

Netzgesteuertes adaptives Videostreaming über Mobilfunknetze

Von der Fakultät für Informatik, Elektrotechnik und Informationstechnik
der Universität Stuttgart zur Erlangung der Würde
eines Doktor-Ingenieurs (Dr.-Ing.) genehmigte Abhandlung

vorgelegt von
Zigmund Orlov
geb. in Riga

Hauptberichter: Prof. em. Dr.-Ing. Dr.-Ing. E. h. Dr. h. c. Paul J. Kühn
1. Mitberichter: Prof. Dr.-Ing. Eckehard Steinbach, TU München
2. Mitberichter: Prof. Dr.-Ing. Andreas Kirstädter

Tag der Einreichung: 8. April 2019
Tag der mündlichen Prüfung: 29. Juli 2019

Institut für Kommunikationsnetze und Rechnersysteme
der Universität Stuttgart

2019

Für Ludmilla und David

Abstract

Today's widespread end-to-end video streaming, to some extent, adapts itself to the prevailing conditions in the network without any assistance from the network itself. The benefit of this solution is that it can be used in any heterogeneous network constellation. The transmission data rate must be estimated by the end systems. However, the limits of this approach are the swiftness and precision of the estimate, which can be harmful either to the quality of the video stream or the efficiency of the resource utilization. To get rid of the latter drawbacks, a network-controlled adaptation can be applied, which can work simultaneously and in conjunction with the end-to-end adaptation process.

This thesis proposes and examines an embedded network-controlled video streaming adaptation in a mobile network. The concept consists of two components. One component makes an adaptation decision with knowledge of the network conditions and capabilities of the end systems and transmits an adaptation recommendation to the video streaming server. In the temporal phases of scarce transmission resources where video data rate adaptation is not possible, the second component applies a proactive video-specific discard of video data, based on its quantitative importance. This changes the video data loss structure, resulting in less video frame loss from the application's point of view.

The proposed techniques will be examined using a specifically developed experimental video streaming environment. To generalize the results, 27 compressed video sequences with different video content and structures are used, generated using three different compression methods. For a vivid assessment of the video quality, a new diagram type developed for this work is proposed.

The results of the procedure for determining the adaptation recommendations show a fast and precise adaptation of the video data rate. The proactive video-specific discarding of video data based on their quantitative importance effects a significant reduction of the video frame loss rate in the temporal phases of scarce transmission resources.

Kurzfassung

Das heute verbreitete Ende-zu-Ende-Videostreaming passt sich ohne jegliche Netzunterstützung selbst bis zu einem gewissen Grade an die im Netz herrschenden Bedingungen an. Der Vorteil dieser Lösung liegt darin, dass sie in jeder heterogenen Netzkonstellation angewandt werden kann. Die Übertragungsdatenrate muss von Endsystemen geschätzt werden. Dies hat aber seine Grenzen in der Schnelligkeit und Präzision der Schätzung, was sich nachteilig entweder in der Qualität des Videostreamings oder in der Effizienz der Ressourcenausnutzung äußert. Die letztgenannten Nachteile können durch eine netzgesteuerte Adaption beseitigt werden, die gleichzeitig und gemeinsam mit dem Ende-zu-Ende-Adaptionsverfahren wirken kann.

In der vorliegenden Arbeit wird eine eingebettete netzgesteuerte Videostreamingadaption in einem Mobilfunknetz vorgeschlagen und untersucht. Das Konzept besteht aus zwei Komponenten. Die eine Komponente trifft mit Kenntnis der Netzgegebenheiten und Fähigkeiten der Endsysteme eine Adaptionentscheidung und übermittelt einen Adaptionsvorschlag an den Videostreamingserver. In den zeitlichen Phasen knapper Übertragungsressourcen, in denen keine Videodatenratenadaption möglich ist, wendet die zweite Komponente ein proaktives videospezifisches Verwerfen von Videodaten basierend auf ihrer quantitativen Wichtigkeit an. Damit wird die Videodatenverluststruktur geändert, was aus Sicht der Anwendung zu weniger Videobildverlusten führt.

Die vorgeschlagenen Verfahren werden unter Einsatz einer eigens dazu entwickelten experimentellen Videostreaming-Untersuchungsumgebung erforscht. Um die Ergebnisse zu verallgemeinern, werden 27 komprimierte Videosequenzen verwendet, die unterschiedliche Videoinhalte und unterschiedliche Videostrukturen aufweisen und mit drei verschiedenen Vorgehensweisen bei der Komprimierung erzeugt wurden. Zur anschaulichen Bewertung der Videoqualität wird ein für diese Arbeit entwickelter Diagrammtyp vorgeschlagen.

Die Ergebnisse des Verfahrens zur Ermittlung der Adaptionsvorschläge zeigen eine schnelle und präzise Adaption der Videodatenrate. Das proaktive videospezifische Verwerfen von Videodaten basierend auf ihrer quantitativen Wichtigkeit bewirkt in den zeitlichen Phasen knapper Übertragungsressourcen eine deutliche Minderung der Videobildverluste.

Inhaltsverzeichnis

Abstract	i
Kurzfassung	iii
Inhaltsverzeichnis	v
Abbildungen	vii
Tabellen	xi
Abkürzungen und Symbole	xv
1 Einführung und Überblick	1
1.1 Phänomen “Videostreaming”	1
1.2 Überblick über diesen Beitrag	4
2 Transport von Streamingvideo über terrestrische Funknetze	5
2.1 Terrestrische Funknetze für Streamingvideo	5
2.2 Terrestrischer Funkkanal	7
2.2.1 Störende Einflüsse im terrestrischen Funkkanal	7
2.2.2 Gängige Funkkanalausprägungen mit Mehrwegeausbreitung	10
2.2.3 Zusammenfassende Diskussion	11
2.3 Datenstreaming	12
2.4 Streamingvideokommunikation als Ausprägung des Datenstreaming	12
2.5 Videostreaming im Schichtenmodell	13
2.5.1 Protokollübersicht der Sitzungs- und der Transportschicht	15
2.5.2 Verfahren der Sicherungs- und Bitübertragungsschicht	19
2.5.3 Die Dienstgüte im Schichtenmodell	21
2.5.4 Zusammenfassende Diskussion	23
2.6 Mobilfunknetz	25
2.6.1 Netzarchitektur	25
2.6.2 Kanalarten	27
2.6.3 Herausforderungen für das Videostreaming	29
3 Grundlagen des adaptiven Videostreaming und der Auswahl von Videosequenzen	31
3.1 Merkmale des Videostreaming	31
3.2 Videoerzeugung	32

3.2.1	Digitalisierung	32
3.2.2	Komprimierung	33
3.2.2.1	Redundanzreduktion	33
3.2.2.2	Irrelevanzreduktion	35
3.2.2.3	Komprimierungsstandards	36
3.2.2.4	MPEG-4-Videostruktur	37
3.3	Videotransport	38
3.3.1	Arten der Videodatenverluste	38
3.3.2	Maßnahmen zur anwendungsgerechten Videodatenauslieferung	39
3.4	Bewertung der Videoqualität	41
3.4.1	Subjektive Messverfahren	42
3.4.2	Objektive Messverfahren	42
3.4.3	Auswahl von Videosequenzen	44
4	Eingebettete netzgesteuerte Videostreamingadaption	47
4.1	Konzept	47
4.2	Netzgesteuerte Datenratenadaption	49
4.2.1	Literaturübersicht	49
4.2.2	Diskussion und Einordnung	52
4.2.3	Implementierung	53
4.3	AQM	56
4.3.1	Literaturübersicht	57
4.3.2	Diskussion und Einordnung	62
4.3.3	Implementierung	64
4.4	Gemeinsame Verwendung der beiden Maßnahmen im VSNC	66
4.4.1	Literaturübersicht	66
4.4.2	Diskussion und Einordnung	66
4.4.3	Implementierung	66
5	Experimentelle Untersuchungsumgebung	67
5.1	Einführung	67
5.2	Systemübersicht	67
5.2.1	Messaufbau	69
5.2.2	PVU-Diagramm	70
6	Untersuchungsergebnisse	73
6.1	Videosequenzen	73
6.1.1	Vorauswahl	73
6.1.2	Komprimierung	76
6.1.3	PSNR	77
6.1.4	Datenumfang in den Videobildarten	77
6.1.5	Datenraten	80
6.1.6	Zusammenfassung	80
6.2	Systemmodell	82
6.2.1	Systembeschreibung	82
6.2.2	Emulationsmodell	82
6.2.3	Videostreamingmodell	82

6.2.4	Videobewertungsmodell	83
6.2.5	Grundscenario	83
6.2.6	Paketierung	84
6.3	Verfahrensergebnisse	85
6.3.1	Netzgesteuerte Datenratenadaption	85
6.3.1.1	Szenario	85
6.3.1.2	Bestimmung des Hysterese-Faktors	85
6.3.1.3	Bewertung der Videoqualität	87
6.3.1.4	Bewertung des Adaptionverhaltens	92
6.3.1.5	Zusammenfassende Diskussion	97
6.3.2	AQM	98
6.3.2.1	Szenario	98
6.3.2.2	Wirkungsweise	98
6.3.2.3	Videobildwichtigkeit	100
6.3.2.4	Vergleich der vorgeschlagenen AQM-Verfahren	103
6.3.2.5	Zeitliche Aufteilung der Videobildverluste	107
6.3.2.6	Zusammenfassende Diskussion	108
7	Zusammenfassung und Ausblick	111
A	Anhang	113
	Literaturverzeichnis	117

Abbildungsverzeichnis

1.1	Die drei treibenden Kräfte für die Videostreaminganwendungen	4
2.1	Mögliche Signaldämpfungswerte im terrestrischen Funkkanal	9
2.2	Arten des Datenstreaming	12
2.3	Videostreaming im ISO-OSI-Referenzmodell	15
2.4	3GP-DASH-Protokollstapel nach [1]	16
2.5	ISMA-Protokollstapel nach [2]	16
2.6	3GPP-PSS-Protokollstapel nach [3]	17
2.7	3GPP-MBMS-Protokollstapel (modifiziert) nach [4] und [5]	18
2.8	DVB-IPTV-Protokollstapel nach [6]	18
2.9	Dienstgüteabschätzung einer Videostreaminganwendung	23
2.10	Videostreamingarten	24
2.11	UTRAN-Architektur (UMTS/HSPA)	26
2.12	EUTRAN-Architektur (LTE)	26
2.13	Kanalarten im RAN	28
3.1	Klassifizierung der Redundanzarten	34
3.2	Redundanzreduktion	34
3.3	Einfluss der zeitlichen Interpixelreduktion	35
3.4	MPEG-4-Videostruktur	37
3.5	Videodatenverlustarten	38
3.6	Objektive Metriken	43
3.7	SI-TI-Diagramm aus [7]	45
4.1	VSNC in einem RAN	48
4.2	Veranschaulichung der Einsatzszenarien der beiden VSNC-Komponenten	48
4.3	Implementierung von VSNC in UTRAN	49
4.4	Netzgesteuerte Datenratenadaption in UTRAN	54
4.5	Videospezifischer AQM-Mechanismus in UTRAN	64
5.1	Experimentelle Untersuchungsumgebung	68
5.2	Messaufbau	70
5.3	Konnektivität im Messaufbau	70
5.4	Neuartiges PVU-Diagramm	71
6.1	Vorauswahl der Videosequenzen (vgl. Tabelle 3.3)	74
6.2	vtc1nw(f)	74
6.3	intros(o)	74

6.4	3inrow(d)	74
6.5	smity1(m)	74
6.6	washdc(c)	74
6.7	football(i)	74
6.8	Der vorausgewählte Videosequenzraum (zur weiteren Eingrenzung)	77
6.9	PSNR, <i>IDK</i>	78
6.10	PSNR, <i>IDKf</i>	78
6.11	PNSR, <i>2DK</i>	78
6.12	Datenanteile, <i>IDK 0B</i>	79
6.13	Datenanteile, <i>IDK 2B</i>	79
6.14	Datenanteile, <i>IDK 11B</i>	79
6.15	Datenanteile, <i>IDKf 0B</i>	79
6.16	<i>IDKf 2B</i>	79
6.17	Datenanteile, <i>IDKf 11B</i>	79
6.18	Datenanteile, <i>2DK 0B</i>	79
6.19	Datenanteile, <i>2DK 2B</i>	79
6.20	Datenanteile, <i>2DK 11B</i>	79
6.21	Der ausgewählte Videosequenzraum für die Verfahrensuntersuchungen	80
6.22	Datenraten, Ausschnitt, <i>IDK</i>	81
6.23	Datenraten, Ausschnitt, <i>IDKf</i>	81
6.24	Datenraten, Ausschnitt, <i>2DK</i>	81
6.25	Kanalverlaufshistogramm	83
6.26	Kanalverlauf	83
6.27	Protokolloverhead	85
6.28	Videobildverluste	85
6.29	Anteil der hohen Qualität	86
6.30	Verhältnis DR_{ef-ef}	86
6.31	Verhältnis DR_{ef-all}	86
6.32	Umschaltvorgänge	86
6.33	Adaptionsvorschläge	86
6.34	<i>football 1DK 2B</i> , Abspielpuffer	88
6.35	<i>football 1DK 2B</i> , empfangene Videobilder	88
6.36	<i>football 1DK 2B</i> , Datenmenge der fehlerfrei empfangenen Videobilder	88
6.37	<i>football 1DK 2B</i> , PSNR-Verlauf	88
6.38	<i>intros 1DK 2B</i> , Abspielpuffer	89
6.39	<i>intros 1DK 2B</i> , empfangene Videobilder	89
6.40	<i>intros 1DK 2B</i> , Datenmenge der fehlerfrei empfangenen Videobilder	89
6.41	<i>intros 1DK 2B</i> , PSNR-Verlauf	89
6.42	<i>vtc1nw 1DK 2B</i> , Abspielpuffer	90
6.43	<i>vtc1nw 1DK 2B</i> , empfangene Videobilder	90
6.44	<i>vtc1nw 1DK 2B</i> , Datenmenge der fehlerfrei empfangenen Videobilder	90
6.45	<i>vtc1nw 1DK 2B</i> , PSNR-Verlauf	90
6.46	<i>football 1DK 2B</i> , PVU	91
6.47	<i>intros 1DK 2B</i> , PVU	91
6.48	<i>vtc1nw 1DK 2B</i> , PVU	91
6.49	<i>football 1DK 0B</i> , PVU	91
6.50	<i>intros 1DK 0B</i> , PVU	91

6.51	<i>vtcInw</i> <i>IDK 0B</i> , PVU	91
6.52	Wirkungsweise für <i>football</i> <i>IDK 11B Pb</i>	99
6.53	Vergleich für verschiedene Abspielpuffergrößen für <i>intros</i> <i>IDK 2B Pb</i>	99
6.54	<i>IDK 0B</i>	101
6.55	<i>IDK 0B P_{6,s}</i>	101
6.56	<i>IDK 2B</i>	101
6.57	<i>IDK 2B P_{6,s}</i>	101
6.58	<i>IDKf 2B</i>	101
6.59	<i>IDKf 2B P_{6,s}</i>	101
6.60	<i>2DK 2B</i>	101
6.61	<i>2DK 2B P_{6,s}</i>	101
6.62	Verfahrensvergleich für <i>intros</i> <i>IDK 2B</i>	103
6.63	<i>IDK 0B Imp_{min} = 4</i>	105
6.64	<i>IDK 2B Imp_{min} = 4</i>	105
6.65	<i>IDK 11B Imp_{min} = 4</i>	105
6.66	<i>IDKf 0B Imp_{min} = 4</i>	105
6.67	<i>IDKf 2B Imp_{min} = 4</i>	105
6.68	<i>IDKf 11B Imp_{min} = 4</i>	105
6.69	<i>IDK 2B Imp_{min} = 0</i>	107
6.70	<i>IDKf 2B Imp_{min} = 0</i>	107
6.71	<i>IDKf 11B Imp_{min} = 4</i>	107
6.72	Histogramm der Unterbrechungsdauern für <i>football</i> <i>IDK 2B BAb</i>	108
A.1	Datenraten, <i>IDK</i>	115
A.2	Datenraten, <i>IDKf</i>	115
A.3	Datenraten, <i>2DK</i>	115
A.4	Örtliche Auflösung	116

Tabellenverzeichnis

3.1	Five-Grade-Scale-Notenauflistung	42
3.2	Abbildung von PSNR auf MOS [8], [9].	44
3.3	Zuordnung der Videosequenzen zu Videokategorien (aus [7])	45
6.1	Beschreibung des vorausgewählten unkomprimierten Videosequenzsatzes.	75
6.2	Unterschiedliche Videostrukturen	76
6.3	Statistikkenwerte der Transportbitraten (kbit/s)	84
6.4	Statistikkenwerte (Quantile) der Transportbitraten (kbit/s)	84
6.5	Anzahl der Ereignisse im Adaptionkontext	93
6.6	Phasen der hohen Qualität (Statistik der Dauern in Sekunden)	95
6.7	Adaptionseffizienz	96
6.8	Videobildwichtigkeiten	100
A.1	Verwendete Komprimierungsparameter für <i>IDKf</i>	113
A.2	Protokolloverhead, 308 kbit/s	114

Abkürzungen und Symbole

Abkürzungen

3G	Dritte Generation (Mobilfunk)
3GP-DASH	3GPP Dynamic Adaptive Streaming over HTTP
3GPP	3rd Generation Partnership Project
3GPP MBMS	3GPP Multimedia Broadcast/Multicast Service
3GPP PSS	3GPP Packet-Switched Streaming Service
4G	Vierte Generation (Mobilfunk)
AL-FEC	Application Layer Forward Error Correction
AQM	Active Queue Management
ARQ	Automatic Repeat Request
ASK	Amplitude-Shift Keying
ATIS	Alliance for Telecommunications Industry Solutions
ATIS IIF	IPTV Interoperability Forum von ATIS
ATM	Asynchronous Transfer Mode
ATSC	Advanced Television Systems Committee
AVC	Advanced Video Coding
BLER	Block Error Rate
CBR	Constant Bit Rate
CCIR	Comité Consultatif International des Radiocommunication
CDMA	Code Division Multiple Access
CIF	Common Intermediate Format (352x288)

COFDM	Coded Orthogonal Frequency Division Multiplexing
CQI	Channel Quality Indicator
DAB	Digital Audio Broadcasting
DASH	Dynamic Adaptive Streaming over HTTP
DCH	Dedicated Transport Channel
DiffServ	Differentiated Services
DMB	Digital Multimedia Broadcasting
DMB-T	Digital Multimedia Broadcasting – Terrestrial
DQPSK	Differential Quadrature Phase-Shift Keying
DSCP	Differentiated Services Code Point
DVB	Digital Video Broadcasting
DVB-H	Digital Video Broadcasting – Handhelds
DVB-IPDC	DVB IP-Datacast
DVB-T	Digital Video Broadcasting – Terrestrial
ECN	Explicit Congestion Notification
EDGE	Enhanced Data Rates for GSM Evolution
ETSI	European Telecommunications Standards Institute
FDMA	Frequency Division Multiple Access
FEC	Forward Error Correction
GPRS	General Packet Radio Service
GSM	Global System for Mobile Communications
HARQ	Hybrid Automatic Repeat Request
HbbTV	Hybrid Broadcast Broadband TV
HDS	HTTP Dynamic Streaming
HEVC	High Efficiency Video Coding
HLS	HTTP Live Streaming
HS-DSCH	High Speed Downlink Shared Channel
HSDPA	High Speed Downlink Packet Access

HSPA	High Speed Packet Access
HTTP	Hypertext Transfer Protocol
HVS	Human Visual System
IBN	Intent Based Networking
IETF	Internet Engineering Task Force
iMB	Integrated Mobile Broadcast
IMT-2000	International Mobile Telecommunications at 2000 MHz
IP	Internet Protocol
IPDC	IP-Datcast
IPTV	Internet Protocol Television
IPv4	Internet Protocol Version 4
IPv6	Internet Protocol Version 6
ISDB	Integrated Services Digital Broadcasting
ISMA	Internet Streaming Media Alliance
ISO	International Organization for Standardization
ITU	International Telecommunication Union
ITU-R	International Telecommunication Union - Radiocommunication Standardization Sector
ITU-T	International Telecommunication Union - Telecommunication Standardization Sector
LAN	Local Area Network
LOS	Line Of Sight
LTE	Long Term Evolution
MAC-d	Medium Access Control For Dedicated Transport Channel
MAC-hs	Medium Access Control For High Speed Downlink Shared Channel
MANET	Mobile Ad Hoc Network
MBMS	Multimedia Broadcast Multicast Service
MDC	Multiple Description Coding
MECN	Multilevel Explicit Congestion Notification

Media-FLO	Media Forward Link Only
MIMO	Multiple Input Multiple Output
MP4	MPEG-4 Dateiformat
MPE	Multiprotocol Encapsulation
MPE-FEC	Forward Error Correction in Multiprotocol Encapsulation
MPEG	Moving Picture Expert Group
MPEG-2	Standard von Moving Picture Expert Group
MPEG-2-TS	MPEG-2-Transportstrom
MPEG-4	Standard von Moving Picture Expert Group
MPEG-DASH	MPEG Dynamic Adaptive Streaming over HTTP
NATO	North Atlantic Treaty Organization
NLOS	Non Line Of Sight
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
OIPF	Open IPTV Forum
OMA	Open Mobile Alliance
OSI	Open Systems Interconnection
PAL	Phase Alternating Line
PC	Personal Computer
PSK	Phase-Shift Keying
PSNR	Peak Signal-To-Noise Ratio
PSS	Packet-Switched Streaming Service
QAM	Quadrature Amplitude Modulation
QoS	Quality of Service
QPSK	Quadrature Phase-Shift Keying
RAN	Radio Access Network
RLC	Radio Link Control
RNC	Radio Network Controller

RNS	Radio Network Subsystem
RTCP	RTP Control Protocol
RTMFP	Secure Real-Time Media Flow Protocol
RTMP	Real-Time Messaging Protocol
RTMPT	Real-Time Messaging Protocol Tunneled
RTP	Real-Time Transport Protocol
RTSP	Real-Time Streaming Protocol
SCTE	Society of Cable Telecommunications Engineers
SDN	Software Defined Networking
SDP	Session Description Protocol
SDTV	Standard Definition Television
SMIL	Synchronized Multimedia Integration Language
SMPTE	Society of Motion Picture and Television Engineers
SR	Segment Routing
SRv6	Segment Routing Version 6
SSIM	Structural Similarity
STANAG 4609	NATO Standardisation Agreement 4609 on Air Reconnaissance Primary Imagery Data Standard
SVC	Scalable Video Coding
T-DMB	Terrestrial Digital Multimedia Broadcasting
TCP	Transmission Control Protocol
TDMA	Time Division Multiple Access
TFRC	TCP-Friendly Rate Control
TISPAN	Telecommunications and Internet converged Services and Protocols for Advanced Networking
TISPAN IPTV	TISPAN Standard for IPTV
TV	Television
UDP	User Datagram Protocol
UE	User Equipment

UMTS	Universal Mobile Telecommunications System
US	United States (of America)
UTRAN	UMTS Terrestrial Radio Access Network
VSNC	Videostreaming-Netz-Controller
WCDMA	Wideband Code Division Multiple Access
WLAN	Wireless LAN

Symbole

A	Parameter der exponentiellen Glättung
$BuffOccup$	Pufferfüllstand
DR_{ef-all}	Verhältnis der Datenmenge von fehlerfrei empfangenen Videobildern im adaptiven Fall zur Datenmenge aller empfangenen Videobildern im adaptionslosen Fall
DR_{ef-ef}	Datenmengenverhältnis von fehlerfrei empfangenen Videobildern im adaptiven und im adaptionslosen Fall
Imp_i	Videobildwichtigkeit
Imp_{min}	Schwellwert für die minderwertigen Wichtigkeitswerte
k_{hys}	Hysterese-Faktor
e_i	Aktueller Schätzwert
e_{i-1}	Vorangegangener Schätzwert
m_i	Aktueller Messwert
$PSNR$	Peak Signal-To-Noise Ratio
r_e	Prognostizierte Übertragungsdatenrate
vr_{avail_i}	Verfügbare diskrete Videodatenrate
$vr_{current}$	Aktuell verwendete Videodatenrate
vr_{maxDec}	Maximale vom Decoder dekodierbare Videodatenrate
δ	Schwellwert für den Pufferfüllstand

1 Einführung und Überblick

1.1 Phänomen “Videostreaming”

Das in den 50er Jahren eingeführte analoge SDTV-Fernsehen (Standard Definition Television) mit seinen Verteilungsstrukturen war lange Zeit der einzige Kanal zur synchronen Verbreitung von Sequenzen bewegter Bilder (*Video*). Die Mediendigitalisierung, moderne Datenkompressionsverfahren und die Entwicklung neuer Kommunikationstechnologien - insbesondere die Erfolgsgeschichte der paketbasierten Kommunikation auf der Basis des Internet-Protokolls (IP) - ermöglichten neuartige Kommunikationskanäle, die im Gegensatz zu dedizierten Kommunikationsnetzen beliebige (digitale) Daten transportieren können. Somit war es nur eine Frage der Zeit, dass das inzwischen global gewordene, interaktive Medium Internet Video- und Audio-Inhalte mittels synchroner Kommunikationsmechanismen anbieten würde.

Die ersten bekannten Versuche einer Audio-Übertragung über das Internet stammen aus dem Jahr 1994 (Firma Real Networks) [10]. Die massenhafte Verbreitung von synchronen Video-Übertragungen setzte um die Jahrtausendwende ein. In Deutschland war es der Fernsehsender RTL2, der die Zuschauer mit einer Live-Übertragung aus dem “Big Brother”-Container im Jahre 2000 auf seine Internet-Seiten lockte [11]. Inzwischen bieten alle große TV- und Radio-Sender speziell aufbereitete Versionen ihrer Sendungen auf den entsprechenden Internet-Portalen an.

Was ist aber der Mehrwert dieser Angebote? Schließlich hat fast jeder Haushalt ein Radio oder ein Fernsehgerät. Diese Frage hat auch der bekannte amerikanische US-Talkshow-Moderator David Letterman am 27. November 1995 in seiner Late Show an Bill Gates gestellt. Hier ein Ausschnitt aus dem Stenogramm [12]:

LETTERMAN: But I can remember a couple of months ago there was like a big break-through announcement that on the Internet or on some computer deal they were going to broadcast a baseball game, you could listen to a baseball game on your computer, and I just thought to myself, does radio ring a bell? You know what I mean? (Audience applauds and cheers wildly.)

GATES: There is a difference.

LETTERMAN: There is a difference?

GATES: It's not a huge difference.

LETTERMAN: What is the difference?

GATES: You can listen to the baseball game whenever you want to.

LETTERMAN: Right. Oh, I see. So it is stored in one of your memory deals?

GATES: Exactly.

Somit wird das für jedermann zugängliche und technologisch offene Internet zu einem Medium, das die Individualisierung der synchronen Kommunikation von Audio- und Video-Inhalten ermöglicht. Dies geht mit einer Änderung des Nutzerverhaltens gegenüber dem klassischen Hörfunk- und Fernsehkonsum einher. D. h., es wird sowohl der Inhalt als auch der Zeitpunkt und ggf. der Ort des Konsums vom Nutzer bestimmt.

Während die in einem Rechnerverbund wie dem Internet eingesetzten Datentransportprotokolle für die Rechner-Rechner-Kommunikation paketbasiert (UDP¹) bzw. bytestrombasiert (TCP²) realisiert sind, werden die Daten in einem Rechner in Form von Dateien gespeichert, was i. d. R. zu einem dateibasierten Datenkonsum führt. Bei dieser (für einen Rechner herkömmlichen) Datenkonsumart kann eine Datei nur dann geöffnet werden, wenn sie komplett im Dateisystem vorliegt. Möchte man die auf einem entfernten System befindlichen Audio- oder Videoinhalte lokal und dateibasiert konsumieren, so muss das Übertragungsende abgewartet werden. Im Falle einer synchronen Audio- oder Videoübertragung wäre ein dateibasierter Konsum nicht anwendungs- bzw. mediumgerecht, was wie folgt begründet werden kann:

- *Verzögerung des Wiedergabebeginns*

Die Audio- und Videodateien können je nach der Komprimierung und der Dauer des Inhaltes sehr groß werden (z. B. dreistellige Megabytewerte für einen Spielfilm). Bei einem Abrufvideodienst (Video-On-Demand) würde der dateibasierte Ansatz zu einer nicht anwendungsgerechten Verzögerung des Wiedergabebeginns führen.

- *Speicherkapazitäten im Zielsystem*

Insbesondere mobile Endgeräte können geringe Speicherkapazitäten aufweisen, so dass eine komplette Speicherung von umfangreichen Audio- bzw. Videoinhalten nicht immer sinnvoll ist.

- *Verletzung der Live-Eigenschaft*

Findet eine Live-Übertragung statt, so kann beim dateibasierten Ansatz die Live-Eigenschaft der Wiedergabe nicht erfüllt werden.

- *Kopierschutz-Problematik*

Auch nichttechnische Gründe können einer Speicherung im Wege stehen. Möchte man eine illegale Vervielfältigung, Verbreitung oder ein weiteres Abspielen unterbinden, so kann eine lokale Speicherung u. U. unerwünscht sein.

Betrachtet man dedizierte digitale Audio- und Videoverteilungskommunikationsnetze³, so stellt man fest, dass die Datenübertragung und der Datenkonsum in diesen Systemen datenstrombasiert stattfindet. Die empfangenen Teildaten können sofort abgespielt werden. Zur Abgrenzung vom dateibasierten Datenaustausch entstand in der Internet-Community der Begriff *Streamingmedia*

¹UDP: User Datagram Protocol

²TCP: Transmission Control Protocol

³Broadcastnetze

für die datenstrombasierte Audio- bzw. Videoübertragung zwischen zwei Endgeräten in den IP-basierten Netzen. Der Pionier der Streamingmediatechnologie, Real Networks, definiert diesen Begriff folgendermaßen [13]:

Streaming Media: Large audio, video, or animation files that are broken up into smaller chunks, or data packets, and delivered by a server to users in such a way that the users can play them immediately, rather than having to wait for the entire files to finish downloading.

Streamingmedia ist der Obergriff für *Streamingvideo*: ein digitales Video, dessen Datenformat ein datenstrombasiertes Abspielen auf dem Zielsystem erlaubt (d. h., ohne dass das Ende der Übertragung eingetreten ist). Der Begriff *Videostreaming* steht für die Übertragung und das gleichzeitige Abspielen eines Streamingvideos im Empfänger.

Die Videodigitalisierung und die Entwicklung der Videokomprimierungsverfahren waren entscheidende Schritte in der Geschichte des Streamingvideos. Im Jahre 1982 erschien die Norm für die digitalen Fernsehsignale im Studiobereich: die Empfehlung 601 von CCIR⁴. Diese Norm liegt mittlerweile in der Version ITU-R BT.601-7 [14] vor und findet seit Anfang der 90er Jahre breite Anwendung [15]. BT.601 spezifiziert ein unkomprimiertes digitales Videosignal, dessen Datenrate enorm hoch ist⁵. Diese Datenrate kann mit Hilfe moderner Komprimierungsverfahren sehr stark reduziert werden, ohne dass man visuelle Qualitätseinbußen in Kauf nehmen muss. Der Komprimierungsfaktor hängt vom Videoinhalt, dem gewählten Videokodierungsstandard und einer konkreten Encoderimplementierung⁶ ab.

Während die Videoinhalte in ihren Datenvolumen dank der fortschrittlichen Komprimierungsverfahren reduziert werden, steigen die Datenraten in den Internetzugangnetzen. 100 Mbit/s bei einem privaten Anschluss sind keine Seltenheit mehr.

Die Entstehung des Phänomens Videostreaming im heutigen Internet wurde maßgebend durch die drei Hauptkomponenten getrieben: Effiziente digitale Videokomprimierungsverfahren, eine hohe Penetration breitbandiger Internetzugänge und zeitsouveräne und selektive Nutzung von Videodiensten. Das Bild 1.1 zeigt diese drei treibenden Kräfte.

Die mittleren Übertragungsdatenraten in den Mobilfunknetzen der 3. und der 4. Generation sind bereits so hoch, dass eine mobile, sprich ortsunabhängige Nutzung von datenintensiven Video-

⁴CCIR: Comité Consultatif International des Radiocommunication. CCIR ist inzwischen in ITU-R (International Telecommunication Union, Radiocommunication Standardization Sector) umbenannt worden.

⁵Beim Digitalisieren wird eine verlustbehaftete, irrelevanzmindernde (Unterabtastung der Chrominanz-Komponenten des Videosignals) Quellencodierung angewandt. Die auf diese Weise erzeugten digitalen Videos weisen in der Endkonsumversion i. d. R. eine effektive Anzahl von 12 bzw. 16 bit pro Pixel auf (je nach Grad der Unterabtastung). Verwendet man die CIF-Auflösung (Common Intermediate Format, 352x288), die der Hälfte der PAL-Auflösung entspricht, so kommt man auf die folgende Videodatenrate (bei einer Videobildwiederholrate von 30 Videobilder/s laut der ursprünglichen CIF-Definition in der ITU-T-Empfehlung H.261 [16]):

$$352 \cdot 288 \cdot 12 \text{ bit} \cdot 30 \frac{\text{Videobilder}}{\text{s}} = 36,49536 \frac{\text{Mbit}}{\text{s}}$$

⁶Ein Videokomprimierungsstandard legt die Encoder-Implementierung nicht fest.

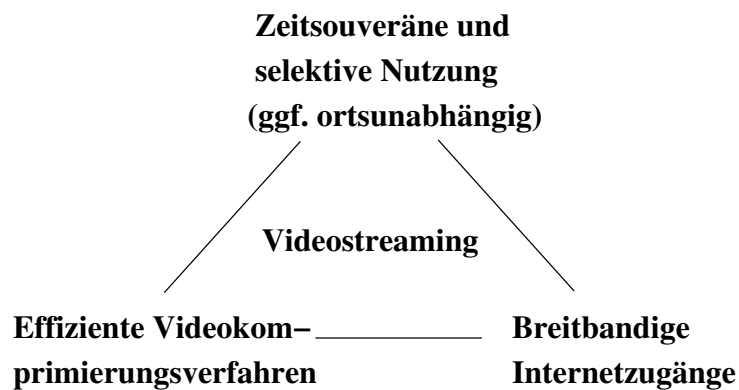


Abbildung 1.1: Die drei treibenden Kräfte für die Videostreaming-Anwendungen

streaming-Anwendungen üblich geworden ist. Aufgrund der Echtzeitnatur von Videostreaming-Anwendungen und der hohen Zeitvarianz von Mobilfunkkanälen ist das Einhalten der medien- bzw. anwendungsgerechten Dienstgüte eine zu bewältigende Herausforderung.

1.2 Überblick über diesen Beitrag

Im Kapitel 2 werden Grundlagen und Mechanismen zum Transport eines Streamingvideos über terrestrische Funknetze vorgestellt. Es wird deutlich, dass der terrestrische Funkkanal eine Videostreaming-Anwendung vor große Herausforderungen stellt. Die relevante Protokollwelt sowie die wichtigen netzseitigen Maßnahmen der Sicherungs- und der Bitübertragungsschicht werden beschrieben. Die Architektur eines Mobilfunknetzes wird in dem für diese Arbeit notwendigen Umfang dargelegt.

Das Kapitel 3 befasst sich mit den für diese Arbeit notwendigen Grundlagen des adaptiven Videostreaming. Eine Übersicht der Maßnahmen der anwendungsgerechten Videodatenlieferung ordnet die in dieser Arbeit vorgeschlagenen Verfahren ein. Nach den Grundlagen der Bewertung der Videoqualität wird auf die Grundlagen der Auswahl der Videosequenzen zur Bewertung des Videostreaming eingegangen.

Im Kapitel 4 wird die in der vorliegenden Arbeit vorgeschlagene eingebettete netzgesteuerte Videostreaming-Adaption vorgestellt, die zwei Funktionskomponenten aufweist. Zu den Verfahren der beiden Komponenten werden ausführliche Literaturübersichten durchgeführt. Nach der Einordnung der Verfahren innerhalb der jeweiligen Verfahrensklasse wird die Implementierung der vorgeschlagenen Verfahren beschrieben.

Das Kapitel 5 beschreibt die eigens für diese Arbeit entwickelte experimentelle Untersuchungsumgebung. Für die Bewertung der Videoqualität wird ein besonders anschaulicher Diagrammtyp vorgeschlagen, der im Rahmen dieser Arbeit entstand.

Im Kapitel 6 werden die zu verwendenden Videosequenzen ausgewählt und die Verfahrensergebnisse für die vorgeschlagene eingebettete Videostreaming-Adaption vorgestellt und bewertet. Die Zusammenfassung und der Ausblick finden sich im Kapitel 7.

2 Transport von Streamingvideo über terrestrische Funknetze

2.1 Terrestrische Funknetze für Streamingvideo

Drahtlose Kommunikationssysteme werden in den Umgebungen und Nutzungsszenarien eingesetzt, in denen eine kabelgebundene Kommunikation nicht möglich oder nicht zweckdienlich ist. Somit werden nicht nur in mobilen, sondern auch in stationären Szenarien drahtlose Kommunikationsformen verwendet, indem informationsmodulierte elektromagnetische Wellen in einen Raum mit Empfängern ausgestrahlt werden. Im Folgenden wird die drahtlose Kommunikation im Funkwellenbereich ¹ über einen terrestrischen Kanal (s. Abschnitt 2.2) betrachtet.

Allgemein wird bei der Videokommunikation zwischen Systemen für *Verteil-*, *Individual-* und *Gruppenrufkommunikation* unterschieden. Systeme für die *Verteilkommunikation* (Broadcast) sind in erster Linie aus dem klassischen linearen Fernsehen und Radio bekannt, d. h., die Informationsübertragung findet ohne Interaktion des Benutzers statt. Ihre Aufgabe ist es, möglichst viele Teilnehmer gleichzeitig mit denselben Informationen zu versorgen. Die Punkt-zu-Alle-Übertragung (Point-to-Anypoint) dieser Kommunikationsart prägt die grundsätzliche Systemarchitektur und die konkrete technische Ausgestaltung. Die Kommunikationsbeziehung ist unidirektional, d. h., es gibt keinen Rückkanal zum Sender und zur Netzinfrastruktur (weder auf der Anwendungsebene noch auf den tieferen Ebenen). Die Sendeeinrichtungen senden ständig mit der gleichen, hohen Leistung. Die Modulations- und Kanalkodierungsverfahren werden bei der Verbreitung nicht an die aktuellen Wellenausbreitungsbedingungen angepasst.

Bei der Verbreitung der digitalen Videoinhalte spielen Systeme für die Verteilkommunikation in den Szenarien mit drahtlosen stationären bzw. portablen Endgeräten eine wichtige Rolle. Für den mobilen Empfang auf einem Smartphone konnten sie sich – zumindest in Europa – bis jetzt nicht durchsetzen. Es existieren mehrere digitale Fernsehstandards, die in den jeweiligen Weltregionen etabliert sind und für den terrestrischen stationären bzw. portablen Empfang erweitert wurden: DVB², ATSC³, ISDB⁴. Eine kompakte Parameterübersicht der digitalen Fernsehstandards findet sich in [17]. Für den mobilen Fernsehempfang (auf einem Smartphone)

¹Funkwellenbereich ist nicht scharf definiert. Die obere Grenze wird bei ca. 3000 GHz angenommen. Darüber befindet sich der optische Bereich.

²DVB: Digital Video Broadcasting. Dieser Standard wird in Europa und in vielen Ländern weltweit eingesetzt.

³ATSC: Advanced Television Systems Committee. Dieser Standard wird in den USA, Kanada, Südkorea und Mexiko eingesetzt.

⁴ISDB: Integrated Services Digital Broadcasting. Dieser Standard wird in Japan verwendet.

wurden folgende Systeme entwickelt: DVB-H⁵ (eine DVB-Erweiterung), DMB⁶ (eine Erweiterung des digitalen Rundfunkübertragungssystems DAB⁷), Media-FLO⁸ (Entwicklung der Firma Qualcomm) und iMB⁹ (eine Mobilfunkerweiterung). Sie sollten Grundlage für ein sogenanntes *Mobile-TV* bilden, was bis jetzt scheiterte.

Um den Einfluss der sich ändernden Wellenausbreitungssituation (s. Abschnitt 2.2) beim portablen bzw. mobilen Fernsehempfang auf die Videoqualität zu mildern, wurde bei Verteilkommunikationssystemen (terrestrischen Fernsehstandards) – neben der klassischen Kanalkodierung – die Möglichkeit einer empfangsseitigen Adaption der Videoqualität vorgesehen. Dies geschieht mit Hilfe der sogenannten hierarchischen Modulation. Auf diesen Adaptionsmechanismus wird im Abschnitt 2.5.2 kurz eingegangen.

Den Gegenpol zur Verteilkommunikation bildet die *Individualkommunikation* (Unicast). Bei dieser Endgerät-zu-Endgerät-Kommunikation ist der Informationsfluss bidirektional. Zur Steuerung der Kommunikationssitzung wird ebenfalls eine bidirektionale Signalisierung benötigt, wobei die Endgeräte u. a. mit Steuerungsinstanzen des Kommunikationsnetzes interagieren. Mobilfunknetze sind Beispiele für solche drahtlosen Systeme.

Mit den steigenden Übertragungsdatenraten durch neuartige Technologien der Funkschnittstelle werden Mobilfunknetze in letzter Zeit zunehmend attraktiver für den Videoempfang – seien es Live-Streams oder Abrufvideodienste.

Im Gegensatz zu den Verteilkommunikationsnetzen werden in einem modernen Mobilfunknetz Übertragungsparameter wie z. B. Sendeleistungen, Modulations- und Kodierungsverfahren in Echtzeit *adaptiert*, um unterschiedlichen Wellenausbreitungsbedingungen während einer mobilen Kommunikationssitzung gerecht zu werden. Diese Mechanismen berücksichtigen allerdings die Anforderungen der Streamingvideokommunikation nicht ausreichend. Vielmehr steht die Integrität der Daten im Vordergrund.

Die *Gruppenrufkommunikation* (Multicast) beschreibt eine Punkt-zu-Mehrpunkt-Kommunikation. Die Gruppenrufkommunikationsverfahren wurden in Systemen für die Individualkommunikation wie Mobilfunk bereits standardisiert. Beispiele sind Multimedia Broadcast Multicast Service (MBMS) in UMTS¹⁰/HSPA¹¹-Netzen und evolved Multimedia Broadcast Multicast Service (eMBMS) in LTE¹²-Netzen. Das Ziel ist, Videoinhalte ressourcenschonend an eine Gruppe von Teilnehmern zu verteilen. Die Dienste haben aber heute noch keine praktische Bedeutung.

⁵DVB-H: H steht für Handheld. Dies ist eine Erweiterung des DVB-T-Standards, der eine terrestrische DVB-Erweiterung beschreibt.

⁶DMB: Digital Multimedia Broadcasting ist eine Erweiterung von DAB zur Übertragung von Videoinhalten.

⁷DAB: Digital Audio Broadcasting. Dies ist ein Standard für digitale Rundfunksübertragungssysteme.

⁸Media-FLO: Dies ist ein von der amerikanischen Firma Qualcomm entwickeltes Übertragungsverfahren von Multimediainhalten an mobile Endgeräte. FLO steht für Forward Link Only.

⁹iMB: Integrated Mobile Broadcast. Die ist eine Broadcasttechnologie für den mobilen Empfang im Frequenzspektrum der Mobilfunksysteme der dritten und vierten Generation.

¹⁰UMTS: Universal Mobile Telecommunications System ist ein Mobilfunkstandard der dritten Generation.

¹¹HSPA: High Speed Packet Access ist eine UMTS-Erweiterung.

¹²LTE: Long Term Evolution ist ein Mobilfunkstandard der vierten Generation.

Im Rahmen dieser Arbeit wird die Streamingvideokommunikation als Individualkommunikation über ein Mobilfunknetz betrachtet.

2.2 Terrestrischer Funkkanal

Der Begriff *terrestrischer Funkkanal* beschreibt einen erdgebundenen Funkwellenübertragungsweg. Dies ist ein schwieriges Übertragungsmedium, dessen Eigenschaften stark von den zahlreichen Umgebungseinflüssen abhängen.

2.2.1 Störende Einflüsse im terrestrischen Funkkanal

Auf die Ausbreitung von elektromagnetischen Wellen wirken Einflüsse wie Topographie/Geomorphographie (Relief, Material), Vegetation, Bebauung, menschliche Aktivitäten (elektromagnetische Störer, sich bewegende Streuobjekte wie fahrende Fahrzeuge) und die momentan herrschenden Witterungsverhältnisse. Ist ein Empfänger mobil, äußert sich der Einfluss in schnellen und langsamen Signalpegelschwankungen.

Die schnellen Signalpegelschwankungen oder *der schnelle Schwund* sind i. d. R. stark frequenzabhängig. Die Hauptgründe für den schnellen Schwund sind:

- *Mehrwegeausbreitung*

Das Signal gelangt zu seinem Empfänger gleichzeitig auf mehreren Wegen, die unterschiedlich lang sind. Dieses Phänomen entsteht durch die Reflexionen und die Streuung der in einem bestimmten Raumwinkel abgestrahlten elektromagnetischen Welle an natürlichen und künstlichen Oberflächen. Die ungünstige Überlagerung der einzelnen Signalpfade am Empfänger kann zu Signaleinbrüchen von bis zu 30 – 40dB führen [18]. Dies ist ein sogenannter *Mehrwegeschwund*¹³.

- *Frequenzverschiebung*

Beim mobilen Empfang kommt es durch die relative Bewegung zum Sender zu einer Frequenzverschiebung der Trägerfrequenz am Empfangsort, die von der Teilnehmergeschwindigkeit abhängt und bis zu mehreren 100Hz betragen kann [19], was zu Pegelbrüchen führt. Dies ist als *Dopplereffekt* bekannt.

Durch die Bewegung eines Empfängers ändert sich die Umgebung (verschiedene Raumbereiche mit unterschiedlichen Ausbreitungssituationen) und der Abstand zum Sender. Dies führt zu Fluktuationen des Signalpegels, die im Vergleich zum Mehrwegeschwund langsam verlaufen. Deshalb nennt man dieses Phänomen *langsamer Schwund*. Neben der Freiraumdämpfung¹⁴ spielt hier die Funkfelddämpfung durch den materiellen Einfluss¹⁵ eine große Rolle. Die folgenden Umstände tragen zum langsamen Schwund bei:

¹³Da er durch die Überlagerung mehrerer Signalpfade zustande kommt, wird er auch als *Interferenzschwund* bezeichnet.

¹⁴Die Empfangsleistung nimmt im materielosen Raum quadratisch mit dem Abstand zum Sender ab.

¹⁵Die auf dem Ausbreitungsweg vorhandene Materie wirkt dämpfend auf das Signal.

- *Dämpfung durch Bodenzusammensetzung, Topographie, und Bebauung*

Während beim nahezu streifenden Welleneinfall auf den Boden die elektromagnetische Energie immer zu fast 100 % reflektiert wird, beeinflusst beim steileren Einfall der Wassergehalt des Bodens und somit seine Leitfähigkeit den Reflexionsfaktor [19]. Dies kann eine Rolle im hügeligen Gelände spielen. Die Bodenfeuchte ändert sich nach einem Niederschlag.

Die Abschattung durch natürliche und künstliche Objekte kann zu großen Signaleinbrüchen führen. In [20] wurden Signalpegelmessungen in einem leicht hügeligen Landgelände (mit Höhenunterschieden von 20–30 m) auf den Mobilfunk- und Fernsehfrequenzen durchgeführt. Gezeigt wurde, dass die Dämpfung im Beugeschatten der Hindernisse mit der zunehmenden Frequenz zunimmt. So war in Schattenbereichen ein bis zu ca. 50 dB niedrigerer Empfangspegel auf 1900 MHz zu beobachten [19].

- *Vegetationsdämpfung*

Durchläuft ein Signal ein Vegetationsgebilde (z.B. einen Park), so erfährt es eine zusätzliche Dämpfung, die sich auf der Basis von empirisch gefundenen Formeln berechnen lässt. Dabei wird ein Teil der Energie in viele Richtungen (Volumenstreuer) abgelenkt. Ein weiterer Teil wird durch die Vegetation absorbiert. Nach [19] ergibt sich ein Dämpfungswert von 11 bzw. 14 dB für die Mobilfunkträgerfrequenzen von 900 bzw. 1800 MHz, wenn ein Weg von ca. 40 m durch ein Waldstück zurückzulegen ist.

- *Gebäudedämpfung*

Es ist schwierig, die Empfangsstärke innerhalb beliebiger Gebäude mit einer beliebigen Inneneinrichtung abzuschätzen. Die mittleren, empirisch ermittelten Eindringverluste für die Außen- und Innenwände hängen von der Frequenz, den Wandparametern und dem Einfallswinkel der Funkwelle ab und liegen zwischen ca. 5 dB und 30 dB [19].

- *Dämpfung durch Witterung*

Witterungsverhältnisse wie Regen, Schnee oder Nebel wirken aufgrund des darin enthaltenen Wassers absorbierend auf das Signal. Darüber hinaus bewirkt ein Niederschlag eine Volumenstreuung ähnlich einem Waldeinfluss [19]. Die Dämpfungsintensität hängt stark von der Niederschlagsrate und der Signalfrequenz ab.

- *Polarisationsverluste*

Wird eine Empfangsantenne falsch (im Bezug auf die notwendige Polarisation) ausgerichtet, führt dies zu zusätzlichen sogenannten Polarisationsverlusten, die im einstelligen dB-Bereich liegen.

Es lassen sich weitere Störeinflüsse identifizieren, die den Signalpegel kurzzeitig oder dauerhaft beeinflussen:

- *Additives Rauschen*

Ein Funkkanal erfährt additives Rauschen, das sich dem Nutzsignal überlagert. Es geht dabei sowohl um weißes (z.B. kosmisches, atmosphärisches oder thermisches) als auch um farbiges (d. h. frequenzabhängiges Rauschen, das von technischen Einrichtungen/-Geräten verursacht wird) Gauß'sches Rauschen. Die Theorie der Rauschprozesse ist sehr umfangreich. Für weiter interessierte Leser sei auf [21] und [22] verwiesen.

- *Impulsförmige Störungen*

Impulsförmige Störungen dauern kurz, treten unregelmäßig auf und sind breitbandig. Sie werden ebenfalls von technischen Geräten aber auch von atmosphärischen Phänomenen wie Gewitter verursacht.

- *Frequenzselektive Störungen durch Funkkommunikationssysteme*

Es geht dabei nicht nur um den Einfluss von unterschiedlichen Funknetzen aufeinander, sondern auch um Interferenzen von benachbarten Zellen eines Mobilfunknetzes und sogar im Spezialfall von UMTS/HSPA um Interferenzen zwischen einzelnen Teilnehmern einer Zelle, was zu einer so genannten Zellatmung führt.

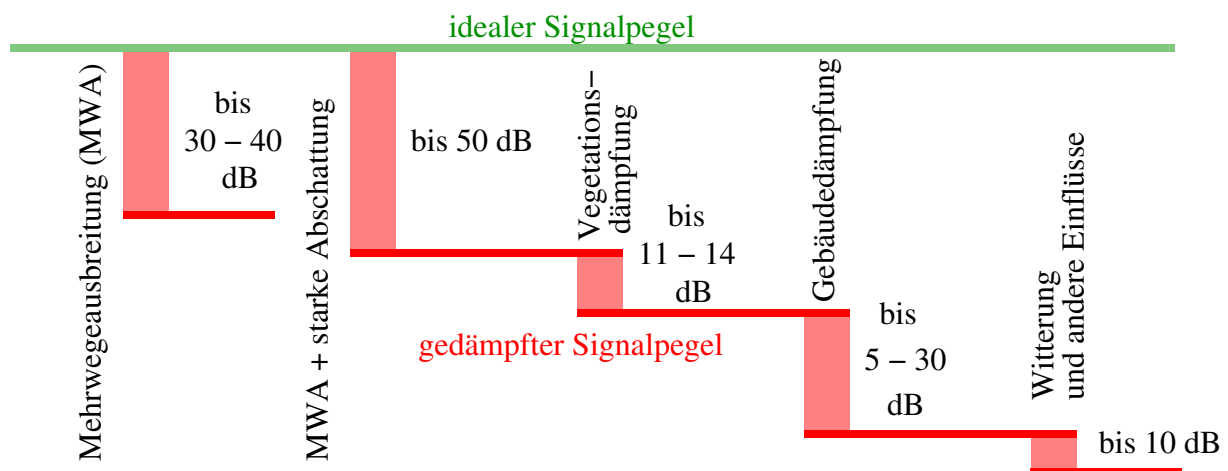


Abbildung 2.1: Mögliche Signaldämpfungswerte im terrestrischen Funkkanal

In der Abbildung 2.1 sind die möglichen Worst-Case-Dämpfungswerte eines terrestrischen Funkkanals zusammengefasst, um seine schwierige Natur anschaulich zu zeigen. Die obere grüne Linie symbolisiert einen idealen Signalpegel, das durch verschiedene Einflüsse z. T. eine sehr stark Dämpfung erfährt (symbolisch: rote Linie). Die Freiraumdämpfung wurde nicht eingezeichnet, da sie vorausberechnet und somit in der Leistungsbilanz eines Funkkommunikationsnetzes berücksichtigt werden kann. Die anderen Dämpfungsarten werden zwar in den Leistungsbilanzen der Netze berücksichtigt, können aber aufgrund der begrenzten Senderleistung und den sehr heterogenen Wellenausbreitungssituationen nicht kompensiert werden.

So wird der langsame Schwund aufgrund der Empfängerbewegung in GSM¹⁶ mit 5 dB und der Mehrwegeschwund mit 3 dB berücksichtigt [19], [23]. Bei UMTS/HSPA sind die entsprechenden Werte 7 bzw. 4 dB [24]. Die LTE-Werte sind den UMTS/HSPA-Werten ähnlich [25], [26]. Die besseren Übertragungsdatenraten lassen sich bei LTE durch den Einsatz von OFDM¹⁷, MIMO¹⁸ und den höheren Modulationsarten¹⁹ erreichen [27], wobei sich bei niedrigen Datenraten UMTS/HSPA und LTE ähnlich verhalten [25], [26]. Bei DVB-T (und dem eingestellten DVB-H) wird eine Reduzierung von 20–30 dB aufgrund der Topographie, der Bebauung und des daraus resultierenden Schwunds bei der Netzdimensionierung berücksichtigt [15].

¹⁶GSM: Global System for Mobile Communications ist ein Mobilfunksystem der zweiten Generation.

¹⁷OFDM: Orthogonal Frequency Division Multiplexing ist ein Mehrträger-Modulationsverfahren.

¹⁸MIMO: Multiple Input Multiple Output bedeutet eine gleichzeitige Übertragung mehrerer unabhängiger Datenströme über denselben Kanal [27].

¹⁹Z. B. 256-QAM

Die Gebäudedämpfung wird i. d. R. extra berücksichtigt. Für die städtischen GSM-Funkzellen werden Eindringverluste von 15 dB für die Zellenkonfigurationen angenommen [23]. Bei UMTS/HSPA wird im städtischen Bereich bei den Eindringverlusten von 25 dB nur ein Empfangsradius von 300 m versprochen [24]. Die LTE-Werte sind den UMTS/HSPA-Werten ähnlich [25], [26]. Die besseren Übertragungsdatenraten lassen sich bei LTE durch den Einsatz der bereits genannten Übertragungsverfahren erreichen [25], [26], [27]. Bei DVB-T und DVB-H, deren Trägersignalfrequenz bis zu 862 MHz reichen kann, geht man von einer ca. 10 dB bis 20 dB starken Gebäudedämpfung aus, was als Abschlag in der Leistungsbilanz benutzt wird [15].

Die Sendeleistung kann den Einfluss des langsamen und des schnellen Schwundes somit nur teilweise kompensieren.

2.2.2 Gängige Funkkanalausprägungen mit Mehrwegeausbreitung

Ein Empfänger kann entweder eine direkte Sicht oder keine direkte Sicht zum Sender haben. Weiterhin kann die Richtcharakteristik einer Antenne bei der direkten Sicht zum Sender so gestaltet werden, dass nur das Signal auf dem direkten Pfad empfangen wird. Es lassen sich deshalb drei Funkkanalausprägungen²⁰ unterscheiden:

- *Rayleigh-Kanal*

Dieser Kanaltyp beschreibt die Überlagerung vieler statistisch unabhängiger Echosignale vergleichbarer Amplitude am Empfänger. Es gibt keinen dominierenden (sprich direkten) Signalpfad. Die Empfangsamplitude des resultierenden Signals ist Rayleigh-verteilt [28].

Ein städtischer Mobilfunk- oder ein DVB-H-Kanal gehören i. d. R. zur Klasse der Rayleigh-Kanäle.

- *Rice-Kanal*

Dieser Kanaltyp beschreibt die Überlagerung vieler statistisch unabhängiger Echosignale vergleichbarer Amplitude und eines dominierenden (direkten) Signales am Empfänger. Die Empfangsamplitude ist Rice-verteilt.

Ein ländlicher Mobilfunkkanal kann von diesem Typ sein. Gleiches gilt auch für einen DVB-T-Kanal bei einer stationären Anwendung.

- *Gauß-Kanal*

Beim diesem Funkkanaltyp wird ein starker Signalpegel beobachtet, der vom Empfang auf dem direkten Signalpfad herrührt. Dieser Kanaltyp ist z. B. bei einem geostationären Satellitenkanal zu beobachten, bei dem eine Antenne mit einem hohen Gewinn von ca. 35 dBi [15] eingesetzt wird.

²⁰Es hat sich eingebürgert, die Hauptausprägungen der Funkkanäle nach den Wahrscheinlichkeitsdichtefunktionen zu benennen, die in guter Näherung die Verteilung des Betrages der komplexen Spannungsamplitude des Empfangssignals beschreiben.

Diese Kanalmodelle beschreiben den schnellen Schwund. Ihre Gültigkeit ist auf kleine Gebiete begrenzt²¹. Der langsame Schwund, der primär durch die Bewegung des mobilen Teilnehmers zustande kommt, ist für die Signalpegelschwankungen in größeren Gebieten verantwortlich.

Für weiter interessierte Leser sei in Bezug auf die Bereiche der Wellenausbreitung und der Funkkanalmodellierung auf [19] und [28] verwiesen.

