zwischen Fahrzeugen muss untersucht werden, ob diese Annahme auch in diesem Anwendungsbereich gilt.

Als schwierig stellte sich die Recherche nach Informationen über die Umsetzung der Zuverlässigkeit in der Praxis dar, da die Firmen ihre Daten nicht der Öffentlichkeit zur Verfügung stellen.

Zusammenfassend kann festgestellt werden, dass eine Übersicht über die technischen Grundlagen der mobilen Kommunikationsnetze geschaffen werden konnte und das es möglich war, die ersten Zuverlässigkeitsmaße zu definieren, die für die Einschätzung der Verfügbarkeit und der Überlebensfähigkeit von Kommunikationsnetzen notwendig sind. Die Forschungsergebnisse, die bezüglich der Zuverlässigkeit veröffentlicht wurden, müssen jedoch auf ihre Übertragbarkeit auf die Elektromobilität überprüft werden.

Literatur

[1] K.K. Aggarwal, J.S. Gupta und K.B. Misra.
A Simple Method for Reliability Evaluation of a Communication System.
IEEE Transactions on Communications, Vol. 23, Nr. 5. 1975.

[2] Mohammed M. Alwakeel. *The Effect of Channel Holding Time Distribution on Handoff Failure Probability and Call Termination Probability*.
Proceedings of the World Congress on Engineering and Computer Science 2011 Vol II (WCECS 2011), San Francisco, USA. 2011.

[3] Tony Ballardie, Paul Francis und Jon Crowcraft. *Core based trees (CBT)*.
SIGCOMM '93 Conference proceedings on Communications architectures, protocols and applications, ACM. New York, 1993.

[4] Frank Beichelt und Peter Franken. *Zuverlässigkeit und Instandhaltung*. 1. Auflage. Berlin: VEB Verlag Technik, 1983.

[5] Josh Broch u. a. *A Performance Comparison of Multi-HopWireless Ad Hoc Network Routing Protocols*.
Proceedings of the 4th annual ACM/IEEE international conference on Mobile computing and networking (MobiCom 98), New York, NY, USA. 1998.

[6] *CAR 2 CAR Communication Consortium Manifesto.*
Overview of the C2C-CC System.
Car 2 Car Communication Consortium, 28. Aug. 2007.

[7] Stephen E. Deering und David R. Cheriton.
Multicast Routing in Datagram Internetworks and Extended LANs.
ACM Trans. on Computer Systems, Vol. 8. 1990.
URL: `http://www.cs.cmu.edu/~srini/15-744/F02/readings/DC90.pdf`
(besucht am 07. 11. 2013).

[8] S. Dharmaraja, Vaneeta Jindal und Upkar Varshney.
Reliability and Survivability Analysis for UMTS Networks: An Analytical Approach.
IEEE Transactions on Network and Service Management, Vol. 5, Nr. 3. 2008.

[9] Felix Graf. »„Car 2 Car/Car 2 X“ Kommunikation. Kommunikation zwischen Fahrzeugen und deren Umgebung«.
Seminar. Universität Koblenz-Landau, 4. Sep. 2009.
URL: `http://www.uni-koblenz-landau.de/koblenz/fb4/ist/AGZoebel/Lehre/ss09/Seminar09/graf.`

[10] Eike Gutt. *LTE Long Term Evolution*. 7. Okt. 2010.

[11] *Heimatregister (HLR, Home Location Register)*. Motorola Solutions. 11. Dez. 2013. URL: http://www.motorolasolutions.com/XC-DE/Business+Product+and+Services/Cellular+Networks/HSPA/HSPA+Core/Home+Location+Register_HSxPA__Loc:XC-DE (besucht am 11. 12. 2013).

[12] Ekbert Hering, Klaus Bressler und Jürgen Gutekunst. *Elektronik für Ingenieure und Naturwissenschaftler*. Berlin Heidelberg: Springer-Verlag, 2005.

[13] *Herzlich Willkommen!* ALM GbR. die medienanstalten. 30. Juni 2012. URL: http://www.klardigital.de/ (besucht am 12. 11. 2013).

[14] ITWissen. *IMSI (international mobile subscriber identity). Internationale Mobilfunk-Teilnehmerkennung*. 28. Okt. 2013. URL: http://www.itwissen.info/definition/lexikon/international-mobile-subscriber-identity-IMSI-Internationale-Mobilfunk-Teilnehmerkennung.html (besucht am 28. 10. 2013).

[15] Shawqi Kharbash und Wenye Wang. *Computing Two-Terminal Reliability in Mobile Ad hoc Networks*. IEEE Wireless Communications and Networking Conference (WCNC 2007), Hongkong, China. 2007.

[16] Raymond Kneip. *Strahlenoptik, Teil 1*. Vorlesungsskript Lycée technique des Arts et Métiers, Département Physique. 4. Okt. 2005. URL: www.ltam.lu/physique/vorlesungen/13GE/cours_13ge/13ge_optik_1a.pdf (besucht am 16. 10. 2013).

[17] Elektronik Kompendium. *Modulation / Modulationsverfahren*. URL: http://www.elektronik-kompendium.de/sites/kom/0211195.htm (besucht am 27. 11. 2013).

[18] *LTE*. 3GPP Mobile Competence Centre. URL: http://www.3gpp.org/technologies/keywords-acronyms/98-lte (besucht am 28. 11. 2013).

[19] LTE-anbieter.info. *Wie die mobile Datenübertragung laufen lernte - die Mobilfunk Geschichte vom A-Netz bis LTE*. URL: http://www.lte-anbieter.info/lte-geschichte.php (besucht am 12. 10. 2013).