Im Rahmen dieser Arbeit wird für die Verfahrensuntersuchungen ein Rayleigh-Kanal mit einem stark ausgeprägten langsamen Schwund – ein städtischer Mobilfunkkanal eines beweglichen Teilnehmers – verwendet.

2.2.3 Zusammenfassende Diskussion

Ein terrestrischer Funkkanal ist für mobile Empfänger in einem hohen Maße zeitvariant. Die Bewegung eines Teilnehmers in einem bebauten Gebiet mit Vegetationsregionen führt zu stark variierenden Signalpegeln am Empfänger. Die Kanalqualität kann im Vergleich zur drahtgebundenen Kommunikation stark schwanken. Die schwankende Kanalqualität beeinflusst die Dienstgüte der nichtelastischen Verkehre (z. B. Verkehre der Streamingvideokommunikation).

In modernen Übertragungssystemen (basierend auf WCDMA²², OFDM bzw. OFDMA²³ und COFDM²⁴) wurde die mittlere Übertragungsrate gegenüber den älteren FDMA²⁵/TDMA²⁶-Systemen (GSM, GPRS²⁷ mit EDGE²⁸) erheblich erhöht. Der Einfluss des schnellen Schwundes auf die Kanalqualität wurde insbesondere bei OFDM-Techniken (LTE) durch das Verwenden mehrerer zueinander orthogonaler Unterträger stark reduziert. Die Zeitvarianz des langsamen Schwundes bleibt aber prinzipbedingt (Teilnehmermobilität) erhalten.

Werden die Übertragungsressourcen nicht exklusiv reserviert, sondern zwischen allen Kommunikationsteilnehmern je nach Bedarf und nach weiteren Kriterien aufgeteilt, wie dies bei den modernen Systemen wie UMTS/HSPA und LTE der Fall ist, so ändert sich je nach dem Verkehrsaufkommen und dem Grad der Teilnehmerbewegung die verfügbare Übertragungsdatenrate innerhalb einer Sitzung der Streamingvideokommunikation. Deshalb sollen Maßnahmen ergriffen werden, die den Mobilfunkkanal trotz der hohen Zeitvarianz für nichtelastische Dienste nutzbar machen.

²¹Die Gültigkeit der Rice- und Rayleigh-Kanaltypen bzw. -modelle ist auf kleine Gebiete begrenzt, die ca. 10 Wellenlängen groß sind [28], was bei 900MHz ca. 3,3m entspricht. Der lokale Mittelwert des Signalpegels in diesem Gebiet kann näherungsweise als konstant betrachtet werden.

²²WCDMA: Wideband Code Division Multiple Access ist ein codebasiertes Vielfachzugriffsverfahren in den Mobilfunknetzen der dritten Generation.

²³OFDMA: Orthogonal Frequency Division Multiple Access ist ein OFDM-basiertes Vielfachzugriffsverfahren, das in den Mobilfunknetzen der vierten Generation Anwendung findet.

²⁴COFDM: Coded Orthogonal Frequency Division Multiplexing ist ein Mehrträgermodulationsverfahren (OFDM), bei dem jedes Modulationssymbol mit einem Fehlerschutz versehen wird. Es wird in den terrestrischen Verteilungskommunikationssystemen verwendet.

²⁵FDMA: Frequency Division Multiple Access ist ein frequenzbasiertes Vielfachzugriffsverfahren.

²⁶TDMA: Time Division Multiple Access ist ein zeitbasiertes Vielfachzugriffsverfahren.

²⁷GPRS: General Packet Radio Service ist eine GSM-Erweiterung.

²⁸EDGE: Enhanced Data Rates for GSM Evolution ist eine GSM-Erweiterung.

2.3 Datenstreaming

Ein ETSI²⁹-Standard [29] beschreibt *Datenstreaming* allgemein als eine stromförmige Datenübertragung. Es wird darin zwischen drei Arten des Datenstreaming unterschieden, die in der Abbildung 2.2 dargestellt sind und im Folgenden besprochen werden.

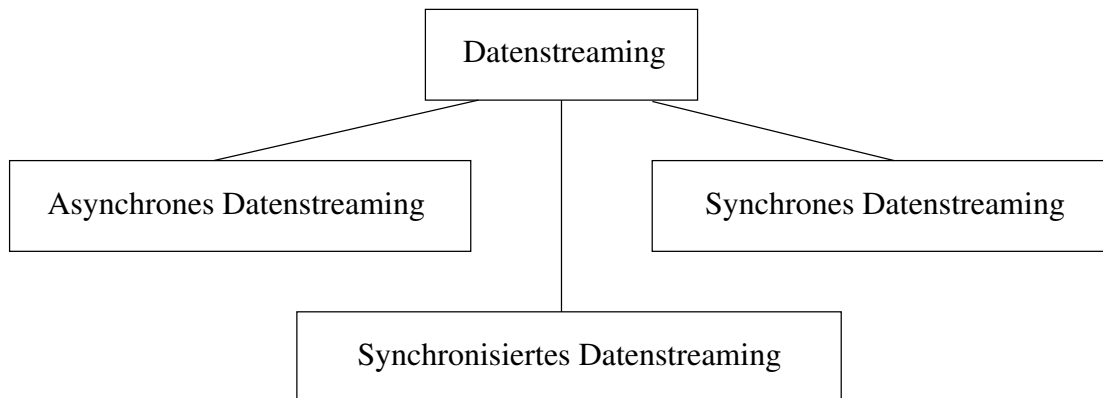


Abbildung 2.2: Arten des Datenstreaming

Bei *asynchronem Datenstreaming* werden einzelne Datenblöcke (bzw. -pakete) eines zusammenhängenden Datenobjektes ohne jegliche einzuhaltende zeitliche Randbedingungen übertragen. Somit stellt nach dieser Definition das Herunterladen von Dateien das asynchrone Datenstreaming dar.

Bei *synchronem Datenstreaming* werden die einzelnen Daten- bzw. Informationsblöcke mit der Daten- bzw. Informationsrate ausgeliefert, mit der sie am Empfänger dargestellt (konsumiert) werden³⁰. Es muss für den Empfänger möglich sein, aus dem empfangenen Datenstrom den Darstellungstakt bzw. die Darstellungszeiten zu bestimmen (*Intra-Datenstromsynchronisation*).

Bei *synchronisiertem Datenstreaming* werden die einzelnen Daten- bzw. Informationsblöcke so ausgeliefert, dass sie am Empfänger *synchron* zu und mit den anderen Datenströmen dargestellt werden können. Man denke dabei an die Synchronisation zwischen den Video- und Audiodaten. Textdaten wie Untertitel sind ein weiteres Beispiel für eine erforderliche *Inter-Datenstromsynchronisation*.

2.4 Streamingvideokommunikation als Ausprägung des Datenstreaming

Je nach Art der Streamingvideokommunikation entstanden unterschiedliche Methoden dafür. Die Methoden der Streamingvideokommunikation lassen sich den zuvor beschriebenen Datenstreamingarten (vgl. Abschnitt 2.3) zuordnen.

²⁹ETSI: European Telecommunications Standards Institute

³⁰Der genau Wortlaut[29] ist: Synchronous data streaming is defined as the streaming of data with timing requirements in the sense that the data and clock can be regenerated at the receiver into a synchronous data stream (e. g. E1, T1).

Progressiver Videodownload war eine vor wenigen Jahren verbreitete TCP-basierte Video-On-Demand-Streamingmethode, die sich als asynchrones Datenstreaming charakterisieren lässt.

Zum synchronen oder ggf. synchronisierten Datentreaming können im weiten Sinne die HTTP³¹/TCP-basierten Methoden wie MPEG-DASH³² [30], Adobe RTMPT³³ [31], Adobe HDS³⁴ [32], Apple HLS³⁵ [33], Microsoft Smooth Streaming [34], das TCP-basierte Adobe RTMP³⁶ [31], das UDP-basierte RTP³⁷-Verfahren (mit oder ohne RTSP³⁸ [35]) sowie das UDP-basierte Adobe RTFMP³⁹ - das bei IETF⁴⁰ standardisiert wurde [36] - gezählt werden.

Ein klarer Fall des synchronisierten Datenstreaming sind die IP-basierten Fernsehproduktionsverfahren, wie sie in der SMPTE⁴¹-ST-2110-Standardfamilie [37] definiert sind. Es gilt dabei, die getrennt zu übertragenden Audio- und Videosignale unterschiedlicher Quellen in einer gemeinsamen Senke untereinander synchronisiert zu verarbeiten. Dafür muss eine allen Quellen und Senken gemeinsame Zeitbasis zur Verfügung gestellt werden.

Die Streamingvideokommunikation wird im weiteren Verlauf der Arbeit Videostreaming genannt.

2.5 Videostreaming im Schichtenmodell

Das ISO⁴²-OSI⁴³-Referenzmodell als Schichtenachitektur wurde von der ISO eingeführt mit dem Ziel, Kommunikationsvorgänge zwischen offenen Systemen zu standardisieren [38] [39]. Das Videostreaming lässt sich darauf abbilden.

Des Weiteren ist festzustellen, dass es alternative, speziell für das Videostreaming erstellte Schichtenmodelle gibt. Zu erwähnen sind das MPEG-4⁴⁴-Terminalmodell und die MPEG-2⁴⁵-Systemschicht. Sie sind ebenfalls auf das ISO-OSI-Referenzmodell abbildbar. Das ISO-OSI-Referenzmodell mit seinen sieben Schichten ist im linken Drittel der Abbildung 2.3 dargestellt. Das MPEG-4-Modell ist zum Vergleich am rechten Abbildungsrand zu sehen. Ebenso ist in der Abbildung die MPEG-2-Systemschicht enthalten.

Die *Verarbeitungsschicht* ist die oberste Schicht des ISO-OSI-Referenzmodells. Deshalb wird sie oft *Anwendungsschicht* genannt. Darin findet der Konsum des Videoinhaltes mit Hilfe un-

³¹HTTP: Hypertext Transfer Protocol

³²MPEG-DASH: MPEG Dynamic Adaptive Streaming over HTTP

³³RTMPT: Real Time Messaging Protocol Tunneled

³⁴HDS: HTTP Dynamic Streaming

³⁵HLS: HTTP Live Streaming

³⁶RTMP: Real Time Messaging Protocol

³⁷RTP: Real-Time Transport Protocol

³⁸RTSP: Real Time Streaming Protocol

³⁹RTMFP: Secure Real-Time Media Flow Protocol

⁴⁰IETF: Internet Engineering Task Force

⁴¹SMPTE: Society of Motion Picture and Television Engineers

⁴²ISO: International Organization for Standardization.

⁴³OSI: Open Systems Interconnection.

⁴⁴MPEG-4: Ein MPEG-Standard zur Video- und Audiokomprimierung und -übertragung.

⁴⁵MPEG-2: Ein MPEG-Standard zur Video- und Audiokomprimierung und -übertragung.

terschiedlicher Anwendungssoftware statt. Die in dieser Arbeit entwickelte Software ermöglicht bei Bedarf das Abspielen eines Ergebnisvideos. Die Bewertung der Ergebnisse der vorliegenden Arbeit wird in den tiefer liegenden Schichten vorgenommen.

Die *Darstellungsschicht* widmet sich allgemein der Datensyntax und -semantik. Speziell bei dem digitalen Video ist hier die Komprimierung bzw. die Dekomprimierung der Videodaten angesiedelt. Aus den Videoinformationen werden hier die Videodaten und umgekehrt. In dieser Arbeit werden verschiedene MPEG-4-basierte Komprimierungsausprägungen eines Videos zwecks Bewertung der vorgeschlagenen Verfahren verwendet. Ein Teil der Bewertung der erzielten Ergebnisse findet in dieser Schicht statt.

Die *Sitzungsschicht* übernimmt das Verwalten einer z. B. multimedialen Sitzung, zu der mehrere Datenströme (-verbindungen) gehören könnten, wobei in jedem Datenstrom die dazugehörigen Unterströme (Audio, Video, Text) enthalten sind. Zu den Aufgaben dieser Schicht gehören z. B. das Multiplexen, die Synchronisation und die Dialogsteuerung aber auch – wie im Falle des Videostreaming – die Datensegmentierung (auch Framing oder Application-Level-Framing genannt). Die in dieser Arbeit entwickelte Software realisiert die für die Untersuchungen notwendigen Aufgaben dieser Schicht - hauptsächlich die Steuerung der Adaption und die Datensegmentierung des Videostroms. Ein Teil der Bewertung der erzielten Ergebnisse findet in dieser Schicht statt.

Im MPEG-4-Terminalmodell sind die Aufgaben der Sitzungsschicht zwischen der für dieses Modell spezifischen Synchronisations- und der Auslieferungsschicht aufgeteilt. Näheres ist in [40] und [41] nachzulesen.

Die *Transport-* und die *Vermittlungsschicht* des ISO-OSI-Referenzmodells sorgen für eine Ende-zu-Ende-Übertragung der zu transportierenden Daten. Ein Teil der Bewertung der erzielten Ergebnisse findet in diesen Schichten statt.

Das MPEG-2-Schichtenmodell weist eine Systemschicht auf, die die Sitzungs-, die Transport- und die Vermittlungsschicht des ISO-OSI-Referenzmodells vereinigt. Den Kern der Systemschicht bildet der MPEG-2-Transportstrom, der entgegen seinem Namen, beliebige Daten transportieren kann, also nicht nur die MPEG-2-komprimierten, sondern auch die MPEG-4-komprimierten Inhalte oder z. B. IP-Pakete. In der Abbildung 2.3 ist die MPEG-2-Systemschicht im mittleren Drittel als eine Realisierung für einen beliebigen Datentransport zu sehen. Näheres ist in [42] nachzulesen. Die MPEG-2-Systemschicht wird z.B. bei DVB verwendet.

Die Funktionen der Transport- und der Vermittlungsschicht des ISO-OSI-Referenzmodells sind von der Auslieferungsschicht in MPEG-4 mit eingeschlossen, die je nach Transportnetz unterschiedliche Auslieferungsmechanismen verwendet (z. B. UDP, MPEG-2-TS⁴⁶).

Die *Sicherungs-* und die *Bitübertragungsschicht* sorgen für den zuverlässigen Transport von Daten über verschiedene Übertragungsmedien. In diesen Schichten existieren allgemeine und videostreamingspezifische Mechanismen für die terrestrischen Funkkanäle, deren Überblick im Abschnitt 2.5.2 nachzulesen ist.

⁴⁶MPEG-2-TS: MPEG-2 Transport Stream

2.5.1 Protokollübersicht der Sitzungs- und der Transportschicht

Es gibt viele Normen und herstellerspezifische Implementierungen für das Videostreaming. Die Abbildung 2.3 fasst die relevanten Entwicklungen in ihrem mittleren Drittel zusammen.

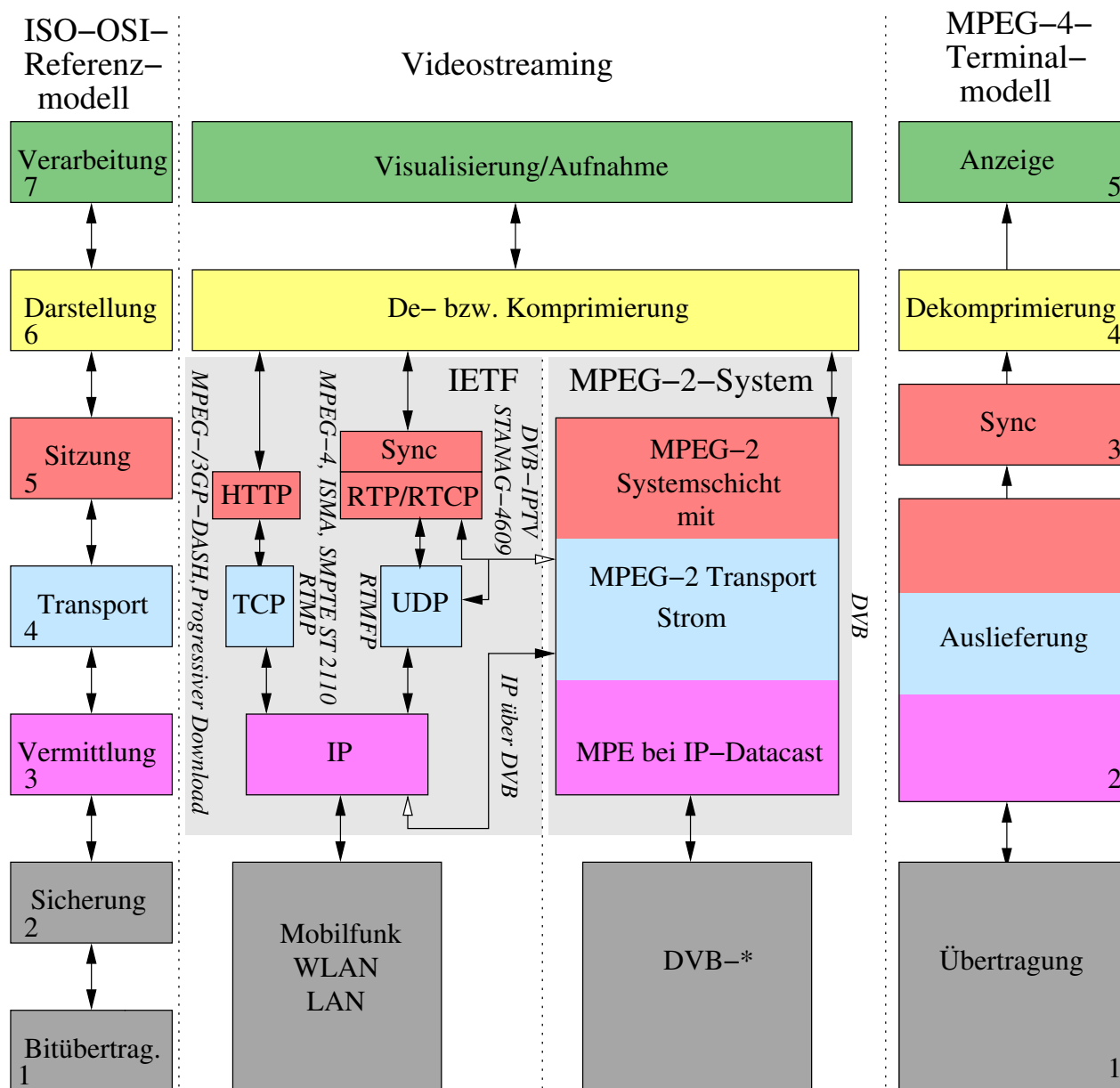


Abbildung 2.3: Videostreaming im ISO-OSI-Referenzmodell

Die herstellerspezifischen HTTP-basierten adaptiven Streamingverfahren (s. Abschnitt 2.4) mündeten in den internationalen Standards *MPEG-DASH* [43], [44], [30] und *3GP-DASH* [1]. Für die Verbreitung dieser Technologie wurde 2012 das DASH-Industry-Forum [45] gegründet. Die Verfahren basieren auf TCP. Trotzdem wird eine anwendungsgerechte Adaption an die verfügbare Übertragungsrate gewährleistet. Ein relevanter Hersteller der Videostreamingserver hat festgestellt, dass die HTTP-basierten Verfahren einen Versatz von 20 bis 30s zu einem Live-Stream aufweisen [46]. Der Protokollstapel dieser Verfahren ist in der Abbildung 2.4 zu sehen.

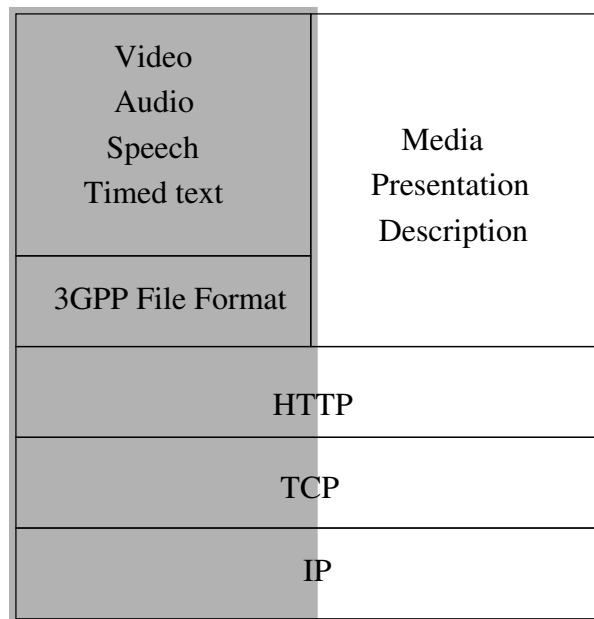


Abbildung 2.4: 3GP-DASH-Protokollstapel nach [1]

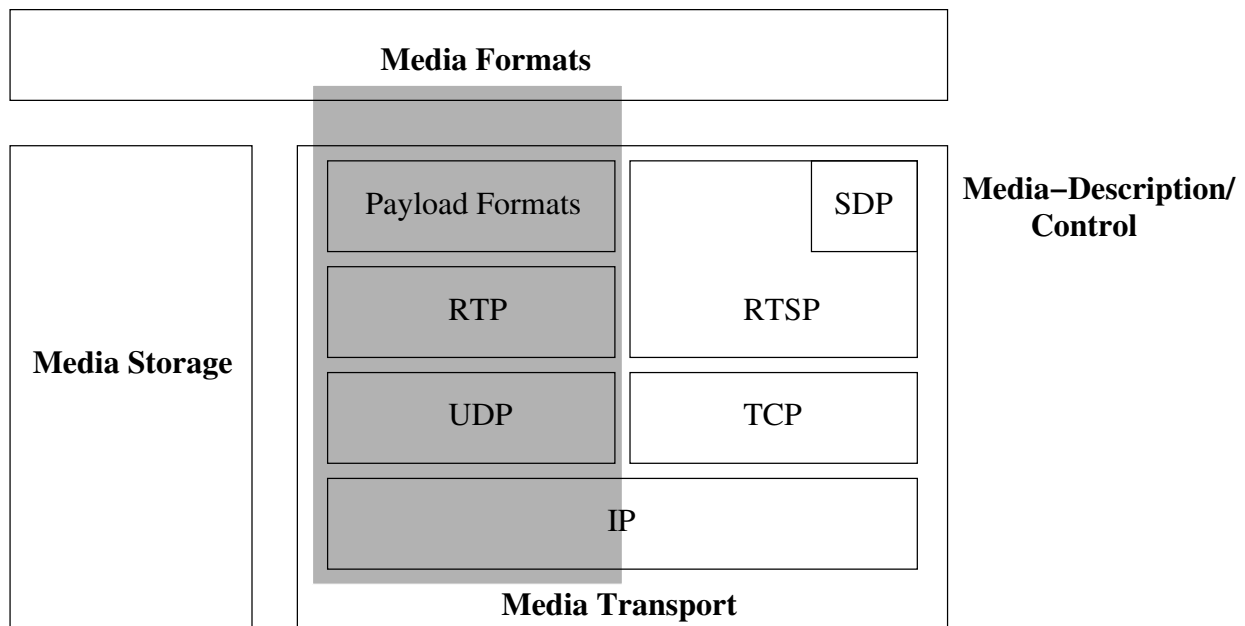


Abbildung 2.5: ISMA-Protokollstapel nach [2]

Die *Internet Streaming Media Alliance (ISMA)* war eine nichtprofitorientierte Industrie-Allianz von namhaften Computer- und Netzwerktechnikherstellern, deren Ziel es war, das Videostreaming auf der Basis der IETF⁴⁷-Protokolle zu normen. Dies ist ihr tatsächlich gelungen. Die ISMA-Hauptdokumente sind die Implementierungsrichtlinien des Videostreaming in der Version 1.0 [47] und in der Version 2.0 [2], wobei die Version 2.0 die Version 1.0 ergänzt und nicht ersetzt. Die Dokumente sind abrufbar bei MPEG-Industry-Forum [48]. Danach sollten Videoinhalte mit RTP/UDP übertragen, mit RTSP “fernbedient” und mit SDP⁴⁸ beschrieben

⁴⁷IETF: Internet Engineering Task Force.

⁴⁸SDP: Session Description Protocol.

Video Audio Speech Timed text Scene description	Capability exchange Scene description Presentation description Still images Bitmap graphics Vector graphics Text Timed text Synthetic audio	Capability exchange Presentation description
Payload formats	HTTP	RTSP
RTP		
UDP	TCP	UDP
	IP	

Abbildung 2.6: 3GPP-PSS-Protokollstapel nach [3]

werden (siehe den Protokollstapel in der Abbildung 2.5). Der ISMA-Protokollstapel bildet eine Grundlage vor allem für die Live-Stream-, die Echtzeit-, die Überwachungs-, die Verteil- und die Gruppenrufkommunikation (die zwei letzteren Kommunikationsformen sind besser als Broadcast und Multicast bekannt).

*3GPP*⁴⁹ *Packet-Switched Streaming Service* (3GPP PSS) ist ein Standard für das IP-Unicast-basierte Videostreaming über Mobilfunknetze [3], [49], [50] – sowohl über UMTS/HSPA als auch über LTE. Der Standard definiert drei Methoden des Videostreaming. Zwei davon sind die bereits erwähnten Verfahren progressiver Videodownload und 3GP-DASH. Eine weitere Methode ist UDP-basiert. Die Abbildung 2.6 zeigt ihren Protokollstapel. Das Streaming erfolgt über RTP. Um möglichen IP-Paketverlusten im Funkkanal, die man mit Hilfe von RTCP feststellen kann, entgegenzuwirken, werden RTP-Paketwiederholungen empfohlen. Um die adaptiven Anwendungen zu ermöglichen, sollen über RTCP fortlaufend Berichte über die Verbindungsqualität versendet werden.

3GPP (evolved) Multimedia Broadcast/Multicast Service (3GPP (e)MBMS) ist ebenfalls ein 3GPP-Standard, der einen Dienst der Verteil- bzw. der Gruppenrufkommunikation beschreibt [4], [5], [51]. Die multimedialen Daten werden gleichzeitig und dabei frequenzeffizient an alle interessierten Teilnehmer verteilt. Der Protokollstapel ist in der Abbildung 2.7 zu sehen. Er gilt sowohl bei MBMS (UMTS/HSPA) als auch bei eMBMS (LTE). Das Videostreaming erfolgt

⁴⁹3GPP: 3rd Generation Partnership Project ist ein Zusammenschluss mehrerer Standardisierungsorganisationen im Mobilfunkbereich.

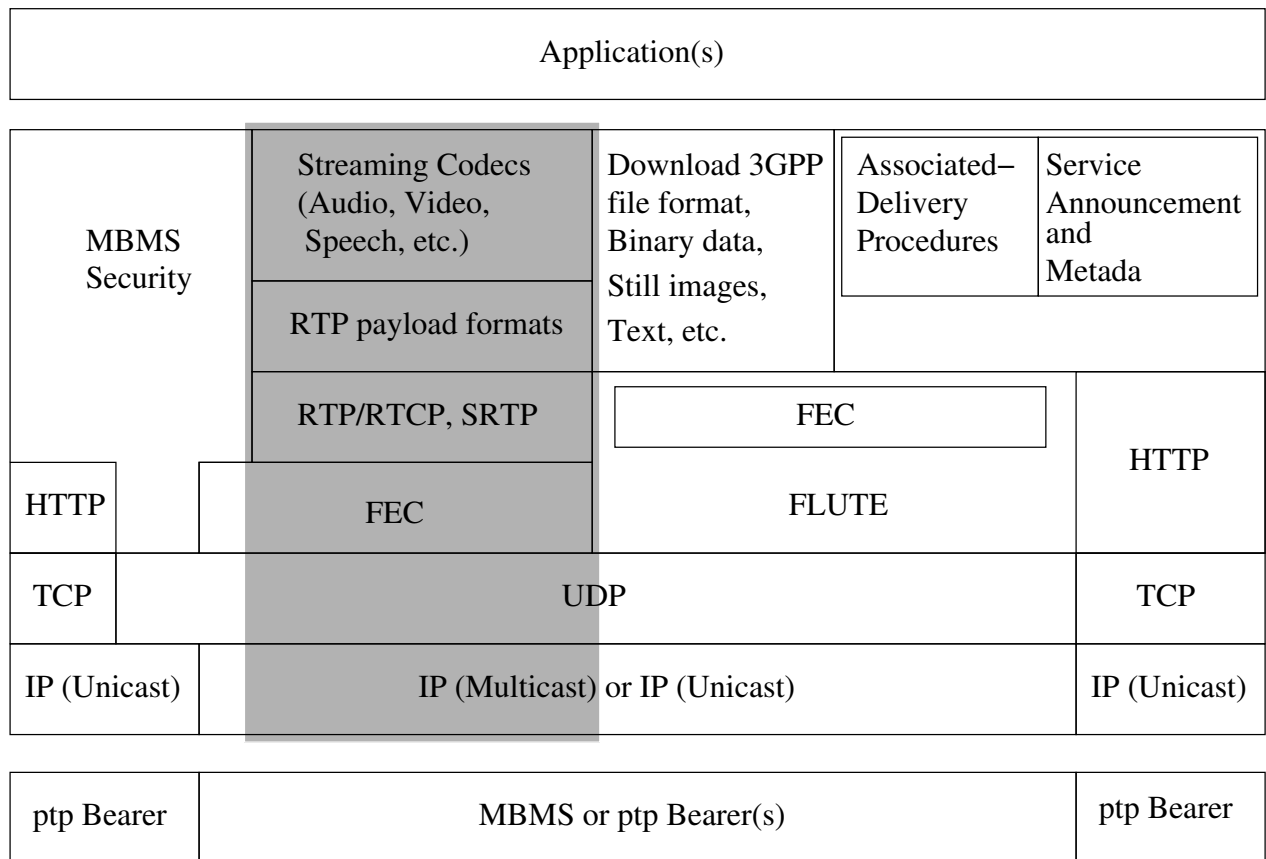


Abbildung 2.7: 3GPP-MBMS-Protokollstapel (modifiziert) nach [4] und [5]

über RTP/UDP [4], [5]. Da der Rückkanal auf der Anwendungsebene für die Adaptivität an den Funkkanal prinzipbedingt nicht gegeben ist, werden RTP-Pakete durch FEC-Verfahren mit einer zusätzlichen Redundanz zwecks Fehlerresistenz versehen.

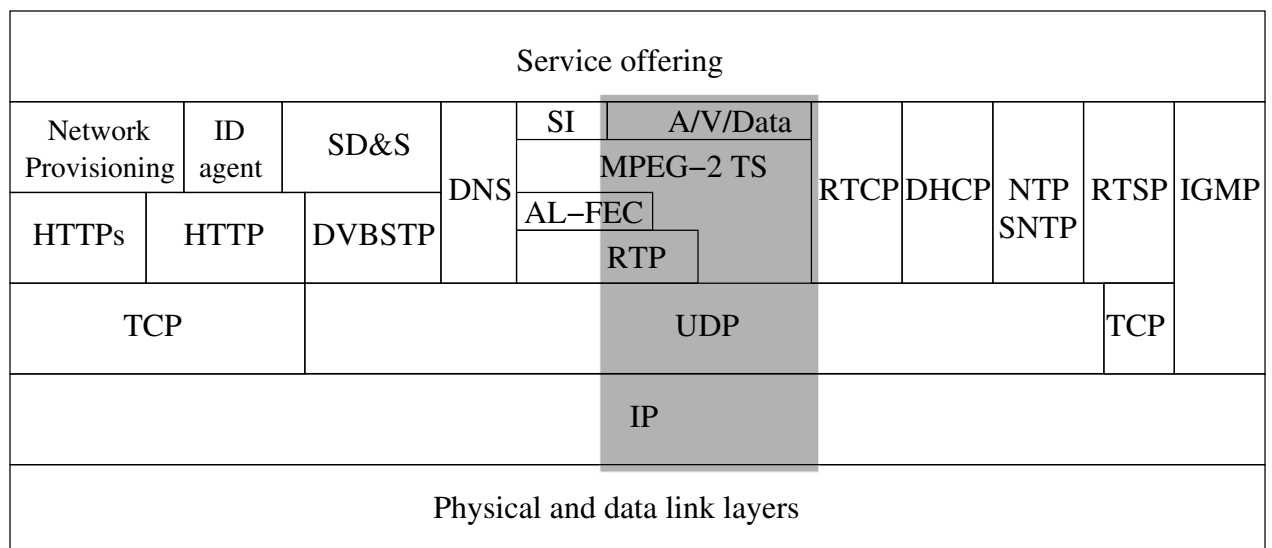


Abbildung 2.8: DVB-IPTV-Protokollstapel nach [6]

Der *DVB-IPTV*-Standard beschreibt die Übertragung des bei DVB verwendeten MPEG-2-Transportstroms, der beliebige Daten enthalten kann aber vor allem die komprimierten Videodaten transportiert, über IP-basierte Netze [6], [52]. Der Protokollstapel ist in der Abbildung 2.8 zu sehen. Danach werden MPEG2-TS-Pakete über RTP/UDP oder direkt über UDP übertragen. Der Standard enthält ein optionales Application-Layer-FEC⁵⁰ (AL-FEC), um den möglichen IP-Paketverlusten entgegenzuwirken. Die IP-Pakete sollen mit den entsprechenden DSCP⁵¹-Werten markiert werden, um den Verkehr im Netz mediengerecht zu behandeln. Die QoS-Anforderungen an ein IP-Netz sind sehr hoch, um die vom klassischen Fernsehen gewohnte Qualität⁵² zu gewährleisten [6].

Adobe RTMP und RTMFP dienen dazu, das Videostreaming direkt über TCP bzw. UDP stattfinden zu lassen. SMPTE-ST-2110 (s. Abschnitt 2.4) basiert auf dem ISMA-Protokollstapel. Der NATO-Standard STANAG-4609⁵³ [53] wird bei der militärischen luftgestützten Videoaufklärung eingesetzt. Der Standard und die Implementierungsrichtlinien [54] schreiben vor, den MPEG-2-TS zu verwenden, der sowohl Video- als auch sogenannte Kollateraldaten (z. B. Geoinformationen) trägt. Die MPEG-2-TS-Pakete werden drahtlos u. a. über IP-Netze übertragen. Wie die MPEG-2-TS-Pakete durch diese übertragen werden sollen, ist nicht festgelegt. In der Praxis verwendet man RTP/UDP dafür.

Es gibt weitere Standards, die entsprechende Protokollstapel aufweisen, aber keine weiteren Transportprotokolle einführen, so dass sie aus den Übersichtsgründen nicht erwähnt werden. Es gilt noch anzumerken, dass bei DVB ein Transport von IP-Paketen in einem MPEG-2-TS vorgesehen ist (DVB-IP-Datacast) [29]. Dabei werden IP-Pakete mittels *Multiprotocol Encapsulation* (MPE) in den MPEG-2-TS eingebettet. MPE stellt Mechanismen zur Empfängeradressierung bereit [15] [55]. Der dazu relevante DVB-H Standard [55], [56], [57] hat sich aber nicht durchgesetzt.

Die im Rahmen dieser Arbeit durchgeführten Untersuchungen befassen sich mit dem RTP/UDP-basierten Videostreaming.

2.5.2 Verfahren der Sicherungs- und Bitübertragungsschicht

Die Leistungsfähigkeit einer Kommunikationsverbindung wird aus der Sicht einer Endanwendung i. A. durch die übertragbare Informationsrate, die Verzögerung und die zeitliche Stabilität beider Parameter charakterisiert. Verschiedene Anwendungen haben verschiedene Anforderungen an diese Parameter.

Die Kapazität eines Funkkanals ist durch physikalische Gegebenheiten begrenzt. Die Kanalqualität kann durch die Sendeleistung beeinflusst werden. Aus Gründen der elektromagnetischen Verträglichkeit und aus gesundheitlichen Gründen sind aber Sendeleistungen durch gesetzliche

⁵⁰FEC: Forward Error Correction.

⁵¹DSCP: Differentiated Services Code Point.

⁵²Es darf maximal ein Artefakt pro Stunde auftreten. Beispielsweise darf ohne AL-FEC die IP-Paketverlustrate bei einem 4Mbit/s-Videostrom nicht höher als $1 \cdot 10^{-6}$ sein [6]. Ein IP-Paket enthält sieben MPEG-2-TS-Pakete. Ein MPEG-2-TS-Paket ist 188 Byte groß.

⁵³NATO Standardisation Agreement (STANAG) 4609 on Air Reconnaissance Primary Imagery Data Standard

Vorschriften beschränkt. Unter diesen Randbedingungen wird in den modernen Kommunikationsnetzen versucht, die optimale Verwendung der vorhandenen Funkressourcen zu erreichen.

Terrestrische Funknetze implementieren unterschiedliche u. a. adaptive Mechanismen auf der Bitübertragungs- und auf der Sicherungsschicht zur Gewährleistung der bestmöglichen Dienstleistung für höhere Schichten. Es gilt dabei, die Auswirkungen des schnellen, aber auch des langsamen Schwunds abzumildern (vgl. Abschnitt 2.2.1).

Die adaptiven Verfahren der Systeme für die Individualkommunikation unterscheiden sich von den Verfahren der Systeme für die Verteilkommunikation, da bei den Letzteren eine senderseitige Adaption nicht möglich ist. Die videospezifischen Verteilkommunikationssysteme bieten aber optional eine wirkungsvolle empfangsseitige Adaptionmöglichkeit an.

Im Folgenden wird eine kurze Übersicht über die grundlegenden Verfahren der zwei unteren Schichten des ISO-OSI-Referenzmodells in den relevanten Kommunikationsnetzen gegeben. Im Wesentlichen sind die folgenden Übertragungsmechanismen zu nennen:

- *Vielfachzugriffsverfahren*

Die Vielfachzugriffsverfahren wurden bereits im Abschnitt 2.2.3 angesprochen. Der Vollständigkeit halber werden sie an dieser Stelle behandelt. Das Ziel der Vielfachzugriffsverfahren ist die bestmögliche Aufteilung der gemeinsamen Funkressourcen unter den vorhandenen Kanälen.

Bei GSM verwendet man dafür ein kombiniertes FDMA/TDMA-Verfahren. Die erreichten Übertragungsdatenraten pro Kanal sind niedrig. Dieses System ist für das Videostreaming kaum geeignet.

Bei UMTS stiegen die Übertragungsdatenraten u. a. durch den Einsatz von WCDMA deutlich. Insbesondere mit der HSPA-Erweiterung ist es möglich geworden, das Videostreaming zu betreiben.

OFDM, worauf OFDMA basiert, sorgt u. a. in Kombination mit MIMO-Technik (LTE) für eine gegen schnellen Schwund resistente Übertragung. Dabei wird das Signal auf mehrere zueinander orthogonal stehende Frequenzträger verteilt. Das System ist für das Videostreaming geeignet.

COFDM ist zwar kein Vielfachzugriffsverfahren, könnte aber dazu erweitert werden. COFDM wird bei DVB-T eingesetzt. Durch die zusätzliche Kanalkodierung innerhalb der Modulationssymbole wird eine noch höhere Robustheit als bei OFDM bzgl. des schnellen Schwunds erreicht.

- *Modulation*

Es gibt zwei adaptive Modulationsarten:

– *Adaptive Modulation*

Wird die Wertigkeit des Modulationsverfahrens während eines Kommunikationsvorganges in Abhängigkeit von der Kanalqualität laufend verändert, so spricht man von der *adaptiven Modulation*. Die adaptive Modulation (in Kombination mit einer ebenfalls adaptiven Kanalkodierung) wird seit der EDGE-Einführung bei GSM verwendet. Auch die Mobilfunksysteme der dritten (UMTS/HSPA) und der vierten Generation (LTE) verwenden sie. Aus der Sicht einer Endanwendung äußert sich die

adaptive Modulation darin, dass die momentan verfügbare Übertragungsdatenrate je nach Kanalsituation schwankt.

– *Hierarchische Modulation*

Bei DVB-T ist eine optionale Möglichkeit der Datenratenadaption an die Übertragungsbedingungen ohne einen Rückkanal vorgesehen. Dies wird mit Hilfe einer so genannten *hierarchischen Modulation* realisiert.

Dabei werden zwei Datenströme unterschiedlicher Datenrate – und unterschiedlich stark fehlergeschützt – auf der Modulationsebene ineinander verschachtelt. Die Verschachtelung wird im Signalraum so durchgeführt, dass sich ein eventuell fehlerhaft empfangenes Modulationssymbol des höherratigen Stromes auf das Modulationssymbol des niederratigen Stromes abbilden lässt.

- *Fehlerkorrektur*

– *Adaptive Kanalkodierung*

Wie vorhin erwähnt, wird die adaptive Modulation zusammen mit der *adaptiven* Kanalkodierung angewendet. Bei einer guten Kanalqualität wird den zu übertragenden Daten weniger Redundanz hinzugefügt als bei einer schlechten Kanalqualität. Somit ist die Nutzdatenrate bei einer guten Kanalqualität höher.

– *Hybrides ARQ*⁵⁴

ARQ ist ein Mechanismus der erneuten Übertragungen⁵⁵ von verloren gegangenen Datenblöcken. *Hybrides ARQ (HARQ)* wird verwendet, um die Anzahl der Sendewiederholungen bei fehlerhaft empfangenen Daten zu minimieren. Die Letzteren werden zunächst nicht verworfen, sondern mit den angeforderten Fehlerschutzdaten kombiniert, um die Fehlerstellen zu finden und zu korrigieren.

– *Ungleichmäßiger Fehlerschutz*

Die Videodaten besitzen aufgrund der gegenseitigen Abhängigkeiten unterschiedliche Wichtigkeiten. Beim üblichen Fehlerschutz werden alle Videodaten mit der gleichen Redundanz versehen. Es wäre denkbar, die Videodaten mit einem *ungleichmäßigen Fehlerschutz* entsprechend der Wichtigkeit der Daten zu versehen, wie dies bei DAB für unterschiedlich wichtige Audiodaten geschieht [15].

Der in der vorliegenden Arbeit verwendete Emulator eines Mobilfunkzugangsnetzes implementiert die relevanten Verfahren der unteren zwei ISO-OSI-Schichten. Für die Details sei auf den Abschnitt 6.2.2 verwiesen.

2.5.3 Die Dienstgüte im Schichtenmodell

Die Dienstgüte einer Videostreaming-Anwendung wird aus der kommunikationstechnischen Sicht im Wesentlichen von den folgenden Faktoren beeinflusst:

⁵⁴ARQ: Automatic Repeat Request

⁵⁵Sendewiederholungen

- *Störender Einfluss des Übertragungswegs*

Damit sind die Datenverluste im Netz und die durch die Übertragung verursachten Verzögerungen bzw. Verzögerungsschwankungen gemeint. Die Details dazu werden im Abschnitt 3.3.1 behandelt.

- *Komprimierungseinstellungen*

Im Rahmen dieser Arbeit wird u. a. die Videoqualität (s. Abschnitt 3.4) geschätzt. Die begleitenden Audiostromdaten werden nicht betrachtet. Somit sind mit den Komprimierungseinstellungen die Einstellungen der Videokomprimierung (s. Abschnitt 3.2.2) gemeint. Aus der kommunikationstechnischen Sicht sind die folgenden Einstellungen relevant:

- *Kompressionsrate*

Die Kompressionsrate⁵⁶ ist das Verhältnis der ursprünglichen Datenmenge zu der komprimierten Datenmenge [59]. Durch eine Veränderung der Kompressionsrate kann die Videodatenrate an die Übertragungsdatenrate angepasst werden.

- *Videostruktur*

Die Videostruktur ist im Abschnitt 3.2.2.4 beschrieben. Gehen durch den störenden Einfluss des Übertragungsweges die besonders wichtigen Videodaten verloren, kann dies zu Folgefehlern bei der Dekomprimierung der fehlerfrei empfangenen Videodaten führen. Die Quantität der Folgefehler hängt von der verwendeten Videostruktur ab.

Die Abbildung 2.9 zeigt, welche Abschätzung der Dienstgüte in welchen Schichten möglich ist. Da Audioströme nicht betrachtet werden, ist in der Abbildung von der Videoqualität die Rede.

Die Videoqualität aus der Sicht eines Zuschauers ist subjektiv und kann nur auf der Anwendungsschicht bestimmt werden. Die Verfahren dazu werden im Abschnitt 3.4 kurz angesprochen.

Auf der Darstellungsschicht sind alle zur Visualisierung benötigten Videodaten verfügbar. Die Videoqualität lässt sich objektiv feststellen. Die objektiv festgestellte Videoqualität lässt sich mit Hilfe der zuvor ermittelten Abbildungen auf die subjektive Videoqualität übertragen.

Auf der Sitzungsschicht ist es möglich, beim Transport verlorene bzw. nur teilweise korrekt empfangene Videobilder sowie verspätete Videobilder festzustellen. Der Umfang der oben erwähnten Folgefehler ist nicht direkt erkennbar.

Auf der Transport- und der Vermittlungsschicht ist es möglich, verlorene Video-UDP-Pakete bzw. verlorene Video-IP-Pakete zu ermitteln. Da die Videobilder je nach der Videodatenrate auf mehrere UDP- bzw. IP-Pakete verteilt sind, ist die Abschätzung des Einflusses der Paketverluste auf die Videoqualität ungenau. Da die Anzahl der verlorenen UDP- und der verlorenen IP-Pakete identisch ist, sind die beiden Schichten in der Abbildung 2.9 verschmolzen dargestellt.

⁵⁶Es existiert auch eine umgekehrte Definition der Kompressionsrate. Danach ist die Kompressionsrate das Verhältnis der komprimierten Datenmenge zu der ursprünglichen Datenmenge. Siehe S.17 in [58].

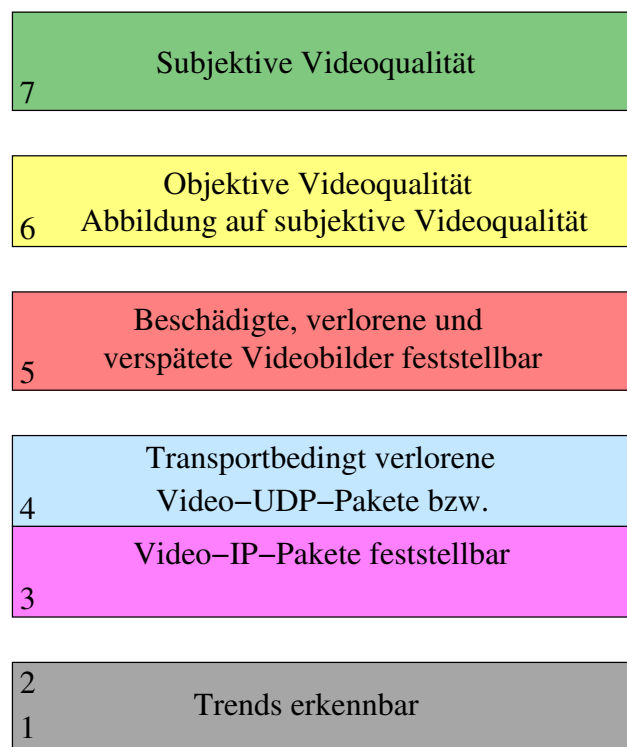


Abbildung 2.9: Dienstgüteabschätzung einer Videostreaming-Anwendung

Die unteren zwei Schichten sind für ein Transportnetz spezifisch. Aus den Statistiken dieser Schichten lassen sich keine direkten Aussagen über die zu erwartende Videoqualität tätigen. Die Trends sind aber erkennbar.

2.5.4 Zusammenfassende Diskussion

Das moderne Videostreaming wird in der Individual-, Gruppenruf- und teilweise in der Verteilkommunikation über IP abgewickelt. Vor allem aber in der Verteilkommunikation verwendet man den MPEG-2-Transportstrom in der Transport- bzw. Vermittlungsschicht (s. Abbildung 2.10). Letzteres ist kein Gegenstand weiterer Betrachtung.

Je nach Anwendungsart verwendet man in der Transportschicht entweder TCP oder UDP.

Die TCP-basierten Varianten sind aufgrund der Protokollnatur nicht für die Gruppenruf- bzw. die Verteilkommunikation geeignet, da TCP ja einen Rückkanal voraussetzt. Auf der anderen Seite garantiert TCP eine fehlerfreie und reihenfolgetreue Zustellung der Daten. Das wird durch die Flusssteuerung und die Staukontrolle erreicht, wobei die Senderate entsprechend angepasst wird. Weil TCP ohne Rücksicht auf die Anwendung agiert, kann dies zu einem nichtanwendungsgerechten Verhalten führen (Unterbrechungen der Wiedergabe). Dies kann man bei dem progressiven Videodownload beobachten.

Der Ansatz des adaptiven Videostreaming über HTTP basiert deshalb darauf, dass ein Videostreamingclient über HTTP/TCP nicht das gesamte Video, sondern nur einen kleinen Teil (ein Videosegment) erhält. Auf dem Videostreamingserver sind unterschiedlich stark komprimierte Varianten des Videos verfügbar. Die Anwendung wählt die Variante, die ihren momentanen

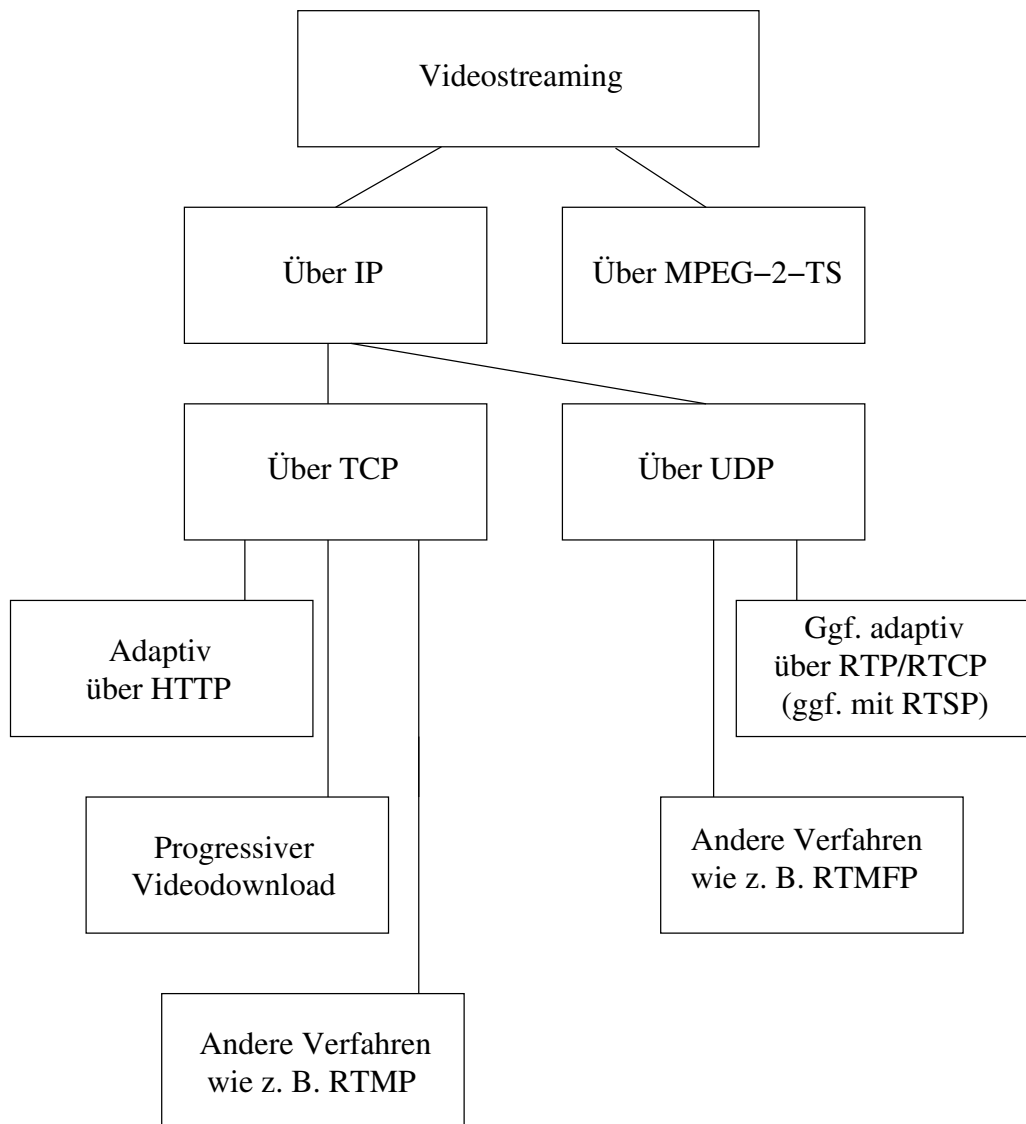


Abbildung 2.10: Videostreamingarten

Empfangsbedingungen entspricht. Die übliche Videosegmentlänge und somit das Zeitintervall bis zum nächsten möglichen Adaptionsschritt liegt laut [60] und [61] bei 10s. Wie bereits erwähnt weisen HTTP-basierte Verfahren einen Versatz von 20 bis 30s zum Live-Stream auf [46].

Andere TCP-basierte Verfahren, wie z. B. Adobe RTMP, implementieren – ähnlich zu den HTTP-basierten Verfahren – eine anwendungsgerechte Sitzungsschicht. Diese Verfahren sind proprietär.

Die UDP-basierten Verfahren sind sowohl für die Individualkommunikation als auch für die Gruppenruf- und die Verteilkommunikation geeignet. UDP verwendet keinen Rückkanal und kann keine fehlerfreie und keine reihenfolgetreue Zustellung garantieren. Dies müssen ggf. Protokolle der höheren anwendungsgerechten Schichten sicherstellen. Dafür aber wird die mittlere Senderate durch die Videostreaminganwendung kontrolliert (und bestimmt). Der Datenstrom wird normalerweise über RTP verschickt, das für die Echtzeitübertragung der Audio- und Videodaten entwickelt wurde. Zusätzlich zu RTP kann RTCP verwendet werden. Damit können Sender und Empfänger Statistiken und Steuerinformationen zu einer RTP-Sitzung aus-

tauschen. Dies ermöglicht die Implementierung von adaptiven Verfahren. Im Gegensatz zu den HTTP/TCP-basierten Verfahren, können die UDP-basierten Verfahren ihre Adaptionentscheidungen auf sehr kurzen Zeitskalen treffen. In [62] wird empfohlen, die Rückmeldung eines Videostreamingclients in einer Mobilfunkumgebung mindestens ein Mal pro Sekunde an den Videosever zu senden. Die besten Adaptionsergebnisse ließen sich mit einer Rückmeldung pro RTT (Round Trip Time) erreichen. Die sehr häufigen Berichte können zu einer höheren Netzbelastung – was bei einer schmalbandigen Verbindung störend ist – und zu einer stärkeren Serverbelastung führen. Deshalb empfiehlt IETF in [63] ein minimales Intervall von 5 s.

Um den fehlerfreien Transport der UDP-basierten Videoströme sicherzustellen, geben die einschlägigen Normen entsprechende Maßnahmen vor. So ist die Anwendung von AL-FEC bei DVB-IPTV auf RTP-Pakete, bei 3GPP (e)MBMS auf UDP-Pakete und bei DVB-IP-Datacast [29] auf IP-Pakete vorgesehen. Bei 3GPP PSS wird die RTP-Paketwiederholung verwendet. Desweiteren wird bei 3GPP PSS optional eine senderseitige Adaptivität empfohlen.

Die modernen Verfahren der Bitübertragungs- und der Sicherungsschicht im Mobilfunk reagieren adaptiv auf die Kanalsituation. Die Kanalinformationen, die sie besitzen, können in den höheren Schichten für die Adaptionentscheidungen verwendet werden.

2.6 Mobilfunknetz

Zur Einordnung des Untersuchungskontexts werden hier die dafür notwendigen Grundlagen der modernen Mobilfunknetze vorgestellt.

2.6.1 Netzarchitektur

Ein modernes Mobilfunknetz weist drei wesentlichen Komponenten auf:

- *Mobiles Gerät UE (User Equipment)*
Mit seiner Hilfe stellt der Mobilfunkteilnehmer eine Verbindung zum Mobilfunknetz her.
- *Mobilfunkzugangsnetz RAN (Radio Access Network)*
Es ist für die Übertragung der Daten über den terrestrischen Funkkanal zuständig.
- *Drahtgebundenes Kernnetz*
Es ist für die Verbindung der einzelnen Mobilfunkzugangsnetzbereiche, die Verwaltung der Dienste und die Weiterleitung der Daten an die externen bzw. die internen Kommunikationsteilnehmer zuständig.

Für die vorliegende Arbeit ist lediglich die Funktion des Mobilfunkzugangsnetzes relevant. Zwei für die Videostreaminganwendungen relevante Architekturen sind in den Abbildungen 2.11 und 2.12 zu sehen. Das Mobilfunkzugangsnetz in UMTS/HSPA trägt den Namen *UTRAN* (Universal Terrestrial Radio Access Network), während es bei LTE *EUTRAN* (Evolved Universal Terrestrial Radio Access) heißt.

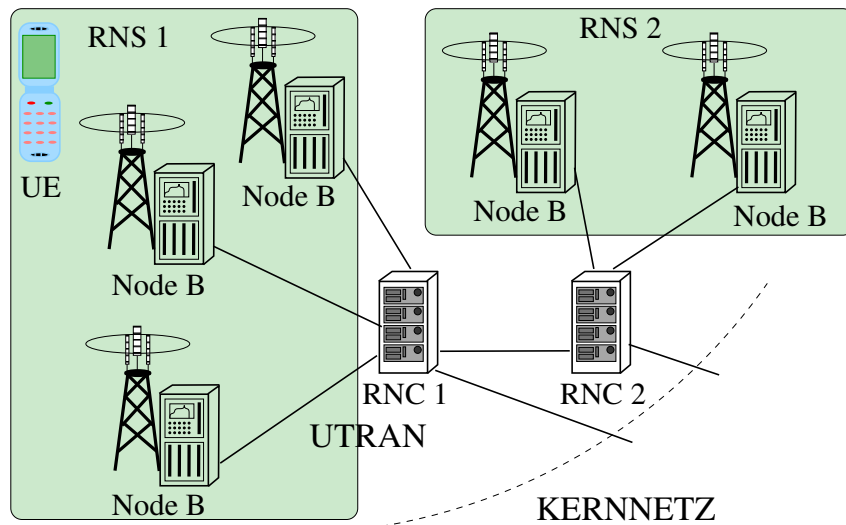


Abbildung 2.11: UTRAN-Architektur (UMTS/HSPA)

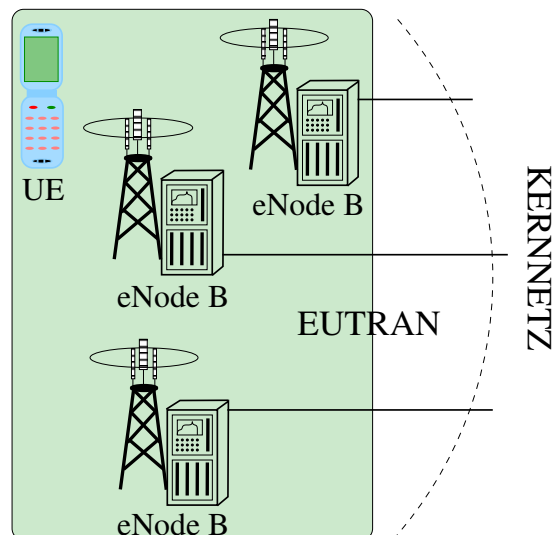


Abbildung 2.12: EUTRAN-Architektur (LTE)

Das Mobilfunkzugangnetz realisiert die Funktionen der Bitübertragungs-, der Sicherungs- und der Vermittlungsschicht.

In einem UTRAN sind die folgenden beiden Komponenten zu unterscheiden:

- *Radio Network Controller (RNC)*
An ein RNC sind mehrere Basisstationen (Node B) angeschlossen, die ein *Radio Network Subsystem (RNS)* bilden. Der RNC verwaltet die Funkressourcen eines RNS und stellt den Übergang zum Kernnetz dar. Darin sind die Funktionen der Sicherungs- und der Vermittlungsschicht implementiert.
- *Node B*
Die Basisstation ist für die Bitübertragungsschicht zuständig. Im Falle der HSPA-

Erweiterung wird hier ein Teil der Sicherungsschicht implementiert. Dies bewirkt eine merkliche Senkung der Verzögerungswerte der übertragenen Datenblöcke.

Ein EUTRAN besteht lediglich aus *eNode-B*-Elementen, die Funktionen aller drei Schichten implementieren. Der Grund für die Verlagerung der RNC-Funktionen in ein eNode-B ist die Eliminierung der Verzögerung aufgrund der verteilten UTRAN-Architektur.

In [64] wurden die Übertragungsverzögerungen in den öffentlichen UMTS/HSPA und LTE-Netzen gemessen und verglichen. Der Median für RTT betrug bei LTE 33 ms. Der entsprechende Wert bei UMTS/HSPA lag bei 42 ms. Hat man nur die Downlink-Richtung vermessen, so betrug der Median der Verzögerung 8 ms bei LTE und 18 ms bei UMTS/HSPA. Diese Werte sind eine enorme Verbesserung gegenüber einer GPRS/EDGE-basierten Übertragung mit den Werten von mehreren Hundert Millisekunden.

Die Verzögerungsunterschiede zwischen UMTS/HSPA und LTE sind für die in dieser Arbeit durchgeführten Untersuchungen irrelevant.

Während UMTS/HSPA-Systeme in der Downlink-Richtung eine Datenrate von bis zu 42 Mbit/s (3GPP Release 7) bieten, sind bei LTE bzw. LTE-Advanced die Datenraten von bis zu 100 Mbit/s bzw. mehrere Hundert Mbit/s theoretisch möglich. Die praktischen Datenraten bei LTE sind aber frequenzbandabhängig und sind deutlich kleiner. Es ist noch anzumerken, dass diese Werte die für alle Mobilfunkteilnehmer gemeinsam zur Verfügung stehenden Ressourcen darstellen.

Für die im Rahmen dieser Arbeit durchgeführten Untersuchungen sind die Datenratenunterschiede zwischen UMTS/HSPA und LTE irrelevant. Es werden hier Videos mit den Videodatenraten von 308 kbit/s bzw. 900 kbit/s verwendet.

Die Untersuchungen der vorliegenden Arbeit werden in einem UTRAN-Kontext durchgeführt.

2.6.2 Kanalarten

In den modernen Mobilfunksystemen wie UMTS/HSPA und LTE lassen sich Kanäle in einem RAN anhand von drei Merkmalen klassifizieren (siehe Abbildung 2.13):

- *Art der zu übertragenden Daten*

Es lässt sich zwischen *Nutz-* und *Steuerdaten* unterscheiden, die jeweils über getrennte Kanäle übertragen werden.

- *Kanalschicht*

Spricht man von Kanälen in einem RAN, so sind darin drei Kanalschichten definiert, um die physikalische Datenübertragung auf der Funkschnittstelle von der logischen Übertragung zu trennen. Im Einzelnen gibt es:

- *Logische Kanäle*

Ein logischer Kanal beschreibt einen Übertragungskanal für die Nutz- oder Steuerdaten innerhalb der Sicherungsschicht des ISO-OSI-Schichtenmodells und dient der

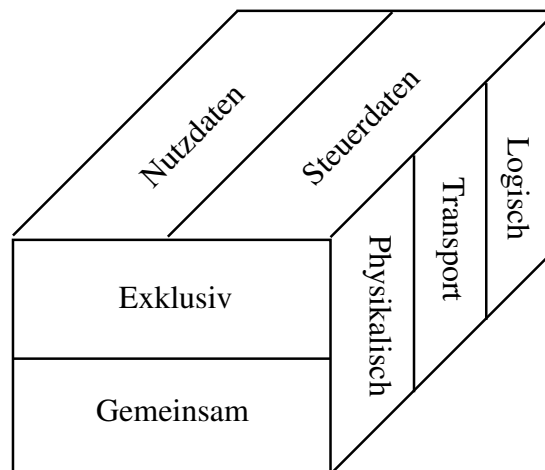


Abbildung 2.13: Kanalarten im RAN

Abstraktion der Datenübertragung von der physikalischen Übertragung. Es existieren unterschiedliche logische Kanalarten für unterschiedliche Informationsflussarten. Logische Kanäle werden je nach Art einzeln oder zusammengefasst auf Transportkanäle durch die Sicherungsschicht abgebildet.

– *Transportkanäle*

Dieser Kanaltyp kann der Sicherungsschicht der ISO-OSI-Schichtenmodells zugeordnet werden. Die Daten aus den logischen Kanälen werden in kleinere Datenblöcke so aufgeteilt, dass eine effiziente Übertragung über die Funkschnittstelle möglich wird. Jeder Datenblock wird mit einem Header versehen, in dem die Parameter der physikalischen Übertragung für den aktuellen Datenblock angegeben sind. Diese Parameter als Ganzes werden *Transportformat* genannt.

Anhand der Parameterwerte des Transportformats kann man unter anderem die aktuelle Datenübertragungsrate zwischen der Sicherungsschicht und der Bitübertragungsschicht feststellen. In den hier vorgestellten Untersuchungen werden die Übertragungsdatenraten im Mobilfunkkanal anhand des für den jeweiligen Datenblock ermittelten Transportformats geschätzt.

Transportkanäle werden einzeln oder zusammengefasst auf physikalische Kanäle abgebildet.

– *Physikalische Kanäle*

Die Übertragung über die Luftschnittstelle erfolgt mittels der physikalischen Kanäle. Auf dieser Schicht findet die Kanalcodierung statt, die im Transportformat beschrieben wird. Ein Transportkanal kann nicht nur auf einen, sondern auch auf mehrere physikalische Kanäle gleichzeitig abgebildet werden, um dessen Inhalt übertragen zu können.

- *Nutzungsart*

Eine weitere Dimension stellt die Nutzungsart der vorhandenen Ressourcen dar. So unterscheidet man zwischen den für einzelne Teilnehmer *exklusiven* Kanälen und den mehreren Teilnehmern *gemeinsamen* Kanälen. Im Gegensatz zu den exklusiven Kanälen mit einer ggf. garantierten Datenrate, kann in den gemeinsam genutzten Kanälen eine ggf. höhere mittlere Übertragungsdatenrate pro Teilnehmer aufgrund der effizienten Ressourcenauslastung erreicht werden. Der Nachteil der gemeinsamen Kanäle ist eine starke Zeitvarianz der Übertragungsdatenrate. Die IP-basierte Kommunikation in den Mobilfunknetzen findet i. d. R. über die gemeinsamen Kanäle statt [27].

Zusätzlich zu den drei Dimensionen wird explizit zwischen den *Downlinkkanälen* (zum Mobilfunkteilnehmer hin) und den *Uplinkkanälen* (vom Mobilfunkteilnehmer weg) unterschieden.

2.6.3 Herausforderungen für das Videostreaming

Aus Sicht der Konstanz der Übertragungsressourcen kann man basierend auf den vorherigen Abschnitten und weiteren bekannten Erkenntnissen die folgenden Aspekte als Herausforderungen für die Videostreaming-Anwendungen in den Mobilfunknetzen identifizieren:

- *Grundsätzliche Zeitvarianz*

- *Zeitvariable Wellenausbreitungssituation durch die Teilnehmermobilität*

Mit der adaptiven Modulation und der adaptiven Kanalcodierung reagiert ein Mobilfunkzugangnetz auf die zeitvariablen Wellenausbreitungssituationen. Dies führt zu den zeitvariablen Übertragungsdatenraten.

- *Zeitvariable Teilnehmeranzahl bzw. zeitvariables Verkehrsaufkommen*

Die zeitvariable Teilnehmeranzahl bzw. das zeitvariable Verkehrsaufkommen eines gemeinsam genutzten Kanals führt zu einer zeitvariablen Ressourcenzuteilung pro Teilnehmer, was zu den zeitvariablen Übertragungsdatenraten führt.

- *Zellatmung bei UTRAN*

Aufgrund des bei UTRAN verwendeten WCDMA kommt es speziell in diesem RAN zu einer so genannten Zellatmung. Es bedeutet, dass sich die Zellgrenzen in Abhängigkeit der Teilnehmerzahl der Zelle verschieben. Bei einer großen Teilnehmerzahl wird das Störgeräusch aufgrund der asynchronen Übertragung der teilnehmerspezifischen orthogonalen Codes merklich, was die Dekodierbarkeit beeinflusst.

- *Zugangnetzheterogenität*

Da die LTE-Versorgung nicht flächendeckend ist, kann ein Videostreamingnutzer nach einem Zellwechsel einem schmalbandigeren UTRAN anstatt einem EUTRAN zugeordnet werden.

- *Entzug der Sendeleistung*

Ein hohes Verkehrsaufkommen, das nach exklusiven Kanälen (z. B. Telefonie) verlangt, kann zum Entzug eines Teils der für die gemeinsamen Kanäle verfügbaren Sendeleistung führen, was zu weniger Übertragungsressourcen in den letzteren Kanälen führen wird.

Diese Herausforderungen gilt es, anwendungsgerecht anzugehen.

3 Grundlagen des adaptiven Videostreaming und der Auswahl von Videosequenzen

3.1 Merkmale des Videostreaming

Das Phänomen Videostreaming wurde im Abschnitt 1.1 dargestellt. Im Rahmen dieser Arbeit wird das Videostreaming zum Konsumieren eines Abrufvideodienstes¹ bzw. eines Live-Stream behandelt.

Das Videostreaming als eine Kommunikationsanwendung und gleichzeitig als eine Übertragungsmethode eines Videos über ein IP-basiertes Datenkommunikationsnetz kann durch die folgenden Merkmale charakterisiert werden:

- *Verteilte Kommunikationsform*
Ein Video wird auf einem Videostreamingserver abgelegt oder in Echtzeit erzeugt. Auf eine Anfrage hin werden die Videodaten an einen entfernten Videoclient gesendet.
- *Datenstrombasierter Inhaltskonsum*
Das Videodaten Speicherformat und die spezielle Anwendungssoftware erlauben einen datenstrombasierten Konsum. Eine Zwischenspeicherung des empfangenen Videodatenstromes auf dem lokalen Datenträger in der Dateiform ist nicht erforderlich. Es ist möglich, die Videobilder ab bestimmten Synchronisationspunkten im Strom zu konsumieren.
- *Gleichzeitiger Datenempfang und Konsum*
Der Konsum erfolgt entsprechend dem Abspielzeitpunkt unmittelbar nach dem Eintreffen der anzuzeigenden Videobilder. Während der Anzeige werden die nächsten Videodaten empfangen.
- *Die Videodatenrate bestimmt die Sendedatenrate.*
Die mittlere Videodatenrate bestimmt die mittlere Sendedatenrate, mit der der Videostreamingserver die Videodaten versendet.
- *Zeitsensitivität*
Nicht nur die Korrektheit der empfangenen Videodaten, sondern auch ihre rechtzeitige

¹Engl: Video-On-Demand

Ankunft im Videoclient tragen zur einem fehlerfreien Verhalten bei. Der Videostreamingverkehr ist unelastisch.

- *Adaptivität über Datenvolumen*

Wird die Adaptivität unterstützt, so wird bei einer sich verändernden Übertragungsdatenrate des Datenkommunikationsnetzes die Videodatenrate und somit die Sendedatenrate an die Übertragungsressourcen über die Veränderung des zu übertragenden Datenvolumens anwendungsgerecht angepasst².

Im Folgenden werden die für diese Arbeit relevanten Sachverhalte beleuchtet.

3.2 Videoerzeugung

Die Erzeugung eines digitalen Videos geschieht durch die Digitalisierung analoger visueller Informationen. Dabei wird eine Reduktion der Irrelevanzen durchgeführt, die den Datenumfang verkleinert. Für das menschliche Wahrnehmungssystem bleibt dies aber unbemerkt. Im Abschnitt 3.2.1 wird das für diese Arbeit relevante digitale Format kurz erklärt.

Um das Video einem örtlich und/oder zeitlich entfernten Nutzer ressourcenschonend zur Verfügung zu stellen, muss durch die anschließende Komprimierung der Datenumfang deutlich reduziert werden (Abschnitt 3.2.2).

3.2.1 Digitalisierung

Die Digitalisierung eines Videos erfolgt bereits in der Kamera. Dabei wird eine lineare Farbtransformation aus dem RGB^3 -Raum in einen YC_1C_2 -Raum durchgeführt. Ein YC_1C_2 -Signal besteht aus einer Helligkeitskomponente Y und zwei Farbdifferenzkomponenten C_1, C_2 . Dieses Format hat psychovisuelle Vorteile bzgl. der Datenreduzierung (Unterabtastung und die kleinere Bitanzahl pro Signalkomponente). Es existieren verschiedene Ausprägungen der linearen Farbraumtransformation, die sich durch die Koeffizientenmatrix der Abbildung unterscheiden. Für diese Arbeit werden Videosequenzen in der sogenannten YUV -Ausprägung verwendet.

Im Rahmen dieser Arbeit werden die unterabgetasteten Videosequenzen mit effektiv 12 Bit pro Pixel verwendet.

Im Zuge der Verfahrensbewertung (Abschnitt 6) erfolgt die objektive Bewertung der Videoqualität auf der Basis der unkomprimierten YUV -Videosequenzen.

²Im Gegensatz zur Anpassung über die Zeit, wie dies bei den elastischen Verkehren (z.B. E-Mail, FTP usw.) der Fall ist.

³RGB: Rot, Grün, Blau

3.2.2 Komprimierung

Ein digitalisiertes Video weist trotz der angewandten Unterabtastung immer noch einen sehr großen Datenumfang auf und stellt dementsprechend große Datenraten- und Speicherplatzanforderungen an die das Video zu verarbeitenden Systeme. Deshalb bedarf es weiterer Datenreduktionsschritte, die zu einer hohen Kompressionsrate (vgl. Abschnitt 2.5.3) führen.

Man unterscheidet zwischen *symmetrischen* Komprimierungsverfahren (der Komprimierungs- und der Dekomprimierungsaufwand sind in etwa gleich groß) und *asymmetrischen* Verfahren (die Dekomprimierung braucht weniger Ressourcen als die Komprimierung). Die modernen Videokomprimierungsverfahren im Videostreamingbereich sind i. d. R. asymmetrischer Natur.

Es lassen sich zwei grundsätzlich unterschiedliche Methoden zur Reduzierung des Datenumfanges identifizieren:

- *Redundanzreduktion*

Das Ziel dieser Maßnahme ist es, ein und dieselbe mehrfach vorhandene Information auf das einmalige Auftreten zu reduzieren. Das Entfernen geschieht ohne einen Informationsverlust. Deshalb ist dies eine *verlustfreie* Komprimierungsmethode.

- *Irrelevanzreduktion*

Hier werden die zu entfernenden Informationen aus psychovisuellen, technischen und/oder semantischen Gründen als unwichtig betrachtet, was den subjektiven Charakter von Verfahren dieser Art unterstreicht. Sie sind deshalb *verlustbehaftet*. Es gibt kein Gegenstück und somit keinen Aufwand bei der Dekomprimierung. Durch den Grad der Irrelevanzreduktion wird ggf. die videostreaminggerechte Anpassung des zu übertragenden Datenvolumens gesteuert.

Im Folgenden wird auf verschiedene Arten der Redundanz und der Irrelevanz und auf Mechanismen und Verfahren zu ihrer Reduzierung näher eingegangen.

3.2.2.1 Redundanzreduktion

Eine Videosequenz weist i. d. R. unterschiedliche Redundanzarten zugleich auf. In der Abbildung 3.1 wird eine Klassifizierung vorgenommen. Den im Bild zu sehenden Begriff *statistische* Redundanz trifft man oft in der englischsprachigen Literatur. Er dient zur Abgrenzung vom dort ebenfalls oft verwendeten Begriff der *psychovisuellen* Redundanz (vgl. [65] und [66]). Der Letztere bezeichnet (nicht ganz zutreffend) die vorhandene Irrelevanz, die aus psychovisuellen Gründen entfernt werden kann.

Die Reduktion der *Kodierungsredundanz* befasst sich mit den mehrfach vorhandenen Informationen, die durch eine nichtoptimale digitale Kodierung hervorgerufen wurden.

Die Korrelation zwischen den Pixelwerten in einem Video ist der Grund für die *Interpixelredundanz*. Tritt diese Redundanz innerhalb eines Videobildes auf, so bezeichnet man sie als

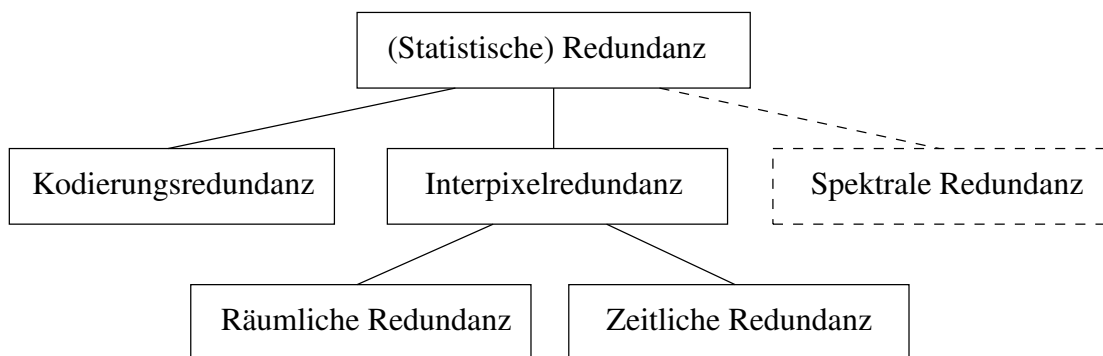


Abbildung 3.1: Klassifizierung der Redundanzarten

die *räumliche* Interpixelredundanz. Korrelieren die Pixelwerte der benachbarten Videobilder, so liegt die *zeitliche* Interpixelredundanz vor.

Die sogenannte *spektrale Redundanz* tritt in den multispektralen Satellitenbildern auf, wird hier nur aus Vollständigkeitsgründen genannt und nicht weiter betrachtet.

In der Praxis werden zwei grundsätzlich unterschiedliche Verfahren zur Interpixelredundanzreduktion angewendet. Das eine Verfahren transformiert das ursprüngliche Signal aus dem Ortsbereich in den Frequenzbereich⁴ [66], [67] und [68]. Es heißt *Signaltransformation*. Das andere Verfahren arbeitet voraussagenbasiert und heißt *Prädiktion*. Es wird zur Reduzierung sowohl der räumlichen als auch der zeitlichen Redundanz angewendet. Nur die Differenz (der Prädiktionsfehler) zwischen der Voraussage und dem tatsächlichen Wert wird kodiert⁵. Kombiniert man beide Verfahrensarten zur Interpixelredundanzreduktion, so heißt das resultierende Verfahren *hybrid*. Alle modernen Videokompressionsverfahren sind hybride Verfahren.

Der grundlegende Ablauf eines Videokompressionverfahrens besteht in der Reduktion der Interpixelredundanz und der Reduktion der Kodierungsredundanz (Entropiekodierung), wie in der Abbildung 3.2 zu sehen ist. Für weitere Details sei auf [66] und [69] verwiesen.

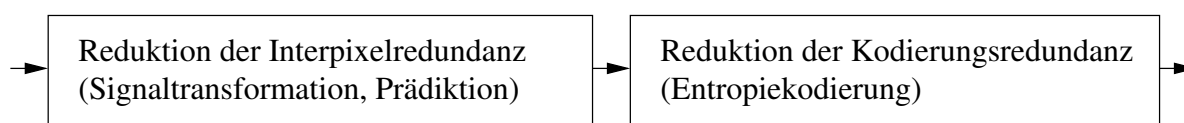


Abbildung 3.2: Redundanzreduktion

Die Grenze der Kompressionsrate der Redundanzreduktionsverfahren wird durch die Entropie⁶ des zu komprimierenden Video bestimmt. Moderne Redundanzreduktionsverfahren erreichen bei der Entfernung der räumlichen Redundanz je nach Videoart eine Kompressionsrate von nur ca. 2 bis 5 [70], [71].

Durch den Einsatz der zeitlichen Interpixelredundanzreduktion werden erhebliche Steigerungen der Kompressionsrate [72] erreicht. Diese Verfahren führen aber zur Entstehung von

⁴Dort hat das Signal eine redundanzärmere Darstellungsform.

⁵Deshalb heißt dieses Verfahren auch *Differenzkodierung* [66].

⁶Hier bedeutet die Entropie den mittleren Informationsgehalt pro Videobildpixel.

Abhängigkeiten zwischen den zeitlich benachbarten Videobildern, die im ursprünglichen digitalisierten nichtkomprimierten Video nicht existierten. Der Vorteil einer besseren Kompressionsrate wird mit den folgenden Nachteilen erkaufte:

- *Prinzipbedingte Komprimierungsverzögerung*
Wird die Reduzierung der zeitlichen Redundanz z. B. eines Live-Videostroms unter der Einbeziehung der zukünftigen Videobilder durchgeführt, so muss die Zukunft abgewartet werden. Dies führt zu der prinzipbedingten Verzögerung für das Versenden der komprimierten Videobilder.
- *Weniger Synchronisationspunkte / Längere Synchronisationszeiten*
Die unabhängig komprimierten Videobilder dienen als Synchronisationspunkte [73] für einen Dekoder, wenn auf eine beliebige Stelle im Videostrom zugegriffen wird. Da nicht jedes Videobild unabhängig komprimiert wird, muss ggf. auf den nächsten Synchronisationspunkt gewartet werden.
- *Höhere Fehleranfälligkeit*
Durch die eingeführten Videobildabhängigkeiten können im Falle von Videodatenverlusten auf dem Übertragungswege – je nach Grad dieser Abhängigkeiten – Folgefehler auftreten. Dies wird *MPEG-Fehlerfortpflanzung* genannt. In der Abbildung 3.3 wird die Gegenläufigkeit der Komprimierungseffizienz und der Fehlerfortpflanzungsresistenz mit dem steigenden Grad der zeitlichen Interpixelredundanzreduktion (die Verläufe sind lediglich aus stilistischen Gründen linear) visualisiert.

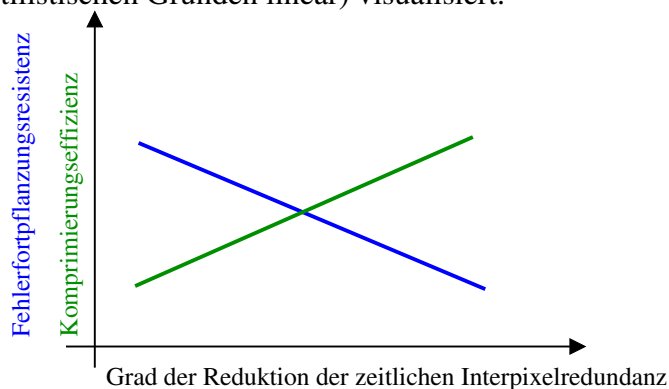


Abbildung 3.3: Einfluss der zeitlichen Interpixelredundanzreduktion

Moderne Videokomprimierungsstandards fordern im Rahmen einer fehlerresistenten Videokomprimierung das regelmäßige Einfügen der unabhängig komprimierten Videobilder und somit auch der Synchronisationspunkte⁷, wie es z. B. in [41] nachzulesen ist.

3.2.2.2 Irrelevanzreduktion

Die Methode der Irrelevanzreduktion ist eine wirksame Methode zur Reduzierung des Videodatenumfangs und kann dort angewendet werden, wo die Redundanzreduktion an ihre natürlichen Grenzen stößt. *Irrelevant* sind laut [59] solche Daten oder Informationen, die ein Empfänger entweder

⁷fixed interval synchronization

- *nicht wahrnehmen kann* oder
- *sich dafür nicht interessiert.*

Die Irrelevanzreduktion ist per Definition eine verlustbehaftete Methode der Datenreduktion. In der Regel wird der Empfänger mit einem menschlichen Betrachter und die irrelevanten Informationen mit der *psychovisuellen Irrelevanz* assoziiert.

Die psychovisuelle Irrelevanz beschreibt Informationsanteile im Videosignal, die vom menschlichen visuellen System nicht wahrgenommen werden können. Subjektiv wirkt das so verschlankte Videosignal verlustlos. Die dabei erzielten Kompressionsraten sind eventuell zu gering, um das Videostreaming ggf. über schmalbandige Kommunikationskanäle stattfinden zu lassen.

Es bedarf ggf. einer stärkeren Komprimierung, die sich zwar u. U. in der Videoqualität bemerkbar macht, aber im Kontext der Nutzung akzeptiert wird. Aus dem Video werden Informationen entfernt, die semantisch keine oder eine irrelevante Rolle spielen. Die einfachsten Methoden der *semantischen Irrelevanzreduktion* sind die Reduzierung der Videobildauflösung, die Reduzierung der Videobildwiederholrate oder die Verwendung von gröberen bzw. ungleichen Quantisierungstufen im Frequenzbereich.

Die modernen Komprimierungsverfahren können unter dem Einsatz der Interpixelredundanz- und der Irrelevanzreduktion bei gleichbleibender Videobildauflösung und bei gleichbleibender Videobildwiederholrate Kompressionsraten von über 150 mit einer noch akzeptablen Videoqualität erzielen [71], [74]. Damit wird eine videostreaminggerechte Adaption möglich, die für ein kontinuierliches fehlerfreies Eintreffen von Videobildern im Videoclient sorgt.

3.2.2.3 Komprimierungsstandards

Es existieren viele unterschiedliche Videokomprimierungsstandards und proprietäre Lösungen. Die Gründe sind zum einen historisch zum anderen anwendungsbereichsbezogen. Es lassen sich im Allgemeinen zwei Bereiche identifizieren:

- *Komprimierungsstandards für eine spätere Nachbearbeitung*
- *Komprimierungsstandards für den Endkonsum*

Die Standards unterscheiden sich grundlegend an den geforderten Merkmalen des resultierenden komprimierten Videos. Die für die spätere Nachbearbeitung erzeugten Videos sind aufgrund ihrer enormen Videodatenraten nicht im Videostreamingkontext verwendbar. Des Weiteren ist der Dekomprimierungsaufwand für ein mobiles Gerät zu groß.

Die bekanntesten modernen Komprimierungsstandards für den Endkonsum sind MPEG-4, MPEG-4 AVC⁸ (H.264) und MPEG-H Part 2 (HEVC⁹ bzw. H.265). MPEG-4 AVC soll im

⁸MPEG-4 AVC: MPEG-4 Advanced Video Coding

⁹HEVC: High Efficiency Video Coding

Bezug auf die erzielten Datenraten bei einer gleich bleibenden Videoqualität im Mittel um 50 % effizienter als MPEG-4 sein. MPEG-H Part 2 soll diesbezüglich im Mittel um 25 % bessere Ergebnisse als MPEG-4 AVC erreichen.

Im Rahmen dieser Arbeit wird eine MPEG-4-Komprimierung verwendet. Dies stellt keine Einschränkung dar, da für die Bewertung der Ergebnisse die vom Videocodec erreichte absolute Videoqualität¹⁰ ohne Belang ist.

3.2.2.4 MPEG-4-Videostruktur

Ein nach MPEG-4-komprimiertes Video weist drei Videobildtypen auf:

- *I-Videobilder*
Die *I*-Videobilder werden unabhängig von den anderen Videobildern komprimiert.
- *P-Videobilder*
Ein *P*-Videobild wird in der Abhängigkeit des direkt vorangegangenen *I*-Videobildes bzw. *P*-Videobildes komprimiert.
- *B-Videobilder*
Ein *B*-Videobild wird in der Abhängigkeit des direkt vorangegangenen *I*- bzw. *P*-Videobildes und des nächsten *I*- bzw. *P*-Videobildes komprimiert.

In der Abbildung 3.4 sind die drei Videobildtypen mit ihren Abhängigkeiten, die durch Pfeile angedeutet sind, dargestellt.

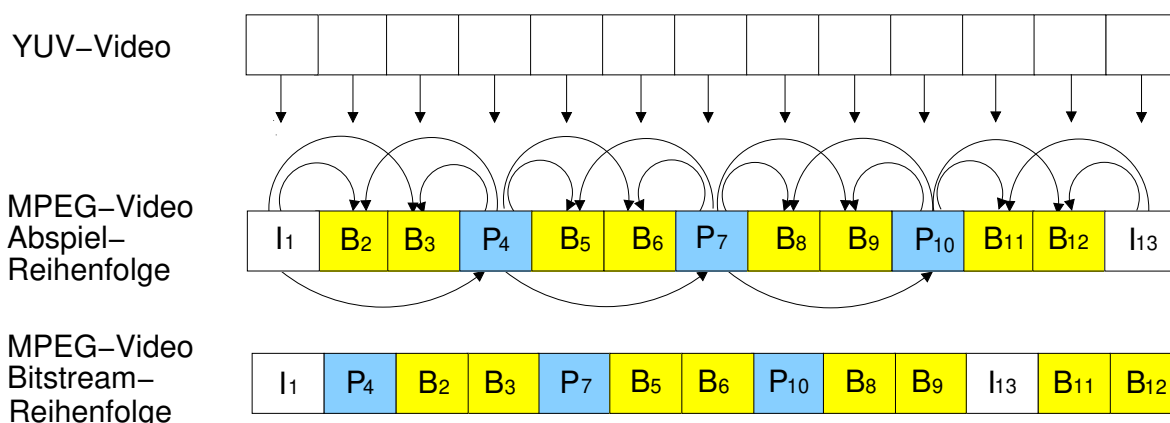


Abbildung 3.4: MPEG-4-Videostruktur

Man unterscheidet nun weiterhin zwischen der Abspielreihenfolge der Videobilder und einer sogenannten Bitstream-Reihenfolge, in der die Videobilder dem Decoder zugeführt bzw. übertragen werden. Damit wird sichergestellt, dass beim Eintreffen der *B*-Videobilder diese sofort dekomprimiert werden können.

¹⁰Für eine vorgegebene Videodatenrate

MPEG-4 AVC und MPEG-H Part 2 erreichen ihre hohe Komprimierungseffizienz u. a. auch durch die Erweiterung der Abhängigkeitsregeln. So können z. B. *P*-Videobilder von mehreren anderen Videobildern abhängen. Diese Erweiterung gilt auch für die *B*-Videobilder, die sogar als Referenzvideobilder fungieren dürfen. Im Rahmen dieser Arbeit wird das klassische Videobildabhängigkeitsmodell nach der Abbildung 3.4 verwendet.

Die Videobildgruppe zwischen zwei *I*-Videobildern inklusive des am Beginn der Gruppe stehenden *I*-Videobildes wird *GoP* (*Group of Pictures*) genannt. So bilden die in der Abbildung 3.4 zu sehenden Videobilder von 1 bis 12 eine GoP der Länge 12.

Eine feste GoP-Länge und die Anzahl der aufeinanderfolgenden *B*-Videobilder sind Parameter eines Videokomprimierungsvorganges¹¹ und beschreiben die Videostruktur eindeutig. Es hat sich eingebürgert, die Videostruktur mit einem Tupel aus den Werten N und M zu beschreiben. N gibt die GoP-Länge an. M gibt den Abstand zwischen zwei Referenzvideobildern an. So beschreibt das Tupel $(N = 12, M = 3)$ die in der Abbildung 3.4 dargestellte Videostruktur.

3.3 Videotransport

Der Transport des Streamingvideos aus der Sicht der Übertragungsprotokolle und des für diese Arbeit relevanten Mobilfunknetzes wurde im Kapitel 2 beschrieben. Es wurde deutlich, dass es auf dem Transportweg zu Videodatenverlusten kommen kann. Deshalb gilt es – nach Möglichkeit – eine anwendungsgerechte Auslieferung der Videostreamingdaten zu ermöglichen.

3.3.1 Arten der Videodatenverluste

Die Videodaten werden in den IP-Paketen übertragen. Die IP-Pakete können im Netz verloren gehen. Des Weiteren werden sie u. U. verzögert an den Videoclient ausgeliefert. Je nach Ausprägung der Anwendung kommen die IP-Pakete zu spät – gemessen am Abspielzeitpunkt – am Videoclient an. Das führt zu Verlusten aufgrund einer zu späten Ankunft. Die Summe der beiden Verluste ergibt die direkten Übertragungsverluste. Die Videobildabhängigkeiten führen zur Folgeverlusten (MPEG-Fehlerfortpflanzung). Dieser Sachverhalt ist in der Abbildung 3.5 dargestellt. Bei der Bewertung der Ergebnisse dieser Arbeit werden alle Verlustarten berücksichtig.

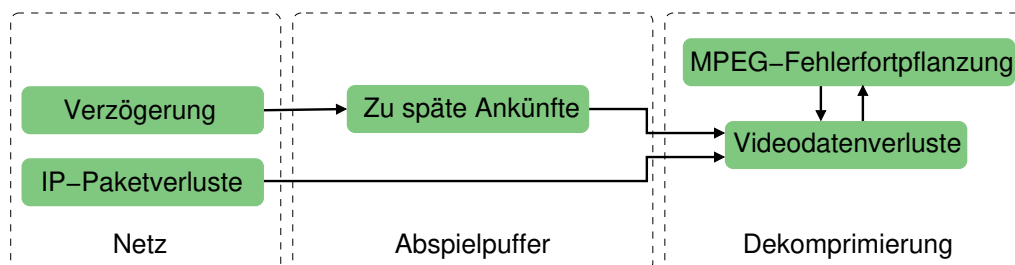


Abbildung 3.5: Videodatenverlustarten

sichtigt.

¹¹Unterschiedlich lange GoP-Längen – angepasst an die Szenenwechsel – werden hier aus den bereits beschriebenen Fehlerresistenzgründen nicht betrachtet.

3.3.2 Maßnahmen zur anwendungsgerechten Videodatenauslieferung

Es lassen sich mehrere relevante Ansätze zur anwendungsgerechten Videodatenauslieferung feststellen:

- *Adaption/Umkomprimierung des zu übertragenden Datenvolumens am Sender*
Nach der Feststellung der Änderung der zur Verfügung stehenden Übertragungsdatenrate wird die Videodatenrate entsprechend angepasst. Die neue Videodatenrate sorgt für eine kontinuierliche Auslieferung der Videobilder bei knappen Übertragungsressourcen. Auf der anderen Seite werden die gestiegenen Übertragungsressourcen zur Erhöhung der Videoqualität – durch die Erhöhung der Videodatenrate - ausgenutzt.
- *Adaption/Umkomprimierung des zu übertragenden Datenvolumens im Netz*
Dieser Ansatz (*Transcoding* genannt) komprimiert das Video entsprechend der im Netz verfügbaren Übertragungsdatenrate in einem Netzelement (*Transcoder* genannt) um.
- *Übertragung eines adaptierbaren Videostroms*
Hier nutzt man eine Komprimierung, die mehrere Ströme aufweist. Werden sie kombiniert, wird eine gute Qualität erzielt. Gehen die Daten in weniger wichtigen Strömen verloren, kann beim fehlerfreien Vorliegen der Daten der Basisqualität, der kontinuierliche Konsum gewährleistet werden.
- *Videostreamingspezifisches Scheduling*
Bei dieser Art von Maßnahmen implementieren die Netzknoten videostreamingspezifische Verfahren der Zuweisung von Übertragungsressourcen.
- *Videostreamingspezifisches AQM (Active Queue Management)*
Bei diesen Verfahren werden in einem IP-Paketpuffer im Falle einer Überlast proaktiv Handlungen an den Videodaten vorgenommen, um die wichtigsten Videodaten – zum Nachteil von weniger wichtigen Videodaten – vor den Verlusten zu schützen.
- *Fehlerkorrigierende Maßnahmen*
Diese Verfahren befassen sich mit der Korrektur der während der Übertragung aufgetretenen Fehler. Sie werden auf verschiedenen Schichten des ISO-OSI-Modells im Kontext des Videostreaming durchgeführt.
- *Disjunkte Wegführung*
Durch die gleichzeitige Auslieferung der Videodaten über die disjunkten Übertragungswege kann die Videoqualität konstant gehalten oder erhöht werden.

Bei der *Adaption/Umkomprimierung des zu übertragenden Datenvolumens* existieren grundsätzlich zwei Verfahrensarten für die Erkennung der Änderung der Übertragungsdatenrate. Die erste Art ist das Ende-Zu-Ende-Verfahren. Es ist sehr verbreitet. Das TCP-basierte MPEG-DASH ist das prominenteste Beispiel [43], [44]. 3GPP PSS (s. Abschnitt 2.5) sieht eine Unterstützung für die Adaptivität bei der UDP-basierten Übertragung vor [3].

Der zweite Ansatz sieht die Einbeziehung der Informationen aus dem Netz vor, um eine schnelle und präzise Feststellung der Änderung der Übertragungsdatenrate durchzuführen. Die vorliegende Arbeit befasst sich mit der netzgesteuerten Adaptivität. Eine ausführliche Literaturübersicht ist im Abschnitt 4.2.1 zu finden.

Die Änderung des zu übertragenden Datenvolumens erfolgt in vielen Fällen durch die Umschaltung (Bitstream-Switching) auf eine andere vorkomprimierte Version des zu konsumierenden Inhaltes [75]. Die andere Variante ist die Umkomprimierung in Echtzeit. Im Rahmen dieser Arbeit wird das Bitstream-Switching verwendet.

Bei der *Adaption/Umkomprimierung des zu übertragenden Datenvolumens im Netz* nutzt man die präzisen Kenntnisse über die verfügbare Übertragungsdatenrate, um die Videodatenrate darauf anzupassen. Beispiele für solche Lösungen sind z. B. in [76] und [77] zu finden. Bei der heutigen Menge an verschiedenen Videokomprimierungsstandards und ihren Ausprägungen bzw. Einstellungsmöglichkeiten scheint die allgemeine Machbarkeit dieser Ansätze schwierig zu sein.

Für die *Übertragung eines adaptierbaren Videostroms* existieren zwei Hauptansätze. Der bekannteste Ansatz heißt *Scalable Video Coding (SVC)* [78]. Das ist eine hierarchische Komprimierung. Ein Videoclient empfängt einen Teilstrom der Basisqualität und einen oder mehrere Teilströme zur Verbesserung der Videoqualität. Der Teilstrom der Basisqualität kann unabhängig von den anderen Teilströmen dekomprimiert werden. Ein Verbesserungsteilstrom kann ohne die Daten im Basisstrom nicht dekomprimiert werden. Gehen bei der Übertragung Daten im Verbesserungsstrom verloren, so beeinflusst dies die Bereitstellung der Basisqualität nicht. Diese Vorgehensweise bietet große Vorteile in der Gruppenrufkommunikation (Multicast). Bei der Individualkommunikation in einem Mobilfunknetz soll überlegt werden, ob es bei schlechten Übertragungsbedingungen nicht besser ist, auf den Verbesserungsteilstrom gänzlich zu verzichten, um den Basisstrom sicher ausliefern zu können. Eine andere Variante wäre die netzseitige Priorisierung des Basisstroms.

Die zweite Hauptausprägung des adaptierbaren Videostromes ist das so genannte *Multiple Description Coding (MDC)* [79]. Auch hier kommen mehrere Ströme zum Einsatz, die allerdings unabhängig voneinander dekomprimiert werden können und eine Basisqualität anbieten. Die gemeinsame Dekomprimierung führt zu einer besseren Videoqualität. Idealerweise überträgt man die Teilströme über disjunkte Wege. Aufgrund einer verminderten Kompressionseffizienz hat das Verfahren bis heute keine große Verbreitung erreicht.

Bei dem *videostreamingspezifischen Scheduling* setzt man Verfahren ein, die die unterschiedliche Wichtigkeit der Videodaten bei der Zuweisung der Übertragungsressourcen berücksichtigen [80], [81], [82]. Manche Verfahren schätzen zusätzlich die Ankunftszeit im Videoclient, um die zu späten Ankünfte zu vermeiden. Die in dieser Hinsicht gefährdeten Videodaten werden bevorzugt behandelt [83]. Bei solchen Verfahren werden gelegentlich komplexe Optimierungsverfahren (sogenannte Rate-Distortion-Verfahren [84]) eingesetzt, die die optimale Zuweisung der Übertragungsressourcen berechnen, um eine anwendungsgerechte Auslieferung von Videodaten für die beste erreichbare Videoqualität zu gewährleisten.

Bei dem *videostreamingspezifischen AQM* werden weniger wichtige Videodaten zugunsten von wichtigeren Videodaten proaktiv in einer Überlastsituation verworfen. Durch die Veränderung der Videodatenverluststruktur wird die Videobildverlustrate gesenkt. Eine ausführliche Literaturübersicht zu diesen Verfahren findet sich im Abschnitt 4.3.1. Die vorliegende Arbeit schlägt Verfahren dieser Art vor und untersucht sie eingehend mit den unterschiedlichen Videostrukturen und anhand einer Auswahl von repräsentativen Videosequenzen.

Bei den *fehlerkorrigierenden Maßnahmen* handelt es sich um die Fehlerkorrektur (FEC), die erneuten Übertragungen, die Maßnahmen zur Erhöhung der Fehlerresistenz beim Komprimieren und um die Fehlerverschleierung bei der Darstellung im Videoclient [85].

FEC-Mechanismen werden bis hoch in die Komprimierungsschicht (Darstellungsschicht) [85] eingesetzt. Bei DVB-IPTV wird FEC auf die RTP-Pakete [6], bei 3GPP (e)MBMS auf die UDP-Pakete [5], bei DVB-IP-Datacast [29] und z. B. in [86] auf die IP-Pakete angewendet. Desweiteren existieren Ansätze für einen unterschiedlich starken FEC-Schutz für unterschiedlich wichtige Videodaten [87]. In [88] wird ein videostreaminggerechter Mechanismus der erneuten Übertragungen auf der Sicherungsschicht des UTRAN vorgeschlagen. In [89] werden die erneuten Übertragungen auf die verloren gegangenen IP-Pakete angewendet. Eine fehlerresistente Maßnahme bei der Komprimierung ist z. B. das bereits beschriebene periodische Einfügen der Synchronisationspunkte (*I*-Videobilder). Für weitere Details sei auf [85] verwiesen. Die Fehlerverschleierungsmechanismen versuchen, den visuellen Einfluss der Übertragungsfehler zu lindern. Es werden dabei die fehlenden bzw. defekten Daten entweder durch die fehlerfrei empfangenen Daten ersetzt, oder es wird interpoliert. Im Rahmen dieser Arbeit wird für ein verloren gegangenes Videobild das letzte korrekt empfangene Videobild verwendet [85].

Bei der *disjunkten Wegeführung* kann man dieselben Videodaten redundant über den jeweiligen Weg versenden. Damit steigt die Wahrscheinlichkeit für die fehlerfreien Ankünfte. Im Falle des bereits beschriebenen MDC werden unterschiedliche Versionen ein und desselben Inhaltes über die verschiedenen Wege versendet. In der Kombination ergibt sich eine bessere Videoqualität als beim Empfang nur eines einzelnen Videostroms. Allerdings ist es nicht einfach bis unmöglich, für die disjunkten Wege in den öffentlichen Netzen zu sorgen. Ganz zu schweigen davon, dass der Videoclient in der Regel mit nur einem Netzanbieter verbunden ist.

3.4 Bewertung der Videoqualität

Es existieren zwei Verfahrensfamilien zur Beurteilung der Videoqualität. Unterschieden wird zwischen den *subjektiven* und den *objektiven* Verfahren.

Bei den subjektiven Verfahren werden die zu bewertenden Videosequenzen von Testpersonen – unter bestimmten vorgegebenen Betrachtungsbedingungen – beurteilt. Die Verfahren sind dementsprechend zeit- und kostenintensiv. Werden die Testbedingungen exakt eingehalten, so liefern diese Verfahren verlässliche Ergebnisse.

Das Ziel der objektiven Verfahren ist, die aufwendigen subjektiven Verfahren durch eine automatisierte rechnergestützte Auswertung zu ersetzen, wobei letztendlich eine Beurteilung ähnlich einer subjektiven Beurteilung durch Testpersonen angestrebt wird. Die objektiv gewonnenen Erkenntnisse werden auf eine subjektive Notenskala abgebildet.

Des Weiteren unterscheiden sich die Verfahren durch die Notwendigkeit einer Referenz für die Beurteilung der Videoqualität. Durch den Einsatz der referenzbasierten Verfahren kann der Einfluss der Übertragungstrecke genauer beurteilt werden. Allerdings steht in der Praxis oft kein Referenzsignal zur Verfügung [15]. Nach verlässlichen objektiven referenzfreien Verfahren wird geforscht, denn sie sind Voraussetzung für die Qualitätsbeurteilung in Netzzwischen-

Tabelle 3.1: Five-Grade-Scale-Notenaufflistung

<i>Five-grade scale</i>			
<i>Quality</i>		<i>Impairment</i>	
5	Excellent	5	Imperceptible
4	Good	4	Perceptible, but not annoying
3	Fair	3	Slightly annoying
2	Poor	2	Annoying
1	Bad	1	Very annoying

elementen auf der Übertragungsstrecke. Ein Beispiel für ein Verfahren dieser Art ist in [15] beschrieben.

3.4.1 Subjektive Messverfahren

Die subjektiven Messverfahren sind von der ITU in [90] standardisiert. Ihre Verwendung im Rahmen dieser Arbeit ist aus den erwähnten Kosten- und Zeitgründen nicht möglich.

Die im selben Standard definierte subjektive Notenskala wird oft in der Literatur für die Abbildung der Auswertungsergebnisse der objektiven Verfahren auf die subjektive Videoqualität eingesetzt. Die Tabelle 3.1 zeigt diese so genannte *Five-Grade-Scale*-Notenaufflistung. Sie wird i. d. R. *Mean Opinion Score (MOS)* genannt. Dies ist eine absolute Notenskala. Es existiert auch eine relative Notenskala, die für vergleichende Bewertungen entwickelt wurde. Für die weiteren Details sei auf [90] verwiesen.

Die MOS-Noten werden in dem in der vorliegenden Arbeit vorgeschlagenen PVU-Diagramm zur Einordnung der Ergebnisse (s. Abschnitt 5.2.2) verwendet.

3.4.2 Objektive Messverfahren

Bei einem objektiven Messverfahren wird das Video automatisch rechnergestützt analysiert und bewertet. Dabei sollen die ggf. auftretenden Artefakte zuverlässig erkannt und entsprechend ihrer Wirkung auf das menschliche Wahrnehmungssystem bewertet werden. Das Ziel ist, ein Ergebnis zu erreichen, das einem mit Hilfe der subjektiven Bewertung erzielten Ergebnis sehr nahe kommt bzw. gleich ist.

Im Folgenden werden die Begriffe Verfahren und Metrik im Bezug auf die objektiven Messverfahren synonym verwendet. Die objektiven Metriken lassen sich in vier allgemeine Kategorien einordnen, wie im Bild 3.6 zu sehen ist [91].

- *Pixelbasierte Metriken*

Für die Berechnung dieser Metriken wird immer eine Referenz benötigt, zu der ein pixelbasierter Vergleich des zu beurteilenden Videos durchgeführt wird. Dafür werden Pixelmerkmale wie Helligkeits- und Chrominanzwerte verwendet.

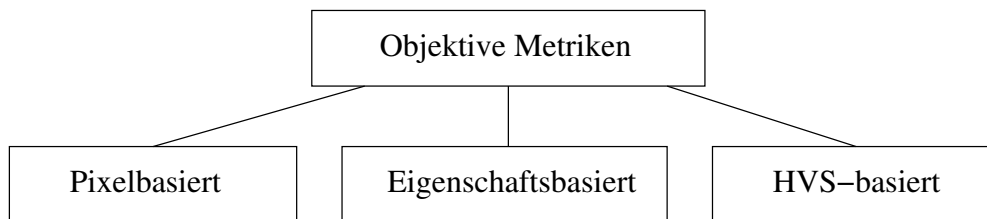


Abbildung 3.6: Objektive Metriken

- *Eigenschaftsbasierte Metriken*

Eigenschaftsbasierte Metriken beurteilen bestimmte Eigenschaften, die für die Wahrnehmung wichtig sind. Das ist z. B. der Kontrast, die Helligkeit, die Schärfe oder bestimmte Komprimierungsartefakte usw.

- *HVS-basierte Metriken*

Diese Metriken versuchen, in einem Modell das menschliche visuelle Wahrnehmungssystem (Human Visual System (HVS)) nachzubilden. Dabei werden z. B. die farbliche Wahrnehmung, die Helligkeits- und die Kontrastbeurteilung eines Menschen modelliert. Darüber hinaus wird die Nichtwahrnehmung bestimmter Bildteile oder -eigenschaften nachgebildet, da die Aufmerksamkeit durch dominierendere Anteile voll in Anspruch genommen wird.

Für die weiteren Details und eine Metrikübersicht sei auf [91] verwiesen.

Im Rahmen dieser Arbeit wird die weit verbreitete pixelbasierte Metrik *Peak Signal-To-Noise Ratio (PSNR)* zur Beurteilung der Videoqualität verwendet. Die Vorteile dieser Metrik liegen in ihrer einfachen Implementierung. Auf der anderen Seite existieren kompliziertere Metriken, die eine bessere Differenzierung bei der Beurteilung der Komprimierungsleistung der Video-codecs erlauben. In der vorliegenden Arbeit stellt der Einsatz von PSNR keine Einschränkung dar, da nicht die Qualität der Komprimierung bewertet wird. Vielmehr geht es um die Bewertung des adaptiven Verhaltens der vorgeschlagenen netzgesteuerten Adaption. Dafür genügt der quantitative Vergleich der PSNR-Werte.

PSNR ist pro Videobild für seine Schwarz-Weiß-Darstellung wie folgt definiert [91], [92]:

$$PSNR = 10 \cdot \log_{10} \left(\frac{S_{max}^2}{MSE} \right)$$

Dabei wird für jedes Videobildpixel der quadratische Fehler zum Referenzbildpixel berechnet, woraus sich dann der mittlere quadratische Fehler (MSE¹²) pro Videobild ergibt. S_{max} steht für den maximalen Signalwert. In den meisten Fällen beträgt $S_{max} = 2^8 - 1$, was für die Acht-Bitdarstellung der Helligkeitswerte (Y im YUV -Format) steht.

Die Tabelle 3.2 zeigt die Abbildung der PSNR-Werte auf die subjektiven MOS-Werte [8], [9].

¹²MSE: Mean Square Error

Bei der anwendungsgerechten Bewertung des Videostreaming bedarf es neben der Beurteilung der Videoqualität auch der Berücksichtigung der Videodatenverluste, der auftretenden Verzögerungen bzw. der zu beobachtenden Verzögerungsschwankungen (s. Abschnitte 2.5.3 und 3.3.1).

Tabelle 3.2: Abbildung von PSNR auf MOS [8], [9].

<i>PSNR, dB</i>	<i>MOS</i>
>37	5 (Excellent)
31 - 37	4 (Good)
25 - 31	3 (Fair)
20 - 25	2 (Poor)
<20	1 (Bad)

3.4.3 Auswahl von Videosequenzen

Die Inhalte von Videos unterscheiden sich naturgemäß. Je nach Videoinhalt kann sich demzufolge die Kompressionsrate für eine konstante vorgegebene Videoqualität unterscheiden. Für eine vorgegebene Videodatenrate kann sich je nach Videoinhalt die Videoqualität unterscheiden. ANSI [93] und ITU [7] definieren quantitative Metriken zur Beschreibung der räumlichen¹³ und zeitlichen¹⁴ Informationen einer Videosequenz. Im Folgenden werden ITU-Definitionen aus [7] verwendet. Die zwei Metriken tragen die Bezeichnungen:

- *Spatial Perceptual Information (SI)*

Zur Berechnung des SI-Wertes einer Videosequenz wird auf jedes Videobild ein Kantendetektor¹⁵ angewendet. Für jedes Videobild erhält man einen Wert, der die Quantität der räumlichen Änderungen beschreibt. Das Maximum dieser Werte wird für den SI-Wert der Videosequenz gewählt.

- *Temporal Perceptual Information (TI)*

Zur Berechnung des TI-Wertes werden pixelweise Differenzbilder zwischen den aufeinanderfolgenden Videobildern berechnet. Für ein Differenzbild erhält man einen Wert, der die Quantität der zeitlichen Änderungen beschreibt. Das Maximum dieser Werte wird für den TI-Wert der Videosequenz gewählt.

Die so berechneten SI- bzw. TI-Werte werden zum Vergleich in einem *SI-TI-Diagramm* dargestellt (s. Abbildung 3.7). In [7] wird darauf hingewiesen, dass bei der Auswahl von Videotestsequenzen möglichst viel Raum des SI-TI-Diagramms abgedeckt werden soll.

In der Abbildung 3.7 sind viele Werte zu sehen, die die in [7] bzw. in [94] standardisierten Videotypen (unterschiedliche Symbole im Diagramm) und Videosequenzen (unterschiedliche Kleinbuchstaben im Diagramm) repräsentieren.

¹³Innerhalb eines Videobildes

¹⁴Änderungen vom Videobild zum Videobild

¹⁵In der Helligkeitsebene

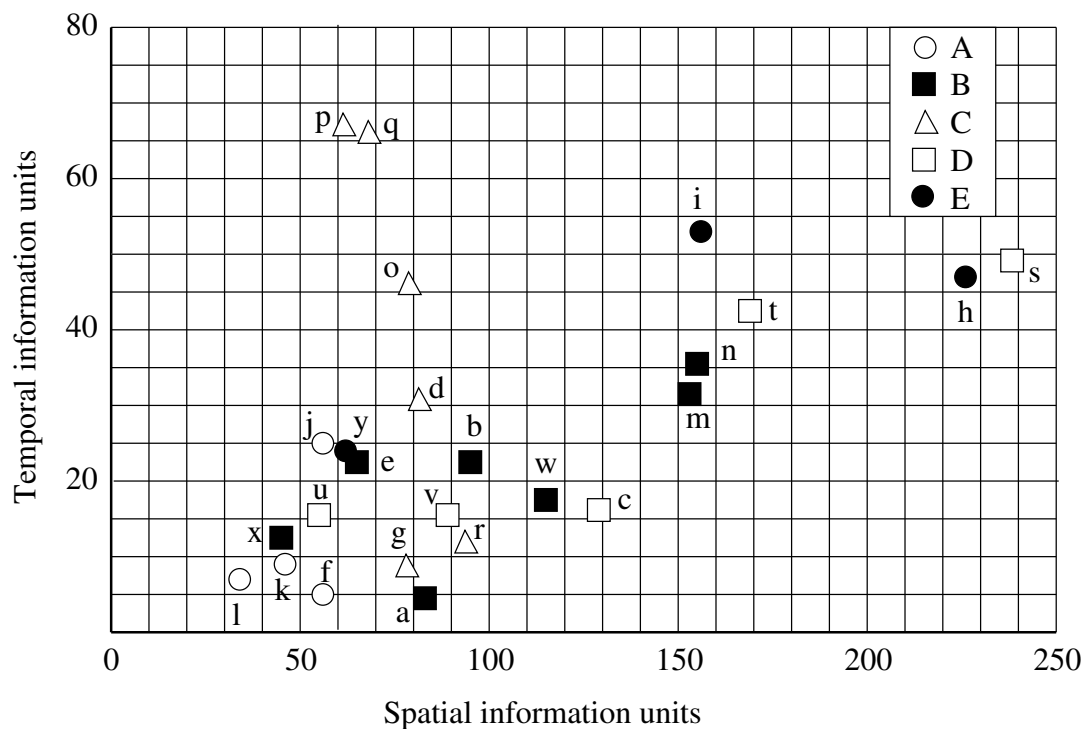


Abbildung 3.7: SI-TI-Diagramm aus [7]

Tabelle 3.3: Zuordnung der Videosequenzen zu Videokategorien (aus [7])

Category	Description	Scene and letter
A	One person, mainly head and shoulder, limited detail and motion	vtc1nw(f), susie(j), disquy(k), disgal(l)
B	One person with graphics and/or more detail	vtc2mp(a), vtc2zm(b), boblec(e), smity1(m), smity2(n), vowels(w), inspec(x)
C	more than one person	3inrow(d), 5row1(g), intros(o), 3twos(p), 2wbord(q), split6(r)
D	Graphics with pointing	washdc(c), cirkits(s), roadmap(t), filter(u), ysmite(v)
E	High object and/or camera motion (examples of broadcast TV)	flogar(h), ftball(i), fedas(y)

Die fünf Videosequenztypen und die dazugehörigen Videosequenzen sind in der Tabelle 3.3 zu sehen. Die Videosequenzen des A-Typs sind typische Videokonferenzsequenzen. Die Videosequenzen des B-Typs würden Nachrichten mit einem Sprecher im Vordergrund oder einem Schulfilm entsprechen. Die Videosequenzen des C-Typs sind Darstellungen von Menschengruppen. Die Videosequenzen des D-Typs entsprechen sehr detailreichen Darstellungen. Die Videosequenzen des E-Typs fassen Sport-, Natur-, Unterhaltungs-, Konzert- usw. Inhalte mit viel Bewegung zusammen.