[20] *LTE-capable transport: A quality user experience demands an end-to-end approach*. Nokia Siemens Networks, 12. Dez. 2011. URL: http://nsn.com/sites/default/files/document/lte_transport_requirements.pdf (besucht am 26. 11. 2013).

[21] Fahreddin Sadikoglu Mammadov. *GSM Architecture*. Near East University. URL: http://staff.neu.edu.tr/~fahri/mobile_L5.pdf (besucht am 22. 10. 2013).

[22] Christoph Meinel und Harald Sack.
WWW. Kommunikation, Internetworking, Web-Technologien. 2004.
SpringerLink: 10.1007/978-3-642-18963-0.

[23] Martin Meyer.
Kommunikationstechnik. Konzepte der modernen Nachrichtenübertragung.
2. Auflage. Braunschweig/Wiesbaden: Vieweg und Sohn, 2002.

[24] Nobelprize.org. *The Nobel Prize in Physics 1909*. Nobel Media AB 2013.
18. Okt. 2013. URL:
http://www.nobelprize.org/nobel_prizes/physics/laureates/1909/
(besucht am 18. 10. 2013).

[25] *Organisation*. 12. Dez. 2013. URL:
http://www.car-to-car.org/index.php?id=22 (besucht am 12. 12. 2013).

[26] Donald O. Pederson und Kartikeya Mayaram.
Analog Integrated Circuits for Communication. Principes, Simulation and Design.
2. Auflage. New York, USA: Springer-Verlag US, 2008.

[27] Charles E. Perkins und Elizabeth M. Royer.
Ad-hoc On-Demand Distance Vector Routing.
Proceedings of the Second IEEE Workshop on Mobile Computer Systems and Applications (WMCSA '99), Washington DC, USA. 1999.

[28] Suresh Rai, Arun Kumar und E.V. Prasad.
Computing terminal reliability of computer network.
Reliability Engineering, Vol. 16, Nr. 2. 1986.

[29] *Releases. Freeze Dates*. 3GPP Mobile Competence Centre.
URL: http://www.3gpp.org/specifications/67-releases (besucht am 28. 11. 2013).

[30] Wolfgang Riggert. *Netzwerktechnologien*. 1. Auflage.
München Wien: Carl Hanser Verlag, 2003.

[31] Wolfgang Riggert. *Rechnernetze*. 3., aktualisierte und erweiterte Auflage.
München Wien: Carl Hanser Verlag, 2005.

[32] Martin Sauter. *Grundkurs Mobile Kommunikationssysteme. UMTS, HSPA und LTE, GSM, GPRS, Wireless LAN und Bluetooth*. 5. Auflage.
Wiesbaden: Springer Vieweg, 2013.

[33] Jochen Schiller. *Mobilkommunikation*. 2., überarbeitete Auflage.
München: Pearson Studium, 2003.

[34] Herbert Schneider-Obermann. *Basiswissen der Elektro-, Digital- und Informationstechnik. Für Informatiker, Elektrotechniker und Maschinenbauer*.
Hrsg. von Otto Mildenberger. 1. Auflage.
Wiesbaden: Vieweg+Teubner Verlag | GWV Fachverlage GmbH, 2006.
SpringerLink: 10.1007/978-3-8348-9122-8.

[35] Andrew P. Snow, Upkar Varshney und Alisha D. Malloy.
Reliability and Survivability of Wireless and Mobile Networks.
Computer, Vol. 3, Nr. 7. 2000.

[36] Andrew S. Tanenbaum und David J. Wetherall. *Computernetzwerke*.
5., aktualisierte Auflage. München: Pearson Deutschland GmbH, 2012.

[37] *The Future is now: Public Safety LTE Communications*. Motorola Solutions, 2012.
URL: http://www.motorolasolutions.com/web/Business/Solutions/Business%20Solutions/Mission%20Critical%20Communications/LTE_for_Government_and_Public_Safety/_Documents/_Static_files/The%20Future%20of%20Public%20Safety_LTE%20Mobile%20Broadband_White%20Paper.pdf (besucht am 25. 11. 2013).

[38] Peter Tittmann. *Die Analyse der Zuverlässigkeit von Kommunikationsnetzen*.
13. März 2000. URL:
https://www.mni.hs-mittweida.de/fileadmin/verzeichnisfreigaben/peter/Vorlesungen/Netzwerkanalyse/ZuvBuch.pdf (besucht am 26. 05. 2013).

[39] Peter Tittmann. *Graphentheorie. Eine anwendungsorientierte Einführung*.
1. Auflage. Leipzig: Carl Hanser Verlag, 2003.

[40] TutorialsPoint.COM. *GSM - Architecture*. 1. Dez. 2013.
URL: http://www.tutorialspoint.com/gsm/gsm_architecture.htm
(besucht am 22. 10. 2013).

[41] Jeanette Wannstrom. *LTE-Advanced*. 3GPP Mobile Competence Centre. 2013.
URL: http://www.3gpp.org/technologies/keywords-acronyms/97-lte-advanced (besucht am 28. 11. 2013).

[42] Wikipedia. *Modulation (Technik)*. 23. Sep. 2013.
URL: http://de.wikipedia.org/wiki/Modulation_%28Technik%29
(besucht am 24. 09. 2013).

[43] Wikipedia. *OSI-Modell*. 2. Okt. 2013.
URL: http://de.wikipedia.org/wiki/OSI-Modell#Schicht_2_.E2.80.93_Sicherungsschicht_.28Data_Link_Layer.29 (besucht am 08. 10. 2013).

Erklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig und nur unter Verwendung der angegebenen Literatur und Hilfsmittel angefertigt habe.

Stellen, die wörtlich oder sinngemäß aus Quellen entnommen wurden, sind als solche kenntlich gemacht.

Diese Arbeit wurde in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegt.

Mittweida, 16. Dezember 2013