Für die Verfahrensuntersuchungen im Rahmen dieser Arbeit wird ein repräsentativer Satz der Videosequenzen anhand ihrer Position im SI-TI-Diagramm ausgewählt (s. Abschnitt 6.1).

4 Eingebettete netzgesteuerte Videostreamingadaption

Das heute verbreitete Ende-zu-Ende-Videostreaming passt sich ohne jegliche Netzunterstützung selbst bis zu einem gewissen Grade an die im Netz herrschenden Bedingungen an. Der Vorteil dieser Lösung liegt darin, dass sie in jeder heterogenen Netzkonstellation angewandt werden kann. Die Übertragungsdatenrate muss von Endsystemen geschätzt werden. Dies hat aber seine Grenzen in der Schnelligkeit und Präzision der Schätzung, was sich nachteilig entweder in der Qualität des Videostreaming oder in der Effizienz der Ressourcenausnutzung äußert. Die letztgenannten Nachteile können durch eine netzgesteuerte Adaption beseitigt werden, die gleichzeitig und gemeinsam mit dem Ende-zu-Ende-Adaptionsverfahren wirken kann.

In diesem Kapitel wird der Vorschlag für eine eingebettete netzgesteuerte Videostreamingadaption vorgestellt. Nach der Konzeptvorstellung werden zwei Konzeptkomponenten sowie das Konzept selbst in die bisherigen weltweiten Forschungsergebnisse eingeordnet und detailliert beschrieben. Außerdem wird das Konzept architektonisch begründet.

4.1 Konzept

Es wird vorgeschlagen, die Videodatenratenanpassung einer Videostreamingsitzung mit dem Netzwissen und den Netzmitteln zu steuern. Die Funktionen kann man in einem Videostreaming-Netz-Controller (VSNC) zusammenbringen. Die Bezeichnung ist stellvertretend für das Konzept gewählt und bedeutet im Allgemeinen keine separaten Hardware-Komponenten.

Der VSNC befindet sich in einem Mobilfunkzugangnetz (RAN), wie der Abbildung 4.1 zu entnehmen ist. Er besteht aus zwei Komponenten:

- *Netzgesteuerte Videodatenratenadaption*

Die netzgesteuerte Videodatenratenadaption ist in der Abbildung 4.1 als netzgesteuerte Adaption dargestellt. Diese Komponente hat das Wissen über die Anpassungsfähigkeiten des Videostreamingsservers und idealerweise über die Verarbeitungsfähigkeiten des Videoclients. Aufgrund ihrer genauen Kenntnisse über die momentane Ressourcensituation und ggf. unter Einhaltung von netzspezifischen Optimierungskriterien und anderweitigen

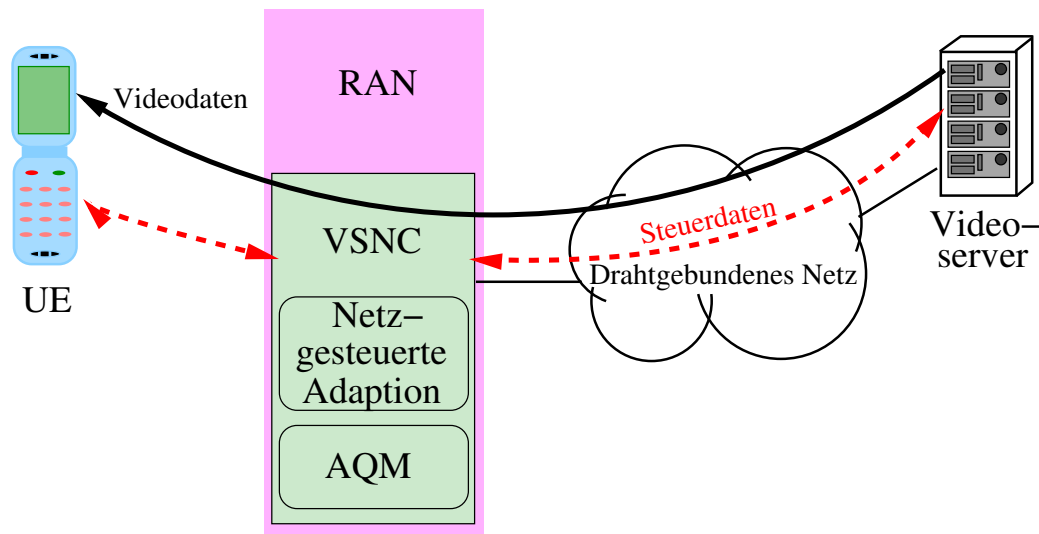


Abbildung 4.1: VSNC in einem RAN

Randbedingungen kann sie Trends präzise und schnell erkennen und eine Videoadaptionsentscheidung in Bezug auf die Anpassung der Videodatenrate treffen. Dem Video-streamingserver wird dies in Form eines *Videoadaptionsvorschlages* mitgeteilt.

- *Active-Queue-Management (AQM)*

Diese Komponente führt das proaktive Verwerfen von Video-IP-Paketen in den Puffern des Mobilfunkzugangsnetzes durch, falls diese aufgrund der Mobilfunkkanalsituation vollzulaufen drohen. Das proaktive Verwerfen findet kontrolliert – mit Kenntnis der unterschiedlichen Wichtigkeiten einzelner Video-IP-Pakete und mit der Fähigkeit, die unterschiedlichen Wichtigkeiten erkennen zu können – statt. Die Komponente agiert ohne Kommunikationskanäle zu den Endsystemen.

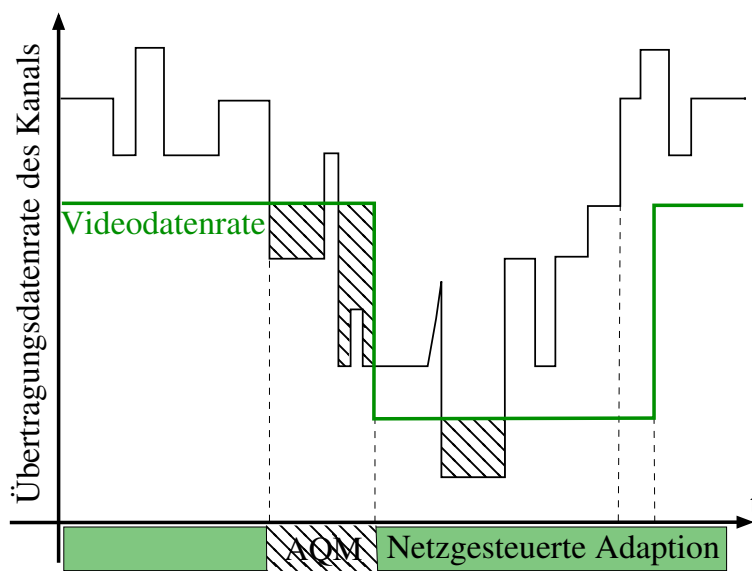


Abbildung 4.2: Veranschaulichung der Einsatzszenarien der beiden VSNC-Komponenten

Die beiden Komponenten agieren unabhängig voneinander und ergänzen sich. Die Einsatzszenarien der beiden VSNC-Komponenten sind in der Abbildung 4.2 veranschaulicht. Die *netzgesteuerte Videodatenratenadaption* (Im Bild *netzgesteuerte Adaption*) steuert die Anpassung der

Videodatenrate an die vom Kanal angebotene Übertragungsdatenrate. Das AQM verwirft die Video-IP-Pakete kontrolliert in den Phasen, in denen die netzgesteuerte Videodatenratenadaption noch nicht greift oder nicht möglich ist. Das Letztere tritt dann ein, wenn keine stärkere Videokomprimierung aufgrund algorithmischer oder organisatorischer Randbedingungen möglich ist.

Im Rahmen dieser Arbeit findet die Implementierung des vorgestellten Konzepts in einem UTRAN in den gemeinsam genutzten Kanälen statt, wie in der Abbildung 4.3 zu sehen ist.

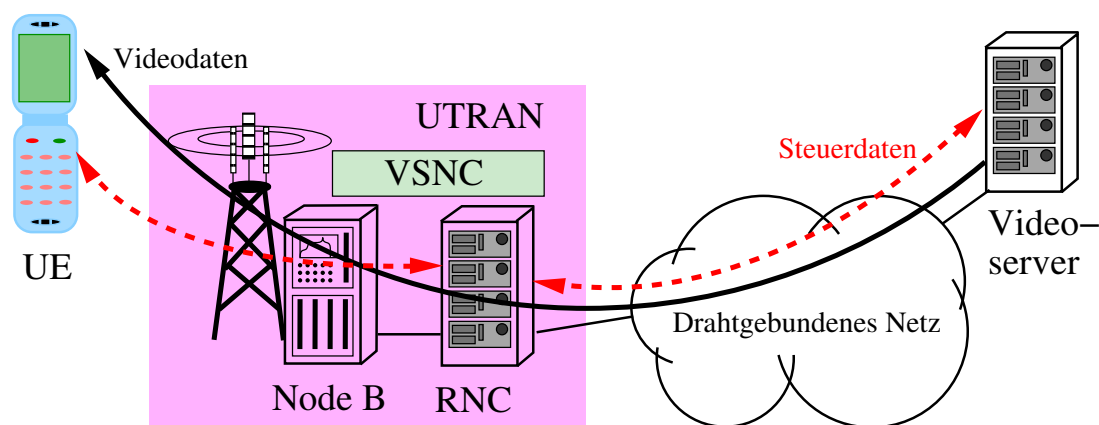


Abbildung 4.3: Implementierung von VSNC in UTRAN

4.2 Netzgesteuerte Datenratenadaption

In diesem Abschnitt wird – nach einer Literaturübersicht – das vorgeschlagene Verfahren eingeordnet und architektonisch begründet. Danach werden die konkreten Implementierungsdetails erklärt.

4.2.1 Literaturübersicht

Es wurden einige Arbeiten auf diesem Gebiet durchgeführt. Die relevanten Entwicklungen sind im Folgenden dargestellt.

Die Autoren von [95] befassen sich mit dem adaptiven IP-basierten Videostreaming, bei dem die Adaption auf Informationen über die Ressourcenauslastung eines drahtgebundenen Netzes basiert. Untersucht wird ein LAN¹ mit mehreren Switches. In der vorgeschlagenen Lösung beobachtet jeder Switch seine Pufferfüllstände fortlaufend. Mehrmals pro Videobildintervall meldet er den absoluten Pufferfüllstand (Anzahl der Video-IP-Pakete) und die Anzahl der weitergeleiteten Video-IP-Pakete seit dem letzten Beobachtungszeitpunkt an die Anwendung. Diese Statistiken werden dem Sender auf zwei alternative Arten mitgeteilt. Einerseits werden sie an die Nutzdaten angehängt und gelangen somit an den Empfänger, der sie dann an den Sender

¹LAN: Local Area Network

zurückschickt. Nur Switches mit schlechteren Statistiken ändern ggf. die bereits an die Video-IP-Pakete angehängten Daten. Die andere Alternative ist eine separate Steuerverbindung vom Switch zum Sender über den Verbindungspfad (die Details sind nicht erklärt). Die Videoadaptionentscheidung findet im Videostreamingserver statt. Der Videostreamingserver muss ggf. eine hohe Zahl von Statistknachrichten verarbeiten. Es wurden mehrere Videos untersucht. Die Zuordnung zu einem SI-TI-Diagramm ist unklar. Die Videoqualitätsbewertung findet PSNR-basiert statt.

In [76] wird vorgeschlagen, einen von dem Videostreamingserver abgesetzten Transcoder zu verwenden, der den Videostrom fortlaufend an die verfügbare Übertragungsdatenrate anpasst. Die Informationen über die verfügbare Übertragungsdatenrate können laut Autoren direkt aus dem Netz bezogen werden. Die Details über die genaue Art, die Häufigkeit und die Auswertung der mitgeteilten Informationen liegen nicht vor. Aus der Beschreibung des Transcodieralgorithmus kann entnommen werden, dass der Transcoder in jedem Videobildintervall eine neue Adaptionentscheidung trifft. Das untersuchte Netz wird nicht beschrieben. Es wurde eine Videosequenz untersucht. Die Videoqualitätsbewertung findet PSNR-basiert statt.

Die Autoren von [96], [97] und [98] schlagen vor, an dem Übergang zwischen dem Mobilfunkzugangszugangnetz und dem drahtgebundenen Netz einen sogenannten Streaming-Agenten einzusetzen. Seine Aufgabe ist u.a., dem Videostreamingserver verbindungsrelevante Statistiken und die Bestätigungsnachrichten für die am Agenten angekommenen Video-IP-Pakete per RTCP mitzuteilen, damit der Server den Zustand des drahtgebundenen Teils der Videostreamingverbindung abschätzt. Das Ziel ist es, dem Videostreamingserver die Unterscheidung zwischen den Stauverlusten im drahtgebundenen Netz und den Paketverlusten im Mobilfunkzugangszugangnetz zu ermöglichen, um die entsprechende Videoadaptionentscheidung zu treffen. Unabhängig davon und gleichzeitig dazu wird eine Ende-zu-Ende-RTCP-Kommunikation verwendet. In [99] und [100] wird das System um die Bestätigungsbenachrichtigungen des Servers für Video-IP-Pakete, die die Basisstation passierten, erweitert. Ein exklusiver Mobilfunkkanal ist der Gegenstand der Betrachtung. Es wurden eine bzw. zwei Videosequenzen untersucht. Die Videoqualitätsbewertung findet PSNR-basiert statt.

In [101] wird das Videostreaming über einen exklusiven Kanal im Mobilfunkzugangszugangnetz (UTRAN) behandelt. Es wird vorgeschlagen, die momentan diesem exklusiven Kanal zugewiesene Übertragungsdatenrate dem Videostreamingserver direkt ohne eine Vorverarbeitung mitzuteilen. Die Videoadaptionentscheidung findet im Videostreamingserver statt. In diesem Artikel passt der Server umgehend seine Videodatenrate an die momentan zugewiesene Übertragungsdatenrate im exklusiven Mobilfunkkanal an. Ein exklusiver Kanal ist viel weniger zeitvariant als ein gemeinsam genutzter Kanal. Es wird eine Videosequenz betrachtet. Es findet keine direkte Videoqualitätsbewertung statt. Es werden die übertragenen Datenmengen verglichen bzw. bewertet.

Die Autoren von [102] und [103] schlagen allgemein vor, einem Videostreamingserver die momentanen Werte der Netzauslastung für die aktuellen Verbindungen mitzuteilen. Es wird von einer Netzheterogenität im Netzzugangsbereich ausgegangen (drahtgebunden und drahtlos). Die Videoadaptionentscheidung wird im Videostreamingserver getroffen. Eine konkrete Implementierung in einem konkreten technischen Kontext und ihre Bewertung liegen nicht vor.

In [104] wird in einem Mobilfunkszenario (UTRAN-ähnlich²) vorgeschlagen, die Management-Information-Base-Daten der Basisstationen für eine Adaptionentscheidung im Videostreamingserver zu verwenden. Die Daten werden periodisch ausgelesen und an den Videostreamingserver gesendet. Gleichzeitig findet eine RTCP-basierte Ende-zu-Ende-Schätzung der Übertragungsbedingungen statt. Der Videostreamingserver kombiniert die Daten, nimmt die Schätzung vor und trifft die Adaptionentscheidung. Die verwendete Videosequenz wird nicht beschrieben. Eine Videoqualitätsbewertung findet nicht statt. Es werden Video-IP-Paketverluste bewertet.

In [105] und in [106] wird vorgeschlagen, in einem Mobilfunkzugangnetz (UTRAN) einen TCP-Proxyserver zu verwenden, um die Kommunikation der Mobilfunkclients mit dem Internet optimal zu gestalten. Der Proxyserver nutzt für die Festlegung der TCP-Übertragungsparameter u. a. die augenblicklichen Attributwerte der Mobilfunkübertragungskanäle (wie die momentane Übertragungsdatenrate und die momentanen Pufferfüllstände). Es ist keine videostreamingspezifische Lösung.

In [107] wird vorgeschlagen, u. a. Multilevel-ECN³ für die Videoadaptionentscheidung in einem Ende-zu-Ende-Adaptionsszenario mit zu verwenden. Es werden drei unterschiedlich starke Staustufen definiert bzw. gemeldet. Die MECN-Informationen gelangen zuerst zum Client, der sie dann an den Videostreamingserver zur weiteren Verwendung sendet. Es wird ein Spielfilm als Untersuchungsgegenstand in einem drahtgebundenen Netzszenario verwendet. Es findet keine Videoqualitätsbewertung statt. Die Bewertung findet anhand der Pufferfüllstände und der Übertragungsdatenraten statt.

In [108] wird ein drahtgebundenes aktives Netz betrachtet. Alle Knoten auf dem Videoverbindungsweg messen ihre Übertragungsdatenraten und melden diese dem am nächsten zum Videostreamingserver stehenden aktiven Netzknoten. Dieser wiederum informiert periodisch (sekündlich) den Videostreamingserver über die momentan verfügbare Übertragungsrate auf dem Verbindungspfad. Die Videoadaptionentscheidung wird im Video Streaming-Server getroffen. Es wurde eine Videosequenz untersucht. Die Videoqualitätsbewertung findet PSNR-basiert statt.

Der Artikel [109] stammt vom Autor der hier vorliegenden Arbeit. Das im Artikel vorgestellte Verfahren und die durchgeführten Untersuchungen bilden die Grundlagen für die weiteren Untersuchungen im Rahmen dieser Arbeit und werden deshalb in den entsprechenden Abschnitten ausführlich beschrieben.

Die Autoren von [86] schlagen vor, die ECN-Meldungen für die Anpassung der Videodatenrate zu verwenden. Zusätzlich werden die Video-IP-Pakete mit einem FEC-Schutz versehen, dessen Stärke von den ECN-Meldungen abhängig ist. Die ECN-Informationen gelangen zuerst zum Client, der sie dann an den Videostreamingserver zur weiteren Verwendung sendet. Es wird ein drahtgebundenes Szenario untersucht. Zwei Videosequenzen werden für die Untersuchung verwendet. Die Videoqualitätsbewertung findet PSNR-basiert statt.

In [77] wird im Kontext eines LTE-Mobilfunkzugangnetzes (EUTRAN) vorgeschlagen, einen Transcoder zu verwenden, der das zu übertragende Video gemäß der Mobilfunkkanalsituation

²Gemeint ist ein RAN von IMT-2000.

³MECN: Multilevel Explicit Congestion Notification

transcodiert. Der Transcoder ist ein Teil des EUTRAN und kann somit auf alle Informationen des EUTRAN zugreifen. Bei dieser Lösung muss EUTRAN viele Ressourcen aufwenden, um zahlreiche mobile Videostreamingnutzer zu bedienen. Eine weitere Schwierigkeit ist, dass der Transcoder mit den zahlreichen Videokomprimierungsformaten und ihren Ausprägungen umgehen muss. Des Weiteren kann der Transcoder die Videodatenrate nicht erhöhen. Es wurde nur eine Videosequenz untersucht. Die Videoqualitätsbewertung findet SSIM⁴-basiert statt.

In [110] wird ein Routingalgorithmus für das Videostreaming in mobilen Ad-Hoc-Netzen (MANET) vorgestellt. Der Algorithmus berechnet – neben den Routen – die Übertragungsdatenraten auf einem Pfad und teilt diese in regelmäßigen Abständen dem Videostreamingserver mit. Die Adaptionentscheidung wird im Videostreamingserver vorgenommen. Untersucht wird eine WLAN⁵-Umgebung. Es wird eine Videosequenz verwendet. Die Videoqualitätsbewertung findet PSNR-basiert statt.

4.2.2 Diskussion und Einordnung

In der Literaturübersicht wurde deutlich, dass alle vorgeschlagenen Lösungsansätze für das netzgesteuerte Videostreaming fortlaufend Daten (Nachrichten), die die Ressourcenauslastung beschreiben, an den Videostreamingserver senden. Die Dateninterpretation, -aufbereitung und die Trenderkennung muss der Videostreamingserver selbst durchführen. Dies ist insbesondere bei der Heterogenität der Netze schwierig.

Bei einer größeren Anzahl der Videoclients kann die Häufigkeit der Nachrichten zu Skalierungsproblemen im Videostreamingserver führen. In den gemeinsam genutzten Mobilfunkkanälen von UTRAN und EUTRAN betragen die kleinsten Transportintervalle in einem Transportkanal 2ms bzw. 1ms. D. h., die Übertragungsdatenrate zwischen der Sicherungsschicht und der Bitübertragungsschicht wird bis zu 500 Mal bzw. 1000 Mal pro Sekunde neu bestimmt, wobei sie starken Amplitudenschwankungen unterliegen kann. Dies kann zu einer sehr großen Häufigkeit von Nachrichten führen, der ein Videostreamingserver zeitnah zu verarbeiten hat.

Des Weiteren ist das Netz der alleinige Übertragungsressourcenverwalter. Obwohl die augenblickliche Übertragungsdatenrate ausreichend für eine Erhöhung der Videodatenrate sein mag, kann das Netz aus anderen Gründen (technisch, optimierungstechnisch, organisatorisch/moneitär) gegen eine Adaption, die zur stärkeren Ressourcenbeanspruchung führt, sein. Es sind bereits heute sogenannte *Intent Based Networks (IBN)* Realität geworden. Dies ist die Fortentwicklung von SDN⁶-Netzen, die von Netz-Controllern gesteuert werden. Eine der IBN-Funktionen ist die sogenannte Assurance, die unter anderem proaktiv Netzprobleme erkennt, beseitigt und Leistungsprognosen für die nächste Zukunft erstellt. Die Implementierungsdetails solcher Netze können zunehmend hinter Kommunikationsschnittstellen verborgen werden, so dass fortlaufende ungefilterte Nachrichten über Ressourcenauslastung der Netzknoten (an den Videostreamingserver gerichtet) u. U. gar nicht möglich sein werden.

Im Rahmen dieser Arbeit wird ein Ansatz vorgestellt, bei dem das Netz einen Videoadaptionsvorschlag bestimmt und an den Videostreamingserver sendet. Das geschieht als eine Teilfunk-

⁴SSIM: Structural Similarity

⁵WLAN: Wireless LAN

⁶SDN: Software Defined Networking

tion eines Videostreaming-Netz-Controllers. Für die Ermittlung des Videoadaptionsvorschlags wird u. a. die Kenntnis der Anpassungsfähigkeit des Videostreamingsservers und idealerweise der Verarbeitungsfähigkeit des Videoclients verwendet. Der Ansatz ist darin begründet, dass nur der Ressourcenverwalter (das Netz) genaue Kenntnisse über die momentane Situation und vorhandene Randbedingungen besitzt und Trends präzise und schnell erkennen kann. Wenn zwischen dem Videoclient und dem Videostreamingserver mehrere Netze liegen, die dieses Konzept implementieren, so würde der Videostreamingserver ggf. mehrere Videoadaptionsvorschläge erhalten. Er würde dann den Vorschlag mit der kleinsten Ressourcenbeanspruchung auswählen. Die Lösung ist präzise und schnell bei ihren Adaptionentscheidungen. Des Weiteren werden eventuelle Skalierungsprobleme im Videostreamingserver vermieden, die von Verarbeitung einer großen Anzahl der Netznachrichten verursacht werden können.

Es seien noch die Transcoder-Lösungsansätze erwähnt. Sie beanspruchen viele Ressourcen in einem RAN, um zahlreiche mobile Videostreamingnutzer zu bedienen. Eine weitere Schwierigkeit ist, dass der Transcoder mit zahlreichen Videokomprimierungsformaten und ihren Ausprägungen umgehen muss. Aus architektonischer Sicht sollte der Videostreamingserver selbst die Daten verarbeiten bzw. ändern, die er versendet. Die Transcoder-Lösungsansätze können die am Eingang des Transcoder anliegende Videodatenrate prinzipbedingt nur senken.

Der in dieser Arbeit beschriebene Lösungsansatz wurde in [109] publiziert.

4.2.3 Implementierung

Die Implementierung der netzgesteuerten Datenratenadaption als Teil von VSNC ist der Abbildung 4.4 zu entnehmen. *RLC*⁷ und *MAC-hs*⁸ sind Komponenten der Sicherungsschicht.

Die Adaptionskomponente des VSNC trifft eine Videoadaptionsentscheidung. Dafür muss sie u. a. das Wissen über die Anpassungsfähigkeiten des Videostreamingsservers und idealerweise über die Verarbeitungsfähigkeiten des Videoclients besitzen. Die dafür notwendigen Protokollerweiterungsvorschläge sind nicht Teil dieser Arbeit. Die Machbarkeit und die konkreten Realisierungen derartiger Erweiterungen sind in manchen der im Abschnitt 4.2.1 betrachteten Publikationen aufgezeigt worden. Im Rahmen dieser Arbeit besitzt die Adaptionskomponente die genaue Kenntnis über die vom Videostreamingserver unterstützten diskreten Videodatenraten. Des Weiteren wird hier davon ausgegangen, dass der Videoclient alle vom Videostreamingserver angebotenen Videodatenraten verarbeiten kann.

Die Datenratenadaptionskomponente weist vier Funktionsblöcke auf:

- *Beobachtung*

Dieser Funktionsblock beobachtet die verfügbaren Ressourcen fortlaufend. Je nach Optimierungszielen und der Algorithmusausprägung für die Schätzung und die Entscheidung kann der Beobachtungsvektor z. B. aus den momentanen Werten der Übertragungsdatenrate auf der Sicherungsschicht, der Anzahl der Nutzer des gemeinsamen Kanals, der verfügbaren Funkleistung, dem Pufferfüllstand des IP-Paketpuffers usw. bestehen. Des

⁷RLC: Radio Link Control

⁸MAC-hs: Medium Access Control for High Speed Downlink Shared Channel

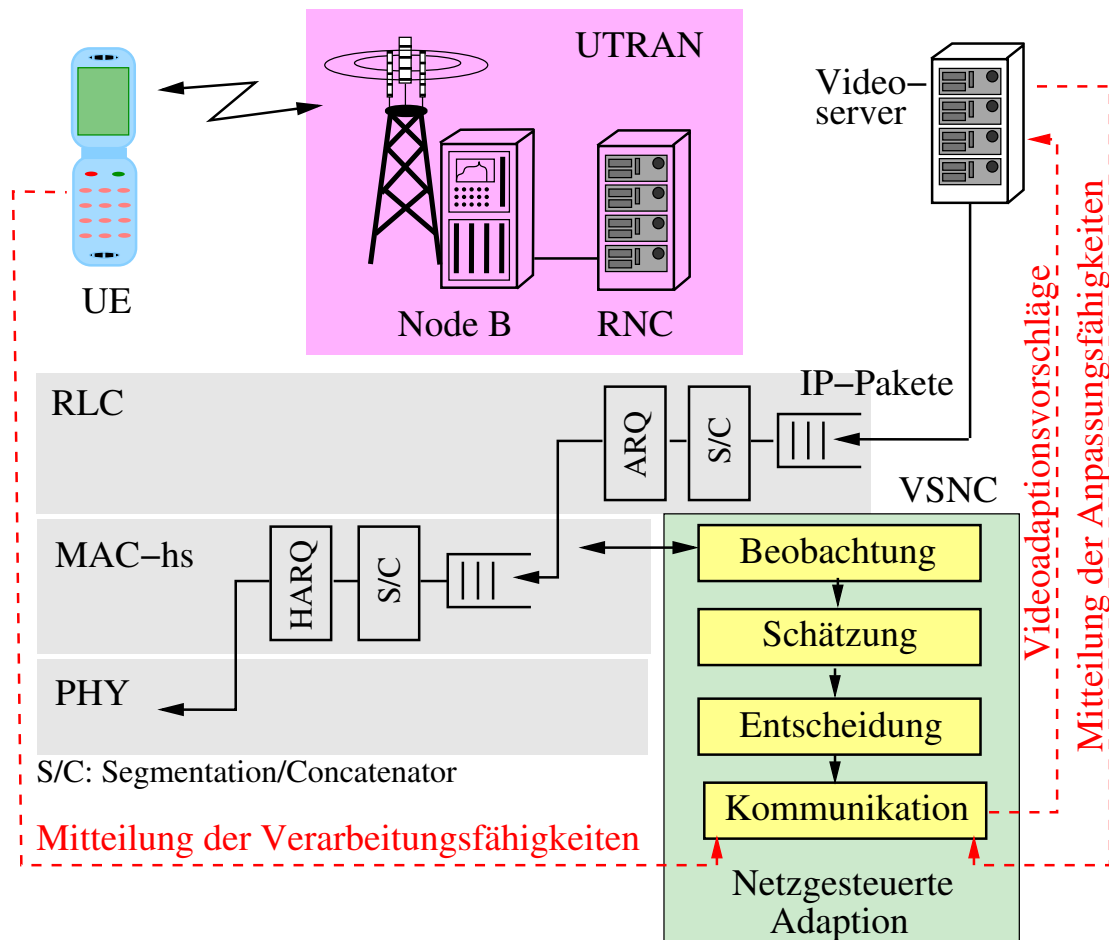


Abbildung 4.4: Netzgesteuerte Datenratenadaption in UTRAN

Weiteren könnte der Funktionsblock zusätzlich den drahtgebundenen Teil der Verbindung per RTP/RTCP überwachen. Die im Rahmen dieser Arbeit implementierte Ausprägung dieses Funktionsblocks wird weiter unten erklärt.

- *Schätzung*

Die Aufgabe dieses Funktionsblockes ist es, auf der Basis der verfügbaren Daten eine Prognose für die quantitative Ressourcensituation für die kurzfristige Zukunft zu treffen. Dies dient als Entscheidungsgrundlage für die einzuleitenden Maßnahmen. Die im Rahmen dieser Arbeit implementierte Ausprägung dieses Funktionsblocks wird weiter unten erklärt.

- *Entscheidung und Festlegung von Maßnahmen*

In diesem Block wird die Entscheidung für die Videoadaption getroffen. Die Entscheidungsfindung hat drei grundsätzliche Randbedingungen: die Adaptionfähigkeiten des Videostreamingserver, die Verarbeitungsfähigkeiten des Videoclients, die netzseitigen Gegebenheiten und Optimierungsziele. Mögliche Maßnahmen können ein Videodatenratenadaptionvorschlag für den Videostreamingserver, das gezielte Einschalten verschiedener AQM-Mechanismen mit z. B. einem Verdrängen des konkurrierenden Verkehrs, Schedulingmaßnahmen, eine zellweite Optimierung usw. sein. Die im Rahmen dieser Arbeit implementierte Ausprägung dieses Funktionsblocks wird weiter unten erklärt.

- *Kommunikation*

Der Kommunikationsblock kommuniziert mit den Systemen außerhalb von VSNC. So wird darüber z. B. der Videoadaptionsvorschlag an den Videostreamingserver mitgeteilt. Dem VSNC können über diesen Funktionsblock die Fähigkeiten der Endsysteme mitgeteilt werden. Dazu gehören die vom Videostreamingserver unterstützten Videodatenraten, die Verarbeitungsfähigkeiten und die Größe des Abspielpuffers des Videoclients. Die im Rahmen dieser Arbeit implementierte Ausprägung dieses Funktionsblocks wird weiter unten erklärt.

Die bei den gemeinsam genutzten Mobilfunkkanälen eingesetzte adaptive Modulation und adaptive Kanalkodierung haben ständige Änderungen der Übertragungsdatenrate auf der Sicherungsschicht als Reaktion auf die zeitliche Mobilfunkkanalvarianz zur Folge. Die in dieser Arbeit implementierte *Beobachtung* fragt bei jedem Scheduling-Intervall das momentan zugewiesene Transportformat in der Sicherungsschicht⁹ ab, und entnimmt diesem die augenblickliche Übertragungsdatenrate¹⁰.

Die hier implementierte *Schätzung* bestimmt die für die Videostreamingverbindung vorhandene Übertragungsdatenrate aus den Werten der augenblicklichen Übertragungsdatenraten (in 2ms-Intervallen) als eine kurzfristige Prognose. Empirisch und basierend auf den Untersuchungen in [111] hat sich herausgestellt, dass die exponentielle Glättung sehr gute Schätzergebnisse bringt. Der Ansatz des Verfahrens basiert auf der Annahme, dass der aktuelle Wert von den vorangegangenen Werten beeinflusst wird. Zum einen werden dabei die Spitzen geglättet. Zum anderen aber macht sich der Vorteil der schnellen Reaktion auf die Änderungen der Werte durch eine zügige Adaptionentscheidung positiv bemerkbar. Der aktuelle Schätzwert (e_i) der Übertragungsdatenrate wird mit dem vorangegangenen Schätzwert (e_{i-1}) und dem aktuellen Messwert (m_i) der augenblicklichen Übertragungsdatenrate wie folgt ermittelt:

$$e_i = (1 - A) \cdot e_{i-1} + A \cdot m_i$$

Die Bestimmung des passenden Wertes des Parameters A ist nicht trivial. Im Rahmen dieser Arbeit wurde $A = 0,04$ empirisch und mit Hilfe von [112] und [113] als eine passende Parametrierung bestimmt.

Die Implementierung des Funktionsblocks *Entscheidung und Festlegung von Maßnahmen* realisiert die Bestimmung eines Videoadaptionsvorschlags für den Videostreamingserver. Die Entscheidung über einen Videoadaptionsvorschlag hat die folgenden Randbedingungen:

⁹Genau gesagt in der MAC-hs-Schicht.

¹⁰Da der Kanal gemeinsam ist, soll die Anzahl der Mobilfunkteilnehmer in der Zelle (n_u) und die verfügbare Funkleistung (als Gewichtung k_{pow} , die eine Funktion von (n_u) ist) bei der Bestimmung der für die Videostreamingverbindung zur Verfügung stehenden Übertragungsdatenrate (r_u) anhand der gemeinsamen augenblicklichen Übertragungsdatenrate (r_s) verwendet werden:

$$r_u = k_{pow} \cdot r_s$$

Allerdings ist bei wenigen Nutzern in der Zelle im hier eingesetzten, simulierten UTRAN-Modell keine Korrektur notwendig.

- Die prognostizierte Übertragungsdatenrate in UTRAN (r_e)
- Die verfügbaren diskreten Videodatenraten des Videostreamingsservers (vr_{avail_i})
- Die aktuell verwendete Videodatenrate $vr_{current}$
- Die maximale vom Videoclient dekodierbare Videodatenrate (vr_{maxDec})
- Die Vermeidung von zu häufigen Umschaltvorgängen zwischen den unterschiedlichen Qualitätsstufen. Ein zu häufiges Umschalten kann als störend empfunden werden. Für die Vermeidung der zu häufigen Umschaltvorgänge wurde ein Hystereseverfahren mittels eines Faktors (k_{hys}) angewendet.

Die Videoadaptionsentscheidung wird wie folgt getroffen:

- *Heraufschalten*

Es wird auf die nächsthöhere verfügbare Videodatenrate vr_{avail_i} heraufgeschaltet, wenn die folgende Bedingung wahr ist:

$$(vr_{current} < vr_{avail_i}) \wedge (r_e \geq (1 + k_{hys}) \cdot vr_{avail_i}) \wedge (vr_{avail_i} \leq vr_{maxDec})$$

In den vorliegenden Untersuchungen wird – wie bereits erwähnt – davon ausgegangen, dass der Videoclient alle vom Videostreamingserver angebotenen Videodatenraten verarbeiten kann.

- *Herunterschalten*

Es wird auf die nächstkleinere Videodatenrate vr_{avail_i} heruntergeschaltet, wenn die folgende Bedingung wahr ist:

$$r_e \leq (1 - k_{hys}) \cdot vr_{current}$$

Die Fairness der Verkehrsströme untereinander muss nicht betrachtet werden, da das Mobilfunkzugangnetz durch das Scheduling für die faire Zuordnung von Übertragungsressourcen an die einzelnen Verbindungen sorgt [114].

Der Funktionsblock *Kommunikation* implementiert in dieser Untersuchung die Benachrichtigung des Videostreamingsservers über den ermittelten Videoadaptionsvorschlag.

4.3 AQM

In diesem Abschnitt wird – nach einer Literaturübersicht – das vorgeschlagene Verfahren eingeordnet. Danach werden die konkreten Implementierungsdetails erklärt.

4.3.1 Literaturübersicht

Es wurden in den letzten drei Jahrzehnten diverse Arbeiten auf diesem Gebiet durchgeführt. Die relevanten Entwicklungen sind im Folgenden dargestellt. Die bei der Einordnung benutzten Begriffe *Videostruktur* bzw. *Videostrukturvielfalt* beziehen sich auf die GoP-Struktur (s. Abschnitt 3.2.2.4). Die Begriffe *Videotyp* bzw. *Videotypvielfalt* beziehen sich auf die im Abschnitt 3.4.3 beschriebenen Videotypen.

Bereits 1993 wird ein kontrolliertes Verwerfen von Video-IP-Paketen beschrieben [95]. Bei einem vollen Puffer eines Switches wird beim Eintreffen eines wichtigen Video-IP-Pakets ein weniger wichtiges Paket aus dem Puffer entfernt. Man definiert zwei Wichtigkeitsstufen: wichtig (Video-IP-Pakete der *I*- und *P*-Videobilder) und unwichtig (Video-IP-Pakete der *B*-Videobilder). Die unterschiedlichen Wichtigkeiten der *P*-Videobilder werden nicht beachtet. Weitere Details sind unbekannt. Die Video-IP-Paket- und Videobildabhängigkeiten werden in der Algorithmusbeschreibung nicht berücksichtigt. Es wird ein drahtgebundenes Netz untersucht. Es wurden mehrere Videos untersucht. Die Zuordnung der Videos zum SI-TI-Diagramm ist unklar. Die untersuchte Videostruktur ist nicht beschrieben. Die Videoqualitätsbewertung findet PSNR-basiert statt.

In [115] wird vorgeschlagen, Video-IP-Pakete einer Videostreamingverbindung in einem Puffer kontrolliert und abhängig von ihrer Wichtigkeit zu verwerfen. Die Wichtigkeit orientiert sich am Videobildtyp. Die unterschiedliche Wichtigkeit der *P*-Videobilder wird nicht beachtet. Der Autor untersuchte ein UDP-basiertes Videokonferenzsystem in einem LAN, dessen Endsysteme kleine Puffer als Konsequenz der Echtzeitanforderung haben. Das Puffermanagement wurde im Empfänger realisiert, um den aufgrund der Verzögerungsschwankung auftretenden Pufferüberlauf kontrolliert zu behandeln. Zusätzlich zum Puffermanagement wurden *I*-Videobilder aus Redundanzgründen doppelt übertragen. Die Video-IP-Paket- und die Videobildabhängigkeiten wurden in der Algorithmusbeschreibung nicht erwähnt. Die Ergebnisse sind allgemeiner qualitativer Natur. Die Videoqualität wurde subjektiv bewertet. Die Videostrukturvielfalt wurde nicht berücksichtigt. Die untersuchte Videostruktur ist ($N = 15, M = 3$). Die Videotypvielfalt – es wurde nur ein Video verwendet – wurde nicht berücksichtigt.

In [116] wird vorgeschlagen, die Videobilder im Sendepuffer eines Videostreamingsservers kontrolliert zu verwerfen, wenn abzusehen ist, dass sie – aufgrund einer Stausituation im Netz – nicht rechtzeitig am Client ankommen würden. Es wird ein CBR¹¹-Netzservice-Szenario untersucht. Ein Stau in so einem Netz kann entstehen, wenn die augenblickliche Videodatenrate schwankt und die Übertragungsdatenrate des Netzes übersteigt. Es werden die Übertragungsressourcen des Netzes und die Abspieelpuffergröße des Empfängers für den Algorithmus als bekannt vorausgesetzt. Es wird ein JPEG-basierter Videostrom verwendet. In [117] erweiterten die Autoren den Algorithmus auf MPEG. Das kontrollierte Verwerfen – je nach Videobildtyp – berücksichtigt die Videobildabhängigkeiten und die Video-IP-Paketabhängigkeiten. Die unterschiedlichen Wichtigkeiten von *P*-Videobildern werden berücksichtigt. Die Autoren verwenden nach ihren Angaben repräsentative Videosequenzen (Spielfilme). Die Zuordnung der untersuchten Videos zu dem SI-TI-Diagramm ist unklar. Die Zuordnung der einzelnen Videotypen zu den erzielten Ergebnissen ist nicht möglich. Die untersuchten Videostrukturtypen sind unklar. Die Ergebnisse sind nicht in Form von Videoqualität dargestellt, sondern beziehen sich auf den ob-

¹¹CBR: Constant Bit Rate

jektiven Algorithmenvergleich (z. B. die Anzahl der verworfenen Videobilder in Abhängigkeit von den untersuchten Algorithmen).

In [118] wird vorgeschlagen, einen sogenannten MPEG-Filter im Netzzugangsbereich (hier DSL) anzubringen. Der Videoclient misst fortlaufend die Video-IP-Paketverlustrate und fordert den Filter ggf. auf, die Videodatenrate zu reduzieren (bzw. zu erhöhen). Die Reduktion geschieht durch das videobildtypabhängige Verwerfen von Videobildern, deren Typ der Filter durch das Dekodieren feststellt. Die Videobildabhängigkeiten werden erwähnt. Der Algorithmus scheint so implementiert zu sein, dass ihre explizite Behandlung nicht notwendig ist. Eine genaue formale Beschreibung ist aber nicht vorhanden. Die unterschiedlichen *P*-Videobildwichtigkeiten werden beim Verwerfen berücksichtigt. Die Untersuchung findet anhand eines nicht weiter spezifizierten Videos statt. Implizit lässt sich erkennen, dass die GoP-Größe 25 Videobilder beträgt. Die Videotyp- und Videostrukturvielfalt sind nicht gegeben. Die Bewertung bezieht sich auf die erzielten Reduzierungen der Videodatenrate. Die erzielte Videoqualität wird subjektiv besser als die Videoqualität ohne Maßnahmen eingeschätzt.

In [119], [120] und in [80] werden Video-IP-Pakete je nach ihrer Wichtigkeit unterschiedlichen Scheduling-Queues eines DiffServ¹²-unterstützenden Routers zugewiesen. Das Hauptaugenmerk liegt auf der Berechnung der Video-IP-Paketwichtigkeiten (basierend auf verschiedenen Merkmalen eines MPEG-komprimierten Videos) und der Abbildung der Wichtigkeit auf die Netz-QoS¹³-Klassen. Es werden keine Video-IP-Pakete proaktiv verworfen. Bei Video-IP-Paketverlusten werden Abhängigkeiten nicht proaktiv behandelt. Das Aufbrechen einer Videoverbindung kann zur falschen Video-IP-Paket- bzw. Videobildreihenfolge führen. Die Untersuchungen beziehen sich auf ein drahtgebundenes Netz. Die Videotypvielfalt ist nicht gegeben. Die Videostrukturvielfalt ist nicht gegeben. Die Videoqualität wird PSNR-basiert geschätzt.

In [121] erweitern die Autoren von [119] ihr Konzept um das Feedback vom Client, um die Schätzung der Video-IP-Paketwichtigkeiten zu verbessern. Die Untersuchungen werden mit drei verschiedenen Videosequenzen durchgeführt. Es ist nicht klar, ob die Videotypvielfalt gegeben ist. Die Videostrukturvielfalt ist nicht gegeben. Die Videoqualität wird PSNR-basiert geschätzt.

In [81] werden die Video-IP-Pakete je nach ihrer Wichtigkeit unterschiedlichen Scheduling-Queues eines DiffServ-unterstützenden Routers zugewiesen. Die Wichtigkeit wird durch die Zugehörigkeit zu einer Komprimierungsschicht (Teilstrom) eines hierarchisch komprimierten Videos bestimmt. Es werden drei Komprimierungsschichten betrachtet. Es werden keine Video-IP-Pakete proaktiv verworfen. Bei Video-IP-Paketverlusten werden Abhängigkeiten nicht proaktiv behandelt. Die Untersuchungen beziehen sich auf ein drahtgebundenes Netz. Als Video wird ein Spielfilm verwendet. Es ist nicht klar, ob die Videovielfalt gegeben ist. Es erfolgt keine Schätzung der Videoqualität. Es werden Paketverluste und Verzögerungen bewertet.

In [82] werden die Video-IP-Pakete je nach ihrer Wichtigkeit zwei Scheduling-Queues eines DiffServ-unterstützenden Routers zugewiesen. Es wird allgemein vorgeschlagen, die Wichtigkeit anhand zweier Randbedingungen zu berechnen, der gewünschten Videoqualität im Empfänger und der Paketverlustrate im Netz. Die konkrete Implementierung orientiert sich am Vorhandensein von MPEG-Headern in der Nutzlast der Video-IP-Pakete. Es werden keine Video-IP-

¹²DiffServ: Differentiated Services

¹³QoS: Quality of Service

Pakete proaktiv verworfen. Bei Video-IP-Paketverlusten werden Abhängigkeiten nicht proaktiv behandelt. Das Aufbrechen einer Videoverbindung kann zur falschen Video-IP-Paket- bzw. Videobildreihenfolge führen. Die Untersuchungen beziehen sich auf ein drahtgebundenes Netz. Es werden unterschiedliche Videotypen untersucht. Die Videos weisen nur *I*-Videobilder auf. Die Videostrukturvielfalt ist deshalb nicht gegeben. Die Bewertung findet PSNR-basiert statt.

In [122] und in [123] wird das Videostreaming über TCP betrachtet. Als Reaktion auf die Stausituationen im Netz schlagen die Autoren vor, die weniger wichtigen Video-IP-Pakete wahrscheinlichkeitsbasiert an der Quelle zu verwerfen, wenn ein bestimmter Senderpufferfüllstand erreicht wurde. Die Wichtigkeit der Video-IP-Pakete wird durch die Zugehörigkeit zu einer Komprimierungsschicht (Teilstrom) eines hierarchisch komprimierten Videos bestimmt. Es werden keine Video-IP-Paket- und Videobildabhängigkeiten behandelt. Die Untersuchungen beziehen sich auf ein drahtgebundenes Netz. Es werden zwei Videos für die Untersuchungen verwendet. Die Videostruktur ist nicht klar. Die Videoqualität wird PSNR-basiert evaluiert.

In [124] schlägt der Autor vor, die zu übertragenden Video-IP-Pakete mit einer sogenannten Life-Time-Information (im IP-Optionsfeld des IP-Headers) zu versehen. Sie wird in der Sicherungsschicht eines Mobilfunksystems verwendet, um im Fehlerfall zu bestimmen, ob eine erneute Übertragung überhaupt sinnvoll wäre. Bei veralteten Video-IP-Paketen soll die Nutzlast entfernt werden und das leere Paket erneut an den Empfänger gesendet werden, damit die Ende-zu-Ende-Staukontrolle nicht reagiert. Es wird ein Mobilfunkzugangnetz (GPRS) untersucht. Das Video wurde durch CBR-, TCP- und TFRC¹⁴-Ströme simuliert. Es werden keine Videodatenabhängigkeiten explizit behandelt. Es findet keine explizite Videoqualitätsbewertung statt. Die Bewertung findet auf der Video-IP-Paketebene statt (z. B. die Anzahl der fehlerfrei empfangenen Pakete).

In [83] wird im Rahmen eines TFRC-basierten Videostreaming vorgeschlagen, die Video-IP-Pakete im Senderpuffer des Videostreamingservers je nach ihrer Wichtigkeit in Sendegruppen von jeweils 3 Paketen zu platzieren. Die Wichtigkeit eines Video-IP-Paketes wird durch die Zugehörigkeit des entsprechenden Videobildes zur Komprimierungsschicht (Teilstrom) und zum Videobildtyp und anhand seines Deadline-Zeitstempels (die bis zum Abspielzeitpunkt verbleibende Zeit) auf z. T. komplexe Weise bestimmt. Die Sendegruppen untereinander sind unterschiedlich wichtig. Die Wichtigkeit einer Sendegruppe wird durch den Deadline-Zeitstempel des wichtigsten Video-IP-Paketes innerhalb der Gruppe bestimmt (die wichtigste Sendegruppe hat den kleinsten Deadline-Zeitstempel bei ihrem wichtigsten Video-IP-Paket). Reichen die Übertragungsressourcen für eine weniger wichtige Sendegruppe nicht aus, werden die entsprechenden Video-IP-Pakete verworfen. Die Behandlung von Video-IP-Paket- und Videobildabhängigkeiten findet nicht statt. Es wird ein drahtgebundenes Netz untersucht. Es sind keine Details zu den untersuchten Videos zu finden. Es findet keine Videoqualitätsbewertung statt. Stattdessen werden algorithmische und übertragungstechnische Metriken verwendet.

In [125] wird ein Scheduling von mehreren Videoströmen vorgeschlagen, bei dem sich die Sendereihenfolge der Video-IP-Pakete innerhalb einer Router-Queue an der Wichtigkeit der Videobilder orientiert. Unterschiedliche *P*-Videobildwichtigkeiten werden nicht berücksichtigt. Die Queues werden nacheinander so bedient, dass alle Ströme statistisch fair behandelt werden. Die Video-IP-Paket- und Videobildabhängigkeiten werden beim Scheduling nicht berücksichtigt.

¹⁴TCP-Friendly Rate Control

Es wird ein drahtgebundenes Netz untersucht. Es findet keine Videoqualitätsbewertung statt. Stattdessen wird die Bewertung anhand der videobildtypspezifischen Paketverlustraten durchgeführt. Es werden drei Videos untersucht. Die Zuordnung zum SI-TI-Diagramm ist unklar. Die Videostruktur ist nicht beschrieben.

In [126] und [127] wird vorgeschlagen, *B*- und *P*-Videobilder im Puffer eines Netzknotens abhängig vom Pufferfüllstand proaktiv zu verwerfen. Dabei werden zwei separate Schwellwerte des Pufferfüllstandes für unterschiedliche Videobildtypen (*B*- und *P*-Videobilder) verwendet. Die unterschiedlichen Wichtigkeiten der *P*-Videobilder werden nicht berücksichtigt. Der Algorithmus verwendet einen dritten dynamischen Schwellwert des Pufferfüllstandes, um – gleichzeitig zum proaktiven Verwerfen der weniger wichtigen Videobilder – die Pakete der Querverkehre proaktiv zu verwerfen. Somit werden andere Verkehrsströme zugunsten des Videostreaming verdrängt. Die Video-IP-Paketabhängigkeiten werden nur teilweise berücksichtigt, denn die bereits im Puffer gespeicherten Video-IP-Pakete werden beim Verlust späterer dazugehöriger Video-IP-Pakete nicht aus dem Puffer entfernt. Die Videobildabhängigkeiten werden nicht berücksichtigt. Es werden drei Videos für die Untersuchung verwendet. Die Zuordnung zum SI-TI-Diagramm ist allerdings unklar. Die Zuordnung der Ergebnisse zu Videotypen ist nicht möglich. Es wird lediglich eine Videostruktur untersucht ($N = 12, M = 3$). Die Nichtberücksichtigung der unterschiedlichen Wichtigkeiten der *P*-Bilder kann sich insbesondere bei reinen ($M = 1$)-Strukturen (IPPPP-Strukturen) ungünstig bemerkbar machen. Die Videoqualität wird anhand der Videobildfehlerrate (FER¹⁵) geschätzt. Das untersuchte Netzwerkmodell ist drahtgebunden.

In [128] wird ein Mobile-IP-Szenario in einer drahtlosen Umgebung untersucht. Der Client hat mehrere drahtlose Netzwerkinterfaces und ist deshalb multihomed angebunden. Ein Home-Agent kann den Client über mehrere Netzwerkpfade erreichen. Er schätzt die Auslieferungszeit der Videobilder über jeden einzelnen Pfad und verteilt den Strom darauf so, dass die Auslieferungszeit minimal ist. Um die rechtzeitige Auslieferung der wichtigen Videobilder zu erzwingen, werden weniger wichtige Videobilder ggf. proaktiv verworfen. Die Wichtigkeit orientiert sich am Videobildtyp. Es ist nicht klar, ob die unterschiedlichen Wichtigkeiten der *P*-Videobilder im Algorithmus berücksichtigt werden. Auch werden Videobilder verworfen, die laut Schätzung nicht rechtzeitig ausgeliefert werden können. Dabei werden zwar die Videobildabhängigkeiten theoretisch berücksichtigt. Es wird aber darauf hingewiesen, dass dies für die *I-B*-Abhängigkeiten nicht realisiert wurde. Es werden zwei Videos untersucht. Die Zuordnung zum SI-TI-Diagramm ist nicht klar. Die Videostruktur ist nicht klar. Es wird eine PSNR-basierte Bewertung vorgenommen.

In [129] wird ein Mobilfunkzugangsnetz (UTRAN) untersucht. Es wird vorgeschlagen, in einer Überlastsituation, die Video-IP-Pakete, die am längsten im Puffer und am wenigsten wichtig sind, zu verwerfen. Die Wichtigkeit soll den MPEG-Daten der Pakete entnommen werden. Weitere Details werden nicht mitgeteilt. Die Video-IP-Paket- und Videobildabhängigkeiten werden bei der Beschreibung der Vorgehensweise nicht erwähnt. Es wird eine Videosequenz untersucht, die zwei Videotypen abbildet. Die Zuordnung der Ergebnisse zu den einzelnen Videotypen ist nicht möglich. Es wird nur eine Videostruktur untersucht ($N = 30, M = 3$). Die Videoqualität wird PSNR-basiert bewertet.

¹⁵FER: Frame Error Rate

In [130] wird vorgeschlagen, dass unterschiedlich wichtige Videobilder mit einem unterschiedlich starken FEC versehen werden. Des Weiteren soll der Videostreamingserver die Video-IP-Pakete unterschiedlich – je nach ihrer Wichtigkeit – markieren, damit sie in den DiffServ-Routern entsprechend behandelt werden. Die Wichtigkeit orientiert sich an der Zugehörigkeit zu einer hierarchischen Komprimierungsschicht (Teilstrom) und an dem Videobildtyp. Die unterschiedlichen Wichtigkeiten der P -Videobilder werden beim Verwerfen nicht berücksichtigt. Die DiffServ-Router haben keine Kenntnisse der Video-IP-Paket- und Videobildabhängigkeiten. Es wird ein drahtgebundenes Netz untersucht. Es wird nur eine Videosequenz untersucht. Die Videostruktur ist nicht angegeben. Die Videoqualität wird PSNR-basiert beurteilt.

In [131] wird eine WLAN-Umgebung mit einem sich darin befindlichen Videostreamingserver untersucht. Als Reaktion auf eine schwankende Übertragungsdatenrate des Funkkanals wird vorgeschlagen, die Videobilder im Sendepuffer je nach ihrer Wichtigkeit zu verwerfen. Die Wichtigkeit orientiert sich am Videobildtyp. Die unterschiedlichen Wichtigkeiten von P -Videobildern werden beim Verwerfen nicht berücksichtigt. Der Sendepuffer besteht aus nur einem einzigen Videobild, das ggf. durch ein wichtigeres Videobild überschrieben wird. Der untersuchte Videotyp wird nicht angegeben. Die untersuchte Videostruktur ist nicht ganz klar (vermutlich $(N = 9, M = 3)$). Die Leistungsbewertung erfolgt anhand von algorithmischen Statistiken (die Anzahl der verworfenen Videobilder der unterschiedlichen Videobildtypen). Es erfolgt keine Bewertung der Videoqualität.

In [132] wird für einen Videostreamingserver – als Ergänzung zu Staukontrollmechanismen – eine Multipufferarchitektur vorgeschlagen, in der Videobilder unterschiedlich je nach ihrer Wichtigkeit behandelt werden. Die Wichtigkeit orientiert sich am Videobildtyp. Unterschiedliche Wichtigkeiten von P -Videobildern werden nicht berücksichtigt. Für jeden Videobildtyp wird ein logischer Sendepuffer angelegt. Die Auslieferungszeiten der Video-IP-Pakete am Client werden geschätzt. Wenn ein Video-IP-Paket nicht rechtzeitig ausgeliefert werden kann, wird berechnet, welche Video-IP-Pakete ggf. proaktiv verworfen werden können. Dabei werden die Video-IP-Paketabhängigkeiten berücksichtigt. Es ist nicht klar, ob die Videobildabhängigkeiten berücksichtigt werden. Es wird ein drahtgebundenes Netzwerk untersucht. Es werden zwei Videosequenzen verwendet. Die untersuchte Videostruktur ist $(N = 16, M = 3)$. Die Videostrukturvielfalt ist nicht gegeben. Die Bewertung der Videoqualität erfolgt PSNR-basiert.

Die Untersuchung in [88] behandelt einen Retransmission-Mechanismus auf der Sicherungsschicht eines Mobilfunkzugangnetzes (UTRAN). Es wird ein exklusiver Kanal betrachtet. Es wird für die Sicherungsschicht vorgeschlagen, die I -Videobilddaten im Fehlerfall viel öfter erneut zu übertragen als die P -Videobilddaten. Die P -Videobilddaten werden ggf. zugunsten der I -Videobilddaten im Fehlerfall verworfen. Die Randbedingung des Verfahrens ist, dass seine Retransmission-Kosten in etwa gleich den üblichen Retransmission-Kosten sind. Die unterschiedlichen Wichtigkeiten der P -Videobilder werden implizit durch die Art des Algorithmus beim Verwerfen berücksichtigt. Es werden zwei Videosequenzen untersucht. Die Videos weisen nur eine $(N = 15, M = 1)$ -Struktur auf. Die Videodaten-, Video-IP-Paket- und Videobildabhängigkeiten werden in der Verfahrensbeschreibung nicht erwähnt. Die Videoqualitätsbewertung findet PSNR-basiert statt.

Der Artikel [133] stammt vom Autor der hier vorliegenden Arbeit. Die im Artikel vorgestellten Verfahren und Untersuchungen bilden die Grundlagen für die weiteren Untersuchungen im

Rahmen dieser Arbeit und werden deshalb in den entsprechend Abschnitten ausführlich beschrieben.

In [134] wird ein Puffermanagementverfahren für WLAN-Umgebungen vorgestellt. Es wird vorgeschlagen, pro Videoverbindung eine separate virtuelle Queue im Access-Point anzulegen und zu verwalten. Die Sicherungsschicht des WLAN wird bei der Initiierung einer Verbindung entsprechend dynamisch konfiguriert. Für jedes Videobild in der Queue wird die Auslieferungszeit beim Client geschätzt. Dazu verwendet man die Informationen aus der Sicherungsschicht. Wird festgestellt, dass ein Videobild wahrscheinlich zu spät ankommt, werden weniger wichtige Videobilder – die in der Queue davor stehen – verworfen. Die Wichtigkeit orientiert sich am Videobildtyp. Die unterschiedlichen Wichtigkeiten der *P*-Videobilder werden nicht berücksichtigt. Die Videobildabhängigkeiten werden berücksichtigt. Das Verfahren der Berechnung der Auslieferungszeit ist komplex und WLAN-spezifisch. Die Videotypvielfalt ist nicht gegeben. Die untersuchte Videostruktur ist ($N = 12, M = 3$). Es ist keine Videostrukturvielfalt gegeben. Die Videoqualität wird PSNR-basiert evaluiert.

In [77] wird im Kontext eines LTE-Mobilfunkzugangsnetzes (EUTRAN) vorgeschlagen, u. a. die Videobilder anhand ihrer Wichtigkeit in einer Überlastsituation kontrolliert zu verwerfen. Es werden zwei Wichtigkeitsstufen definiert (hoch für *I*- und *P*-Videobilder und niedrig für *B*-Videobilder). Die unterschiedlichen Wichtigkeiten der *P*-Videobilder werden somit beim Verwerfen nicht berücksichtigt. Es werden keine weiteren Details angegeben. Es werden keine Untersuchungsergebnisse bzgl. des kontrollierten Verwerfens präsentiert.

In [135] wird ein Puffermanagementverfahren für Videostreamingserver vorgestellt, die TCP-basiert sind (z. B. DASH). Beim Erkennen einer Stausituation werden weniger wichtige Video-IP-Pakete proaktiv aus dem Sendepuffer entfernt. Die Wichtigkeit orientiert sich dabei an der Zugehörigkeit zu einer hierarchischen Komprimierungsschicht (Teilstrom). Die Videobildtypen werden nicht berücksichtigt. Die Berücksichtigung der Video-IP-Paket- und Videobildabhängigkeiten wird erwähnt. Die Details sind aber unbekannt. Es wird ein drahtgebundenes Netz simuliert. Es wird nur ein Video untersucht. Die Videostruktur ist unbekannt. Die Videotyp- und die Videostrukturvielfalt sind nicht gegeben. Es wird ein drahtgebundenes Netz untersucht. Die Videoqualität wird PSNR-basiert evaluiert.

In [136] wird im Kontext des LTE-Mobilfunkzugangsnetzes (EUTRAN) ein UDP-basiertes Videostreaming vorgeschlagen. Die Autoren bemängeln, dass das TCP-basierte DASH oft zu langsam und zu ungenau auf die schnellen und starken Mobilfunksignalschwankungen reagiert. Stattdessen soll ein auf der hierarchischen Komprimierung basierendes Videostreaming verwendet werden. Die Teilströme sollten mit Hilfe von LTE-QoS-Mechanismen ggf. unterschiedlich behandelt werden. Somit soll sichergestellt werden, dass zumindest der Teilstrom der Basisqualität am Empfänger ankommt. Es wird nur ein Video untersucht. Die Videotyp- und Videostrukturvielfalt ist nicht gegeben. Die Videoqualitätsbewertung wird mit SSIM durchgeführt und auf MOS abgebildet.

4.3.2 Diskussion und Einordnung

Wie man der Literaturübersicht entnehmen kann, haben sich zahlreiche Autoren mit dem proaktiven Verwerfen von weniger wichtigen Videodaten zugunsten wichtigerer Videodaten befasst

(sei es im Netz, sei es im Videostreamingserver, sei es im Videoclient). Die Verfahren unterscheiden sich in der Definition der Wichtigkeit, der Vorgehensweise und den Optimierungszielen. Alle Autoren sind sich einig, dass der Lösungsansatz eines proaktiven kontrollierten Verwerfens basierend auf der Wichtigkeit der Videodaten eine Verbesserung in einer Überlastsituation bringt.

Die meisten Verfahren des kontrollierten Verwerfens entfernen Videodaten videobildtypabhängig. Alle diesbezüglichen Publikationen haben gemein, dass ihre Ergebnisse nicht eindeutig quantitativ eingeordnet werden können, denn Videobildumfangen variieren zum Teil stark, je nach Videotyp, Videostruktur und Vorgehensweise bei der Komprimierung (Ein-Durchgangs- bzw. Zwei-Durchgangs-Kodierung). Diese Vielfalt wird in den Untersuchungen nicht betrachtet, d. h., der Untersuchungsgegenstand (Video) wird innerhalb einer Publikation nicht vollständig untersucht und ist zudem publikationsspezifisch. Die unterschiedlichen Untersuchungskontexte verschiedener Veröffentlichungen erschweren die Ergebniseinordnung zusätzlich.

Die Videobildwichtigkeit orientiert sich in den meisten Fällen am Videobildtyp und an der Zugehörigkeit zu einer Schicht eines hierarchisch kodierten Videos. Oft kommt noch der Aspekt der rechtzeitigen Auslieferung am Client dazu. Die Definitionen der Wichtigkeiten unterscheiden sich. Es gibt einfache Vorschläge wie die Unterscheidung zwischen *I*-/*P*-Videobildern und *B*-Videobildern. Auf der anderen Seite gibt es komplexere Lösungsvorschläge, die Wichtigkeit anhand von MPEG-Daten des Video zu berechnen. In den Publikationen mit der videobildtypbasierten Wichtigkeit wird i. d. R. die unterschiedliche Wichtigkeit der *P*-Videobilder beim proaktiven Verwerfen nicht berücksichtigt, was sich bei ($M = 1$)-Strukturen negativ bemerkbar machen würde. Des Weiteren berücksichtigen viele Publikationen keine Video-IP-Paket- und/oder Videobildabhängigkeiten beim proaktiven Verwerfen.

Die vorgeschlagenen Algorithmen sind zum Teil komplex und beinhalten oft weitere Komponenten. Die gleichzeitig und z. T. vorrangig wirkenden Verbesserungsmechanismen, wie das Verwerfen anhand von Deadlines, die Verwendung von FEC, die doppelte bzw. erneute Übertragung, oder das Verdrängen des konkurrierenden Verkehrs, überlagern die Wirkung des proaktiven Verwerfens. Es fehlt an der grundsätzlichen Erkenntnis, was ein proaktives videobildtypabhängiges Verwerfen quantitativ bringt. Es fehlt an der grundsätzlichen Erkenntnis, wie groß die Unterschiede zwischen der Berücksichtigung und der Nichtberücksichtigung der Videodatenabhängigkeiten sind.

Im Rahmen dieser Arbeit wird ein – auf dem Schwellwert des Pufferfüllstandes basierendes – Verfahren vorgeschlagen, bei dem die Komplexität mehrerer Wichtigkeitsstufen – und des damit verbundenen algorithmischen Handlungszwangs ggf. mit einem großen Parameterraum – durch die Einführung einer quantitativen Wichtigkeit der Videobilder umgangen wird. Damit werden auch unterschiedliche Wichtigkeiten der *P*-Videobilder berücksichtigt. Durch die Entkopplung vom Videobildtyp, die einfache Bestimmung bzw. Erkennung der Wichtigkeit und die schwellwertbasierte Vorgehensweise lässt sich das Verfahren in den heutigen Netzknoten implementieren und selektiv auf die ankommenden Video-IP-Pakete anwenden. Die Mittel zur selektiven Anwendung bieten die schon heute verfügbaren Techniken wie das sogenannte Network-Programming in Kombination mit dem Einsatz von Segment-Routing (SRv6). Dabei lassen sich in jedem IPv6-Paketheader die sogenannten Netzfunktionen mitteilen, die vom Netzknoten auf das IP-Paket anzuwenden sind. Im Rahmen dieser Untersuchung wird das beschriebene Verfahren in UTRAN und auf der Basis von IPv4 implementiert.

Das vorgeschlagene Verfahren wird mit mehreren komprimierten Videosequenzen untersucht, die unterschiedlichen Videotypen angehören und unterschiedliche Videostrukturen aufweisen. Des Weiteren wurden sie mit unterschiedlichen Vorgehensweisen komprimiert. Somit lässt sich das Ergebnis einordnen und verallgemeinern. Der Effekt der Berücksichtigung der Videodatenabhängigkeit wird quantitativ herausgearbeitet. Die optimale Parametrierung des Verfahrens wird durch umfangreiche Messreihen ermittelt.

Die Untersuchungen, auf welchen diese Arbeit aufbaut, wurden vom Autor der hier vorliegenden Ausarbeitung in [133] publiziert.

4.3.3 Implementierung

Die Implementierung des hier vorgeschlagenen videospezifischen AQM-Mechanismus als Teil von VSNC erfolgt in der vorliegenden Arbeit in UTRAN und ist der Abbildung 4.5 zu entnehmen.

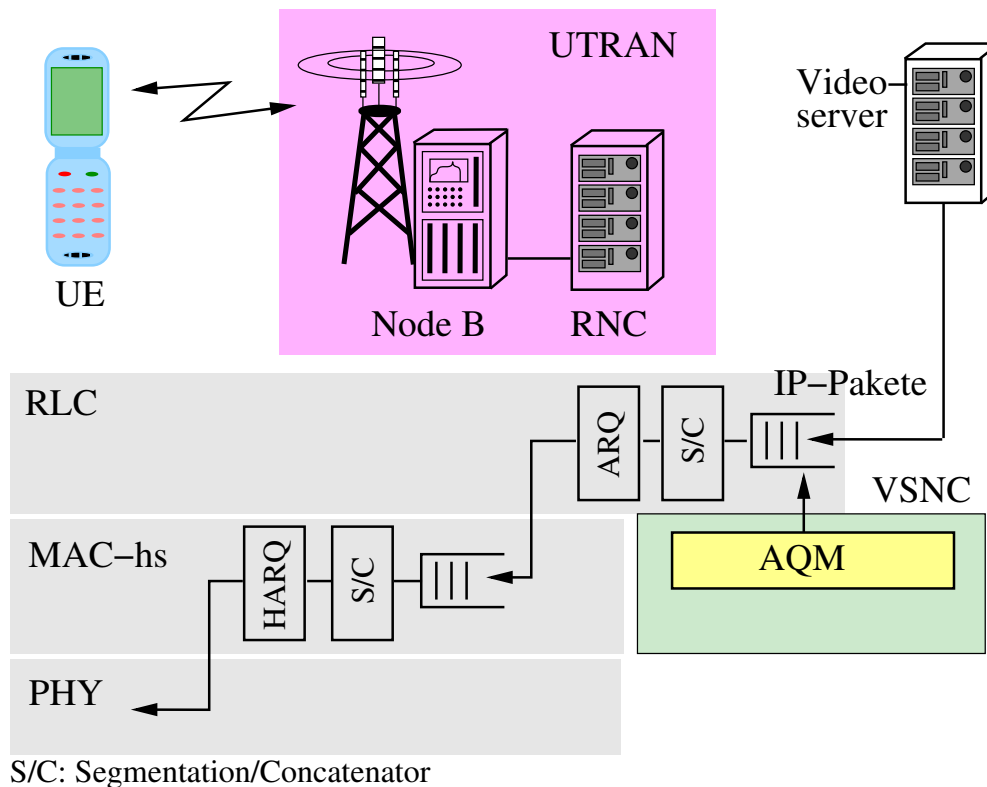


Abbildung 4.5: Videospezifischer AQM-Mechanismus in UTRAN

Das videospezifische AQM-Verfahren wird auf den für eine Videostreamingverbindung dedizierten Eingangspuffer für IP-Pakete oberhalb der Sicherungsschicht – wie es das hier eingesetzte UTRAN-Modell vorsieht – angewendet.

Die Video-IP-Pakete werden in einer abzusehenden Überlastsituation in der Abhängigkeit ihrer Wichtigkeit proaktiv zugunsten wichtigerer Video-IP-Pakete verworfen. Um die Überlastsituation zu erkennen, wird der Pufferfüllstand *BufOccup* überwacht. Beim Überschreiten eines Schwellwertes δ wird das Verwerfen aktiv.

Die Wichtigkeit der Videobilder und somit der Video-IP-Pakete wird basierend auf dem Umfang der Videodatenabhängigkeiten quantitativ berechnet. Konkret gibt die hier implementierte Wichtigkeit eines Videobildes (Imp_i) die Anzahl der von ihm abhängigen Videobilder an. Ist bei einem Videobild $Imp_i = 0$, bedeutet dies, dass bei seinem Verwerfen keine weiteren Videobilder beim Dekomprimieren betroffen sind. Ist bei einem Videobild $Imp_i = 7$, so sind durch sein Verwerfen sieben weitere Videobilder beim Dekomprimieren betroffen. Die Bestimmung der Wichtigkeit basiert hier auf den Regeln der Abhängigkeiten zwischen drei Videobildtypen (I -Videobilder, P -Videobilder, B -Videobilder), die im Abschnitt 3.2.2.4 beschrieben wurden.

Durch das Einführen der quantitativen Wichtigkeit wird die Komplexität mehrerer videobildtypbasierter Wichtigkeitsstufen – und des damit verbundenen algorithmischen Handlungszwangs ggf. mit einem großen Parameterraum – umgangen. Damit werden die unterschiedlichen Wichtigkeiten von P - und sogar von I -Videobildern erfasst, was in den meisten videobildtypbasierten Verfahren nicht der Fall ist (kein Verfahren berücksichtigt die unterschiedlichen Wichtigkeiten der I -Videobilder). Die unterschiedlichen Wichtigkeiten von I -Videobildern treten in den Videostrukturen mit den variablen GoP-Längen auf. Des Weiteren lässt sich das Verfahren durch die Entkopplung vom Videobildtyp auf *beliebige* Ausprägungen der Videobildabhängigkeiten anwenden.

Einfachheit ist das Ziel des vorgeschlagenen Verfahrens, damit die Realisierung in den heute üblichen Komponenten möglich ist. Die Trennung der weniger wichtigen Video-IP-Pakete von wichtigeren Video-IP-Paketen erfolgt binär. Dafür wird der Schwellwert (Imp_{min}) eingeführt, der die Grenze für die minderwertigen Wichtigkeitswerte festlegt. Das Verfahren verwirft proaktiv die Video-IP-Pakete mit der Wichtigkeit Imp_i , wenn die folgende Aktivierungsbedingung wahr ist:

$$(BufOccup > \delta) \wedge (Imp_i \leq Imp_{min})$$

Die Bestimmung der optimalen Werte für die beiden Schwellwerte wird im Rahmen dieser Untersuchung in zahlreichen Messungen mit der notwendigen Videotyp- und Videostrukturvielfalt bestimmt.

In Bezug auf die Berücksichtigung der Videodatenabhängigkeiten beim proaktiven Verwerfen werden folgende Strategien untersucht:

- *Verwerfen auf Basis der Video-IP-Pakete*
Es werden die einzelnen Video-IP-Pakete verworfen, wenn die Aktivierungsbedingung wahr ist. Es werden keine Abhängigkeiten der Videobilddaten berücksichtigt. Das ist die einfachste Vorgehensweise für die Implementierung in den heute üblichen Komponenten.
- *Verwerfen auf Basis der Videobilder*
Es werden komplette Videobilder aus dem Puffer entfernt, wenn für irgendein Video-IP-Paket, das zum betroffenen Videobild gehört, die Aktivierungsbedingung gilt. Konkret wird wie folgt vorgegangen: Ein neu eingetroffenes Video-IP-Paket wird verworfen, wenn ein vorangegangenes Video-IP-Paket desselben Videobildes verworfen wurde. Wird ein neu eingetroffenes Video-IP-Paket aufgrund des proaktiven Verwerfens oder aufgrund des Pufferüberlaufs verworfen, so werden alle sich bereits im Puffer befindlichen Video-IP-Pakete desselben Videobildes entfernt.

- *Verwerfen auf Basis der Videobilder mit anschließendem Verwerfen der abhängigen Videobilder*

Das Verwerfen eines kompletten Videobildes wird um das Verwerfen sämtlicher davon abhängiger Videobilder erweitert. Aufgrund der Bitstream-Reihenfolge der Videobilder bei der Übertragung werden alle betroffenen Videobilder nach und nach verworfen.

Im Rahmen der Untersuchungen werden die vorgeschlagenen Verfahren untereinander verglichen.

4.4 Gemeinsame Verwendung der beiden Maßnahmen im VSNC

In diesem Abschnitt wird – nach einer Literaturübersicht – das vorgeschlagene Verfahren eingeordnet. Danach werden die konkreten Implementierungsdetails erklärt.

4.4.1 Literaturübersicht

Die Literaturrecherche ergab nur zwei relevante Behandlungen dieser Ausprägung.

Die Idee, beide Maßnahmen gemeinsam zu verwenden, wurde vom Autor der hier vorliegenden Arbeit in [109] geäußert.

Zu einem späteren Zeitpunkt wurde in [77] vorgeschlagen, ein proaktives Verwerfen von Videobildern als unterstützende Maßnahme für das Transcoding in einem LTE-Mobilfunkzugangnetz (EUTRAN) zu verwenden. Die Details der Verfahren sind in den Abschnitten 4.2.1 und 4.3.1 nachzulesen.

4.4.2 Diskussion und Einordnung

Die Vorstellung des Konzepts, die Diskussion und die Einordnung der einzelnen Maßnahmen sowie die architektonische Begründung haben in den Abschnitten 4.1, 4.2.2 und 4.3.2 stattgefunden.

Des Weiteren lässt sich anmerken, dass die beiden Maßnahmentearten in der Literatur ggf. eher miteinander verglichen werden. Sie sollten aber als sich ergänzende Maßnahmen eingesetzt werden. Dies wurde auch in [77] grundsätzlich erkannt.

4.4.3 Implementierung

Das Konzept wurde im Abschnitt 4.1 vorgestellt. Die Implementierungen der beiden Maßnahmen wurden in den Abschnitten 4.2.3 und 4.3.3 beschrieben. Die beiden Maßnahmeninstanzen im Rahmen dieser Arbeit sind voneinander unabhängig und benötigen keinen Kommunikationskanal untereinander.

5 Experimentelle Untersuchungsumgebung

5.1 Einführung

Die in dieser Arbeit vorgestellten Ergebnisse wurden mit einer eigens dazu entwickelten experimentellen Videostreaming-Untersuchungsumgebung erzielt. Die Grundidee dafür stammt aus der EvalVid-Software [8].

Der Videoverkehr wird tracebasiert [137] emuliert. Auf der Ebene der IP-Pakete entspricht der emulierte Videoverkehr einem realen MPEG-4-basierten Videostrom, der mit einem UDP-basierten Videostreaming zustande käme. Im erweiterten IP-Header werden die wichtigsten videobildbasierten Daten mit übertragen (im Folgenden *Metadaten* genannt). Dazu gehören u. a. die Zugehörigkeit der Videodaten zu einem Videobild, die Videobildart, die Informationen über die Videobildabhängigkeiten und die Videobildwichtigkeit. Der Videosender ist adaptiv. Dafür wurde die Bitstream-Switching-Methode [75] für die Umschaltung an den *I*-Videobildern implementiert.

Im Folgenden wird die Untersuchungsumgebung näher vorgestellt.

5.2 Systemübersicht

Die Untersuchungsumgebung wurde mit C++ entwickelt und läuft unter Linux¹. Die Abbildung 5.1 zeigt die Komponenten des entwickelten Systems. Die hellblau dargestellten Komponenten symbolisieren Dateien. Die weißen Komponenten repräsentieren Programme.

Der Ausgangspunkt sind die unkomprimierten *YUV*-Videosequenzen. Sie werden mit der passenden Länge, Bildwiederholrate und Bildauflösung dem Komprimierungsvorgang unterzogen. Beim Komprimierungsvorgang werden die Vorgaben für die Videostruktur und die mittlere Videodatenrate eingestellt. Anschließend wird mit Hilfe eines MPEG-4-Parsers ein MPEG4-Trace erzeugt, der die Informationen über die Videobildtypen und Videobildumfänge enthält. Die Komprimierungssoftware *ffmpeg* stammt aus [138]. Der MPEG-4-Parser stammt aus [8].

¹Ein Unix-basiertes Betriebssystem

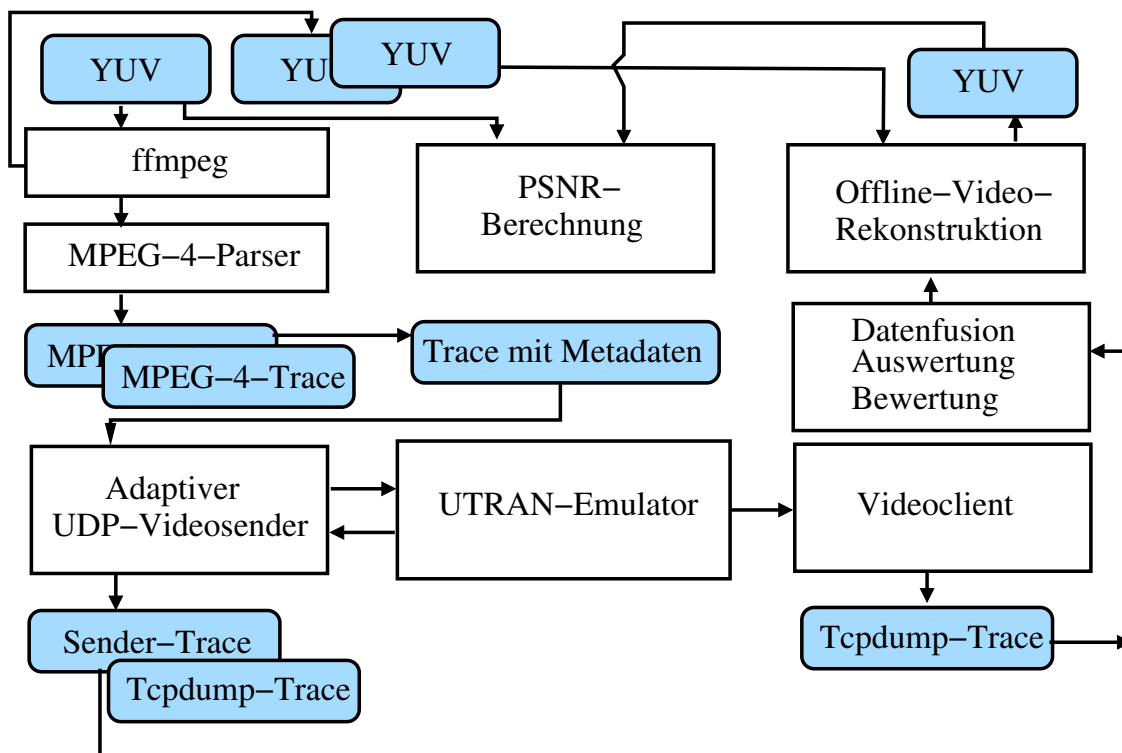


Abbildung 5.1: Experimentelle Untersuchungsumgebung

Für die Untersuchungen wird der MPEG-4-Trace in eine andere Form umgeschrieben. Diese enthält die benötigten Metadaten wie z. B. die Zugehörigkeit der Videodaten zu einem Videobild, die Videobildart, die Informationen über die Videobildabhängigkeiten, die Videobildwichtigkeit und die Abspielreihenfolge. Der Trace, der die Metadaten enthält, liegt in der Bitstream-Reihenfolge (Sendereihenfolge) vor.

Anhand des videobildbasierten, Metadaten enthaltenden Trace erzeugt und versendet der Videosender die Video-IP-Pakete mit der emulierten² Nutzlast an den Videoclient. Die Videobildwiederholrate und der IP-Paketierungs-Parameter (die maximale Größe eines IP-Pakets) werden beim Aufruf eingestellt. Die IP-Paketierung folgt den Regeln³ aus RFC 6416 [63]. Somit entsprechen die Verkehrscharakteristiken des so erzeugten emulierten Videostroms den Verkehrscharakteristiken eines realen Videostromes.

Wird ein adaptives Szenario untersucht, so erhält der Videosender mehrere Traces. Der Videosender weist einen Rückkanal zum Empfang von Adaptionsvorschlägen auf.

Alle Video-IP-Pakete werden mit den bereits besprochenen Metadaten im Optionsfeld (insgesamt 16byte) des IPv4-Paketheaders versehen. Das RTP wird nicht verwendet. Dies stellt für die durchgeführten Untersuchungen keine Einschränkung dar, da die Funktionalität des Protokolls nicht verwendet wird. Der Gesamtprotokolloverhead beträgt 44byte⁴. Zum Vergleich

²Die Anzahl der Bytes entspricht einem realen Video-IP-Paket. Die Semantik der Nutzlast entspricht nicht dem Videoinhalt.

³Pro RTP/UDP-Paket sollten die Videodaten nur eines einzigen Videobildes übertragen werden.

⁴ $20(IP - Paketheader) + 16(IP - Optionsfeld) + 8(UDP - Datagrammheader) = 44$

beträgt der minimale Gesamtprotokolloverhead eines realen RTP/UDP-basierten Videostroms 40 byte⁵.

Der UTRAN-Emulator wurde um die für diese Untersuchungen notwendigen Komponenten erweitert. Die Grundfunktion des UTRAN-Emulators [139], [140] ist nicht der Bestandteil dieser Arbeit.

Der Sendevorgang wird im Videosender in Form einer Datei protokolliert (Sender-Trace). Desweiteren werden die IP-Paketheader aller versendeten Video-IP-Pakete mit den Sendezeitstempeln unter Einsatz der Paketmitschnitt-Software *tcpdump* [141] aufgezeichnet. Auf der Empfangsseite werden ebenfalls die IP-Paketheader aller empfangenen Video-IP-Pakete mit den Empfangszeitstempeln aufgezeichnet. Die Zeitbasis ist exakt dieselbe, denn der Videosenderrechner besitzt einen separaten Netzanschluss im Videoclientnetz, worüber die empfangenen IP-Paketheaderdaten aufgenommen werden. Zusätzlich werden die IP-Paketheaderdaten im Videoclientrechner aufgezeichnet.

Nach dem Ende der Übertragung werden die Daten der Sende- und der Empfangsseite zusammengeführt und detailliert ausgewertet. Dabei werden automatisierte Bewertungen bezüglich der Überschreitung der Grenzwerte vorgenommen (z. B. die Berechnung der Wahrscheinlichkeit für eine Videoabspielunterbrechungsdauer, die länger als eine Sekunde dauert). Gleichzeitig werden weitere Daten erzeugt wie z. B. die Abspiel-Traces.

Mit Hilfe eines Abspiel-Trace können die Empfangsvideos offline rekonstruiert werden (sowohl im adaptiven als auch im adaptionslosen Fall). Insbesondere im adaptiven Fall ist von Interesse, die PSNR-Werte berechnen zu lassen. Die geschieht mit den offline rekonstruierten Videos anhand des PSNR-Berechnungstool von [8].

5.2.1 Messaufbau

Der Messaufbau besteht aus drei Rechnern, wie in der Abbildung 5.2 zu sehen ist. Die Rechner besitzen mehrere Netz-Interfaces. Die Konnektivität im Messaufbau ist der Abbildung 5.3 zu entnehmen.

Der Clientrechner und der Serverrechner befinden sich jeweils in einem eigenen Netz. In der Abbildung sind die Netze als Servernetz und als Clientnetz gekennzeichnet. Die Übertragung des Videos findet über den an beiden Netzen angeschlossenen UTRAN-Rechner statt (der dunkelgraue Bereich im Bild). Desweiteren weist der Server einen separaten Netzanschluss im Clientnetz auf. Darüber werden die im Clientnetz angekommenen Video-IP-Pakete aufgezeichnet. Damit wird die exakt gleiche Zeitbasis zur Berechnung der Verzögerung sichergestellt.

Das hellgrau Netzsegment ist vom Untersuchungsbereich getrennt und dient der Steuerung der Messvorgänge. Ebenfalls darüber findet der Zugriff auf den Messaufbau von aussen statt.

⁵ $20(IP - \text{Paketheader}) + 8(UDP - \text{Datagrammheader}) + 12(RTP - \text{Sequenzheader}) = 40$



Abbildung 5.2: Messaufbau

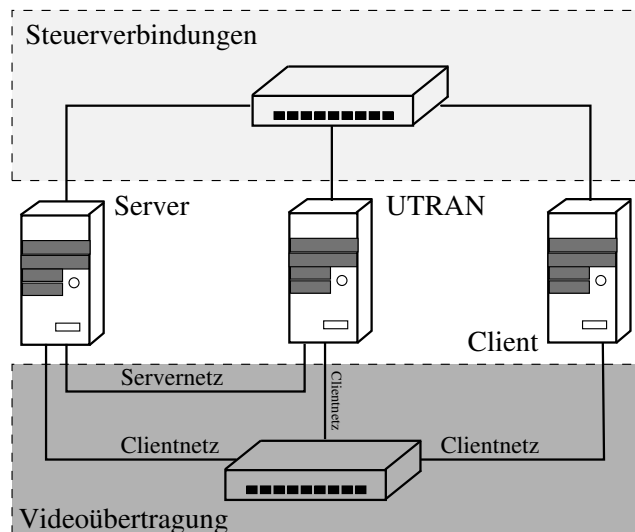


Abbildung 5.3: Konnektivität im Messaufbau

5.2.2 PVU-Diagramm

Bei der Verwendung von PSNR als Videoqualitätsmetrik für die Bewertung einer Videostreamingsitzung beurteilen viele Autoren ihre Ergebnisse anhand von PSNR-Mittelwerten über alle Videobilder. Eine Verteilung der PSNR-Werte würde aber viel mehr Informationen geben. Allerdings überdecken sich beim Vergleich mehrerer Sitzungen die Verteilungen u. U., so dass die visuelle Bewertung in ein und demselbem Diagramm anstrengend sein kann.

Nicht nur PSNR sondern auch die Dauern der Abspielunterbrechungen spielen eine wesentliche Rolle bei der Wahrnehmung der Qualität einer Videostreamingsitzung. Damit ist die Dauer der Verluste der unmittelbar aufeinander folgenden Videobilder gemeint. Für diese Zeit kann das zuletzt fehlerfrei empfangene Videobild eingeblendet werden. Das Videobild ist eingefroren, sagt man. Auch hier gilt, dass eine Verteilung aussagekräftiger als der Mittelwert ist.

In [142] wurde eine subjektive Bewertung von zahlreichen Videostreamingsitzungen unter dem Aspekt der Unterbrechungsdauern durchgeführt. Es nahmen 15 Personen an der Bewertung teil. Zusammenfassend lässt sich feststellen, dass Unterbrechungen, die eine Sekunde überschreiten, als störend empfunden werden.

In dieser Arbeit wird vorgeschlagen, die PSNR-Verteilungen und Unterbrechungsdauern-Verteilungen in Form von abgewandelten Box-Plot-Darstellungen in ein und demselben Diagramm darzustellen. Ein Box-Plot-Diagramm zeigt das 1.Quartil (0,25-Quantil, im Folgenden Q_1 genannt), das 2.Quartil (0,5-Quantil oder Median, im Folgenden Q_2 genannt) und das 3.Quartil (0,75-Quantil, im Folgenden Q_3 genannt) einer Verteilung. Des Weiteren werden Ausreißer angezeigt. Die Definition der Ausreißer ist bei Box-Plot-Diagrammen nicht einheitlich. Es wird hier vorgeschlagen, für die Ausreißer den Minimum- und den Maximumwert zu verwenden.

Die Abbildung 5.4 zeigt das neuartige *PVU-Diagramm* (PSNR-VideoUnterbrechungs-Diagramm) am Beispiel einer bewerteten Videostreamingsitzung. Die Unterbrechungsdauern-

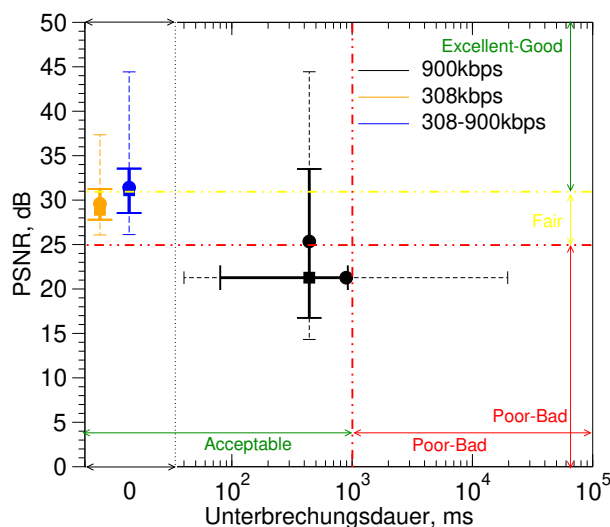


Abbildung 5.4: Neuartiges PVU-Diagramm

Achse ist logarithmisch. Das schwarze Kreuz zeigt die PSNR-Verteilung (der vertikale Balken) und die Unterbrechungsdauern-Verteilung (der horizontale Balken). Der Kreuzpunkt markiert den Medianwert (Q_2) der jeweiligen Verteilung und wird mit einem Quadrat gekennzeichnet. Der jeweilige Mittelwert wird mit einem Kreis markiert.

Die Querstriche an den Enden der durchgezogenen Linien eines jeden Balkens markieren die jeweiligen Q_1 -Werte (unten bzw. links) und die jeweiligen Q_3 -Werte (oben bzw. rechts). Zwischen Q_3 und Q_1 liegen die mittleren 50% der Werte der jeweiligen Verteilung.

Die Querstriche an den Enden der gestrichelten Linien eines jeden Balkens markieren die jeweiligen Minimumwerte (unten bzw. links) und die jeweiligen Maximumwerte (oben bzw. rechts). Zwischen Q_3 und dem Maximumwert liegen die oberen 25% der Werte der jeweiligen Verteilung. Zwischen dem Minimumwert und Q_1 liegen die unteren 25% der Werte der jeweiligen Verteilung.

Das PVU-Diagramm besteht aus vier Quadranten. Die Grenzen zwischen den Quadranten werden durch die Abbildung von PSNR-Werten auf MOS (s. Tabelle 3.2) und durch die subjektive Bewertung der Unterbrechungsdauern in [142] bestimmt. Die MOS-Noten sind im Diagramm auf Englisch, weil sie ursprünglich so festgelegt wurden und sollen zwecks richtiger Einordnung beibehalten werden. Eine Videostreamingsitzung, die eine gute Qualität aufweist, wird mit ihrem Kreuz komplett im ersten Quadranten (oben links) dargestellt. Der horizontale Versatz zwischen dem blauen und dem orangen Kreuz dient nur dem Zweck, dass die beiden Verteilungen auseinander gehalten werden können. Der Punkt 0 der x -Achse wurde zu diesem Zweck zu einem Intervall für die Darstellungszwecke ausgedehnt.

Das schwarze Kreuz stellt eine schlechte Qualität dar. Die beste Qualität erfährt die Videostreamingsitzung, die das blaue ausgeartete Kreuz darstellt (es gibt keine Unterbrechungen). Das orange ausgeartete Kreuz (es gibt keine Unterbrechungen) hat schlechtere PSNR-Werte als das blaue Kreuz, befindet sich aber trotzdem im ersten Quadranten.

Auf diese Art visualisiert das PVU-Diagramm anschaulich und vergleichend die Qualität einer adaptiven Videostreamingsitzung.

6 Untersuchungsergebnisse

Die im Kapitel 4 vorgestellten Vorschläge werden hier quantitativ und qualitativ untersucht und bewertet. Im Abschnitt 6.1 wird die Auswahl der zu verwendenden Videosequenzen begründet. Der Abschnitt 6.2 stellt das Systemmodell vor. Im Abschnitt 6.3 werden die Ergebnisse der Untersuchungen dargestellt.

6.1 Videosequenzen

Im Abschnitt 3.4.3 wurde gezeigt, dass sich die Videosequenzen unterschiedlichen Inhaltes nach dem Umfang der örtlichen und der zeitlichen Informationen in einem zweidimensionalen Diagramm (Abbildung 3.7) differenzieren lassen können. In diesem Abschnitt soll nun untersucht werden, ob und wie der unterschiedliche Videoinhalt und die verschiedenen Komprimierungsausprägungen einen Einfluss auf die Datenumfänge bzw. die Datenraten der Videobilder hat. Es werden repräsentative Videosequenzen für die Verfahrensuntersuchungen ausgewählt.

6.1.1 Vorauswahl

Die Vorauswahl basiert auf den standardisierten Sequenzen aus [7] und [94], die im Abschnitt 3.4.3 genannt werden.

Etliche Forschungsgruppen weltweit bieten frei zugängliche Datenbanken von Videotestsequenzen [143], [144], [145], [146], [147], [148] an. Teilweise enthalten diese Datenbanken die erwähnten standardisierten Videosequenzen. Das angebotene Testmaterial weist unterschiedliche Orts- und Zeitauflösungen auf. Zudem sind unterschiedliche Farbunterabtastungen anzutreffen¹. Dies muss bei der Erzeugung des vergleichbaren komprimierten Videotestmaterials beachtet bzw. berücksichtigt werden.

Bei der Auswahl aus den im freien Internet verfügbaren standardisierten Videosequenzen wurde darauf geachtet, dass alle vier Quadranten des TI-SI-Diagramms berücksichtigt werden. Der so entstandene Videosequenzsatz ist mit den grünen Kreisen in der Abbildung 6.1 markiert.

Die vorausgewählten Videosequenzen wurden ggf. zu Sequenzen mit der 4 : 2 : 0-Farbabtastung und 25 Videobildern/s konvertiert. Die örtliche Auflösung aller Sequenzen ist CIF (s. Abbildung

¹Im Wesentlichen 4 : 2 : 2 und 4 : 2 : 0

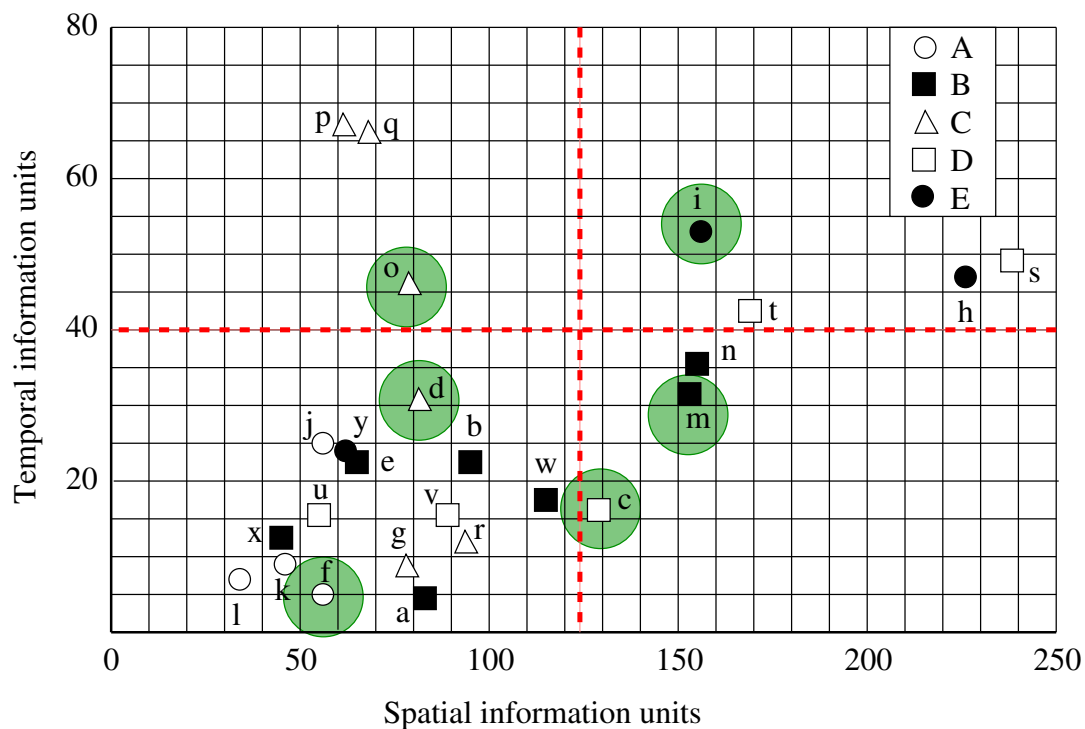


Abbildung 6.1: Vorauswahl der Videosequenzen (vgl. Tabelle 3.3)

im Anhang A.4). Die Originalsequenzen besitzen eine kurze Länge zwischen 5 und 15 s. Deshalb wurden durch ein mehrfaches Aneinanderhängen ein und derselben Videosequenz längere, und deshalb für die eingehenden Untersuchungen besser geeignete Videosequenzen erzeugt. Die Bildabzüge der vorausgewählten Videosequenzen sind in den Abbildungen 6.2 bis 6.7 zu sehen. Ihre Beschreibung ist in der Tabelle 6.1 zu finden.



Abbildung 6.2: vtc1nw(f)



Abbildung 6.3: intros(o)



Abbildung 6.4: 3inrow(d)



Abbildung 6.5: smity1(m)



Abbildung 6.6: washdc(c)



Abbildung 6.7: football(i)

Tabelle 6.1: Beschreibung des vorausgewählten unkomprimierten Videosequenzsatzes.

<i>Eigenschaften</i>	<i>vtc1nw(f)</i>	<i>intros(o)</i>	<i>3inrow(d)</i>	<i>smity1(m)</i>	<i>washdc(c)</i>	<i>football(i)</i>
<i>Inhalt</i>	Teilnehmerin einer Videokonferenz	Mehrere Personen stellen sich vor, Kamerabewegungen	Drei Personen am Tisch, Kamerabewegungen	Ein Händler im TV-Shop	Stadtplan, Hand- und Stiftbewegungen	American-Football-Spiel
<i>Originale Videotestsequenzen in den Datenbanken</i>						
<i>Länge, Videobilder</i>	360	360	360	449	360	260
<i>Zeitliche Auflösung, Videobilder/s</i>	30	30	30	30	30	30
<i>Räumliche Auflösung, Pixel</i>	CIF	CIF	CIF	CIF	CIF	CIF
<i>Farbunterabtastungs- und Speicherformat</i>	uyvy422	yuv422p	uyvy422	yuv420p	uyvy422	yuv420p
<i>Bezugsquelle</i>	[149], [150] ^a	[151]	[152]	[153]	[154], [155] ^b	[156], [157] ^c
<i>Für die Untersuchungen vorbereitete Videosequenzen (Grundsequenzen)</i>						
<i>Länge, Videobilder</i>	301	301	301	374	301	130 ^d
<i>Zeitliche Auflösung, Videobilder/s</i>	25	25	25	25	25	25
<i>Räumliche Auflösung, Pixel</i>	CIF	CIF	CIF	CIF	CIF	CIF
<i>Farbunterabtastungs- und Speicherformat</i>	yuv420p	yuv420p	yuv420p	yuv420p	yuv420p	yuv420p
<i>Zu verwendende Videosequenzen (Aneinanderhängen der Grundsequenzen)</i>						
<i>Länge, Videobilder</i>	33110	33110	33110	32912	33110	33150
<i>Länge, s</i>	1324,4	1324,4	1324,4	1316,48	1324,4	1326

^aAlternative Quelle^bAlternative Quelle^cAlternative Quelle mit einem anderen, aber mit [156] überlappenden Ausschnitt.^dBei der Konvertierung zur Bildrate von 25 Videobildern/s ergab sich eine Länge von 217 Videobildern. In dieser Ausprägung wies das Video einen enormen TI-Anteil auf, so dass eine Komprimierung mit einer kleinen Videodatenrate (308 kbit/s) nicht möglich war. Es wurde deshalb ein Ausschnitt verwendet, der sich entsprechend komprimieren ließ. Andere Autoren verwenden bei Bedarf ebenfalls Ausschnitte, z. B. in [143].

Tabelle 6.2: Unterschiedliche Videostrukturen

B-Serien	1	2	3	4	5	6	7	8	9	10	11	12	13	GoP
ohne <i>B</i> -Serien (<i>OB</i>)	<i>I</i>	<i>P</i>	<i>P</i>	<i>P</i>	<i>P</i>	<i>P</i>	<i>P</i>	<i>P</i>	<i>P</i>	<i>P</i>	<i>P</i>	<i>P</i>	<i>I</i>	12
<i>2B</i> -Serien (<i>2B</i>)	<i>I</i>	<i>B</i>	<i>B</i>	<i>P</i>	<i>B</i>	<i>B</i>	<i>P</i>	<i>B</i>	<i>B</i>	<i>P</i>	<i>B</i>	<i>B</i>	<i>I</i>	12
<i>3B</i> -Serien	<i>I</i>	<i>B</i>	<i>B</i>	<i>B</i>	<i>P</i>	<i>B</i>	<i>B</i>	<i>B</i>	<i>P</i>	<i>B</i>	<i>B</i>	<i>B</i>	<i>I</i>	12
<i>4B</i> -Serien	<i>I</i>	<i>B</i>	<i>B</i>	<i>B</i>	<i>B</i>	<i>P</i>	<i>B</i>	<i>B</i>	<i>B</i>	<i>B</i>	<i>I</i>			10
<i>5B</i> -Serien	<i>I</i>	<i>B</i>	<i>B</i>	<i>B</i>	<i>B</i>	<i>B</i>	<i>P</i>	<i>B</i>	<i>B</i>	<i>B</i>	<i>B</i>	<i>B</i>	<i>I</i>	12
<i>8B</i> -Serien	<i>I</i>	<i>B</i>	<i>B</i>	<i>B</i>	<i>B</i>	<i>B</i>	<i>B</i>	<i>B</i>	<i>B</i>	<i>I</i>				9
<i>11B</i> -Serien (<i>11B</i>)	<i>I</i>	<i>B</i>	<i>B</i>	<i>B</i>	<i>B</i>	<i>B</i>	<i>B</i>	<i>B</i>	<i>B</i>	<i>B</i>	<i>B</i>	<i>B</i>	<i>I</i>	12

6.1.2 Komprimierung

Die für die Untersuchungen erstellten Videosequenzen wurden mit einem MPEG-4-Encoder aus dem ffmpeg-Projekt komprimiert [138]. Die Komprimierung wurde mit einer mittleren Videodatenrate als Optimierungskriterium vorgenommen. Pro Videosequenz wurden unterschiedlich stark komprimierte Versionen erzeugt (mit den mittleren Videodatenraten von 308kbit/s bzw. 900kbit/s).

Des Weiteren wurden unterschiedliche Vorgehensweisen bei der Komprimierung verwendet. Es wurden alle sechs Videosequenzen mit einer Ein-Durchgangs-Kodierung (im weiteren Verlauf *IDK* genannt), mit einer Zwei-Durchgangs-Kodierung² (im weiteren Verlauf *2DK* genannt) und mit einer Ein-Durchgangs-Kodierung mit einer wenig variablen Videodatenrate (im weiteren Verlauf *IDKf* genannt) komprimiert.

Die Länge der Videobildgruppe (GoP-Länge) wurde beim Encoder auf der Basis von DVB-Vorgaben [158] und von Angaben in [15] und [159] auf einen konstanten Wert von 12 eingestellt. Dies entspricht einer Zeitspanne von 0,5 s.

Die Struktur der Videobildgruppe kann unterschiedlich gestaltet werden. Eine häufige Praxis ist die Verwendung der zwei unmittelbar aufeinander foldenden *B*-Videobilder (im folgenden *2B* genannt) [15], [159]. Auf der anderen Seite bieten führende Hersteller von Netzwerküberwachungskameras Profile ohne *B*-Videobilder [160]. Der Grund ist eine kleinere Encoder-Verzögerung (aufgrund der nicht vorhandenen bidirektionalen Videobildabhängigkeiten und des kleineren Rechenaufwands). Diese Videostruktur wird im Folgenden *OB* genannt. Eine Videostruktur ohne *P*-Videobilder ist ihrerseits dadurch interessant, dass bei einem Videobildverlust zwischen zwei *I*-Videobildern keine weiteren Videobilder aufgrund der Videobildabhängigkeiten Schaden nehmen. Diese Videostruktur wird im weiteren Verlauf *11B* genannt. Für jede Videosequenz wurden komprimierte Versionen mit unterschiedlich langen *B*-Serien erstellt. Die sich ergebenden Videostrukturen sind in der Tabelle 6.2 dargestellt. Man beachte, dass für bestimmte *B*-Serien-Längen eine GoP-Länge von 12 nicht möglich ist.

Die Abbildung 6.8 zeigt zusammenfassend, welche Komprimierungsdimensionen für die Videosequenzen im Rahmen der Vorauswahl verwendet wurden.

²Eine im Vergleich zur Ein-Durchgangs-Kodierung effizientere Methode im Bezug auf die erzielte Videoqualität.

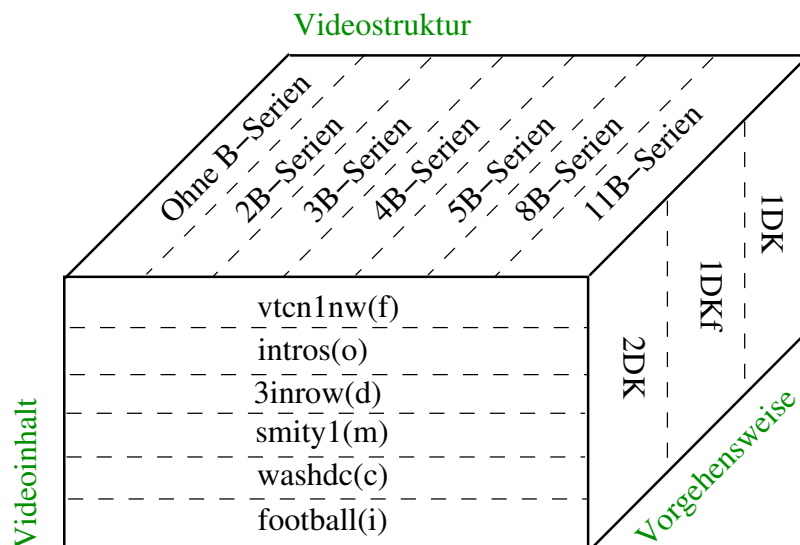


Abbildung 6.8: Der vorausgewählte Videosequenzraum (zur weiteren Eingrenzung)

6.1.3 PSNR

Die Abbildungen 6.9, 6.10 und 6.11 sind Box-Plot-Diagramme, die die PSNR-Verteilungen der vorausgewählten komprimierten Videosequenzen mit Hilfe des ersten, zweiten und dritten Quartils und der entsprechenden Minimum- und Maximumwerte in Abhängigkeit von der Länge der *B*-Serien darstellt. Die mittlere Videodatenrate beträgt 308 kbit/s.

Es lässt sich feststellen, dass der Videosequenzsatz ein breites Videoqualitätsspektrum aufweist. Die Videoqualität ist im Allgemeinen bei *2DK* am größten und bei *1DKf* am kleinsten. Grob kann man insbesondere anhand der Abbildung 6.11 drei Qualitätsgruppen erkennen: *vtc1nw*, *intros-3inrow-smity1* und *washdc-football*. Die Videosequenz *vtc1nw* wird somit für die Verfahrensuntersuchungen ausgewählt.

Des Weiteren lassen sich in vielen Fällen keine signifikanten Unterschiede zwischen *2B*-, *3B*-, *4B*-, *5B*- und *8B*-Kodierung erkennen, so dass man sich auf die verbreitete *2B*-Variante bei den weiteren Untersuchungen beschränken kann. Insbesondere in der Abbildung 6.10 sieht man, dass die *0B*-Kodierung ebenfalls zu untersuchen ist, zumal sie im Überwachungskamerabereich vorkommen kann. Für die Verfahrensuntersuchungen werden somit die Kodierungen *0B*, *2B* und *11B* (vgl. Abschnitt 6.1.2) ausgewählt.

6.1.4 Datenumfang in den Videobildarten

Die Abbildungen 6.12 bis 6.20 zeigen die Verteilung der Daten (*y*-Achse) auf einzelne Videobildtypen (*x*-Achse). Die mittlere Videodatenrate beträgt 308 kbit/s. Insbesondere anhand der *2DK*-Verhältnisse (Abbildungen 6.18, 6.19 und 6.20) lässt sich – neben der bereits ausgewählten *vtc1nw*-Sequenz – *football* als die zweite zu untersuchende Videosequenz auswählen.

Mit Hilfe der Abbildung 6.13 lässt sich die Videosequenz *intros* aufgrund ihrer besonderen Datenaufteilung aus dem vorher bestimmten PSNR-Cluster (*intros-3inrow-smity1*) für die Verfahrensuntersuchungen bestimmen.

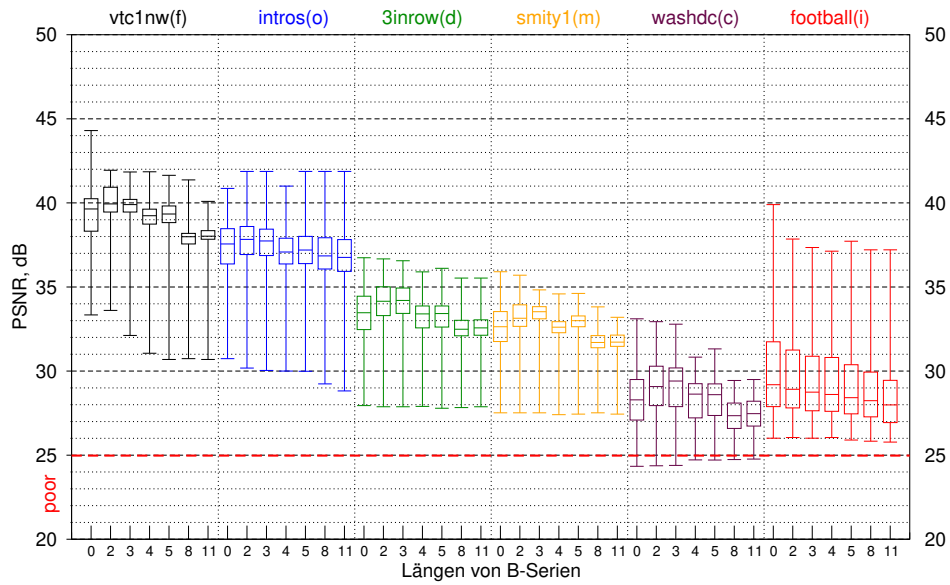


Abbildung 6.9: PSNR, IDK

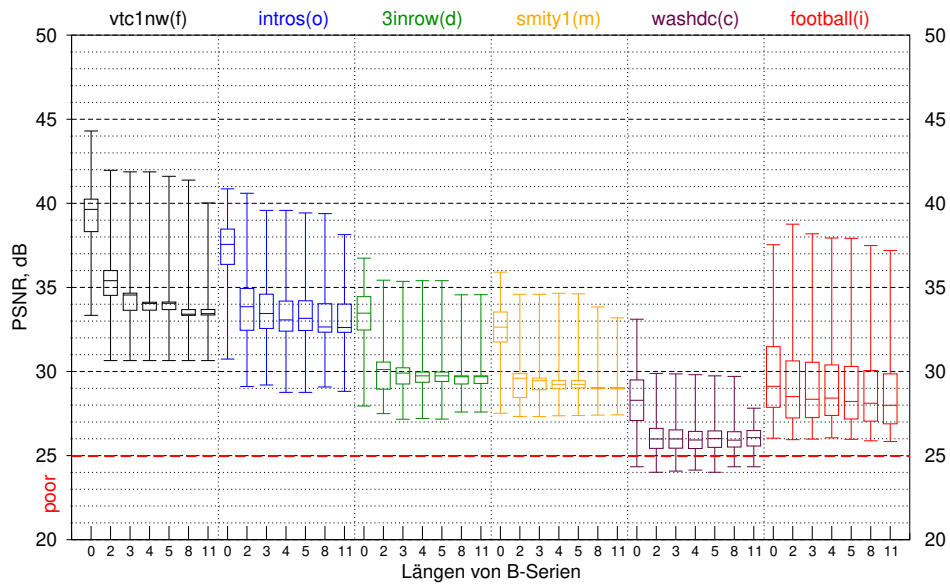


Abbildung 6.10: PSNR, IDKf

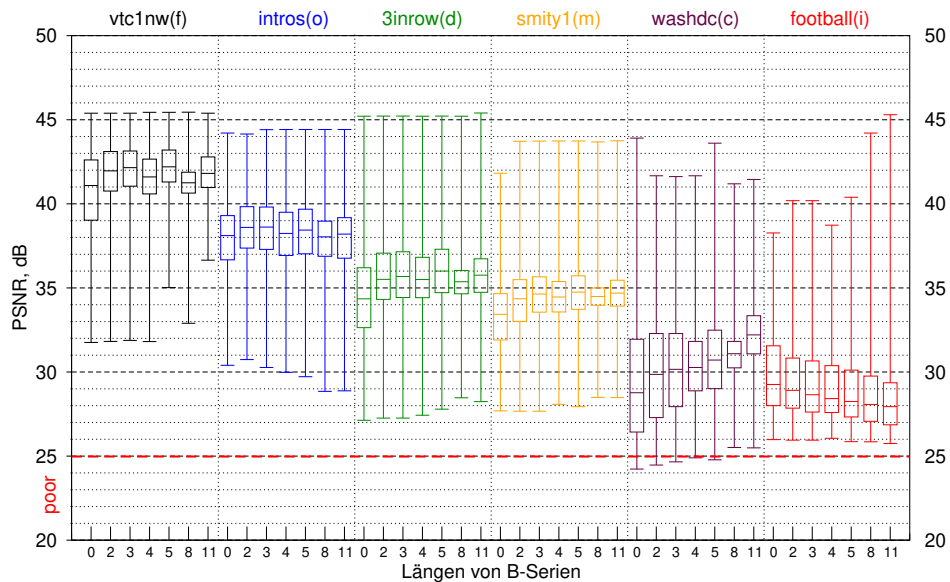


Abbildung 6.11: PSNR, 2DK

6.1.5 Datenraten

In den Abbildungen 6.22 bis 6.24 sind Verteilungen von Videobildraten dargestellt. Die mittlere Videodatenrate beträgt 308kbit/s. Im *2DK*-Fall (Abbildung 6.24) sind die Datenraten der *P*- und *B*-Videobilder viel kleiner als die Datenraten der *I*-Videobilder (die bis zu ca. 10000kbit/s reichen). Deshalb liegen die mittleren 50% der Daten viel tiefer als bei *IDK* und *IDKf*. Das bedeutet, dass die abhängigen Videobilder bei *2DK* wenig Daten transportieren. Aus Platzgründen zeigen Diagramme keine Maximalwerte. Für die vollständigen Darstellungen sei auf den Anhang verwiesen.

In den beiden *IDK*-Fällen sind die Datenratenunterschiede zwischen den unabhängigen und abhängigen Videobildern weniger stark als bei *2DK*. Der *IDK*-Fall weist breitere Verteilungen der mittleren 50% der Daten als der *IDKf*-Fall auf. Das bedeutet, dass bei *IDKf* die Daten gleichmäßiger auf die Videobildgruppe verteilt sind. Die Videosequenz *football* bildet aufgrund ihrer hohen *SI*- und *TI*-Werte eine Ausnahme. Bei höheren mittleren Datenraten weist sie ähnliche Merkmale wie die anderen hier untersuchten Videosequenzen.

6.1.6 Zusammenfassung

Für die Verfahrensuntersuchungen werden die drei in der Abbildung 6.21 zu sehenden Videosequenzen ausgewählt. Sie wurden anhand des *SI-TI*-Raumes und ihrer Datenmerkmale als repräsentativ identifiziert.

Die Strukturen *OB* (u. a. Überwachungsnetzwerkcameras), *2B* (Broadcastsbereich), *11B* (Reduzierung der Videobildabhängigkeiten und somit der Fehlerfortpflanzung) wurden als die wichtigen zu untersuchenden Strukturen ausgewählt.

Die *IDK*-, *IDKf*- und *2DK*-Vorgehensweisen sollen als weitere Dimension untersucht werden. *IDK* und *2DK* können bei Live-Streaming bzw. Abrufvideodiensten vorliegen. *IDKf* reduziert die Variabilität der Videodatenraten innerhalb einer Videobildgruppe.

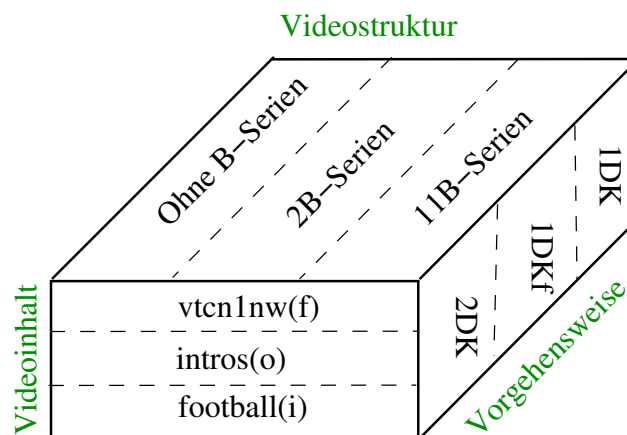


Abbildung 6.21: Der ausgewählte Videosequenzraum für die Verfahrensuntersuchungen

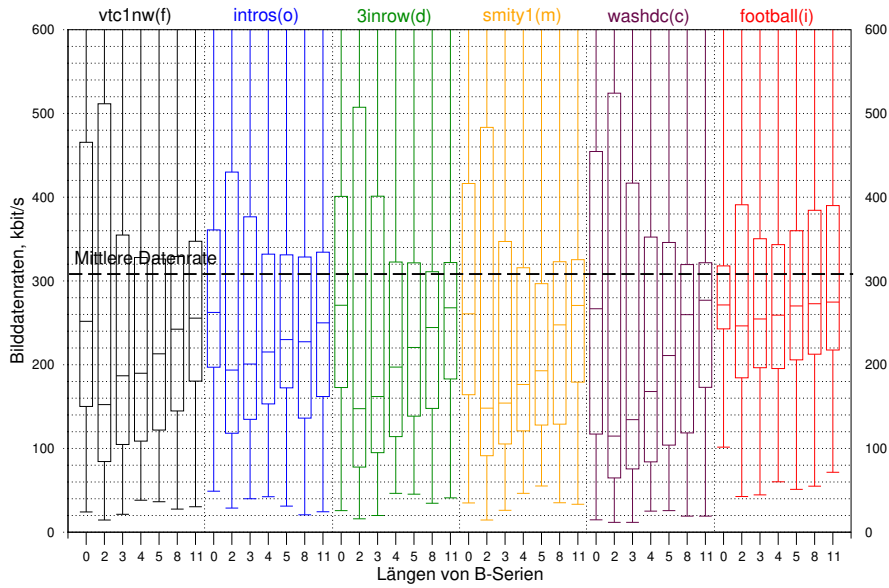


Abbildung 6.22: Datenraten, Ausschnitt, *IDK*

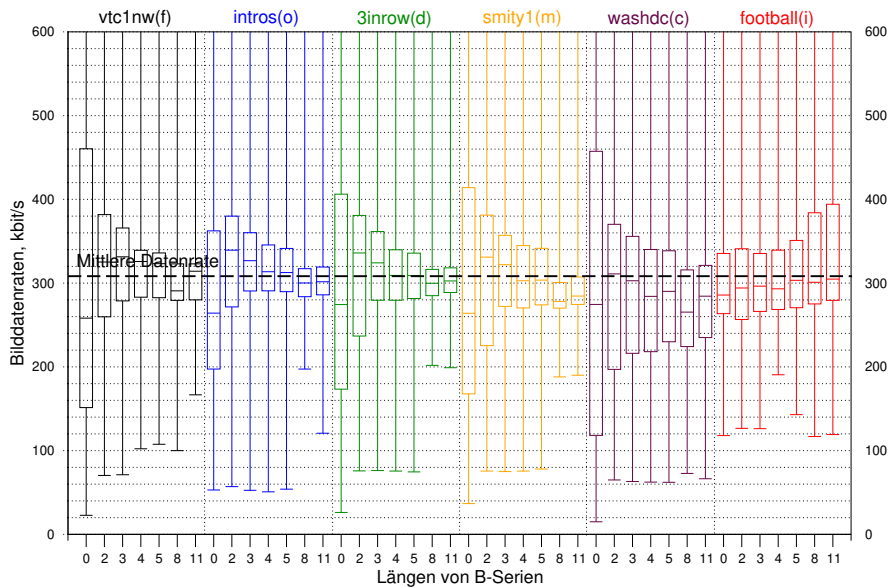


Abbildung 6.23: Datenraten, Ausschnitt, *IDKf*

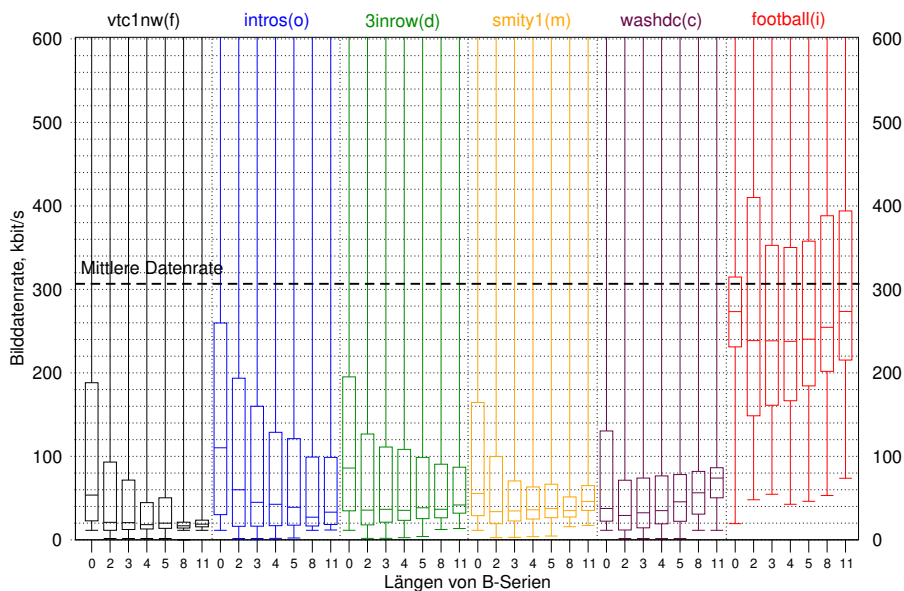


Abbildung 6.24: Datenraten, Ausschnitt, *2DK*

6.2 Systemmodell

6.2.1 Systembeschreibung

Für die Verfahrensuntersuchungen wird ein UTRAN-Emulator für eine Zelle verwendet. Die Untersuchungsszenarien weisen mehrere mobile Endgeräte auf, die mit der Basisstation über einen gemeinsamen Kanal (HS-DSCH³-Kanal) in der Downlink-Richtung und über einen exklusiven Kanal (DCH⁴-Kanal) in der Uplink-Richtung verbunden sind. Das RNC-Modell weist einen Eingangspuffer für die eintreffenden IP-Pakete auf. Das Mobilfunknetz und das Internet werden mit einer konstanten Verzögerung von 20 ms und verlustlos nachgebildet.

6.2.2 Emulationsmodell

Das emulierte UTRAN-Modell ist stark detailliert. Es sind alle relevanten Sicherungsschichtprotokolle (RLC, MAC-d⁵, und MAC-hs) nachgebildet. Das Modell der physikalischen Schicht beruht auf BLER⁶-Kurven eines realitätsnahen urbanen UTRAN-Kanals und wurde im Rahmen eines früheren Forschungsprojektes von einem namhaften Industriepartner zur Verfügung gestellt. Die sich ergebenden Transportformate der Sicherungsschicht (MAC-hs) basieren auf den BLER-Werten von ca. 10 %. Für den Transport der CQI⁷-Nachrichten von mobilen Geräten zur Basisstation wurden ideale, verzögerungsfreie Bedingungen angenommen. Die Flusskontrolle zwischen RNC und der Basisstation (Iub-Flusskontrolle) verwendet kurze Update-Intervalle und weist keine Totzeiten auf [161]. Die maximale Anzahl der MAC-hs-Retransmissions wurde auf 4 begrenzt. Die RLC-Schicht wurde im unacknowledged-mode betrieben. In der Untersuchung wurde der Einfluss der Convergence-Schicht (Teil der Sicherungsschicht oberhalb von RLC) vernachlässigt, da sie in einer Einzellumgebung nur einen geringen Overhead erzeugt. Die detaillierte Beschreibung des verwendeten UTRAN-Modells ist in [139] nachzulesen.

Die Simulation wurde mit Hilfe der IKR-Simulationsbibliothek (IKR-SimLib) [162] implementiert. Die Emulation entstand mit Hilfe von IKR-EmuLib [140]. Die Implementierung des UTRAN-Emulators ist nicht der Bestandteil dieser Arbeit. Im Rahmen dieser Untersuchungen wurde der Emulator um die Komponenten der eingebetten netzgesteuerten Adaption erweitert.

6.2.3 Videostreamingmodell

Der Videoverkehr wird tracebasiert [137] emuliert (s. Kapitel 5). Auf der Ebene der IP-Pakete entspricht der emulierte Videoverkehr einem realen MPEG-4-basierten Videostrom, der mit einem UDP-basierten Videostreaming zustande käme. Die für die AQM-Verfahren notwendigen Metadaten werden in den erweiterten IP-Paketheadern übertragen. Das Umschalten zwischen den verschiedenen Videoqualitätsstufen findet mit der Bitstream-Switching-Methode [75] an den *I*-Videobildern statt.

³HS-DSCH: High Speed Downlink Shared Channel

⁴DCH: Dedicated Transport Channel

⁵MAC-d: Medium Access Control For Dedicated Transport Channel

⁶BLER: Block Error Rate

⁷CQI: Channel Quality Indicator

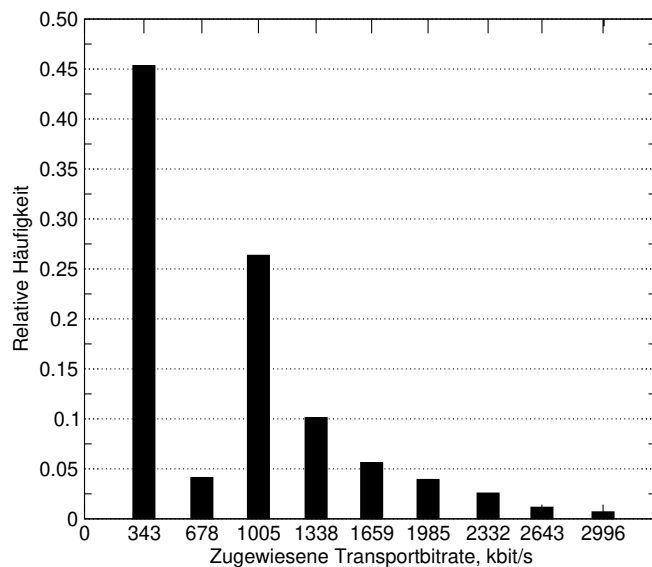


Abbildung 6.25: Kanalverlaufshistogramm

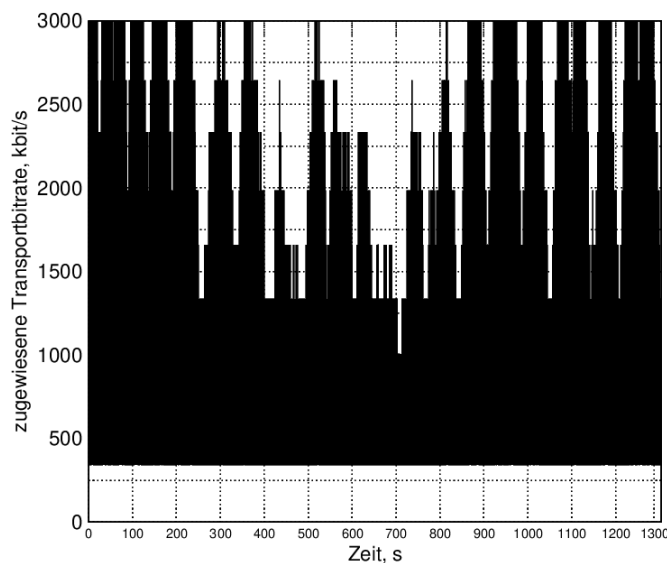


Abbildung 6.26: Kanalverlauf

6.2.4 Videobewertungsmodell

Ein Videobild wird als verloren angenommen, wenn ein beliebiges dazu gehörendes IP-Paket verloren geht. Die ist eine verbreitete Auswertungsmethode bei tracebasierten Untersuchungen [163]. Die Autoren von [8], [164], [127] und [163] betonen, dass das Dekomprimieren nicht-vollständiger Videobilder von der Art der nichtempfangenen Daten und der Implementierung des Decoders abhängt. Insofern bedeutet die Annahme eines Videobildverlustes beim Verlust der Videobildteile eine Worst-Case-Abschätzung. Da beim proaktiven Verwerfen auf Basis der Videobilder ganze Videobilder entfernt werden, entspricht die besagte Annahme den tatsächlichen, in diesem Fall aufgetretenen Videobildverlusten. Alle vom verlorenen Videobild abhängigen Videobilder werden beim Dekomprimieren ebenfalls als verloren angenommen. Anstelle eines verlorenen Videobildes wird das letzte korrekt empfangene Videobild verwendet. Die Rechtzeitigkeit der Videobildankünfte im Empfänger wird mit Hilfe des kumulativen Jitter [165], [165], [8] bewertet. Die durch das adaptive Umschalten verursachten Videobildverluste aufgrund der Videobildabhängigkeiten werden herausgerechnet, da sie implementierungsbedingt sind und durch spezielle Maßnahmen in der Videokodierung bzw. in der Videobildübertragung eliminiert werden können.

6.2.5 Grundszenario

Der mobile Teilnehmer, für den der emulierte Videostream bestimmt ist, bewegt sich mit 30 km/h. Je nach Untersuchung werden weitere CBR⁸-erzeugte Videostreams als konkurrierender Videoverkehr von mobilen Teilnehmern simuliert. Sie starten an unterschiedlichen Stellen des Mobilfunkkanal-Trace. Bei mehreren Teilnehmern wird der Proportional-Fair-Scheduler (MAC-hs-Schicht) [166] verwendet.

⁸CBR: Constant Bit Rate

Der Mobilfunkkanal-Trace wird vom mobilen Teilnehmer innerhalb eines Emulationslaufes (22 Minuten) in Echtzeit zwei Mal durchlaufen (hin und zurück). Dabei werden je nach Untersuchungsvideosequenz die Grundsequenz *football* 255 Mal, die Grundsequenzen *intros* und *vtcInw* jeweils 110 Mal übertragen.

In der Abbildung 6.25 ist das Histogramm der Transportbitraten für $2ms\text{-TTI}$ ⁹ zu sehen. Das zeitliche Auftreten der Transportbitraten ist in der Abbildung 6.26 dargestellt. Der resultierende Kanalverlauf ist aufgrund der beschriebenen Hin- und Rückbewegung in etwa spiegelsymmetrisch. Die Statistikkenwerte der Transportbitratenverteilung sind Tabellen 6.3 und 6.4 zu entnehmen. Das Minimum beträgt 343 kbit/s, der Mittelwert liegt bei 867,6 kbit/s. Die Standardabweichung beträgt 601,5 kbit/s. Der Medianwert liegt bei 1005 kbit/s. Die Nettodatenraten sind um ca. 10% kleiner. Daher sind die Videosequenzen mit 308 kbit/s und 900 kbit/s komprimiert.

Tabelle 6.3: Statistikkenwerte der Transportbitraten (kbit/s)

<i>Minimum</i>	<i>Maximum</i>	<i>Mittelwert</i>	<i>Standardabweichung</i>	<i>Varianz</i>	<i>Variationskoeffizient</i>	<i>quadrierter Variationskoeffizient</i>
343	2996	867,6	601,5	$3,618 \cdot 10^5$	0.6933	0.4806

Tabelle 6.4: Statistikkenwerte (Quantile) der Transportbitraten (kbit/s)

Q_1	Q_2 (Median)	Q_3	$Q_{0,05}$	$Q_{0,95}$	$Q_{0,99}$	$Q_{0,999}$
343	1005	1005	343	1985	2643	2996

6.2.6 Paketierung

Im LAN-Umfeld beträgt die maximale IP-Paketgröße in der Regel 1500 Byte. Mit dieser Paketgröße werden viele Video Streaminguntersuchungen durchgeführt (z. B. [126] und [127]). Auf der anderen Seite ist es naheliegend, dass bei kleineren Paketen weniger Bytes beim IP-Paketverlust verloren gehen. Deshalb verwenden manche Autoren für die Untersuchungen im drahtlosen Bereich kleinere Paketgrößen (z. B. 1000 Byte in [128]).

Für alle hier verwendeten Videosequenzen wurde der Protokolloverhead (s. Abschnitt 5.2) für die maximalen Nutzlasten von 512, 1024 bzw. 1456 Byte berechnet (s. Tabelle A.2 im Anhang). Der Auszug daraus ist für den Fall *IDK 2B* in der Abbildung 6.27 graphisch dargestellt.

Es wurden Testmessungen in einer Überlastsituation durchgeführt. Die relativen Videobildverluste für den Fall *IDK 2B* sind der Abbildung 6.28 zu entnehmen. Aufgrund des größeren Overheads bei der gleichbleibenden Videodatenrate steigt die Brutto-Videodatenrate, was zu mehr IP-Paketverlusten führt. Das führt seinerseits zur größeren Videobildverlusten.

Für die Verfahrensuntersuchungen wurde die maximale IP-Paketgröße auf 1500 Byte festgelegt.

⁹TTI: Transmission Time Interval

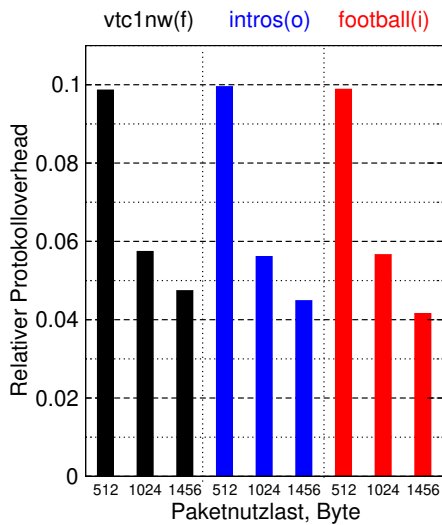


Abbildung 6.27: Protokolloverhead

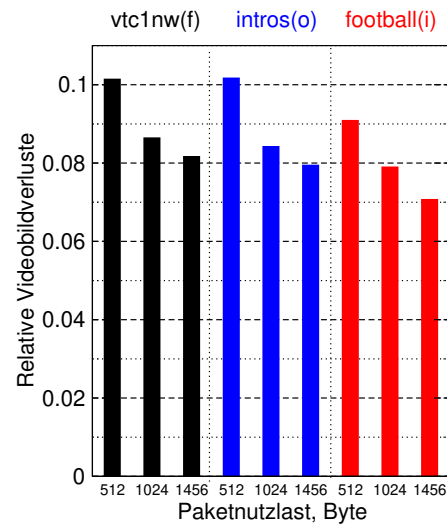


Abbildung 6.28: Videobildverluste

6.3 Verfahrensergebnisse

6.3.1 Netzgesteuerte Datenratenadaption

6.3.1.1 Szenario

Es soll die Fähigkeit zur Adaption bewertet werden, wenn der Videostreamingserver die Adaptionsvorschläge aus dem Netz erhält. Untersucht wird die Adaption an die übertragungstechnischen Auswirkungen der sich schnell ändernden Wellenausbreitungssituation, d. h. an den langsamen und schnellen Schwund¹⁰.

Für die Bewertung wird angenommen, dass der Videoclient einen Abspielpuffer von 6 s besitzt [167]. Der Eingangspuffer im RNC ist 212000 Byte groß. Die in den Puffer passende Datenmenge entspricht somit einem Abspielintervall zwischen 5,25 s und 5,3 s (je nach sequenzspezifischem Overhead) für eine 308 kbit/s-Videsequenz.

6.3.1.2 Bestimmung des Hysterese-Faktors

Der Hysterese-Faktor k_{hys} (vgl. Abschnitt 4.2.3) wurde mit dem Ziel eingeführt, ein zu häufiges Hin- und Herschalten zwischen den Videoqualitätstufen zu vermeiden. Das System darf trotzdem nicht zu träge sein, um die vorhandenen Übertragungsressourcen effizient auszunutzen, wenn diese zunehmen. Des Weiteren soll es schnell genug sein, um auf den Entzug der Ressourcen zu reagieren. Sonst drohen Videodatenverluste.

Die Untersuchung zur Bestimmung des zu verwendenden Hysterese-Faktors wurde mit den drei Videosequenzen *football*, *intros* und *vtc1nw* in der gängigsten Ausprägung *IDK 2B* durchgeführt. Die Videosequenzen standen dem Videostreamingserver in zwei Qualitätsstufen zur Verfügung: 308 kbit/s-Variante (Basisqualität) und 900 kbit/s-Variante (hohe Qualität).

¹⁰Ein dynamisches Zuschalten von konkurrierenden mobilen Teilnehmern wird vom Emulator nicht unterstützt.

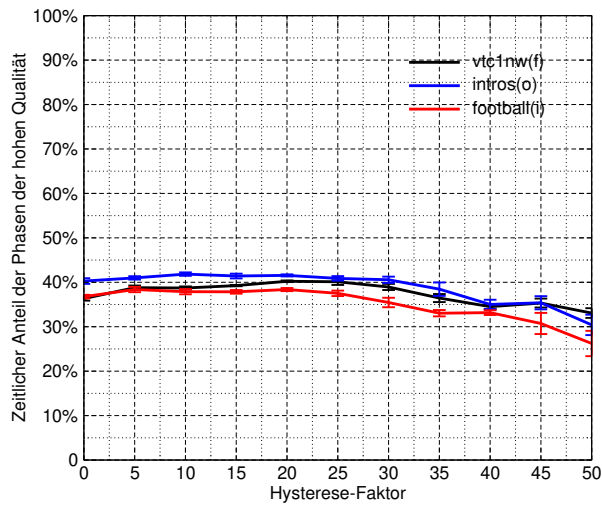


Abbildung 6.29: Anteil der hohen Qualität

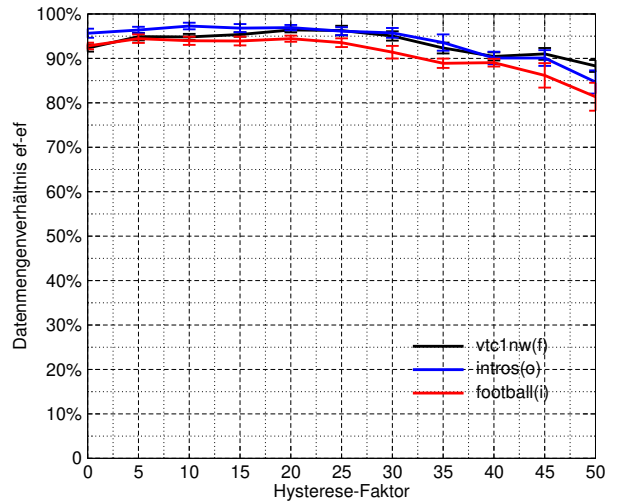


Abbildung 6.30: Verhältnis DR_{ef-ef}

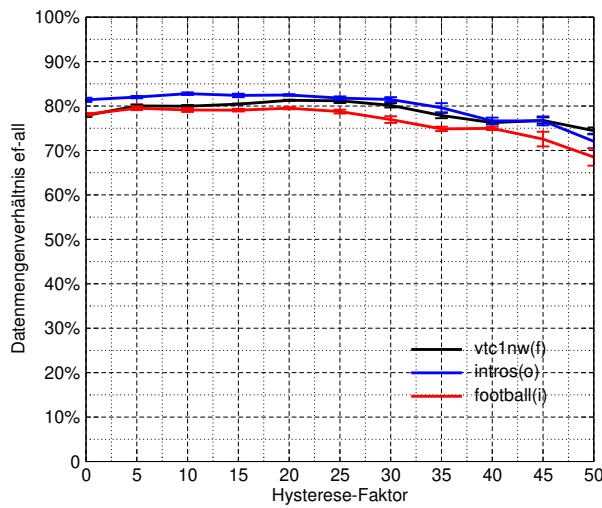


Abbildung 6.31: Verhältnis DR_{ef-all}

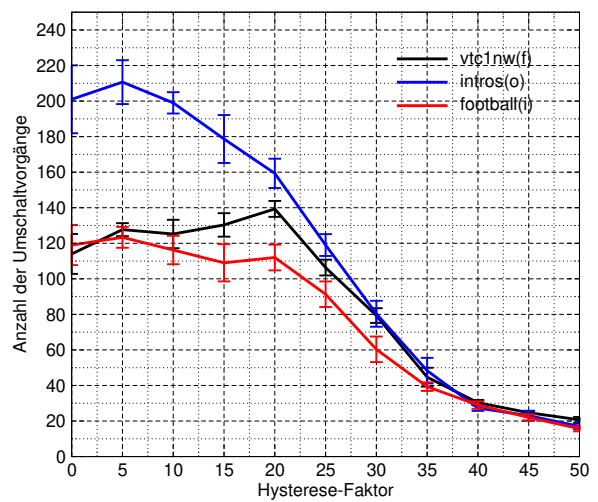


Abbildung 6.32: Umschaltvorgänge

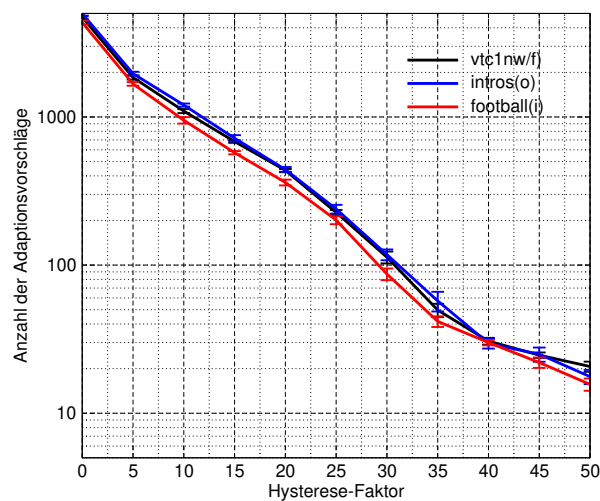


Abbildung 6.33: Adaptionsvorschläge

Die Abbildung 6.29 zeigt den erzielten zeitlichen Anteil der Phasen der hohen Qualität. Er beginnt tendenziell ab $k_{hys} = 30$ zu sinken. Auch treten ab $k_{hys} = 30$ vereinzelt (im Promillebereich) Videobildverluste im Netz auf.

Die Abbildung 6.30 visualisiert das Verhältnis der Datenmenge der *fehlerfrei* empfangenen Videobilder im Adaptionenfall zur Datenmenge der *fehlerfrei* empfangenen Videobilder im adaptionslosen Fall DR_{ef-ef} . Im adaptiven Fall treten im Allgemeinen bis zu $k_{hys} = 30$ keine Videobilddatenverluste im Netz auf. Der Videoinhalt wird fehlerfrei abgespielt. Im adaptionslosen Fall treten massive Videobildverluste auf, da viele Videobilddaten bei der Übertragung verloren gehen (um die 35 % der Videobilder gehen im Netz verloren). Das Verhältnis der Datenmengen liegt bis zu $k_{hys} = 35$ über 90 %.

Die Abbildung 6.31 zeigt das Verhältnis der Datenmenge der *fehlerfrei* empfangenen Videobilder im Adaptionenfall zur Datenmenge *aller* im adaptionslosen Fall empfangenen Videobilddaten DR_{ef-all} . Im adaptionslosen Fall gehen ca. 25 % der Videobildbytes im Netz verloren. Im Adaptionenfall kommen die Daten im Allgemeinen bis zu $k_{hys} = 30$ fehlerfrei an. DR_{ef-all} liegt bis $k_{hys} = 30$ bei ca. 80 %.

Die Abbildung 6.32 veranschaulicht die stattgefundenen Umschaltvorgänge zwischen der Phasen der hohen Qualität und der Basisqualität. Bis $k_{hys} = 20$ liegt die Zahl für *vtc1nw* und *football* um die 110. Bei *intros* ist sie höher. Die Ursache dafür wird in den späteren Abschnitten erläutert.

Allgemein lässt sich feststellen, dass der Hysterese-Faktor zwischen 0 und 30 zu wählen ist. In den Grenzbereichen (0 bis 9 bzw. 20 bis 30) kann es vereinzelt zu sehr seltenen Videobildverlusten kommen. Für die weitere Untersuchung wurde $k_{hys} = 10$ festgelegt.

Der Abbildung 6.33 ist die Anzahl der Adaptionen vorschläge in logarithmischer Darstellung zu entnehmen. Die Kurvenverläufe sind negativ exponentiell. Bei $k_{hys} = 10$ beträgt der Wert ca. 1000 (für die Übertragungsdauer einer Videosequenz von 22 Minuten).

Es lässt sich feststellen, dass das Umschalten an den *I*-Videobildern zu einer viel kleineren Anzahl an Umschaltvorgängen führt, als es Adaptionen vorschläge gibt. Beträgt die GoP-Länge 12 und sind die Videostrukturen der unterschiedlichen Qualitätsstufen durchgehend identisch, so kann eine Umschaltung bei diesem Verfahren alle 0,48 s stattfinden. In dieser Zeit können mehrere Adaptionen vorschläge ankommen, die sich gegenseitig aufheben. Bei einer entsprechenden Dimensionierung des RNC-Eingangspuffers stellt das für den Fall eines schnellen Entzugs der Übertragungsressourcen keine Einschränkung dar.

6.3.1.3 Bewertung der Videoqualität

Die Bewertung der Videoqualität im Videostreamingkontext soll neben der Bildqualität auch die Kontinuität des Abspielens betrachten.

Die Kontinuität des Abspielens wird im Folgenden exemplarisch anhand der gängigen *IDK-2B*-Ausprägung beurteilt. Die anderen Ausprägungen weisen das gleiche Verhalten auf.

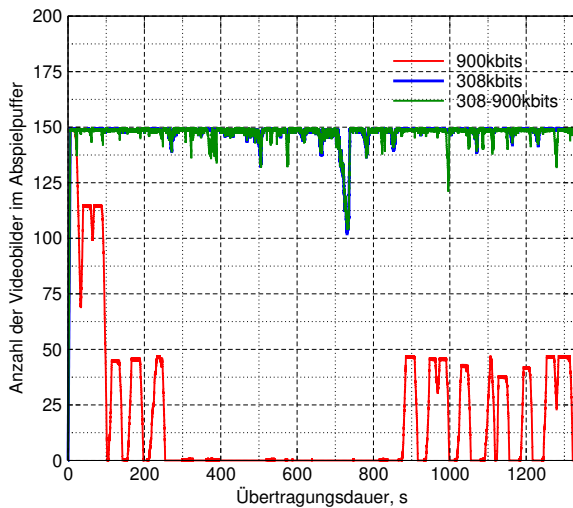


Abbildung 6.34: *football 1DK 2B*, Abspielpuffer

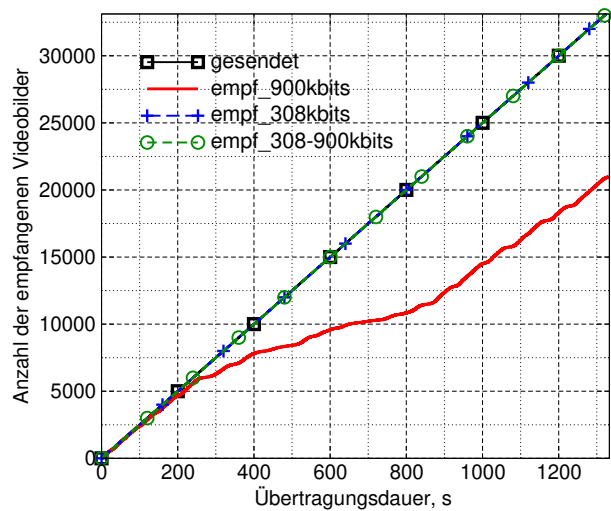


Abbildung 6.35: *football 1DK 2B*, empfangene Videobilder

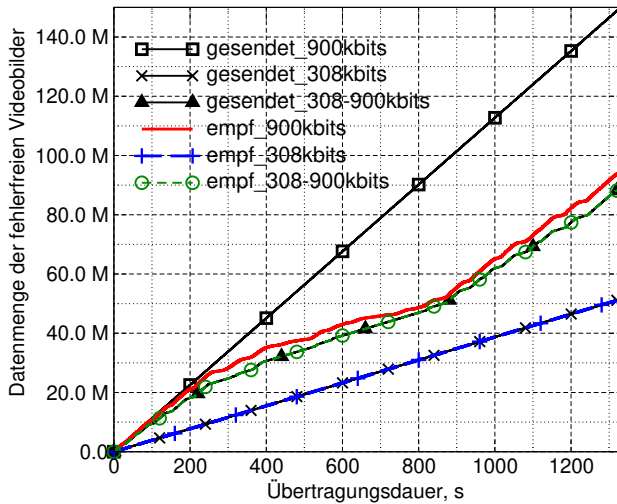


Abbildung 6.36: *football 1DK 2B*, Datenmenge der fehlerfrei empfangenen Videobilder

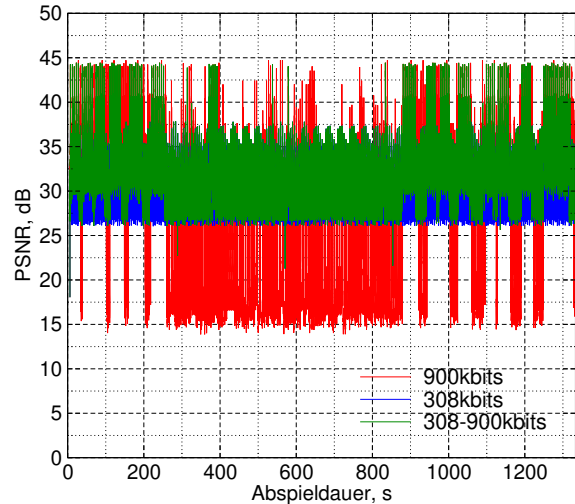


Abbildung 6.37: *football 1DK 2B*, PSNR-Verlauf

Die Abbildungen 6.34, 6.38 und 6.42 zeigen den Füllstand des Abspielpuffers im mobilen Endgerät. 150 Videobilder entsprechen einem Abspielzeitintervall von 6 s bei der Videobildrate von 25 Videobildern/s. Es ist der Verlauf für die zwei adaptionlosen Fälle – 308 kbit/s-Variante (Basisqualität) und 900 kbit/s-Variante (hohe Qualität) – und für den Adaptionfall (im Bild 308 – 900 kbit/s) zu sehen. Es treten keine Pufferleerläufe für den adaptionlosen 308 kbit/s-Fall auf. Beim adaptionlosen 900 kbit/s-Fall kann man drei Hauptphasen erkennen: Von 0 s bis ca. 250 s und von knapp 900 s bis fast zum Ende wird der Client zeitweise mit den Videobildern regelmäßig versorgt. Die maximale Größe der Pufferfüllstände in diesen Phasen entspricht der Speicherkapazität¹¹ des RNC-Eingangspuffers. In der mittleren Phase (von ca. 250 s - bis knapp 900 s) lässt die Kanalqualität keine regelmäßige Versorgung mit den Videobildern hoher

¹¹Bei 900 kbit/s ohne Berücksichtigung des Protokolloverheads ergibt sich eine $Videobildanzahl = ((212000 \text{ byte} \cdot 8 \text{ bit}) : 900000 \text{ byte/s}) \cdot 25 \text{ Videobilder/s} = 47,11 \text{ Videobilder}$

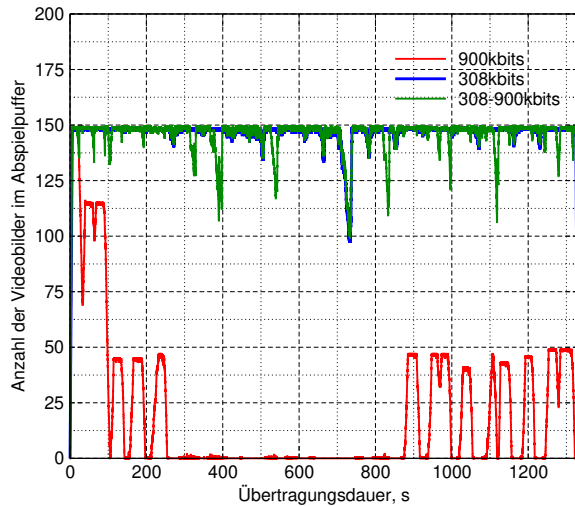


Abbildung 6.38: *intros* 1DK 2B, Abspielpuffer

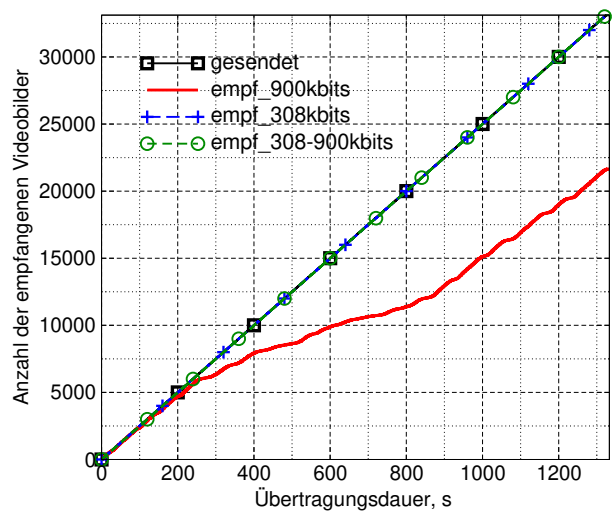


Abbildung 6.39: *intros* 1DK 2B, empfangene Videobilder

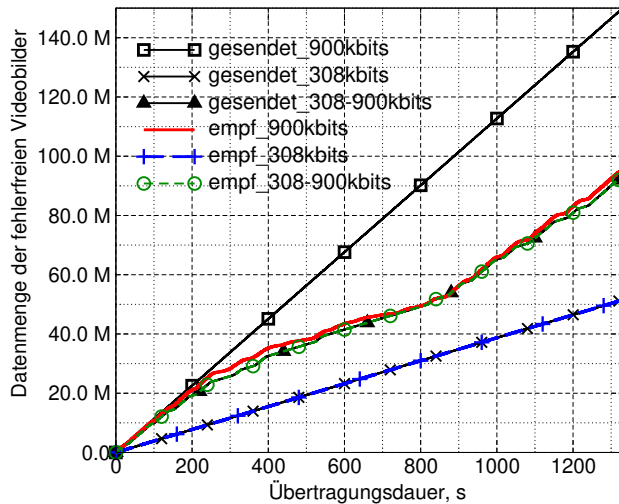


Abbildung 6.40: *intros* 1DK 2B, Datenmenge der fehlerfrei empfangenen Videobilder

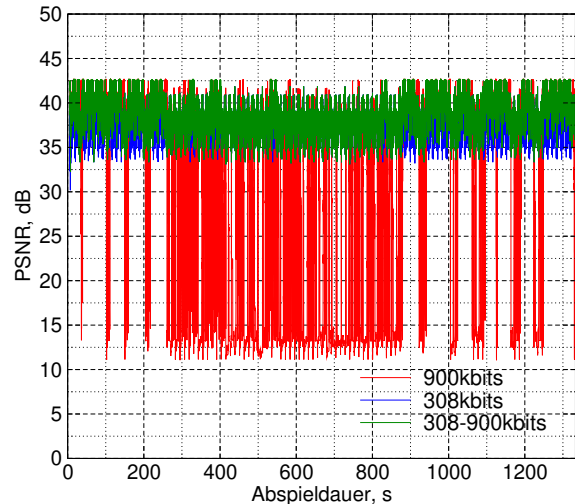


Abbildung 6.41: *intros* 1DK 2B, PSNR-Verlauf

Qualität zu. Im adaptiven Fall treten keine Pufferleerläufe auf. Der adaptive Fall überdeckt in den Diagrammen den adaptionslosen 308kbit/s-Fall (die blaue Kurve wird überdeckt).

Die Abbildungen 6.35, 6.39 und 6.43 visualisieren die Anzahl der fehlerfrei empfangenen Videobilder. Im adaptionslosen 308kbit/s-Fall werden alle gesendeten Videobilder empfangen. Die Kurve des adaptionslosen 900kbit/s-Falls zeigt die drei zuvor festgestellten Phasen. Die Verflachung der Steigung bedeutet Videobildverluste. Am rechten Rand (das Ende der Übertragung) beträgt die Differenz zwischen den gesendeten Videobildern und den fehlerfrei empfangenen Videobildern ca. 35%. Vor allem aufgrund der Fehlerfortpflanzung beträgt die Gesamtvideobildverlustrate knapp 55%. Im adaptiven Fall werden alle Videobilder fehlerfrei ausgeliefert.

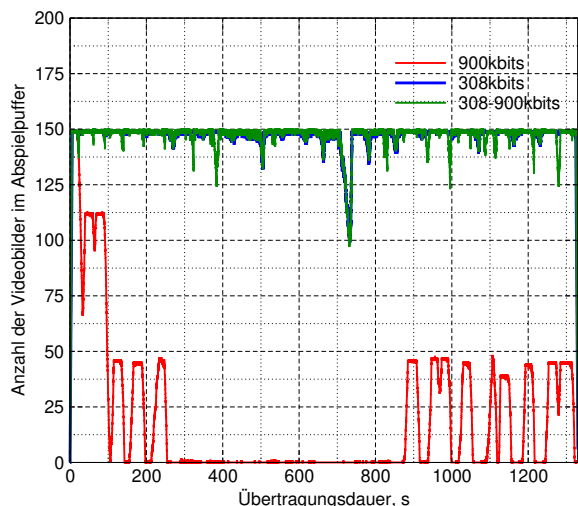


Abbildung 6.42: *vtc1nw IDK 2B*, Abspielpuffer

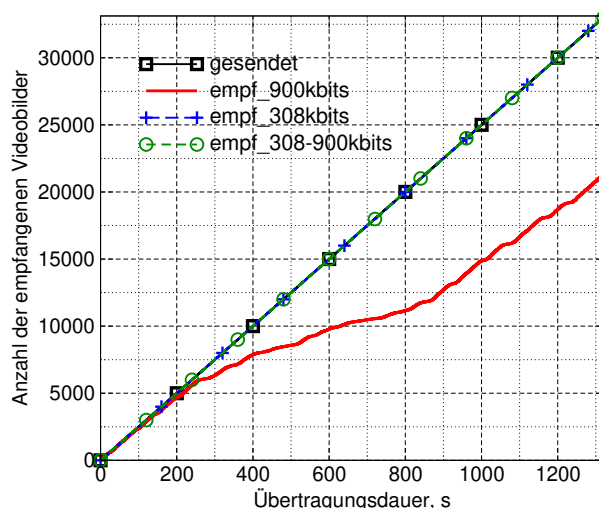


Abbildung 6.43: *vtc1nw IDK 2B*, empfangene Videobilder

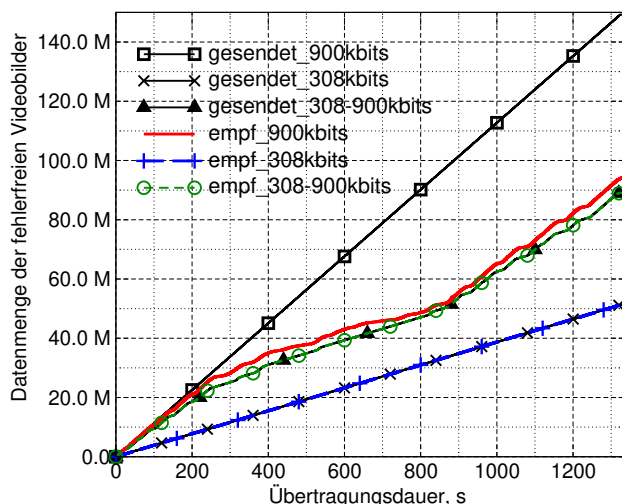


Abbildung 6.44: *vtc1nw IDK 2B*, Datenmenge der fehlerfrei empfangenen Videobilder

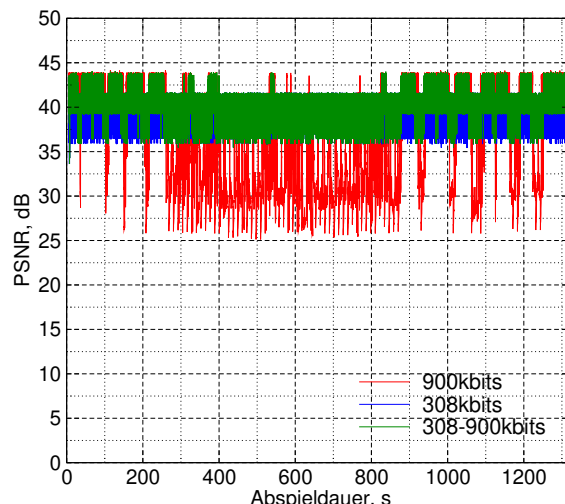


Abbildung 6.45: *vtc1nw IDK 2B*, PSNR-Verlauf

Die Abbildungen 6.36, 6.40 und 6.44 stellen die Datenmengen der fehlerfrei empfangenen Videobilder über die Übertragungszeit dar. Im adaptionslosen 308 kbit/s-Fall werden alle Videobilder fehlerfrei empfangen. Im adaptionslosen 900 kbit/s-Fall treten Videobildverluste auf. Im Adaptionsfall verläuft die Übertragung fehlerfrei. Das Verhältnis DR_{ef-ef} zwischen der 900 kbit/s-Empfangskurve und der Empfangskurve des adaptiven Falls liegt über 90%. Im Abschnitt 6.3.1.4 wird das DR_{ef-ef} gesondert betrachtet.

Die Abbildungen 6.37, 6.41 und 6.45 zeigen den PSNR-Verlauf. Im adaptionslosen 308 kbit/s-Fall ist er durchgängig gleich. Des Weiteren ist er periodisch mit der Dauer der Grundsequenz. Im adaptionslosen 900 kbit/s-Fall sind PSNR-Einbrüche durch das verlustbedingte Einfrieren des zuletzt fehlerfrei empfangenen Videobildes zu erklären. Die Tiefe der Einbrüche ist unterschiedlich, worauf gleich bei der Betrachtung von PSNR-Verteilungen eingegangen wird.

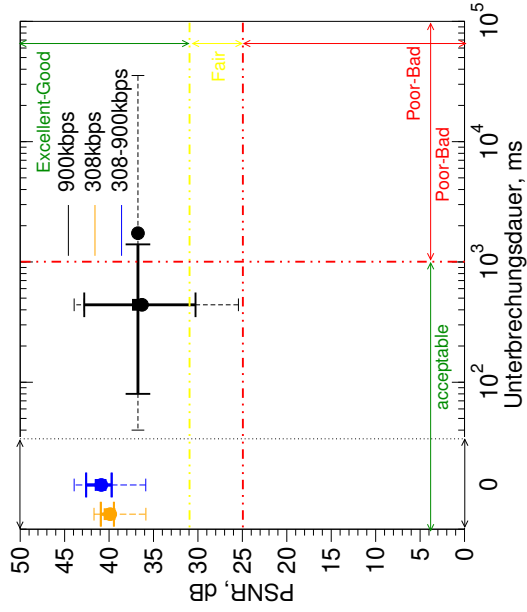


Abbildung 6.46: *football* IDK 2B, PVU

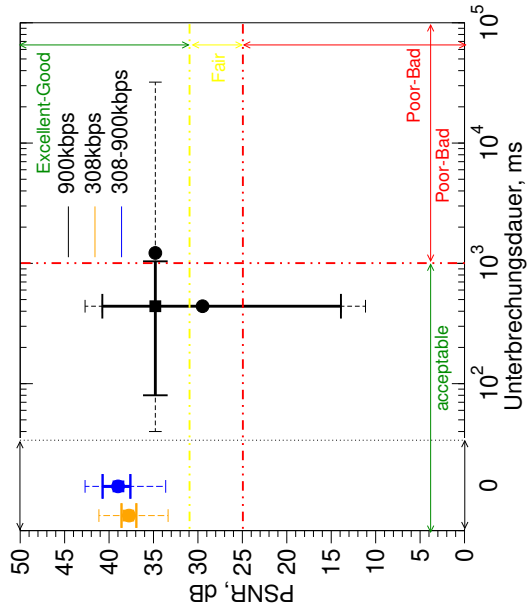


Abbildung 6.47: *intros* IDK 2B, PVU

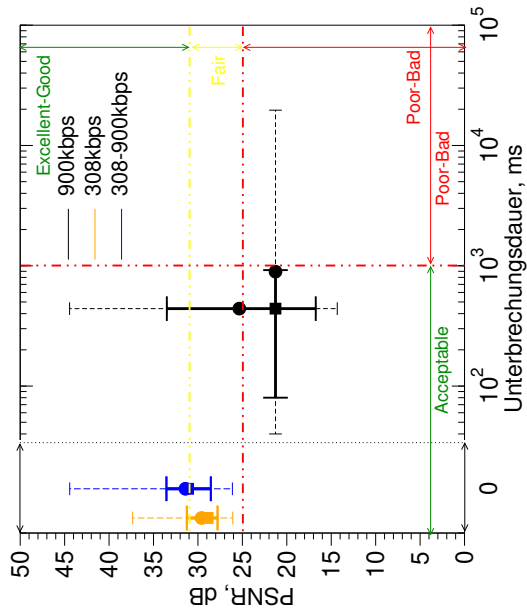


Abbildung 6.48: *vtc I_{nw}* IDK 2B, PVU

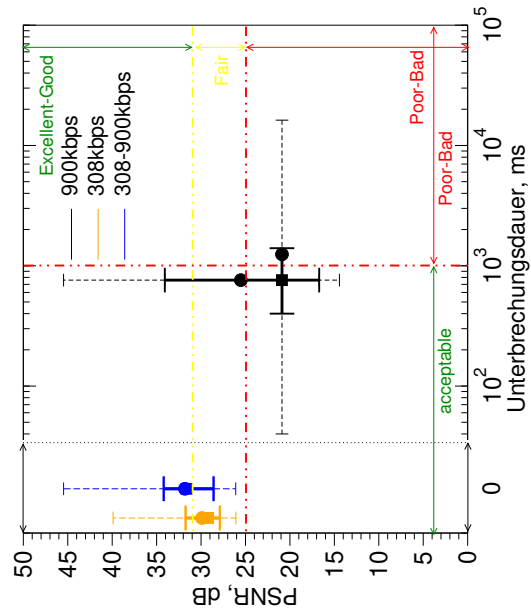


Abbildung 6.49: *football* IDK 0B, PVU

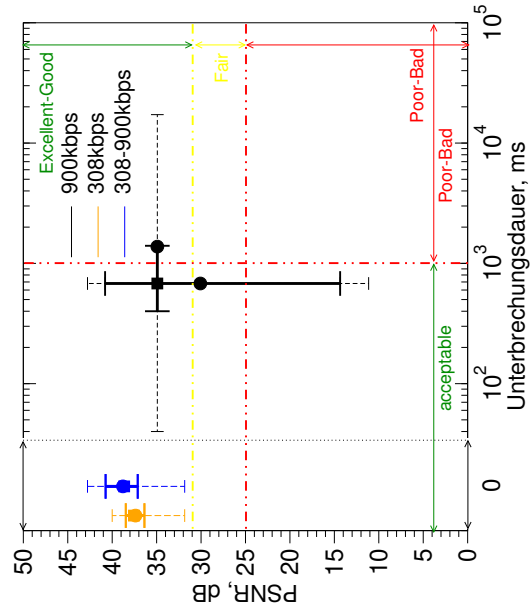


Abbildung 6.50: *intros* IDK 0B, PVU

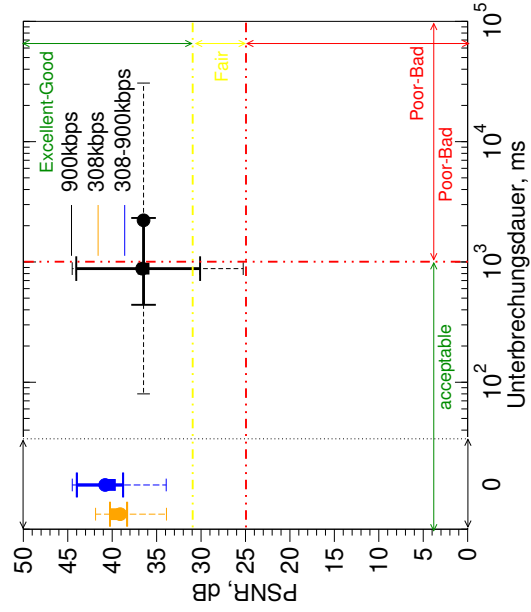


Abbildung 6.51: *vtc I_{nw}* IDK 0B, PVU

Die Abbildungen 6.46 bis 6.51 sind PVU-Diagramme (s. Erklärungen im Abschnitt 5.2.2) für die meist verbreiteten Videoausprägungen. Im adaptionslosen 308 kbit/s-Fall treten keine Videounterbrechungen (Einfrieren des Videobildes) auf. Deshalb gibt es keinen vertikalen Balken (das Kreuz ist ausgeartet). Es ist deutlich zu erkennen – wie bereits in der Abbildung 6.9 festgestellt –, dass sich die PSNR-Werte im fehlerfreien Fall je nach Videoinhalt deutlich unterscheiden (Vergleich derselben Komprimierungsausprägung: wie z. B. *IDK 2B* in den Abbildungen 6.46, 6.47 und 6.48). Das orange ausgeartete Kreuz hat je nach Videoinhalt eine andere Position auf der PSNR-Achse. Im adaptionslosen, fehlerbehafteten 900 kbit/s-Fall ist festzustellen, dass sich der Einfluss der Fehler im PSNR am stärksten bei *football* und am schwächsten bei *vtcInw* zeigt. Der Grund dafür ist, dass *vtcInw* im SI-TI-Diagramm 6.1 deutlich weniger an zeitlichen Änderungen als *football* aufweist. Das äußert sich beim Einfrieren des Videobildes in den besseren PSNR-Werten bei *vtcInw*. Dies ist die Ursache für unterschiedlich tiefe, fehlerbedingte PSNR-Einbrüche in den Abbildungen 6.37, 6.41 und 6.45.

Die Verteilungen der Unterbrechungsdauern sind innerhalb einer Komprimierungsausprägung (z. B. *IDK 2B*) in etwa gleich. Beim Vergleich von *IDK 2B* mit *IDK 0B* stellt man fest, dass der Medianwert der Dauern bei *IDK 0B* zunimmt (das Kreuz verschiebt sich nach rechts). Das liegt daran, dass *IDK 0B* mehr Videobilder mit Abhängigkeiten pro Videobildgruppe als *IDK 2B* aufweist (vgl. Tabelle 6.2). Das hat zur Folge, dass bei Videobildverlusten die Fehlerfortpflanzung stärker ausgeprägt ist.

Im adaptiven Fall treten keine Unterbrechungen auf. Die PSNR-Verteilungen liegen komplett im ersten Quadranten. Die Videoqualität (PSNR-Werte) ist höher als im fehlerfreien adaptionslosen 308 kbit/s-Fall und im fehlerbehafteten adaptionslosen 900 kbit/s-Fall.

6.3.1.4 Bewertung des Adaptionungsverhaltens

Der Videostreamingserver erhält die vom VSNC ermittelten Videoadaptionsvorschläge. Daraufhin führt er bei dem nächsten *I*-Videobild ggf. einen Umschaltvorgang auf die andere Videoqualitätsstufe durch. Die letzten zwei Spalten der Tabelle 6.5 zeigen die Anzahl der Adaptionsvorschläge bzw. der Umschaltvorgänge samt den Vertrauensintervallen¹². Für die anderen Spalten sind aus den Übersichtsgründen keine Vertrauensintervalle angegeben.

Jedes IP-Paket wird im RAN in ggf. mehreren TTIs übertragen. Die Anzahl der Änderungen der Übertragungsdatenrate von TTI zu TTI ist in der vierten Spalte zu sehen. Fiktive Vorschläge – ohne exponentielle Glättung als Schätzung der Übertragungsdatenrate – basierend auf den augenblicklichen Werten der Übertragungsdatenrate mit Kenntnis der Adaptionfähigkeiten des Videostreamingsservers sind in der fünften Spalte zu finden. Die darauffolgende Spalte zeigt den Einfluss der exponentiellen Glättung auf die Anzahl der Nachrichten. Sie wird dramatisch um zwei Größenordnungen reduziert. Die Hysterese führt zu einer weiteren deutlichen Abnahme der Nachrichtenanzahl. Die Anzahl der Umschaltvorgänge entspricht nicht der Anzahl der Adaptionsvorschläge. Das liegt darin begründet, dass die Umschaltvorgänge nur zu einem periodischen *I*-Videobild stattfinden. Innerhalb einer GoP-Periode können sich die Adaptionsvorschläge gegenseitig aufheben. Somit ist die Anzahl der Umschaltvorgänge kleiner als die Anzahl der Adaptionsvorschläge.

¹²95 %-Vertrauensintervall

Tabelle 6.5: Anzahl der Ereignisse im Adaptionkontext

<i>Video</i>	<i>IP-Pakete</i>	<i>TTIs</i>	<i>Datenraten-änderungen</i>	<i>Vorschläge o. Glättung</i>	<i>Vorschläge m. Glättung</i>	<i>Vorschläge m. Glättung m. Hysterese</i>	<i>Umschaltvorgänge</i>
<i>IDK</i>							
<i>football 0B</i>	74630	532123	252171	119086	4128	889 ±27	114 ±4
<i>football 2B</i>	76006	534403	253992	119269	4348	950 ±48	116 ±8
<i>football 11B</i>	76379	535212	254226	120301	4342	904 ±44	122 ±7
<i>intros 0B</i>	78301	545337	259296	118011	4887	1145 ±26	180 ±5
<i>intros 2B</i>	81676	547942	261764	119451	4983	1204 ±29	199 ±6
<i>intros 11B</i>	76739	531760	252200	117698	4386	997 ±27	126 ±8
<i>vtcInw 0B</i>	79036	532615	252803	117709	4449	955 ±39	115 ±8
<i>vtcInw 2B</i>	82931	534404	254083	116890	4727	1096 ±34	125 ±8
<i>vtcInw 11B</i>	76246	535323	253935	118165	4341	993 ±35	125 ±11
<i>IDKf</i>							
<i>football 0B</i>	75775	533004	252311	119708	4048	811 ±50	108 ±12
<i>football 2B</i>	77573	538377	255139	120636	4214	884 ±29	115 ±6
<i>football 11B</i>	80837	557258	263074	125561	4493	914 ±36	123 ±9
<i>intros 0B</i>	76485	534681	253145	116276	4469	999 ±27	118 ±8
<i>intros 2B</i>	82244	545688	259412	119852	4591	1031 ±42	169 ±12
<i>intros 11B</i>	79236	534243	252282	119594	3986	770 ±47	109 ±6
<i>vtcInw 0B</i>	79048	532438	252710	117809	4417	954 ±51	117 ±9
<i>vtcInw 2B</i>	79337	531702	251189	117684	4031	814 ±41	102 ±9
<i>vtcInw 11B</i>	76023	534564	252195	118894	3996	784 ±48	114 ±8
<i>2DK</i>							
<i>football 0B</i>	75286	528487	249120	115835	4337	910 ±31	114 ±10
<i>football 2B</i>	75807	526958	247656	110437	4364	984 ±25	123 ±8
<i>football 11B</i>	73552	520318	242901	109780	4455	872 ±42	136 ±9
<i>intros 0B</i>	85880	555212	264044	113882	6194	1583 ±19	276 ±9
<i>intros 2B</i>	86921	547655	260317	112297	5862	1500 ±29	246 ±5
<i>intros 11B</i>	82013	509589	237227	103478	5166	1379 ±32	163 ±5
HQ nicht konform							
<i>vtcInw 0B</i>	71887	459811	206453	101232	4824	1247 ±15	155 ±7
<i>vtcInw 2B</i>	67448	419418	177041	91698	4772	1312 ±17	165 ±6
<i>vtcInw 11B</i>	66040	403860	164744	84114	4689	1506 ±19	145 ±12
<i>vtcInw*0B</i>	82460	550762	255939	125638	7201	1778 ±15	225 ±6
<i>vtcInw*2B</i>	76475	518750	226383	123312	7226	1729 ±18	221 ±6
<i>vtcInw*11B</i>	67718	492786	210664	117714	6591	1554 ±14	243 ±6

Die Tabelle ist in drei Bereiche *IDK*, *IDKf* und *2DK* aufgeteilt. Der *2DK*-Bereich besteht aus zwei Teilen. Der untere Teil des *2DK*-Bereiches fasst Videos (*vtcInw*) zusammen, deren Video- datenrate der höheren Qualitätsstufe aufgrund der Implementierung des verwendeten Video- co-

dec unter der einheitlichen Videodatenrate von 900kbit/s liegt. D. h., die Ausprägungen der hohen Qualität sind videodatenratenbezogen nicht konform zu den anderen Videosequenzen.

Es werden zuerst die Bereiche *IDK*, *IDKf* und der 900kbit/s-konforme Teil von *2DK* diskutiert. In jedem der drei Bereiche existieren Videosequenzen, die besonders viele Umschaltvorgänge erfahren (grün hinterlegt). Der Grund dafür wird weiter unten erläutert. Mehr Umschaltvorgänge resultieren in diesem Fall (wie später zu sehen ist) in mehr zeitlichen Anteilen der hohen Qualität. Dies bedeutet wiederum die Erhöhung der Anzahl der IP-Pakete (s. zweite Tabellenspalte), die in TTIs bedient werden. Die Erhöhung der TTI-Anzahl kann ihrerseits zu einer höheren Anzahl an Adaptionsvorschlägen führen. Denn die zur Schätzung der Übertragungsdatenrate verwendeten Werte der augenblicklichen Übertragungsdatenrate werden pro TTI bestimmt.

Im nicht 900kbit/s-konformen *2DK*-Bereich sind auffallend niedrige TTI-Zahlen gelb hinterlegt. Das ist wie folgt begründet. Die *vtcInw*-Videogrundsequenz besitzt laut der Abbildung 6.1 die kleinste Menge an räumlichen und zeitlichen Informationen. Deshalb erzielt der Videocodec eine sehr gute Qualität bei niedrigen Videodatenraten. In diesem Fall hat der Videocodec – trotz der Vorgaben – Videodatenraten unter 900kbit/s verwendet (bei *vtcInw OB* 692kbit/s, bei *vtcInw 2B* 511kbit/s und bei *vtcInw 11B* 459kbit/s). In diesen Fällen hat der mit 30km/h fahrende Teilnehmer andere Mobilfunkkanalbedingungen erfahren als die Videostreamingsitzungen aus dem oberen Teil der Tabelle. Die mit einem Stern (*) gekennzeichneten *vtcInw*-Zeilen zeigen die Ergebnisse für die Adaption an die tatsächliche Videodatenrate der Stufe der hohen Qualität.

Die Tabelle 6.6 zeigt die Statistik der Phasen der hohen Qualität bei der adaptiven Übertragung der jeweiligen Videosequenz. Die erste Datenspalte enthält die Gesamtdauer der Phasen der hohen Qualität (mit Vertrauensintervallen) innerhalb einer Videostreamingsitzung. Die genauen Längen der Videosequenzen sind in der Tabelle 6.1 zu finden (*football*: 1326s, *intros* und *vtcInw*: 1324,4s). Die dunkelgrün hinterlegten Spalten entsprechen den bereits identifizierten Videosequenzen mit mehr Umschaltvorgängen als bei den anderen 900kbit/s-konformen Videosequenzen. Nun stellt man fest, dass sie mehr zeitliche Anteile der hohen Qualität aufweisen als die anderen 900kbit/s-konformen Videosequenzen. Der Grund dafür kann aus den Werten der blau hinterlegten Zellen wie folgt ermittelt werden.

Bei der Umschaltung an den GoP-Grenzen würde man den minimalen Wert einer Phase bei 0,48s erwarten¹³. In den betroffenen Tabellenzellen (blaue Zellen in der dritten Tabellenspalte) ist dies nicht der Fall. Des Weiteren ist die Wahrscheinlichkeit für eine unter 0,48s liegende Dauer deutlich über 0 (die blauen Zellen in der fünften Tabellenspalte). Auch das 1. Quartil liegt unter 0,48s (die blauen Zellen in der sechsten Tabellenspalte). Der Grund dafür liegt darin, dass der Videocodec in seltenen Fällen die GoP-Länge abweichend von der Vorgabe verändert. Dadurch entsteht ein GoP-Versatz zwischen den unterschiedlich stark komprimierten Videosequenzen. Somit sind die *I*-Videobilder in der Videosequenz der Basisqualität und in der Videosequenz der hohen Qualität versetzt zueinander. Ein Umschaltung kann somit schneller und öfter stattfinden. Die Videosequenzen mit den versetzten *I*-Videobildern erzielen – durch die schnellere Reaktion auf einen Adaptionsvorschlag und das Nutzen der kurzen (unter einer GoP-Periode) Phasen der hohen Übertragungsdatenraten – bessere Ergebnisse.

¹³Die hier verwendete GoP-Länge beträgt 12 Videobilder.

Im nicht 900 kbit/s-konformen 2DK-Bereich sieht man in den hellgrün hinterlegten Zellen, dass die Adaption an die tatsächliche Videodatenrate dramatische Verbesserungen (im Vergleich zu den Werten in den gelben Zellen) erzielt.

Tabelle 6.6: Phasen der hohen Qualität (Statistik der Dauern in Sekunden)

Video	Dauer insgesamt	Minimum	Maximum	$p(D < 0,48s)$	Q1	Q2	Q3	Q _{0,95}
<i>IDK</i>								
<i>football 0B</i>	490,70 ±6,31	0,48	38,4	0	0,48	1,92	20,16	31,68
<i>football 2B</i>	502,37 ±7,67	0,48	67,68	0	0,96	1,92	7,68	31,2
<i>football 11B</i>	495,55 ±8,51	0,48	36,96	0	0,96	1,92	11,04	31,68
<i>intros 0B</i>	545,5 ±5,94	0,04	43,48	0,31	0,4	1,12	2,84	32,04
<i>intros 2B</i>	554,17 ±5,11	0,12	43,32	0,29	0,36	0,96	3,72	29,64
<i>intros 11B</i>	497,86 ±5,00	0,48	36,48	0	0,96	1,44	8,64	31,68
<i>vtcInw 0B</i>	501,17 ±7,51	0,48	40,32	0	0,48	2,4	20,16	32,16
<i>vtcInw 2B</i>	512,78 ±4,62	0,48	66,24	0	0,96	1,92	7,20	31,20
<i>vtcInw 11B</i>	508,51 ±4,07	0,48	37,92	0	0,96	1,92	10,56	31,68
<i>IDKf</i>								
<i>football 0B</i>	487,92 ±9,53	0,48	38,4	0	0,96	1,92	20,64	33,60
<i>football 2B</i>	499,87 ±5,91	0,48	67,68	0	0,96	1,92	11,04	31,2
<i>football 11B</i>	502,17 ±5,46	0,48	36	0	0,96	2,40	11,04	31,20
<i>intros 0B</i>	512,84 ±6,09	0,04	73,20	0,003	0,48	1,92	17,36	31,32
<i>intros 2B</i>	535,48 ±6,04	0,12	35,64	0,25	0,36	1,32	4,92	29,64
<i>intros 11B</i>	487,25 ±8,31	0,48	58,08	0	0,96	2,4	18,24	31,68
<i>vtcInw 0B</i>	499,49 ±9,32	0,48	37,92	0	0,96	1,92	17,76	31,68
<i>vtcInw 2B</i>	490,94 ±7,87	0,48	65,76	0	0,96	1,92	21,6	31,20
<i>vtcInw 11B</i>	490,75 ±7,56	0,48	37,92	0	0,96	2,4	16,8	31,68
<i>2DK</i>								
<i>football 0B</i>	489,07 ±5,01	0,48	37,92	0	0,96	1,92	20,16	31,68
<i>football 2B</i>	507,79 ±3,98	0,48	71,04	0	0,96	1,44	6,24	31,20
<i>football 11B</i>	482,77 ±6,46	0,48	58,56	0	0,48	1,92	7,68	29,28
<i>intros 0B</i>	579,86 ±6,18	0,04	68,60	0,38	0,32	0,76	2,44	29,08
<i>intros 2B</i>	575,54 ±4,39	0,12	36,96	0,31	0,36	0,84	2,88	27,96
<i>intros 11B</i>	481,20 ±7,67	0,48	36,48	0	0,48	1,44	5,76	26,88
HQ nicht konform								
<i>vtcInw 0B</i>	424,80 ±4,41	0,48	30,24	0	0,96	1,92	6,24	24,48
<i>vtcInw 2B</i>	390,24 ±2,58	0,48	31,68	0	0,48	1,44	5,28	23,04
<i>vtcInw 11B</i>	344,06 ±2,23	0,48	27,84	0	0,96	1,92	5,28	21,60
<i>vtcInw*0B</i>	790,99 ±4,58	0,48	84,96	0	0,48	0,96	2,40	45,12
<i>vtcInw*2B</i>	1064,22 ±1,23	0,48	255,36	0	0,48	0,96	2,4	58,56
<i>vtcInw*11B</i>	1143,64 ±1,30	0,48	261,6	0	0,48	0,96	1,92	56,64

Die Tabelle 6.7 zeigt die Adaptionseffizienz in Bezug auf den zeitlichen Anteil der Phasen der hohen Qualität und die übertragenen Datenmengen (mit Vertrauensintervallen). Die letzte Spal-

Tabelle 6.7: Adaptionseffizienz

Video	Zeitlicher Anteil der Phasen der hohen Qualität	DR_{ef-all}	DR_{ef-ef}
<i>IDK</i>			
<i>football 0B</i>	37,01 % $\pm 0,48$	78,35 % $\pm 0,43$	91,37 % $\pm 0,80$
<i>football 2B</i>	37,88 % $\pm 0,58$	79,11 % $\pm 0,51$	93,96 % $\pm 0,93$
<i>football 11B</i>	37,37 % $\pm 0,64$	78,91 % $\pm 0,56$	91,96 % $\pm 0,97$
<i>intros 0B</i>	41,19 % $\pm 0,44$	82,00 % $\pm 0,41$	96,37 % $\pm 0,80$
<i>intros 2B</i>	41,84 % $\pm 0,39$	82,78 % $\pm 0,35$	97,28 % $\pm 0,74$
<i>intros 11B</i>	37,59 % $\pm 0,37$	78,82 % $\pm 0,34$	92,32 % $\pm 0,71$
<i>vtcInw 0B</i>	37,84 % $\pm 0,57$	79,13 % $\pm 0,84$	92,13 % $\pm 0,90$
<i>vtcInw 2B</i>	38,72 % $\pm 0,35$	79,96 % $\pm 0,32$	94,82 % $\pm 0,69$
<i>vtcInw 11B</i>	38,39 % $\pm 0,31$	79,45 % $\pm 0,28$	93,48 % $\pm 0,64$
<i>IDKf</i>			
<i>football 0B</i>	36,80 % $\pm 0,71$	78,19 % $\pm 0,64$	90,96 % $\pm 1,05$
<i>football 2B</i>	37,70 % $\pm 0,44$	79,16 % $\pm 0,40$	92,57 % $\pm 0,78$
<i>football 11B</i>	37,87 % $\pm 0,41$	81,70 % $\pm 0,35$	95,15 % $\pm 0,73$
<i>intros 0B</i>	38,72 % $\pm 0,45$	79,78 % $\pm 0,41$	93,86 % $\pm 0,80$
<i>intros 2B</i>	40,43 % $\pm 0,45$	81,30 % $\pm 0,41$	95,93 % $\pm 0,81$
<i>intros 11B</i>	36,79 % $\pm 0,61$	78,09 % $\pm 0,55$	91,63 % $\pm 0,95$
<i>vtcInw 0B</i>	37,71 % $\pm 0,70$	79,03 % $\pm 0,63$	92,40 % $\pm 1,05$
<i>vtcInw 2B</i>	37,07 % $\pm 0,59$	78,32 % $\pm 0,53$	91,82 % $\pm 0,93$
<i>vtcInw 11B</i>	37,05 % $\pm 0,57$	78,31 % $\pm 0,51$	91,89 % $\pm 0,91$
<i>2DK</i>			
<i>football 0B</i>	36,88 % $\pm 0,37$	78,22 % $\pm 0,35$	91,35 % $\pm 0,71$
<i>football 2B</i>	38,29 % $\pm 0,30$	79,53 % $\pm 0,28$	102,67 % $\pm 0,70$
<i>football 11B</i>	36,41 % $\pm 0,49$	78,32 % $\pm 0,49$	93,94 % $\pm 0,90$
<i>intros 0B</i>	43,78 % $\pm 0,47$	86,03 % $\pm 0,49$	102,09 % $\pm 0,93$
<i>intros 2B</i>	43,46 % $\pm 0,33$	84,78 % $\pm 0,27$	107,16 % $\pm 0,70$
<i>intros 11B</i>	36,33 % $\pm 0,58$	77,66 % $\pm 0,54$	90,77 % $\pm 0,94$
HQ nicht konform			
<i>vtcInw 0B</i>	32,07 % $\pm 0,33$	72,90 % $\pm 0,26$	79,73 % $\pm 0,55$
<i>vtcInw 2B</i>	29,46 % $\pm 0,19$	76,10 % $\pm 0,08$	80,53 % $\pm 0,36$
<i>vtcInw 11B</i>	25,98 % $\pm 0,17$	78,27 % $\pm 0,05$	81,28 % $\pm 0,33$
<i>vtcInw*0B</i>	59,72 % $\pm 0,34$	90,85 % $\pm 0,27$	99,36 % $\pm 0,62$
<i>vtcInw*2B</i>	80,35 % $\pm 0,09$	97,02 % $\pm 0,06$	102,66 % $\pm 0,41$
<i>vtcInw*11B</i>	86,35 % $\pm 0,10$	97,75 % $\pm 0,08$	101,52 % $\pm 0,42$

te beinhaltet das bereits eingeführte Verhältnis der Datenmengen der fehlerfrei empfangenen Videobilder im Adaptions- und adaptionslosen Fall DR_{ef-ef} (vgl. Abschnitt 6.3.1.2). Abgesehen vom bereits diskutierten gelben Bereich werden bei DR_{ef-ef} durchgängig Werte von über 90% erreicht. Die DR_{ef-ef} -Werte der dunkelgrünen Bereiche liegen über 95%. Dies ist die

Folge des bereits besprochenen agilen Adaptionverhaltens innerhalb einer GoP-Periode. Die Videosequenz *intros OB* weist den höchsten zeitlichen Anteil der hohen Qualität auf und erreicht einen Wert für das DR_{ef-ef} von über 100 %.

Die dunkelblau hinterlegten Zellen zeigen erhöhte Werte von DR_{ef-ef} im Verhältnis zu ihren zeitlichen Anteilen der hohen Qualität (zweite Tabellenspalte). Dies liegt darin begründet, dass diese Videosequenzen im adaptionslosen Fall aufgrund ihrer Struktur besonders an den ungünstigen Übertragungsverhältnissen gelitten haben und deshalb kleinere fehlerfreie Datenmengen im adaptionslosen Fall aufweisen als andere Sequenzen. Das lässt sie im adaptiven Fall – zusätzlich zu den guten Adaptionsergebnissen – als besonders gut auffallen.

Die grau hinterlegten Zellen weisen erhöhte Werte der beiden Datenmengenverhältnisse im Vergleich zu ihrem zeitlichen Anteil der hohen Qualität auf. Das liegt daran, dass diese Videosequenz in der Basisqualität trotz einer Vorgabe an den Videocodec eine leicht erhöhte Videodatenrate von 335 kbit/s aufweist.

Die nicht 900kbit/s-konformen Videosequenzen erreichen bei der sequenzgerechten Adaption sehr gute Werte für DR_{ef-ef} um die 100 % (hellgrün hinterlegt).

Das Verhältnis DR_{ef-all} (die dritte Tabellenspalte) setzt die Datenmenge der *fehlerfrei* empfangenen Videobilder im Adaptionfall zur Datenmenge aller im adaptionslosen Fall empfangenen Videobilddaten in Beziehung (vgl. Abschnitt 6.3.1.2). Im letzteren Fall sind viele Videobilddaten lediglich Teile eines Videobildes. Die erreichten Werte im 900kbit/s-konformen Tabellenbereich liegen um die 80 %. Die besonders adaptionsagile *intros-OB*-Videosequenz weist einen Wert von 86 % auf.

Die nicht 900kbit/s-konformen Videosequenzen erreichen bei der sequenzgerechten Adaption sehr gute Werte für DR_{ef-all} von 90 bis knapp 98 % (hellgrün hinterlegt).

6.3.1.5 Zusammenfassende Diskussion

Die vorgestellten Ergebnisse zeigen eine vom Mobilfunkzugangsnetz gesteuerte Adaption, die in allen untersuchten Fällen zu einer fehlerfreien Videostreamingsitzung führt. Das Verfahren wurde auf die 27 unterschiedlichen Videosequenzen angewendet.

Die im Mobilfunkzugangsnetz ermittelten Adaptionsvorschläge werden ohne einen Zeitverzug an den Videostreamingserver versandt. Die Reaktionsschnelligkeit der Adaption hängt von der Dauer der Übertragung des Adaptionssvorschlages und von der Implementierung der Videostreaminganwendung ab. Theoretisch ist es möglich, eine Umschaltung in jedem Videobildintervall (alle 40 ms) durchzuführen. Im Vergleich dazu ermittelten die Autoren von [101] – für den Fall einer RTP/UDP-basierten Ende-Zu-Ende-Lösung – eine Adaptionsverzögerung von 2,2 s für das Herunterschalten und 9,8 s für das Heraufschalten. Bei der HTTP-basierten DASH-Lösung wird die Adaptionszeitskala durch die eingestellte Videosegmentlänge bestimmt, was zu einer Verzögerung von mehreren Sekunden führen kann.

Beim Herunterschalten zeigt sich die Agilität und die Präzision der hier untersuchten netzgesteuerten Adaption u. a. in einer fehlerfreien Übertragung. Beim Herunter- und Heraufschalten

während der gesamten Videostreamingsitzung sind diese Eigenschaften durch das Erreichen von hohen DR_{ef-ef} -Werten um bis zu 100 % und darüber hinaus belegt.

Das vorgestellte Konzept der Adaptionsvorschläge erspart dem Videostreamingserver die Interpretation, die Aufbereitung und die Bewertung von netzspezifischen Daten. Im Vergleich zu den anderen netzbasierten Lösungen, die den Videostreamingserver mit den ungefilterten relevanten Netznachrichten zu versorgen vorschlagen, reduziert die hier untersuchte Vorgehensweise die Anzahl von relevanten Nachrichten um zwei Größenordnungen (von mehr als 200000 auf ca. 1000 für eine 22-minütige Videostreamingsitzung). Somit werden potenzielle Skalierungsprobleme eines Videostreamingservers – verursacht durch das Verarbeiten einer hohen Anzahl von Netznachrichten – vermieden.

6.3.2 AQM

6.3.2.1 Szenario

Die vorgeschlagenen videospezifischen AQM-Verfahren werden hinsichtlich ihrer proaktiven Handlung in einer Situation schwankender und knapper Übertragungsressourcen bewertet. Es werden die 308 kbit/s-Videsequenzen verwendet. Dies stellt somit den Fall dar, dass nach einer erfolgreichen netzgesteuerten Adaption keine weitere niedrigere Qualitätstufe des Videos im Videostreamingserver vorliegt.

Um die Grundknappheit der Ressourcen zu erzeugen, werden bei jeder Untersuchung drei weitere CBR-erzeugte Videoströme innerhalb des UTRAN-Emulators als konkurrierender Videoverkehr mit der Videodatenrate von 308 kbit/s simuliert.

Für die Bewertung wird angenommen, dass der Videoclient einen Abspielpuffer von 6 s besitzt [167]. Der Eingangspuffer im RNC ist 212000 Byte groß. Die in den Puffer passende Datenmenge entspricht somit einem Abspielintervall zwischen 5,25 s und 5,3 s (je nach sequenzspezifischem Overhead) für eine 308 kbit/s-Videsequenz.

Es werden drei im Kapitel 4.3.3 vorgeschlagene, auf quantitativer Wichtigkeit basierende Verfahren untersucht. Zwecks einfacherer Ergebnisbeschreibung wird das paketbasierte Verfahren im Folgenden als Pb abgekürzt. Für das videobildbasierte Verfahren wird die Abkürzung Bb verwendet. Das videobildbasierte Verfahren mit proaktivem Verwerfen der abhängigen Videobilder wird mit BAb gekennzeichnet.

6.3.2.2 Wirkungsweise

Die vorgeschlagenen Verfahren verwerfen in einer Überlastsituation proaktiv die Videodaten der weniger wichtigen Videobilder zugunsten von wichtigen Videobildern. Die direkte Wirkung wird anhand von Pb am Beispiel von *football IDK 11B* in der Abbildung 6.52 visualisiert. Es wurde zwecks Veranschaulichung der Wirkung angenommen, dass der Videoclient keine Abspielfristen für die Videobilder setzt. D. h., es gibt keine Videobildverluste aufgrund einer verspäteten Videobildankunft. Dargestellt sind Videobildverluste in der Abhängigkeit vom

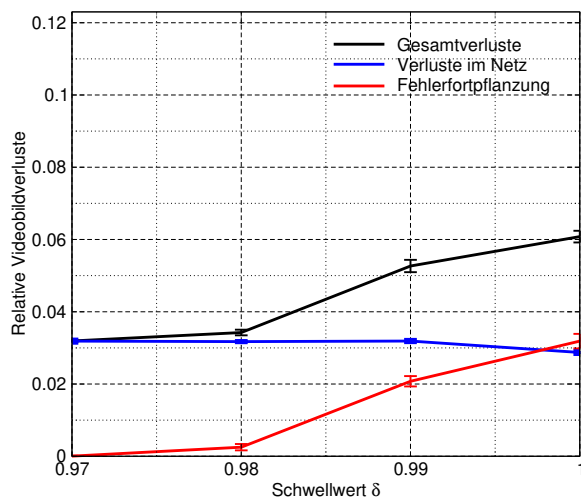


Abbildung 6.52: Wirkungsweise für *football 1DK 11B Pb*

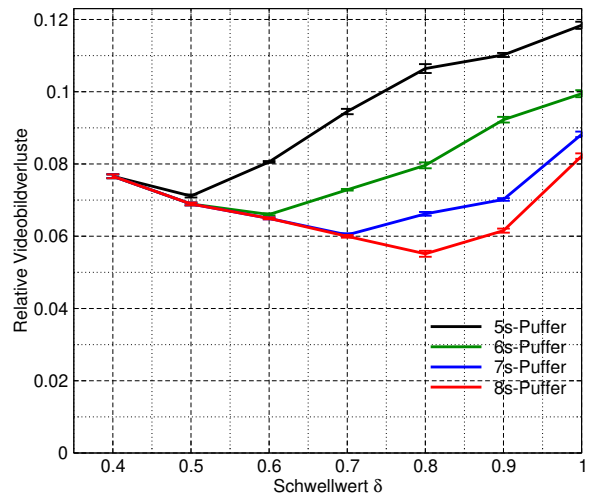


Abbildung 6.53: Vergleich für verschiedene Abspielpuffergrößen für *intros 1DK 2B Pb*

Schwellwert des RNC-Pufferfüllstandes δ , bei dessen Überschreitung das proaktive Verwerfen von Videobildern durchgeführt wird.

Die blaue Kurve zeigt die im Netz verworfenen Videobilder. Beim Schwellwert $\delta = 1$ greift das proaktive Verwerfen nicht. Die Gesamtverlustrate setzt sich aus den Verlusten im Netz und den Verlusten aufgrund der Videobildabhängigkeiten (im Bild *Fehlerfortpflanzung*) zusammen. Das proaktive Verwerfen von *B*-Videobildern zugunsten von *I*-Videobildern eliminiert die Verluste aufgrund der Videobildabhängigkeiten bereits bei $\delta = 0,97$ komplett¹⁴. Durch die Veränderung der Videobildverluststruktur sinkt somit die Gesamtverlustrate.

Ein Videostreamingclient verwendet einen Abspielpuffer, um die Verzögerungsschwankungen bei der Auslieferung der Videobilder auszugleichen. Nach dem Beginn der Wiedergabe müssen die neu ankommenden Videobilder rechtzeitig zur ihrer Wiedergabe vorliegen. In einer Überlastsituation können die im RNC-Puffer befindlichen Videobilder eventuell nicht rechtzeitig ausgeliefert werden. Der Schwellwert δ kann abhängig von der Größe des Abspielpuffers des Videoclients so eingestellt werden, dass die Videobildverluste aufgrund einer zu späten Auslieferung minimiert bzw. eliminiert werden. Die Abbildung 6.53 zeigt am Beispiel von *intros 1DK 2B Pb* die Videobildverluste in Abhängigkeit vom Schwellwert δ für unterschiedlich große Abspielpuffer im Videoclient. Die jeweiligen Minimumwerte der Videobildverlustrate entsprechen der Grenze für den Schwellwert, unter der keine Verluste aufgrund einer zu späten Ankunft auftreten. Liegt z. B. das Minimum der Videobildverlustrate bei $\delta = 0,6$, so gibt es für $\delta < 0,6$ keine Verluste aufgrund einer zu späten Ankunft.

¹⁴Der im Puffer frei gewordene Raum entspricht $0,03 \cdot 212000 = 6360$ Byte. Die mittlere *I*-Videobildgröße liegt für diese Videosequenz bei $2880 \text{ Byte} \pm 32$. Da finden im Mittel zwei *I*-Videobilder Platz.

Tabelle 6.8: Videobildwichtigkeiten

	1	2	3	4	5	6	7	8	9	10	11	12	13
<i>0B</i>	<i>I/11</i>	<i>P/10</i>	<i>P/9</i>	<i>P/8</i>	<i>P/7</i>	<i>P/6</i>	<i>P/5</i>	<i>P/4</i>	<i>P/3</i>	<i>P/2</i>	<i>P/1</i>	<i>P/0</i>	<i>I/11</i>
<i>2B</i>	<i>I/13</i>	<i>B/0</i>	<i>B/0</i>	<i>P/10</i>	<i>B/0</i>	<i>B/0</i>	<i>P/7</i>	<i>B/0</i>	<i>B/0</i>	<i>P/4</i>	<i>B/0</i>	<i>B/0</i>	<i>I/13</i>
<i>11B</i>	<i>I/22</i>	<i>B/0</i>	<i>B/0</i>	<i>B/0</i>	<i>B/0</i>	<i>B/0</i>	<i>B/0</i>	<i>B/0</i>	<i>B/0</i>	<i>B/0</i>	<i>B/0</i>	<i>B/0</i>	<i>I/22</i>

6.3.2.3 Videobildwichtigkeit

Die Tabelle 6.8 zeigt die ermittelten quantitativen Videobildwichtigkeiten für *0B*, *2B* und *11B*, die auf dem Umfang der Videobildabhängigkeiten basieren (vgl. Abschnitt 4.3.3). So gibt der Wert *I/11* an, dass dieses *I*-Videobild eine Wichtigkeit von 11 besitzt¹⁵.

Die hier vorgeschlagenen Verfahren führen das proaktive Verwerfen anhand der Videobildwichtigkeit Imp_i als Hauptkriterium (vgl. Abschnitt 4.3.3) durch. Ein Video-IP-Paket wird verworfen, wenn Imp_i der darin enthaltenen Videobilddaten dem Schwellwert Imp_{min} gleich ist oder darunter liegt. Ist der Schwellwert beispielsweise $Imp_{min} = 4$, so können bei *0B* ggf. alle *P*-Videobilder mit $Imp_i = 4$ und kleiner verworfen werden. Bei *2B* können ggf. alle *B*-Videobilder und alle *P*-Videobilder mit $Imp_i = 4$ verworfen werden. Im Falle von *11B* können ggf. alle *B*-Videobilder verworfen werden. Entscheidend für die Wirkung der Verfahren ist neben der Anzahl der potenziell zu verwerfenden Videobilder auch der Datenumfang, der in diesen Videobildern enthalten ist.

Die Abbildungen 6.54 bis 6.61 zeigen die Videobildverluste in Abhängigkeit vom Schwellwert für die Videobildwichtigkeit Imp_{min} für das Verfahren *Pb*. Die waagerechten gestrichelten Referenzlinien zeigen zum Vergleich die jeweiligen Videobildverluste, wenn kein AQM-Verfahren eingesetzt wird. Jeder Punkt einer Kurve stellt einen optimalen Wert dar, der aus den Messungen mit dem Schwellwert des RNC-Pufferfüllstandes δ von 0,35 bis 1,0 mit einer in den relevanten Bereichen geringen Schrittweite ermittelt wurde.

Die Abbildungen 6.54, 6.56, 6.58 und 6.60 visualisieren den Fall, in dem der Videoclient keine Abspielfristen für die Videobilder setzt. D. h., es gibt keine Videobildverluste aufgrund von verspäteten Videobildankünften. Die restlichen Abbildungen stellen den Fall dar, in dem der Videoclient einen Abspielpuffer von 6s (im Bild P_{6s}) aufweist.

Über alle Videosequenzen betrachtet kann festgestellt werden, dass der optimale Schwellwert (für den die Videobildverluste minimal wird) je nach Sequenz $Imp_{min} = 4$ bzw. $Imp_{min} = 0$ ist. Der Fall $Imp_{min} = 4$ ist optimal im Allgemeinen für die gängigsten Ausprägungen *IDK 0B*, *IDK 2B* und ggf. *2DK 2B*. Der Fall $Imp_{min} = 0$ ist optimal für die Ausprägungen, die besonders viele Daten in den *B*-Videobildern aufweisen, wie *IDKf 2B*.

Die aus Platzgründen nicht dargestellte Ausprägung *IDKf 0B* zeigt ein ähnliches Verhalten wie *IDK 0B*. Die ebenfalls aus Platzgründen nicht dargestellte Ausprägung *2DK 0B* zeigt ein ähnliches Verhalten wie *2DK 2B*. Alle *11B*-Ausprägungen weisen nur einen Punkt für $Imp_{min} = 0$ auf. Sie und alle anderen Ausprägungen werden gesondert im Vergleich der *Pb*-, *Bb*- und *BAb*-Verfahren im Abschnitt 6.3.2.4 behandelt.

¹⁵Von diesem *I*-Videobild hängen 11 weitere Videobilder ab.

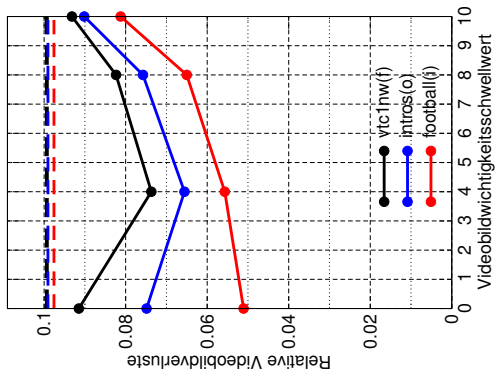


Abbildung 6.57: *IDK 2B P_{6s}*

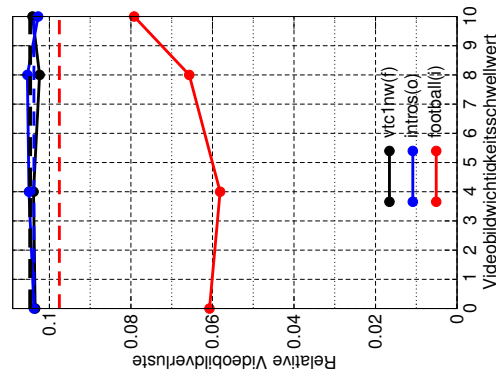


Abbildung 6.61: *2DK 2B P_{6s}*

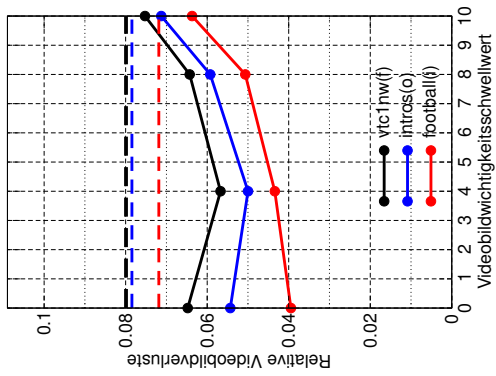


Abbildung 6.56: *IDK 2B*

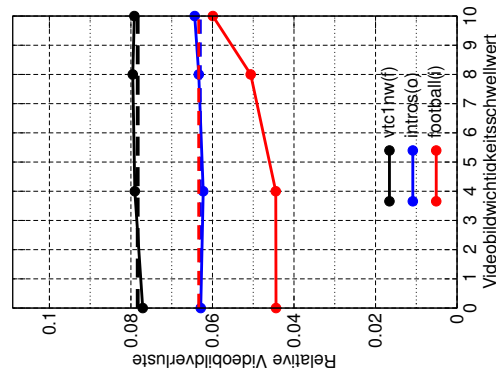


Abbildung 6.60: *2DK 2B*

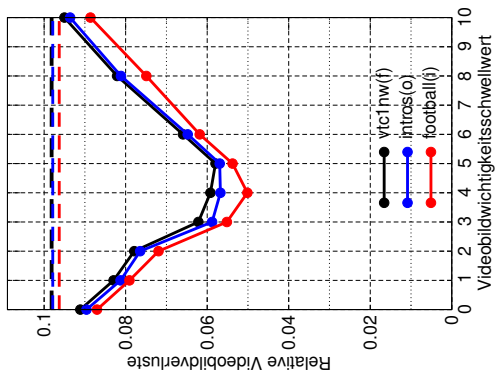


Abbildung 6.55: *IDK 0B P_{6s}*

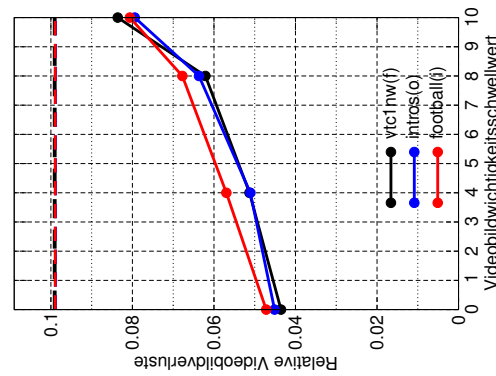


Abbildung 6.59: *IDKf 2B P_{6s}*

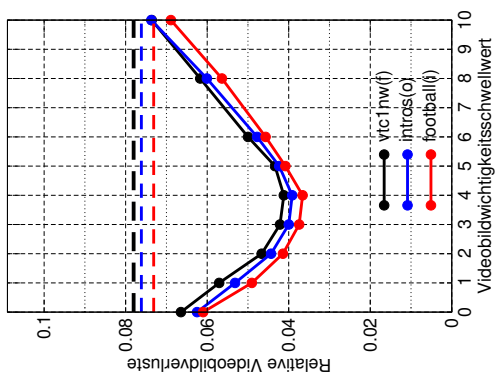


Abbildung 6.54: *IDK 0B*

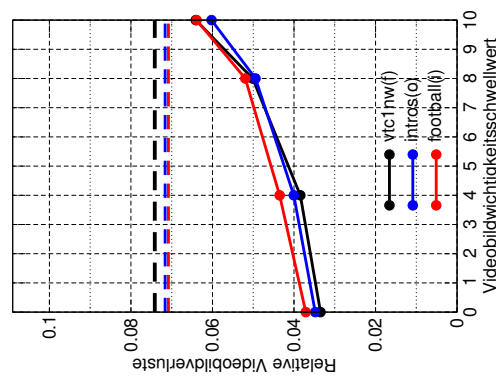


Abbildung 6.58: *IDKf 2B*

Im Fall, in dem der Videoclient keine Abspielfristen für die Videobilder setzt, liegen die Referenzlinien i. d. R. auseinander. Dies korreliert mit dem Datenumfang in den *I*- und *P*-Videobildern (vgl. Abschnitt 6.1.4). Je größer er ist, desto größer ist die Verlustwahrscheinlichkeit für die *I*- und *P*-Videobilder, wenn kein AQM-Verfahren eingesetzt wird. Dadurch steigen dann die Verluste aufgrund der Videobildabhängigkeiten. Interessant ist der Fall *2DK 2B*, in dem die Referenzlinien von *intros* und *football* zusammenfallen. Die Videobildverluststrukturen (Verluste im Netz und Verluste aufgrund von Videobildabhängigkeiten) unterscheiden sich bei beiden Sequenzen. In der Summe ergibt sich aber in etwa derselbe Wert. Es ist wichtig zu betonen, dass es hier um Videobildverluste geht. Die Videobyteverluste sind bei allen Ausprägungen in etwa gleich.

Wenn der Videoclient einen Abspielpuffer von 6s aufweist, sind die Referenzlinien i. d. R. eng beieinander. Dies ist auf die Verluste aufgrund von verspäteten Einkünften zurückzuführen, die den Vorteil der eventuellen kleineren Verluste aufgrund der Videobildabhängigkeiten schmälern.

Bei *IDK 0B* (s. Abbildung 6.54) verlaufen die Kurven in einem engen Abstand. Gemäß der Abbildung 6.12 sind die Sequenzen sehr ähnlich. Allerdings würde man erwarten, dass *vtcInw* bessere Ergebnisse als *intros* erzielt, da die erstere Sequenz etwas mehr Daten in den *P*-Videobildern aufweist. Betrachtet man aber die zu verwerfenden *P*-Videobilder für $Imp_{min} = 4$, so stellt man fest, dass die mittlere Videobildgröße und das 1. und 2. Quartil der Videobildgrößenverteilung bei *vtcInw* am kleinsten unter den drei Sequenzen sind. Insofern für $Imp_{min} = 4$ hat das Verfahren im Mittel bei *vtcInw* weniger Daten zum proaktiven Verwerfen.

Bei *IDK 2B* verlaufen die Kurven in einem größeren Abstand. Betrachtet man die Abbildung 6.13, so wird es deutlich, dass der Datenumfang in den *B*-Bildern bei *football* am größten und bei *vtcInw* am kleinsten ist. Die Unterschiede in den Datenumfängen sind groß. Das Verfahren hat im Falle von *football* deutlich mehr Daten zum proaktiven Verwerfen als bei *vtcInw*. Die Sequenz *intros* liegt dazwischen. Das Verhalten von *football* bei $Imp_{min} = 0$ wird in Abschnitt 6.3.2.4 behandelt.

Bei *IDKf 2B* fallen die Kurven für *intros* und *vtcInw* beinahe zusammen. Anhand der Abbildung 6.16 stellt man fest, dass die beiden Sequenzen sehr viele Daten in den *B*-Videobildern aufweisen, so dass das proaktive Verwerfen ausreichend Daten zur Verfügung hat. Die Sequenz *football* hat etwas weniger Daten in den *B*-Videobildern.

Bei *2DK 2B* stellt man fest, dass das Verfahren nicht greift. Auch *Bb*- und *BAb*-Verfahren greifen nicht. Genauso verhält es sich auch bei *2DK 0B* und *2DK 11B*. Durch die besonders effiziente Videokomprimierung werden die Daten stark in die *I*-Videobilder verlagert (s. Abbildungen 6.18 bis 6.20). Dadurch stehen nicht genügend Daten für das proaktive Verwerfen zur Verfügung. Die Videosequenz *football* ist eher eine Ausnahme. Sie lässt sich für 308 kbit/s über das *2DK*-Verfahren in ihrer Datenumfangsstruktur nicht wesentlich verändern (s. Abbildungen 6.18 bis 6.20). Bei höheren Videodatenraten besitzt sie eher die üblichen *2DK*-Eigenschaften.

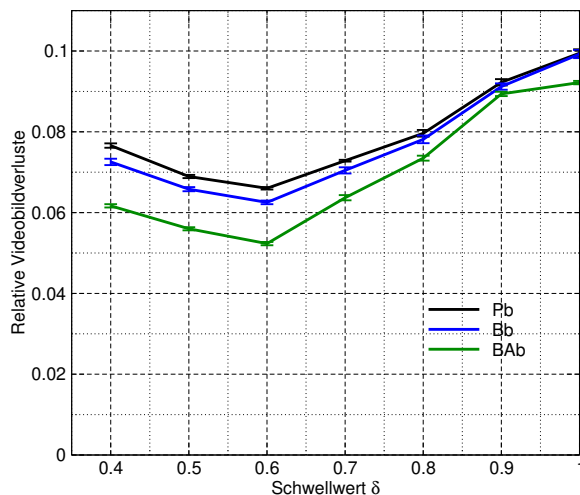


Abbildung 6.62: Verfahrenvergleich für *intros 1DK 2B*

6.3.2.4 Vergleich der vorgeschlagenen AQM-Verfahren

Anhand der Erkenntnisse aus dem Abschnitt 6.3.2.3 sollen nun die drei Verfahren *Pb*, *Bb* und *BAb* verglichen werden. Es wird $Imp_{min} = 4$ bzw. $Imp_{min} = 0$ entsprechend verwendet. Die *2DK*-Ausprägungen werden nicht weiter untersucht. Die optimalen Einstellungen für den Schwellwert δ liegen¹⁶ für die *OB*-Videosequenzen bei $\delta = 0,5$ und für die *2B*- bzw. *11B*-Videosequenzen bei $\delta = 0,6$. Das gilt für alle drei Verfahren.

Die videostrukturspezifischen Schwellwerte δ können durch die unterschiedliche Anzahl der Videobilder erklärt werden, die potenziell zum proaktiven Verwerfen zur Verfügung stehen. Wie im Abschnitt 6.3.2.3 erörtert und in der Tabelle 6.8 zu sehen ist, sind bei $Imp_{min} = 4$ im *OB*-Fall maximal 5 Videobilder pro Videobildgruppe zu verwerfen. Im *2B*-Fall sind 9 und im *11B*-Fall sind 11 Videobilder betroffen. Bei ein und demselben Schwellwert δ würde dies in einer andauernden Überlastsituation zu größeren Werten des RNC-Pufferfüllstandes bei *OB*-Videosequenzen führen, was u. a. zu höheren Verlusten aufgrund von verspäteten Ankünften führen kann.

Für alle ab jetzt diskutierten Ergebnisse gilt, dass der Videoclient einen Abspielpuffer von 6s aufweist.

Die Abbildung 6.62 zeigt die Videobildverlustrate in Abhängigkeit vom Schwellwert δ für drei Verfahren *Pb*, *Bb* und *BAb* bei $Imp_{min} = 4$ im Falle einer gängigen Ausprägung *intros 1DK 2B*. Bei $\delta = 1$ wird kein proaktives Verwerfen durchgeführt, *Bb* und *BAb* führen trotzdem ihre Aktionen videobildweit ggf. unter Berücksichtigung der Videobildabhängigkeiten bei den aufgrund des RNC-Pufferüberlaufs verloren gehenden Videobilddaten aus.

Der enge Abstand zwischen den *Pb*- und *Bb*-Kurven ist dadurch zu erklären, dass die proaktiv zu verwerfenden *B*-Videobilder dieser Sequenz aufgrund ihrer Größe meistens in einem einzigen

¹⁶Für die zu verwendenden Werte von $Imp_{min} = 4$ bzw. $Imp_{min} = 0$

Video-IP-Paket¹⁷ übertragen werden. Die Größe der P -Videobilder liegt meistens unter zwei maximalen IP-Paketgrößen. Verkleinert man die maximale IP-Paketgröße auf 512 Byte steigt der Abstand. Er steigt ebenfalls, wenn mehr Daten in die abhängigen Videobilder (B - und P -Videobilder) verlagert werden, wie später zu sehen ist. Bei $\delta = 1$ sieht man den Unterschied kaum (die Bb -Kurve liegt aber trotzdem unter der Pb -Kurve). Bei einer Videosequenz mit vielen Daten in den I -Videobildern ist der Unterschied bei $\delta = 1$ (AQM nicht aktiv) sehr deutlich. So beträgt der Abstand bei der nicht weiter betrachteten *vtcInw-2DK-11B*-Sequenz (vgl. die Abbildung 6.20) 0,01.

Eine eindeutige Verbesserung bringt BAb . Durch das proaktive Verwerfen der nicht zu verwendenden abhängigen Videobilder eines proaktiv verworfenen P -Videobildes mit $Imp_i = 4$ wird das System entlastet. Bei $\delta = 1$ beobachtet man, dass BAb auch beim RNC-Pufferüberlauf durch die Erkennung und das gezielte Entfernen aller zum verlorenen Videobild gehörenden Datenteile und seiner Videobildabhängigkeiten ebenfalls für eine Entlastung sorgt.

Die Abbildungen 6.63 bis 6.68 zeigen den Vergleich von Pb , Bb und BAb in Bezug auf die erzielte Minderung der Videobildverluste bei $Imp_{min} = 4$. OA bedeutet Ohne AQM und stellt einen jeweiligen Referenzpunkt dar.

Im Falle von *IDK 0B* (s. Abbildung 6.63) wurde der Abstand zwischen Pb -Werten bereits im Abschnitt 6.3.2.3 diskutiert. In allen drei Videosequenzen sind die potenziell zu verwerfenden P -Videobilder mit $Imp_i \leq 4$ im Mittel unter der maximalen IP-Paketgröße. Deshalb greift das Bb -Verfahren kaum. Bei *football* ist die Videobildgrößenverteilung am schmalsten¹⁸. Bei *vtcInw* ist sie am breitesten¹⁹. Die Sequenz *intros* liegt dazwischen. Das BAb -Verfahren greift bei allen Videosequenzen. Die Minderung der Videobildverluste durch den Einsatz von Pb , Bb , bzw. BAb gegenüber OA beträgt bei *vtcInw* ca. 40%, 41% bzw. 48%, bei *intros* ca. 42%, 44% bzw. 51% und bei *football* ca. 48%, 48% bzw. 58%.

Die Verlagerung der Daten aus den I -Videobildern in die P -Videobilder fand bei *IDKf 0B* (s. Abbildung 6.66) in einem sehr geringen Umfang statt (vgl. Abbildung 6.18). Bei *intros* wurden um knapp 0,5% mehr Daten als bei *vtcInw* verlagert, was die Absenkung der *intros*-Kurve erklärt. Die Minderung der Videobildverluste durch den Einsatz von Pb , Bb bzw. BAb gegenüber OA beträgt bei *vtcInw* ca. 40%, 41% bzw. 48%, bei *intros* ca. 43%, 45% bzw. 53% und bei *football* ca. 49%, 50% bzw. 59%.

Bei *IDK 2B* (s. Abbildung 6.64) werden die bereits diskutierten Abstände der Pb -Werte auch für andere Verfahren in etwa beibehalten. Im Vergleich zu $0B$ -Sequenzen zeigt das Bb -Verfahren ausgeprägtere Unterschiede zum Pb . Das Gleiche gilt für BAb im Bezug auf Bb . Die Minderung der Videobildverluste durch den Einsatz von Pb , Bb bzw. BAb gegenüber OA beträgt bei *vtcInw* ca. 26%, 31% bzw. 41%, bei *intros* ca. 34%, 37% bzw. 47% und bei *football* ca. 42%, 45% bzw. 55%.

¹⁷Die verwendete maximale IP-Paketgröße beträgt hier 1500 Byte.

¹⁸Der Variationskoeffizient beträgt 0.21

¹⁹Der Variationskoeffizient liegt bei 0.62

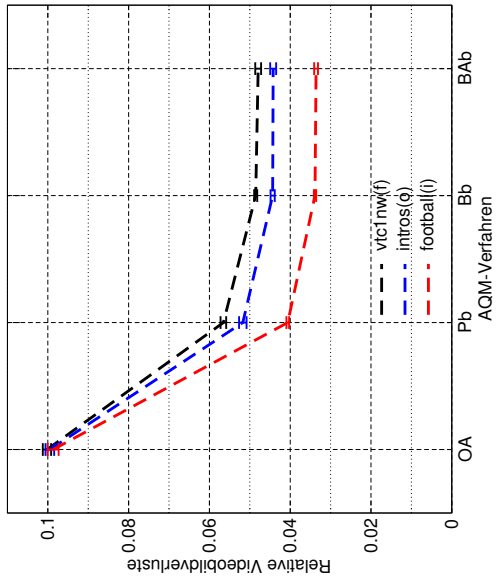


Abbildung 6.65: *IDK 11B Imp_{min} = 4*

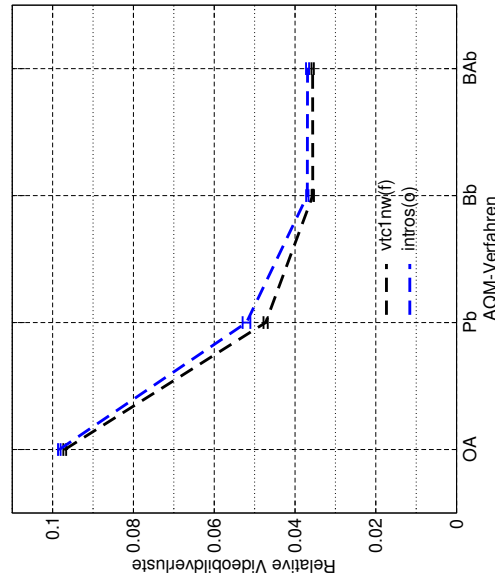


Abbildung 6.68: *IDKf 11B Imp_{min} = 4*

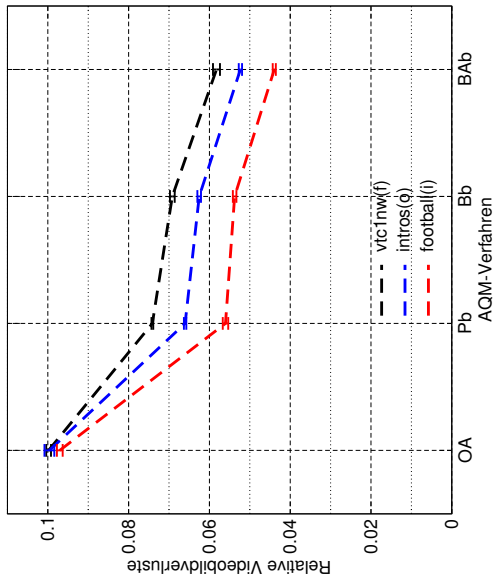


Abbildung 6.64: *IDK 2B Imp_{min} = 4*

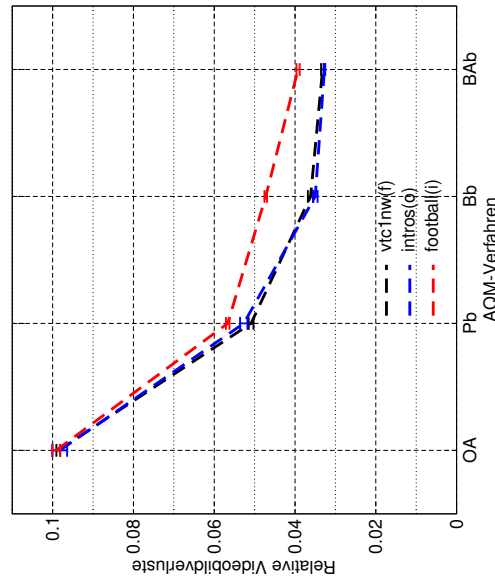


Abbildung 6.67: *IDKf 2B Imp_{min} = 4*

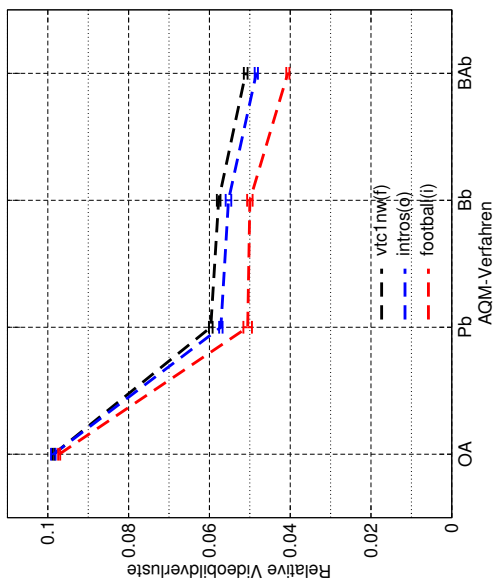


Abbildung 6.63: *IDK 0B Imp_{min} = 4*

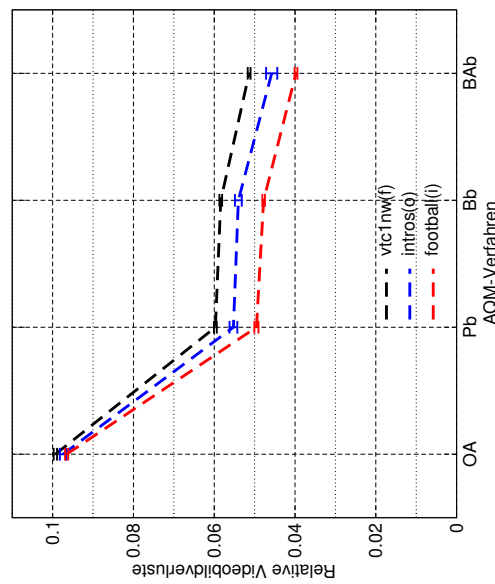


Abbildung 6.66: *IDKf 0B Imp_{min} = 4*

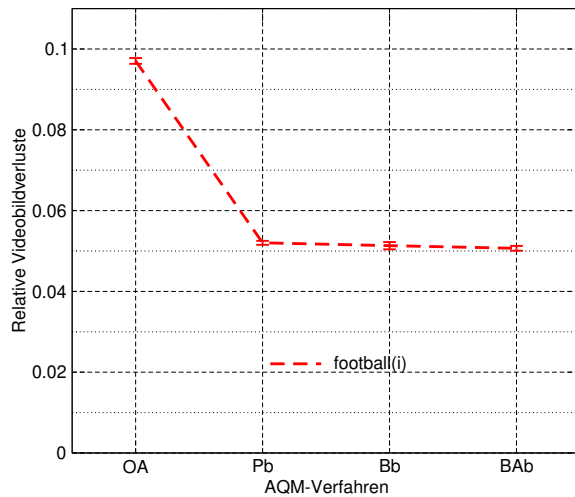
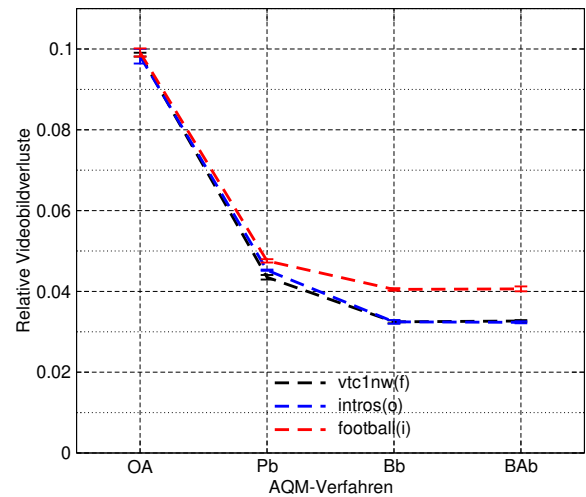
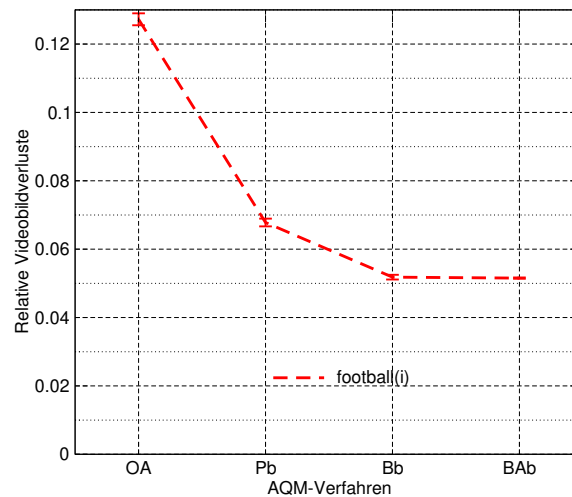
Bei *IDKf 2B* hat die Verlagerung der Daten in die *B*- und *P*-Videobilder insbesondere bei *vtcInw* und *intros* (s. Abbildung 6.16) – wie bereits im Abschnitt 6.3.2.3 besprochen – enorme Vorteile gebracht (s. Abbildung 6.67). Alle Verfahren profitieren davon. Die Minderung der Videobildverluste durch den Einsatz von *Pb*, *Bb* bzw. *BAb* gegenüber *OA* beträgt bei *vtcInw* ca. 48%, 63% bzw. 66%, bei *intros* ca. 46%, 64% bzw. 66% und bei *football* ca. 43%, 52% bzw. 60%.

Bei *IDK 11B* (s. Abbildung 6.65) kann *BAb* seine Stärke nicht zeigen, da es keine *P*-Videobilder gibt. Deshalb erzielt *BAb* dieselben Werte wie *Bb*. Die relative Lage von *football* korreliert mit den Erkenntnissen aus der Abbildung 6.14. Denn *football* besitzt den größten Anteil von Daten in den *B*-Videobildern. Auffallend ist, dass *vtcInw* über *intros* liegt. Gemäß der Abbildung 6.14 würde man umgekehrte Verhältnisse oder ein Zusammenfallen der beiden Kurven erwarten. Denn *vtcInw* trägt marginal mehr Daten in den *B*-Videobildern als *intros*. In den beiden Fällen werden im Netz proaktiv in etwa dieselbe Anzahl von Bytes verworfen. Die Anzahl der verworfenen Videobilder ist bei *vtcInw* höher. Das Phänomen ist auf die Videobildgrößenverteilung der *B*-Videobilder zurückzuführen, die bei *intros* breiter ist. Bei größeren *B*-Videobildern reicht ggf. der Datenumfang eines Videobilds, um den RNC-Pufferfüllstand unter den Schwellwert δ zu senken. Die Minderung der Videobildverluste durch den Einsatz von *Pb*, *Bb* bzw. *BAb* gegenüber *OA* beträgt bei *vtcInw* ca. 42%, 52% bzw. 52%, bei *intros* ca. 48%, 55% bzw. 55% und bei *football* ca. 58%, 66% bzw. 66%.

Bei *IDKf 11B* (s. Abbildung 6.68) korrelieren die relativen Lagen der Kurven mit den Erkenntnissen aus der Abbildung 6.14. Denn *vtcInw* besitzt etwas mehr Daten in den *B*-Videobildern als *intros*. Die *football*-Sequenz weist – wie bereits im Abschnitt 6.3.1.4 erwähnt – eine erhöhte Videodatenrate auf. Deshalb sind ihre Ergebnisse in der separaten Abbildung 6.71 zu sehen. Die Verfahren profitieren von der stärkeren Verlagerung der Daten in die *B*-Videobilder. *BAb* kann seine Stärken nicht ausspielen. Denn die *B*-Videobilder fehlen. Die Minderung der Videobildverluste durch den Einsatz von *Pb*, *Bb* bzw. *BAb* gegenüber *OA* beträgt bei *vtcInw* ca. 51%, 63% bzw. 63%, bei *intros* ca. 47%, 62% bzw. 62% und bei *football* ca. 47%, 59% bzw. 59%.

Die Abbildung 6.69 zeigt *football IDK 2B* für $Imp_{min} = 0$. *Pb* erzielt hier bessere Ergebnisse als bei $Imp_{min} = 4$ (vgl. Abbildung 6.64), denn die Sequenz trägt ausreichend viele Daten in den *B*-Videobildern. Es ist deshalb für *Pb* schädlich, *P*-Videobilder verwerfen zu dürfen. Denn dann werden ggf. Folgefehler verursacht. *Bb* zeigt kaum Wirkung, weil die *B*-Videobilder sich in diesem Fall eher selten auf mehrere IP-Pakete verteilen lassen. *BAb* kann nicht zur Entfaltung kommen, weil es keine *P*-Videobilder zum proaktiven Verwerfen zur Verfügung stehen. Für *BAb* ist $Imp_{min} = 4$ die bessere Wahl (vgl. Abbildung 6.64).

Die Abbildung 6.70 zeigt *IDKf 2B* für $Imp_{min} = 0$. Das sind Sequenzen, die besonders viele Daten in den *B*-Videobildern tragen. Der Anteil liegt gemäß der Abbildung 6.16 über 60%. *BAb* zeigt seine Stärke nicht, weil keine *P*-Videobilder zum proaktiven Verwerfen zur Verfügung stehen. *Pb* erzielt bessere Ergebnisse als bei $Imp_{min} = 4$. *Bb* erzielt sogar dieselben Ergebnisse wie *BAb* bei $Imp_{min} = 4$. Die Minderung der Videobildverluste durch den Einsatz von *Pb*, *Bb* bzw. *BAb* gegenüber *OA* beträgt bei *vtcInw* ca. 56%, 67% bzw. 67%, bei *intros* ca. 54%, 67% bzw. 67% und bei *football* ca. 52%, 59% bzw. 59%. Dieser Fall hat gegenüber $Imp_{min} = 4$ Vorteile in der Aufteilung der Videoverluste über die Zeit. Die Unterbrechungsdauern bzw. die Dauern des Einfrierens des Videobildes werden kürzer. Denn es werden keine Videobilder mit den Videobildabhängigkeiten verworfen.

Abbildung 6.69: 1DK 2B $Imp_{min} = 0$ Abbildung 6.70: 1DKf 2B $Imp_{min} = 0$ Abbildung 6.71: 1DKf 11B $Imp_{min} = 4$

6.3.2.5 Zeitliche Aufteilung der Videobildverluste

Durch das proaktive Verwerfen werden nicht nur die Videoverluste gemindert. Es werden auch die Unterbrechungsdauern (die Zeiten des Videobildeinfrierens aufgrund der Videobildverluste) kleiner. Die Abbildung 6.72 veranschaulicht den Sachverhalt. Dargestellt sind Histogramme der Unterbrechungsdauern für *OA* und für *football 1DK 2B BAb* mit $Imp_{min} = 4$ in doppeltlogarithmischer Darstellung.

Im *OA*-Fall erkennt man mehrere lange Unterbrechungen. Im *BAb*-Fall sind die Unterbrechungen kurz. Man erkennt die Funktionsweise des Verfahrens. Die drei *BAb*-Hauptspalten bei 0,04 s, 0,08 s und 0,20 s repräsentieren zusammen 99,7 % aller Unterbrechungen. Die Unterbrechungen von 0,04 s bzw. 0,08 s entsprechen dem Verwerfen eines bzw. zweier *B*-Videobilder. Denn die Videobildrate beträgt 25 Videobilder/s. Die Unterbrechungen von 0,20 s entsprechen dem Verwerfen eines *P*-Videobildes mit $Imp_{min} = 4$ samt Videobildabhängigkeiten, da es insgesamt genau 5 Videobilder sind (vgl. Tabelle 6.8).

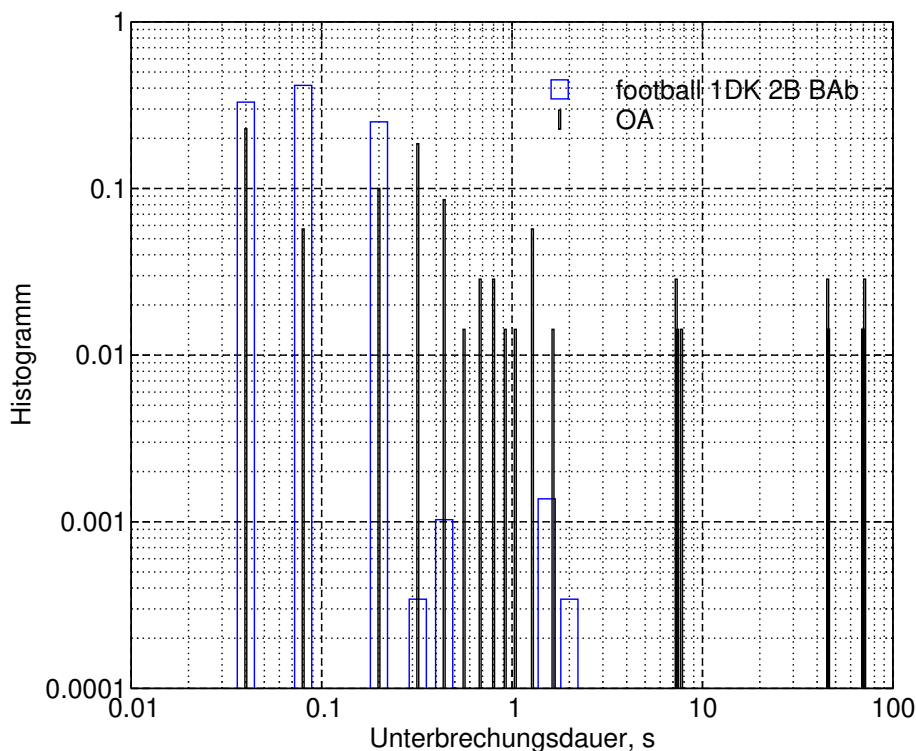


Abbildung 6.72: Histogramm der Unterbrechungsdauern für *football 1DK 2B BAb*

6.3.2.6 Zusammenfassende Diskussion

Die Ergebnisse zeigen die Wirkung von drei in dieser Arbeit vorgeschlagenen videospezifischen AQM-Verfahren in einer RAN-Umgebung. Für die Untersuchungen wurden 27 unterschiedliche Videosequenzen verwendet. Mit den Verfahren können die Videobildverlustraten erheblich gesenkt werden.

Das proaktive *Verwerfen auf Basis der Video-IP-Pakete* erzielt eine videosequenzspezifische Minderung der Videobildverlustrate von 26 % bis 52 %.

Das proaktive *Verwerfen auf Basis der Videobilder* erzielt eine videosequenzspezifische Minderung der Videobildverlustrate von 31 % bis 67 %. Die Verbesserung gegenüber dem paketbasierten Verfahren hängt stark von der Videobildgröße ab. Zu einer durch das paketbasierte Verfahren erzielten Minderung kommt eine weitere additive Minderung von 0 % bis 18 % hinzu. Die Bezugsgröße ist die Videobildverlustrate ohne AQM.

Das proaktive *Verwerfen auf Basis der Videobilder mit anschließendem Verwerfen der abhängigen Videobilder* erzielt eine videosequenzspezifische Minderung der Videobildverlustrate von 41 % bis 67 %. Die Verbesserung gegenüber dem reinen videobildbasierten Verfahren hängt stark von der Quantität der vorhandenen Videobildabhängigkeiten ab. Zu einer durch das reine videobildbasierte Verfahren erzielten Minderung kommt eine weitere additive Minderung von 0 % bis 10 % dazu. Die Bezugsgröße ist die Videobildverlustrate ohne AQM.

Für die übliche GoP-Länge von 12 Videobildern werden optimale Ergebnisse beim Verwerfen der Videodaten erzielt, deren Anzahl der von ihnen abhängigen Videobilder vier oder weniger beträgt.

Der Schwellwert des Füllstandes des IP-Paketpuffers im RAN, bei dessen Überschreitung das proaktive Verwerfen durchgeführt wird, hängt von der Größe des Puffers selbst und von der Größe des Abspielpuffers im Videoclient ab. Desweiteren besteht eine Abhängigkeit von der Videostruktur (z. B. *0B*, *2B*, *11B*). Letztere bietet dem proaktiven Verwerfen je nach ihrer Ausprägung – bei einer konstanten Schwelle für minderwertige Videobildwichtigkeiten – eine unterschiedliche Maximalanzahl der zu verwerfenden Videobilder.

Das proaktive Verwerfen reduziert die Unterbrechungsdauern (Einfrieren des Videobildes aufgrund von Videobildverlusten) beim Abspielen. D. h., neben der Minderung der Videobildverluste, wird die zeitliche Blockbildung von Videobildverlusten erheblich reduziert.

7 Zusammenfassung und Ausblick

In der vorliegenden Arbeit wird eine eingebettete netzgesteuerte Videostreamingadaption in einem Mobilfunknetz vorgeschlagen und untersucht. Das Konzept besteht aus zwei Komponenten. Die eine Komponente trifft mit Kenntnis der Netzgegebenheiten und Fähigkeiten der Endsysteme eine Adaptionentscheidung und übermittelt einen Adaptionsvorschlag an den Videostreamingserver. In den zeitlichen Phasen knapper Übertragungsressourcen, in denen keine Videodatenratenadaption möglich ist, wendet die zweite Komponente ein proaktives videospezifisches Verwerfen von Videodaten basierend auf ihrer quantitativen Wichtigkeit an. Damit wird die Videodatenverluststruktur geändert, was aus Sicht der Anwendung zu weniger Videodatenverlusten führt.

Die vorgeschlagenen Verfahren werden unter Einsatz einer eigens dazu entwickelten experimentellen Videostreaming-Untersuchungsumgebung erforscht. Um die Ergebnisse zu verallgemeinern, werden 27 komprimierte Videosequenzen verwendet, die unterschiedliche Videoinhalte und unterschiedliche Videostrukturen aufweisen und mit drei verschiedenen Vorgehensweisen bei der Komprimierung erzeugt wurden. Zur anschaulichen Bewertung der Videoqualität wird ein für diese Arbeit entwickelter Diagrammtyp vorgeschlagen.

Die Ergebnisse des Verfahrens zur Ermittlung der Adaptionsvorschläge zeigen eine schnelle und präzise Adaption der Videodatenrate. Es gehen keine Videodaten durch eine stark schwankende Mobilfunkkanalqualität verloren. Die ggf. steigenden Übertragungsressourcen werden für die Steigerung der Videoqualität effizient ausgenutzt. Die Anzahl der Adaptionsvorschläge ist um zwei Größenordnungen kleiner als die Anzahl der relevanten Netznachrichten, die eine Änderung der Übertragungsdatenrate signalisieren. Somit werden kommunikationstechnische Skalierungsprobleme bei einem Videostreamingserver vermieden. Durch die Einbettung in das Mobilfunknetz können zahlreiche netzspezifische Randbedingungen bei der Ermittlung von Adaptionsvorschlägen berücksichtigt werden.

Das vorgeschlagene proaktive videospezifische Verwerfen der Videodaten basierend auf ihrer quantitativen Wichtigkeit bewirkt in den zeitlichen Phasen knapper Übertragungsressourcen eine starke Minderung der Videobildverlustrate bis zu 67%. Die unterschiedliche Wirkung auf verschiedene Videosequenztypen wird herausgearbeitet. Ebenfalls wird der quantitative Wirkungsunterschied zwischen den drei untersuchten Verfahrensausprägungen festgestellt. Die Parametrierung der Verfahren wird dank des repräsentativen Videosequenzsatzes verallgemeinert ermittelt. Es wird gezeigt, dass neben der deutlichen Minderung der Videobildverlustrate die Blockbildung von Videobildverlusten erheblich reduziert wird. Durch die Einführung der quantitativen Wichtigkeit und die Entkopplung vom Videobildtyp lässt sich das Verfahren auf beliebige Ausprägungen der Videodatenabhängigkeiten anwenden.

Ausblickend könnte die eingebettete netzgesteuerte Videostreamingadaption mit Protokollen für die Kommunikation mit den Endsystemen (Videostreamingserver und -client) ausgestattet werden. Die darüber zu beziehenden Informationen erlauben eine passende Parametrierung der vorgeschlagenen Verfahren. Zukünftige Untersuchungen müßten mit anderen Ausprägungen des Mobilfunkkanals durchgeführt werden, um weitere Verfahren der Schätzung der Übertragungsdatenrate zu untersuchen und die Parametrierungsrichtlinien des Hysterese-Faktors ggf. allgemeingültig zu ermitteln. Es ist denkbar, dass die Hysterese-Faktoren für das Herauf- und Herunterschalten unterschiedlich sein können. In weiteren Untersuchungen des proaktiven Verwerfens könnte auf die Videobildabhängigkeiten auf der Ebene der Makroblöcke (Teile eines Videobildes) eingegangen werden.

A Anhang

Tabelle A.1: Verwendete Komprimierungsparameter für *IDKf*

<i>Video</i>	<i>IDKf-Parameter</i>		<i>Bemerkung</i>
	<i>bq</i>	<i>iq</i>	
<i>vtcInw 0B</i>	0.4	0.4	
<i>vtcInw nB</i>	0.4	nicht verwendet	
<i>intros 0B</i>	0.4	0.4	
<i>intros nB</i>	0.4	nicht verwendet	
<i>3inrow 0B</i>	0.4	0.4	
<i>3inrow nB</i>	0.4	nicht verwendet	
<i>smity1 0B</i>	0.4	0.4	
<i>smity1 nB</i>	0.4	nicht verwendet	
<i>washdc 0B</i>	0.4	0.4	
<i>washdc nB</i>	0.4	nicht verwendet	
<i>football 0B</i>	2	2	
<i>football 5B</i>	0.4	nicht verwendet	mittlere Datenrate 320 kbit/s bei dem eingestellten Wert von 308 kbit/s
<i>football 8B</i>	0.4	nicht verwendet	mittlere Datenrate 330 kbit/s bei dem eingestellten Wert von 308 kbit/s
<i>football 11B</i>	0.4	nicht verwendet	mittlere Datenrate 335 kbit/s bei dem eingestellten Wert von 308 kbit/s
<i>football nB</i>	0.4	nicht verwendet	

Tabelle A.2: Protokolloverhead, 308 kbit/s

Video	Protokolloverhead																	
	512 / 40			512 / 44			1024 / 40			1024 / 44			1456 / 40			1456 / 44		
	1DK	2DK	1DKf	1DK	2DK	1DKf	1DK	2DK	1DKf	1DK	2DK	1DKf	1DK	2DK	1DKf	1DK	2DK	1DKf
<i>vtcInw 0B</i>	9.1	9.3	9.1	10	10.3	10	5.2	5.7	5.2	5.7	6.3	5.7	4	4.6	4	4.4	5.1	4.4
<i>vtcInw 2B</i>	9	9.6	9.1	9.9	10.6	10	5.2	5.9	5.1	5.7	6.5	5.6	4.3	4.9	4.3	4.7	5.3	4.7
<i>vtcInw 11B</i>	9	9.8	9.1	9.9	10.8	10.1	5.2	6.2	5.2	5.8	6.8	5.7	4	5.1	4	4.4	5.6	4.4
<i>intros 0B</i>	9.2	9.3	9.1	10	10.2	9.9	5.2	5.6	5.2	5.7	6.2	5.7	3.9	4.4	3.9	4.2	4.9	4.3
<i>intros 2B</i>	9	9.4	9.1	9.9	10.4	10	5.1	5.7	4.9	5.6	6.3	5.4	4.1	4.6	4.5	4.5	5.1	4.9
<i>intros 11B</i>	9.1	9.4	9	10	10.3	9.9	5.3	5.8	5.3	5.8	6.4	5.8	4	4.8	4.4	4.4	5.2	4.8
<i>football 0B</i>	9.1	9.1	8.9	10	10	9.8	5.5	5.3	5.3	6	5.9	5.9	3.7	3.9	3.8	4	4.3	4.2
<i>football 2B</i>	9	9.2	9	9.9	10.1	9.9	5.1	5.1	5.3	5.7	5.6	5.8	3.8	3.9	3.9	4.2	4.3	4.3
<i>football 11B</i>	9.1	9.1	8.9	10	10	9.8	5.3	5.3	5.2	5.8	5.8	5.7	3.8	3.8	4	4.2	4.2	4.4

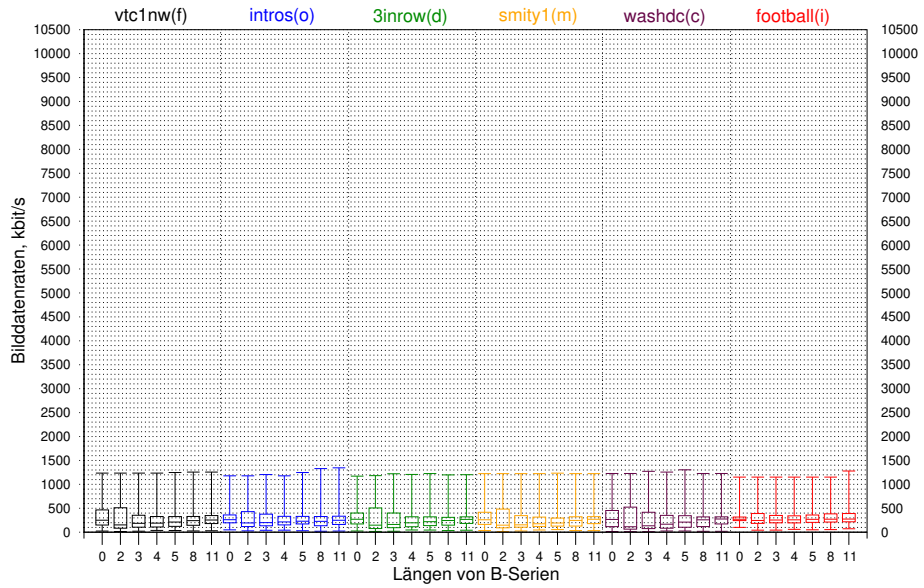


Abbildung A.1: Datenraten, *IDK*

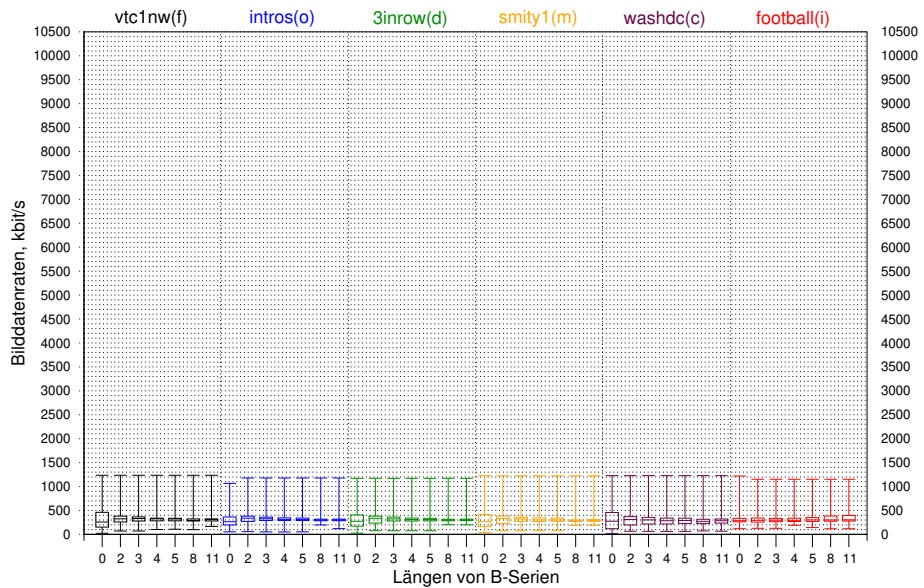


Abbildung A.2: Datenraten, *IDKf*

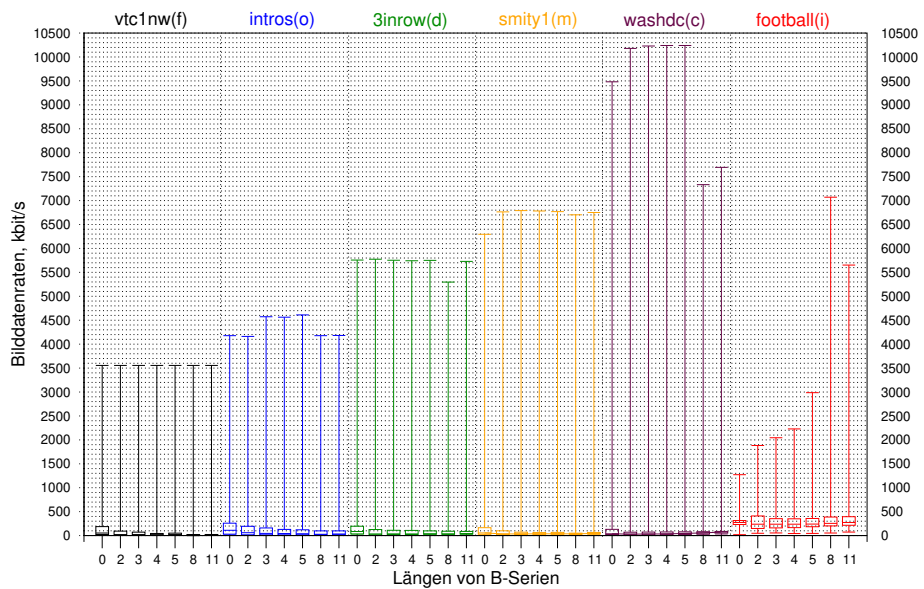


Abbildung A.3: Datenraten, *2DK*

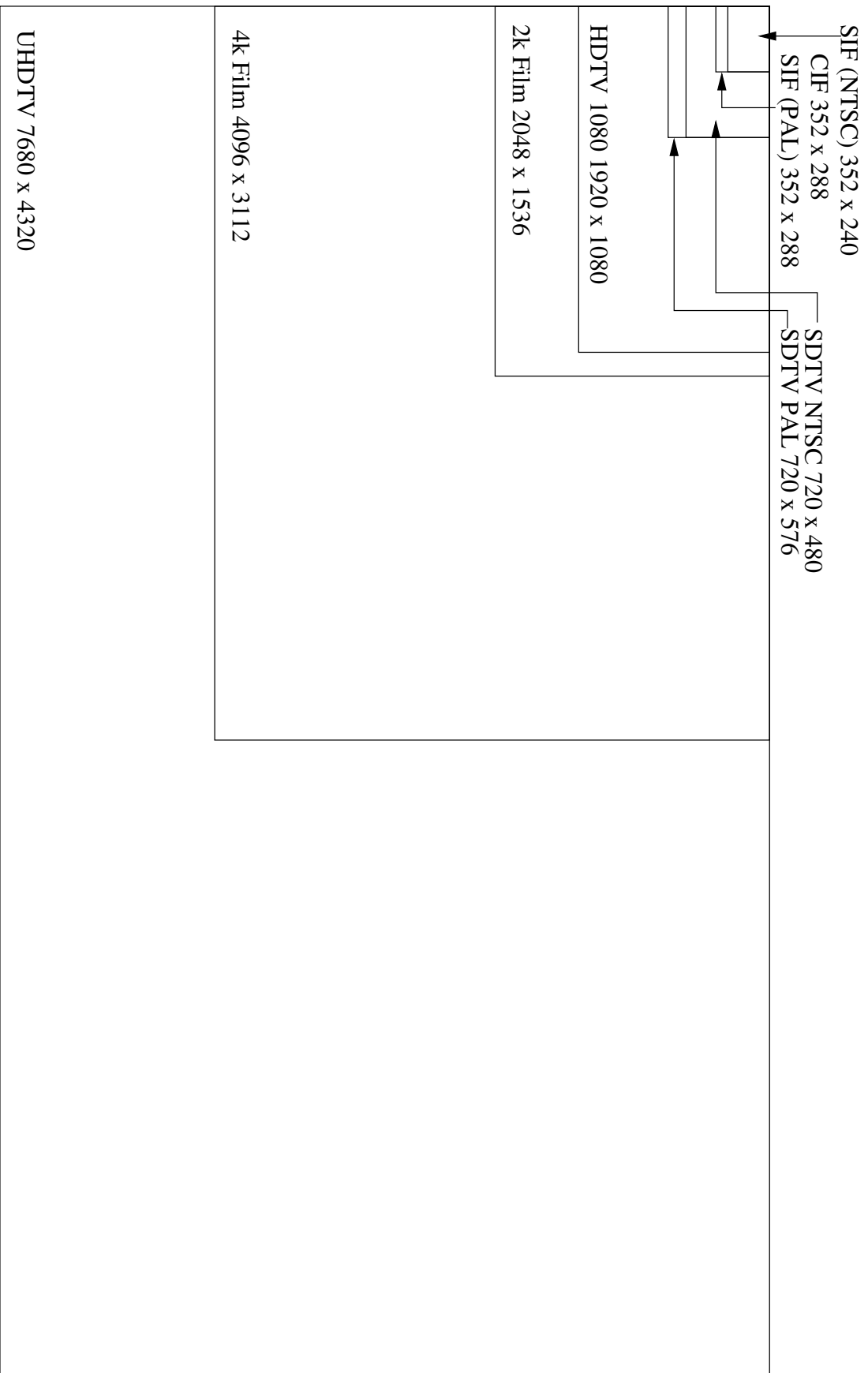


Abbildung A.4: Örtliche Auflösung

Literaturverzeichnis

- [1] ETSI 3GPP TS 26.247, Universal Mobile Telecommunications System (UMTS); LTE; Transparent end-to-end Packet-Switched Streaming Service (PSS); Progressive Download and Dynamic Adaptive Streaming over HTTP (3GP-DASH), July 2012.
- [2] ISMA. Internet Streaming Media Alliance, Implementation Specification, Version 2.0, April 2005.
- [3] 3GPP TS 126 234 v15.1.0, Transparent end-to-end Packet-switched Streaming Service (PSS); Protocols and codecs (3GPP TS 26.234 version 15.1.0 Release 15), September 2018.
- [4] ETSI TS 126 346, Universal Mobile Telecommunications System (UMTS); LTE; Multimedia Broadcast/Multicast Service (MBMS); Protocols and codecs, 2009.
- [5] 3GPP TS 26 346, Multimedia Broadcast/Multicast Service (MBMS); Protocols and codecs, 2018.
- [6] ETSI TS 102034, Digital Video Broadcasting (DVB); Transport of MPEG-2 DVB Services over IP Based Networks, 2007.
- [7] ITU-T Recommendation P.910, Subjective video quality assessment methods for multimedia applications, April 2008.
- [8] J. Klaue, B. Rathke, A. Wolisz. EvalVid - A Framework for Video Transmission and Quality Evaluation. In *Proc. 13th Conference on Modelling Techniques and Tools for Computer Performance Evaluation*, September 2003.
- [9] Jens-Rainer Ohm. Bildsignalverarbeitung für Multimedia-Systeme, Skript zur Vorlesung, 1999.
- [10] Wikipedia. http://de.wikipedia.org/wiki/Streaming_Audio. 2007.
- [11] M. Ruhnke. Neue Distributionskanäle für Fernsehsender durch Streaming Video. Master's thesis, FH Wiesbaden, 2001.
- [12] Lettermann. <http://users.abac.com/ksitterley/billgate.htm>, 1995.
- [13] Real Networks. <http://service.real.com/help/library/whitepapers/rproxy/htmlfiles/proxy2gl.htm>, 2000.

- [14] ITU-R Recommendation BT.601-7 (03/2011), Studio encoding parameters of digital television for standard 4:3 and wide screen 16:9 aspect ratios , 2011.
- [15] W. Fischer. *Digitale Fernseh- und Hörfunktechnik in Theorie und Praxis*. Springer, 2009.
- [16] ITU-T Recommendation H.261, Video Codec For Audiovisual Services at p x 64 kbits, 1993.
- [17] K. Jack. *Video Demystified*. Elsevier, 2007.
- [18] K. Debes. Minimierung elektromagnetischer Felder des Mobilfunks, UMTS, DECT, Powerline und Induktionsfunkanlagen. Abschlussbericht, Juli 2004.
- [19] N. Geng and W. Wiesbeck. *Planungsmethoden für die Mobilkommunikation*. Springer-Verlag Berlin Heidelberg, 1998.
- [20] J. T. Hviid, J. B. Andersen, J. Toftgard, and J. Bojer. Terrain-based propagation model for rural area-an integral equation approach. *Antennas and Propagation, IEEE Transactions on*, 43(1):41–46, Jan 1995.
- [21] G. Vasilescu. *Electronic Noise And Interfering Signals*. Springer-Verlag Berlin Heidelberg, 2005.
- [22] V. Tuzlukov. *Signal Processing Noise*. CRC Press, 2002.
- [23] ETSI ETR103, European digital cellular telecommunication system; Radio network planning aspects, February 1995.
- [24] UMTS world. WCDMA Link Budget, <http://www.umtsworld.com/technology/linkbudget.htm>.
- [25] LTE Encyclopedia: LTE Radio Link Budgeting and RF Planning, <https://sites.google.com/site/lteencyclopedia/lte-radio-link-budgeting-and-rf-planning>.
- [26] LTE Encyclopedia: LTE Link Budget Comparison, <https://sites.google.com/site/lteencyclopedia/lte-radio-link-budgeting-and-rf-planning/>.
- [27] Martin Sauter. *Grundkurs Mobile Kommunikationssysteme*. Springer Vieweg, 2018.
- [28] M. Pätzold. *Mobilfunkkanäle*. Vieweg, 1999.
- [29] ETSI EN 301192, Digital Video Broadcasting(DVB); DVB specification for data broadcasting, 2008.
- [30] I. Sodagar. The MPEG-DASH Standard for Multimedia Streaming Over the Internet. *IEEE Multimedia*, pages 62–67, October-December 2011.
- [31] Adobe. Adobe’s Real Time Messaging Protocol, December 2012.
- [32] Adobe. HTTP Dynamic Streaming Specification, version 3.0 final, 2013.
- [33] Apple. HTTP Live Streaming, https://developer.apple.com/documentation/http_live_streaming.
- [34] Microsoft. Smooth Streaming, <http://www.iis.net/downloads/microsoft/smooth-streaming>.

- [35] IETF RFC 7826, Real Time Streaming Protocol 2.0 (RTSP), December 2016.
- [36] IETF RFC 7016, Adobe's Secure Real-Time Media Flow Protocol, November 2013.
- [37] SMPTE ST 2110, Professional Media Over Managed IP Networks Suite of Standards, September 2018.
- [38] ISO/IEC 7498-1, Information technology - Open Systems Interconnection - Basic Reference Model: The Basic Model, 11 1994.
- [39] Andrew S. Tanenbaum. *Computernetzwerke*. Pearson Studium, 3 edition, 2000.
- [40] ISO/IEC 14496-1, Information technology - Coding of audio-visual objects - Part 1: Systems, 11 2004.
- [41] ISO/IEC JTC1/SC29/WG11, MPEG-4 Overview, July 2000.
- [42] ISO/IEC 13818-1, Information technology - Generic coding of moving picture and associated audio information: Systems, 12 2000.
- [43] ISO/IEC 23009-1, Information technology - Dynamic adaptive streaming over HTTP (DASH) - Part 1: Media presentation description and segment formats, April 2012.
- [44] T. Stockhammer. Dynamic Adaptive Streaming over HTTP - Design Principles and Standards. In *Proceedings of the second annual ACM conference on Multimedia systems*, pages 133–144. ACM Press, September 2011.
- [45] DASH Industry Forum. <http://dashif.org>.
- [46] WOWZA. Using Network Cameras for 24x7 Live Streaming: Wowza + Axis, 2019.
- [47] ISMA. Internet Streaming Media Alliance, Implementation Specification, Version 1.0.1, June 2004.
- [48] ISMA. [http://www.mpegif.org/m4if/bod/ISMA/ISMA Index.php](http://www.mpegif.org/m4if/bod/ISMA/ISMA%20Index.php).
- [49] ETSI TS 126 233, Universal Mobile Telecommunications System (UMTS); LTE; Transparent end-to-end Packet-switched Streaming Service (PSS); General description, Juni 2011.
- [50] ETSI TS 126 234 v10.1.0, Universal Mobile Telecommunications System (UMTS); LTE; Transparent end-to-end Packet-switched Streaming Service (PSS); Protocols and codecs (3GPP TS 26.234 version 10.1.0 Release 10), September 2011.
- [51] ETSI TS 123 246, Universal Mobile Telecommunications System (UMTS); LTE; Multimedia Broadcast/Multicast Service (MBMS); Architecture and functional description, 2009.
- [52] ETSI TR 102 033 v1.1.1, Digital Video Broadcasting (DVB); Architectural framework for the delivery of DVB-services over IP-based networks, April 2002.
- [53] NATO Standardisation Agreement (STANAG) 4609 on Air Reconnaissance Primary Imagery Data Standard, 2006.

- [54] NATO Motion Imagery (MI) STANAG 4609, Implementation Guide, 2006.
- [55] ETSI TS 102 470 v1.2.1, "Digital Video Broadcastin (DVB); IP Datacast: Program Specific Information (PSI)/Service Information (SI); Part 1: IP Datacast over DVB-H", March 2009.
- [56] ETSI TR 102 469 v1.1.1, "Digital Video Broadcasting (DVB); IP Datacast over DVB-H: Architecture", May 2006.
- [57] ETSI TS 102 472 v1.2.1, "Digital Video Broadcasting (DVB); IP Datacast over DVB-H: Content Delivery Protocols", December 2006.
- [58] W. Dankmeier. *Grundkurs Codierung: Verschlüsselung, Kompression, Fehlerbeseitigung*. Vieweg, 2006.
- [59] Roland Schmitz, Roland Kiefer, Johannes Maucher, Jan Schulze, and Thomas Suchy. *Kompendium Medieninformatik, Mediennetze*. Springer-Verlag Berlin Heidelberg, 2006.
- [60] WOWZA media systems. MPEG-DASH segmenter property reference, 2019.
- [61] BITMOVIN. Optimal Adaptive Streaming Formats MPEG-DASH & HLS Segment Length, 2019.
- [62] P. Froejdh, U. Horn, M. Kampmann, A. Nohlgren, and M. Westerlund. Adaptive Streaming within the 3GPP Packet-Switched Streaming Service. *IEEE Network*, pages 34–40, March/April 2006.
- [63] IETF RFC 3550, RTP: A Transport Protocol for Real-Time Applications, July 2003.
- [64] M. Laner, P. Svoboda, P. Romirer-Maierhofer, N. Nikaein, F. Ricciato, and M. Rupp. A comparison between one-way delays in operating HSPA and LTE networks. In *2012 10th International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt)*, pages 286–292, May 2012.
- [65] H. Wu and K. Rao, editors. *Digital Video Image Quality and Perceptual Coding*. CRC Press, 2006.
- [66] Y. Q. Shi and H. Sun. *Image and Video Compression for Multimedia Engineering: Fundamentals, Algorithms, and Standards*. CRC Press, 2000.
- [67] K. R. Rao and P. C. Yip, editors. *The Transform And Data Compression Handbook*. CRC Press, 2001.
- [68] J. Bergh, F. Ekstedt, and M. Lindberg. *Wavelets mit Anwendungen in Signal- und Bildverarbeitung*. Springer, 2007.
- [69] B. Jähne. *Practical Handbook on Image Processing for Scientific and Technical Applications*. CRC Press, 2004.
- [70] Q. Cai, L. Song, G. Li, and N. Ling. Lossy and lossless intra coding performance evaluation: HEVC, H.264/AVC, JPEG 2000 and JPEG LS. In *Proceedings of The 2012 Asia Pacific Signal and Information Processing Association Annual Summit and Conference*, pages 1–9, Dec 2012.

- [71] U. Schmidt. *Professionelle Videotechnik*. Springer-Verlag Berlin Heidelberg, 2009.
- [72] Pinnacle Systems. MPEG-2 White Paper. Technical report, Pinnacle Systems, Februar 2000.
- [73] M. Pettersson, R. Sjöberg, and J. Samuelsson. Dependent random access point pictures in hevc. In *2015 IEEE International Conference on Image Processing (ICIP)*, pages 867–871, September 2015.
- [74] J. Ohm, G. J. Sullivan, H. Schwarz, T. K. Tan, and T. Wiegand. Comparison of the Coding Efficiency of Video Coding Standards - Including High Efficiency Video Coding (HEVC). *IEEE Transactions on Circuits and Systems for Video Technology*, 22(12):1669–1684, Dec 2012.
- [75] J. G. Apostolopoulos, W. Tan, and S. J. Wee. Video Streaming: Concepts, Algorithms and Systems. Technical report, Mobile and Media Systems Laboratory, HP Laboratories Palo Alto, Hewlett-Packard Company, September 18th, 2002.
- [76] Alexandros Eleftheriadis and Dimitris Anastassiou. Meeting arbitrary QoS constraints using dynamic rate shaping of coded digital video. In *International Workshop on Network and Operating Systems Support for Digital Audio and Video*, pages 88–100. Springer, 1995.
- [77] R. B. Miller, J. P. Hearn, C. Purzynski, F. Cuervo, and M. Scheutzow. Mobile video delivery using network aware transcoding in an LTE network. *Bell Labs Technical Journal*, 16(4):43–61, March 2012.
- [78] Fraunhofer Heinrich-Hertz-Institut. Scalable Video Coding in H.264/AVC, <https://www.hhi.fraunhofer.de/en/departments/vca/research-groups/image-video-coding/research-topics/scalable-video-coding-in-h264avc.html>.
- [79] Vivek K. Goyal. Multiple description coding: Compression meets the network. *IEEE Signal processing magazine*, 18(5):74–93, 2001.
- [80] Jitae Shin, Jong Won Kim, and C. C. Kuo. Quality-of-service mapping mechanism for packet video in differentiated services network. *IEEE Transactions on Multimedia*, 3(2):219–231, 2001.
- [81] Toufik Ahmed, Guillaume Buridant, and Ahmed Mehaoua. Encapsulation and marking of MPEG-4 video over IP differentiated services. In *Computers and Communications, 2001. Proceedings. Sixth IEEE Symposium on*, pages 346–352. IEEE, 2001.
- [82] Davide Quaglia and Juan Carlos De Martin. Delivery of MPEG video streams with constant perceptual quality of service. In *Multimedia and Expo, 2002. ICME'02. Proceedings. 2002 IEEE International Conference on*, volume 2, pages 85–88. IEEE, 2002.
- [83] Jie Huang, Charles Krasic, Jonathan Walpole, and Wu-chi Feng. Adaptive live video streaming by priority drop. In *Proceedings of the IEEE Conference on Advanced Video and Signal Based Surveillance*, page 342. IEEE Computer Society, 2003.
- [84] Mark Kalman and Bernd Girod. Techniques for improved rate-distortion optimized video streaming. *ST Journal of Research*, 2(1):45–54, 2005.

- [85] D. Wu, Y. T. Hou, and Y. Zhang. Transporting Real-Time Video over the Internet: Challenges and Approaches. *Proceedings of the IEEE*, Vol. 66, No. 12:1855 – 1875, December 2000.
- [86] X. Zhu, R. Pan, N. Dukkupati, V. Subramanian, and F. Bonomi. Layered Internet Video Engineering (LIVE): Network-Assisted Bandwidth Sharing and Transient Loss Protection for Scalable Video Streaming. In *2010 Proceedings IEEE INFOCOM*, pages 1–5, March 2010.
- [87] Uwe Horn, K. Stuhlmüller, M. Link, and B. Girod. Robust internet video transmission based on scalable coding and unequal error protection. *Signal Processing: Image Communication*, 15(1):77 – 94, 1999.
- [88] Hao Liu, Wenjun Zhang, Songyu Yu, and Xiaokang Yang. Channel-Aware Frame Dropping for Cellular Video Streaming. In *2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings*, July 2006.
- [89] Nick Feamster and Hari Balakrishnan. Packet loss recovery for streaming video. In *12th International Packet Video Workshop*, pages 9–16. PA: Pittsburgh, 2002.
- [90] ITU-R Recommendation BT.500-11, Methodology for the subjective assessment of the quality of television pictures, 2002.
- [91] Hong Ren Wu and Kamisetty Ramamohan Rao. *Digital video image quality and perceptual coding*. CRC press, 2017.
- [92] ITU-T Recommendation P.930, Principles of a reference impairment system for video, 08 1996.
- [93] ANSI T1.801.03-2003, Digital Transport of One-Way Signals - Parameters for Objective Performance Assessment, 2003.
- [94] ANSI T1.801.01-1995 (R2001), Digital Transport of Video Teleconferencing/Video Telephony Signals - Video Test Scenes for Subjective and Objective Performance Assessment, 1995.
- [95] Hemant Kanakia, Partho P. Mishra, and Amy Reibman. An Adaptive Congestion Control Scheme for Real-time Packet Video Transport. In *Conference Proceedings on Communications Architectures, Protocols and Applications, SIGCOMM '93*, pages 20–31, New York, NY, USA, 1993. ACM.
- [96] T. Yoshimura. QoS control architecture with RTP monitoring agent for mobile multimedia streaming. In *IPSSJ Symposium on Multimedia, Distributed, Cooperative and Mobile Systems (DICO 2001)*, 2001.
- [97] T. Yoshimura, T. Ohya, T. Kawahara, and M. Etoh. Rate and robustness control with RTP monitoring agent for mobile multimedia streaming. In *2002 IEEE International Conference on Communications. Conference Proceedings. ICC 2002 (Cat. No.02CH37333)*, volume 4, pages 2513–2517 vol.4, April 2002.

- [98] Gene Cheung and Takeshi Yoshimura. Streaming agent: A network proxy for media streaming in 3G wireless networks. In *Packet Video Workshop*, volume 1, pages 529–532, 2002.
- [99] Gene Cheung, Wai-Tian Tan, and Takeshi Yoshimura. Double feedback streaming agent for real-time delivery of media over 3G wireless networks. *IEEE Transactions on Multimedia*, 6(2):304–314, April 2004.
- [100] Gene Cheung, Wai-Tian Tan, and Takeshi Yoshimura. Real-time video transport optimization using streaming agent over 3G wireless networks. *IEEE Transactions on Multimedia*, 7(4):777–785, Aug 2005.
- [101] S. Chemiakina, L. D’Antonio, F. Forti, R. Lalli, J. Petersson, and A. Terzani. QoS enhancement for adaptive streaming services over WCDMA. *IEEE Journal on Selected Areas in Communications*, 21(10):1575–1584, Dec 2003.
- [102] Toufik Ahmed, Youssef Iraqi, Raouf Boutaba, and Ahmed Mehaoua. Streaming MPEG-4 Audio-Visual Objects with Quality Adaptation.
- [103] Toufik Ahmed and Ahmed Mehaoua. Adaptive MPEG-4 streaming based on AVO classification and network congestion feedback. In *Packet Video*, 2003.
- [104] Y. Onoe, H. Hagino, Y. Atsumi, Y. Miyazaki, H. Komaki, T. Kushida, and N. Yamanouchi. Network information based rate controls on multimedia streaming servers. In *23rd International Conference on Distributed Computing Systems Workshops, 2003. Proceedings.*, pages 543–548, May 2003.
- [105] N. Moller, I. C. Molero, K. H. Johansson, J. Petersson, R. Skog, and A. Arvidsson. Using Radio Network Feedback to Improve TCP Performance over Cellular Networks. In *Proceedings of the 44th IEEE Conference on Decision and Control*, pages 7434–7439, Dec 2005.
- [106] Inés Cabrera Molero, Niels Möller, Justus Petersson, Robert Skog, Åke Arvidsson, Oscar Flårdh, and Karl H Johansson. Cross-layer adaptation for TCP-based applications in WCDMA systems. In *IST Mobile & Wireless Communications Summit, Dresden, Germany*, 2005.
- [107] A. Raghuv eer, E. Kusmier ek, and D. H. C. Du. Network-aware rate adaptation for video streaming. In *2004 IEEE International Conference on Multimedia and Expo (ICME) (IEEE Cat. No.04TH8763)*, volume 2, pages 1039–1042 Vol.2, June 2004.
- [108] Rim Hammi and Ken Chen. Some experience on video flow regulation with an active network approach. In *2005 Dagstuhl Seminar*, 2005.
- [109] Zigmund Orlov. Network-driven adaptive video streaming in wireless environments. In *2008 IEEE 19th International Symposium on Personal, Indoor and Mobile Radio Communications*, pages 1–6, September 2008.
- [110] Wilder E. Castellanos, Juan C. Guerri, and Pau Arce. A QoS-aware routing protocol with adaptive feedback scheme for video streaming for mobile networks. *Computer Communications*, 77:10–25, 2016.

- [111] Lars Burgstahler and Martin Neubauer. Trend-based Modifications of the Exponential Moving Average Algorithm for Bandwidth Estimation. In *2nd Polish-German Teletraffic Symposium PGTS*, 2002.
- [112] P. Mertens, S. Raessler. *Prognoserechnung*. Physica-Verlag, 6-te auflage edition, 2005.
- [113] R. G. Brown. *Smoothing, Forecasting and Prediction of Discrete Time Series*. Englewood Cliffs, 1963.
- [114] N. Challa and H. Cam. Cost-aware downlink scheduling of shared channels for cellular networks with relays. In *IEEE International Conference on Performance, Computing, and Communications, 2004*, pages 793–798, April 2004.
- [115] I. Joe. Packet loss and jitter control for real-time MPEG video communications. *COMPUTER COMMUNICATIONS*, 19(11):901–914, September 1996.
- [116] Z. Zhanga, S. Nelakuditia, R. Aggarwala, R. Tsang. Efficient Selective Frame Discard Algorithms for Stored Video Delivery Across Resource Constrained Networks. In *IEEE*, (*Konferenz ist nicht bekannt*), 1999.
- [117] Z. Zhanga, S. Nelakuditia, R. Aggarwala, R. Tsang. Efficient Selective Frame Discard Algorithms for Stored Video Delivery Across Resource Constrained Networks. *Real-Time Imaging*, 7(3):255–273, June 2001.
- [118] M. Hemy. MPEG system streams in best-effort networks. *Proceedings of Packet Video99*, Apr., 1999.
- [119] Jitae Shin, Jong Won Kim, and C. C. Jay Kuo. Content-based packet video forwarding mechanism in differentiated service networks. In *IEEE Packet Video Workshop 00 (May 2000)*, 2000.
- [120] Jitae Shin, JongWon Kim, Daniel C. Lee, and C. C. Kuo. Adaptive packet forwarding for relative differentiated services and categorized packet video. In *Communications, 2001. ICC 2001. IEEE International Conference on*, volume 3, pages 763–767. IEEE, 2001.
- [121] Jitae Shin, Jin-Gyeong Kim, JongWon Kim, D. C. Lee, and C. C. Kuo. Aggregated QoS mapping framework for relative service differentiation-aware video streaming. In *Proc. Packet Video Workshop, Korea*, 2001.
- [122] Anand Balan, Omesh Tickoo, Ivan Bajic, Shivkumar Kalyanaraman, and John Woods. Integrated buffer management and congestion control for video streaming. In *Proc. of the 10th IEEE International Conference on Image Processing, Barcelona, Spain*, 2003.
- [123] I. V. Bajic, O. Tickoo, A. Balan, S. Kalyanaraman, and J. W. Woods. Integrated end-to-end buffer management and congestion control for scalable video communications. In *Proceedings 2003 International Conference on Image Processing (Cat. No.03CH37429)*, volume 3, pages III–257, September 2003.
- [124] Andrei Gurtov and Reiner Ludwig. Lifetime Packet Discard for Efficient Real-time Transport over Cellular Links. *SIGMOBILE Mob. Comput. Commun. Rev.*, 7(4):32–45, October 2003.

- [125] Y. Bai and M. R. Ito. Packet scheduling to support loss guarantee for video traffic. In *10th International Conference on Telecommunications, 2003. ICT 2003.*, volume 2, pages 1343–1348 vol.2, Februar 2003.
- [126] M. R. Ito and Y. Bai. A packet discard scheme for loss control in IP networks with MPEG video traffic. In *iccs*, pages 497–503. IEEE, 2002.
- [127] Y. Bai and M. R. Ito. Application-aware buffer management: new metrics and techniques. *IEEE Transactions on Broadcasting*, 51(1):114–121, March 2005.
- [128] K. Chebrolu, R. R. Rao. Selective frame discard for interactive video. In *2004 IEEE International Conference on Communications*, July 2004.
- [129] Günther Liebl, Hrvoje Jenkac, Thomas Stockhammer, Christian Buchner, and Axel Klein. Radio link buffer management and scheduling for video streaming over wireless shared channels. In *In Proceedings of the Packet Video Workshop*, 2004.
- [130] T. Ahmed, A. Mehaoua, R. Boutaba, and Y. Iraqi. Adaptive packet video streaming over IP networks: a cross-layer approach. *IEEE Journal on Selected Areas in Communications*, 23(2):385–401, Februar 2005.
- [131] S. Kozlov, P. van der Stok, and J. Lukkien. Adaptive scheduling of MPEG video frames during real-time wireless video streaming. In *Sixth IEEE International Symposium on a World of Wireless Mobile and Multimedia Networks*, pages 460–462, June 2005.
- [132] Hongli Luo, Mei-Ling Shyu, and Shu-Ching Chen. A multi-buffer scheduling scheme for video streaming. In *2005 IEEE International Conference on Multimedia and Expo*, pages 1218–1221, July 2005.
- [133] Zigmund Orlov and Marc C. Necker. Enhancement of video streaming QoS with active buffer management in wireless environments. In *European Wireless Conference*, 2007.
- [134] Kyu-Hwan Lee, Hyun-Jin Lee, and Jae-Hyun Kim. An active buffer management based on the virtual transmission delay for video streaming service. In *2011 - MILCOM 2011 Military Communications Conference*, pages 816–821, November 2011.
- [135] Amparito A. Morales Figueroa and Lorenzo Favalli. Buffer Management for Scalable Video Streaming. In *Proceedings of the 9th EAI International Conference on Mobile Multimedia Communications*, MobiMedia '16, pages 18–23, ICST, Brussels, Belgium, Belgium, 2016. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering).
- [136] Y. Nagai, T. Okamawari, and T. Fujii. A novel streaming method using QoS control function of LTE to prevent video freezing. In *2016 IEEE Wireless Communications and Networking Conference*, pages 1–6, April 2016.
- [137] Patrick Seeling, Frank Fitzek, and Martin Reisslein. *Video traces for network performance evaluation: a comprehensive overview and guide on video traces and their utilization in networking research*. Springer Science & Business Media, 2007.
- [138] FFmpeg Documentation, <http://www.ffmpeg.org/ffmpeg-doc.html>.

- [139] M. C. Necker and A. Weber. Protocol Interference Between UP-and Downlink Channels in HSDPA. In *2006 IEEE 17th International Symposium on Personal, Indoor and Mobile Radio Communications*, pages 1–5, September 2006.
- [140] M. C. Necker, C. M. Gauger, S. Kiesel, and U. Reiser. IKREmuLib: A Library for Seamless Integration of Simulation and Emulation. In *13th GI/ITG Conference - Measuring, Modelling and Evaluation of Computer and Communication Systems*, pages 1–18, March 2006.
- [141] <https://www.tcpdump.org/>.
- [142] Inwheel Joe. Packet loss and jitter control for real-time MPEG video communications. *Computer Communications*, 19(11):901 – 914, 1996. Recent advances in networking technology.
- [143] EvalVid Framework, <http://www.tkn.tu-berlin.de/research/evalvid/>.
- [144] Video Quality Experts Group (VQEG), <http://www.its.bldrdoc.gov/vqeg/>.
- [145] Xiph.Org Foundation, <http://media.xiph.org/>.
- [146] Video Trace Library of Arizona State University, <http://trace.eas.asu.edu/index.html>.
- [147] Center for Image Processing Research of Rensselaer Polytechnic Institute, <http://www.cipr.rpi.edu/>.
- [148] Institut für Informationsverarbeitung, Leibniz Universität Hannover, <ftp://ftp.tnt.uni-hannover.de/pub/svc/testsequences/>.
- [149] ftp://vqeg.its.bldrdoc.gov/MM/cif/ANSI_SRC_vtc1nw_cif.avi.
- [150] http://media.xiph.org/video/derf/y4m/vtc1nw_422_cif.y4m.
- [151] http://media.xiph.org/video/derf/y4m/intros_422_cif.y4m.
- [152] ftp://vqeg.its.bldrdoc.gov/MM/cif/ANSI_SRC_3inrow_cif.avi.
- [153] <http://ivms.stanford.edu/~esetton/salesman/salesman.cif>.
- [154] ftp://vqeg.its.bldrdoc.gov/MM/cif/ANSI_SRC_washdc_cif.avi.
- [155] http://media.xiph.org/video/derf/y4m/washdc_422_cif.y4m.
- [156] http://media.xiph.org/video/derf/y4m/football_cif.y4m.
- [157] http://media.xiph.org/video/derf/y4m/football_422_cif.y4m.
- [158] DVB SMI 029. Recommended Encoding Practices for Storage Media Interoperability.
- [159] U. Reimers. *DVB - Digitale Fernsehtechnik*. Springer, 2008.
- [160] Axis Communications. H.264 video compression standard, New possibilities within video surveillance. Technical report, 2008.

- [161] M. C. Necker and A. Weber. Impact of Iub flow control on HSDPA system performance. In *2005 IEEE 16th International Symposium on Personal, Indoor and Mobile Radio Communications*, volume 3, pages 1703–1707 Vol. 3, September 2005.
- [162] IKR. IKR Simulation Library, <http://www.ikr.uni-stuttgart.de/IKRSimLib>.
- [163] B. Kulapala P. Seeling, M. Reisslein. Network Performance Evaluation Using Frame Size and Quality Traces of Single-Layer and Two-Layer Video: A Tutorial. *IEEE Communications Surveys & Tutorials*, vol. 6, no. 3, Third Quarter 2004.
- [164] Frank Fitzek, Patrick Seeling, and Martin Reisslein. VideoMeter tool for YUV bitstreams. *Technical Report acticom-02-001 acticorn-mobile networks*, 2002.
- [165] F. Hartanto and H. R. Sirisena. Cumulative inter-ADU jitter concept and its applications. In *Proceedings Tenth International Conference on Computer Communications and Networks (Cat. No.01EX495)*, pages 531–534, October 2001.
- [166] N. Challa and H. Cam. Cost-aware downlink scheduling of shared channels for cellular networks with relays. In *IEEE International Conference on Performance, Computing, and Communications, 2004*, pages 793–798, April 2004.
- [167] John G. Apostolopoulos, Wai-tian Tan, and Susie J. Wee. Video streaming: Concepts, algorithms, and systems. *HP Laboratories, report HPL-2002-260*, 2002.