

Contents

Abbreviations & Mathematical Notation	ix
1 Introduction	1
2 Packet Based Speech and Audio Transmission	5
2.1 System Description	6
2.2 Session Setup and Quality of Service Control	8
2.2.1 Signaling Protocols for Managing a Session of Multimedia Packet Transmission	9
2.2.2 Traffic Management and Quality of Service	9
2.3 Speech and Audio Codecs for Packet Networks	9
2.3.1 Demands and Constraints for Packet Transmission	10
2.3.2 Speech and Audio Coding Standards and their Suitability for Packet Transmission	10
2.3.2.1 Waveform Speech Codecs	11
2.3.2.2 Hybrid Speech Coding Schemes	11
2.3.2.3 Multiple Description Coding	12
2.3.2.4 Embedded (Layered) Coding Schemes	13
2.3.2.5 Coding Schemes for Audio Signals	14
2.3.3 Speech and Audio Codecs Considered in This Work	14
2.4 Packet Structure for IP based Multimedia Transmission	14
2.4.1 Packetization of Media Frames and Assembly of IP Packets	15
2.4.1.1 RTP Payload Formats for Important Speech & Audio Codecs	15
2.4.1.2 RTP Payload Formats for Forward Error Correction	16
2.4.1.3 Calculation of Packet Size and Packet Data Rate	17
2.4.2 Robust Header Compression for Wireless Links	18
2.5 Packet Transmission over Wired and Wireless Networks	19

2.5.1	Local Area Networks (LAN)	20
2.5.2	Broadband Access Networks	20
2.5.3	Wireless and Mobile Access Networks	21
2.5.4	Public Internet	21
2.6	Transmission Impairments in Packet Networks	22
2.6.1	Packet Forming and Transmission Delay	22
2.6.2	Packet Losses	24
2.6.3	Bit Errors on Wireless Transmission Channels	24
2.7	Techniques for Combating Packet Loss	25
2.7.1	Sender-driven Packet Loss Recovery	25
2.7.2	Receiver-based Packet Loss Concealment	27
2.7.3	Sender-assisted Packet Loss Concealment	28
2.7.4	Receiver Buffer for Compensation of Jitter	29
2.7.5	Utilization of Packets with Residual Bit Errors	30
2.7.5.1	Discussion of Concepts	30
2.7.5.2	Unequal Error Detection for Packet-Switched Channels	31
3	Model of the Packet Transmission Channel	33
3.1	Modeling Packet Loss: Gilbert(-Elliott) Models and Alternatives . .	35
3.1.1	Notation for Describing Packet Loss Distributions	36
3.1.2	The Bernoulli Model for Independent Packet Losses	37
3.1.3	Gilbert(-Elliott) Models for Bursty Packet Losses	38
3.1.3.1	Generalized Gilbert Model (Gilbert-Elliott Model) .	39
3.1.3.2	Gilbert Model	39
3.1.3.3	Simplified Gilbert Model	40
3.1.4	Alternative Channel Model: 4-State Markov Model	41
3.2	Extended Gilbert-Elliott Model Considering Various Transmission Time Intervals and Packet Sizes	42
3.2.1	Model Adaptation for Multiples of the Transmission Time Interval	43
3.2.2	Model Adaptation for Arbitrary Packet Sizes	44
3.2.3	Examples for the Channel Model Adaptation	46
3.3	Modeling Varying Transmission Delay (Jitter)	46
3.3.1	The Weibull Distribution for Modeling Jitter	47
3.3.2	Packet Loss Due to Jitter Depending on Receiver Buffer Length	48

3.3.3	Incorporating Jitter Losses into the Gilbert-Elliott Packet Loss Model	48
3.4	Deriving Probabilities of Loss Patterns from Model Parameters . . .	49
3.4.1	Generalized Gilbert-Elliott Model	49
3.4.2	Extended Gilbert-Elliott Model	50
4	Analysis of Forward Error Correction Capabilities on Packet Transmission Channels	53
4.1	Theoretical Determination of Residual Losses for Different FEC Schemes	55
4.1.1	Interleaved Transmission of Media Frames	56
4.1.2	Overview of Considered FEC Schemes	58
4.1.3	Data Rate and Delay Constraints	66
4.1.4	No Forward Error Correction	67
4.1.5	Repetition Code	69
4.1.5.1	Separate Transmission of FEC Frames	69
4.1.5.2	Piggybacked Transmission of FEC Frames	71
4.1.6	Exclusive Disjunction (XOR) Codes	72
4.1.6.1	Separate Transmission of FEC Frames	72
4.1.6.2	Piggybacked Transmission of FEC Frames	74
4.1.7	Block Codes (e.g., Reed-Solomon Codes)	77
4.1.7.1	Separate Transmission of FEC Frames	77
4.1.7.2	Piggybacked Transmission of FEC Frames	80
4.2	Theoretical Determination of Residual Losses after Retransmission .	83
4.3	Forward Error Correction on Channels with Varying Transmission Delay (Jitter)	88
5	System Optimization for Speech and Audio Transmission over Packet Networks	91
5.1	Optimization Problem: Criteria, Parameters, and Constraints	92
5.1.1	Optimization Criterion	92
5.1.2	Variable Parameters in Packet-Based Multimedia Transmission	92
5.1.2.1	Media codec and packetization of media frames . . .	93
5.1.2.2	Forward Error Correction (FEC)	93
5.1.2.3	Receiver Buffer and Frame Loss Concealment	94
5.1.3	Application Demands: Audio Quality and Delay	94
5.1.3.1	Speech Conversation (IP telephony, Voice over IP) .	94
5.1.3.2	Music Streaming	94

5.1.4	Network Constraints: Transmission Delay, Errors, and Capacity	95
5.2	General Questions and Considerations	95
5.2.1	Choice of Frame Length per Packet	96
5.2.2	Forward Error Correction versus Retransmission	99
5.2.3	Choice of Forward Error Correction Scheme	102
5.2.4	Separate vs. Piggybacked Transmission of FEC Data	104
5.2.5	Forward Error Correction to Reduce Jitter Buffer Length	105
5.2.6	Adaptation of System Parameterization for Changing Channel Characteristics	106
5.3	Multicast Music Streaming on Wireless LAN	106
5.4	Voice over IP on Wireless LAN (VoWLAN)	108
5.4.1	Optimal Frame Length with Layer 2 Retransmissions	110
5.4.2	Joint Optimization of Frame Length and Forward Error Correction	112
5.4.3	Optimal Parameterization for Heterogeneous Networks with WLAN Access	118
5.5	Voice over IP on UMTS Packet Channels	124
5.6	Voice over IP on IP Channels with Varying Transmission Delays	127
5.6.1	Optimal Choice of Jitter Buffer Length	127
5.6.2	Joint Optimization of Jitter Buffer and Forward Error Correction	128
5.7	Conclusions	130
6	Packet Loss Concealment	133
6.1	State-of-the-Art in Packet Loss Concealment	134
6.2	CELP Codec Parameters and their Significance for Quality	137
6.2.1	PESQ Measurements of the Separate Concealment of AMR Codec Parameters	139
6.2.2	Conclusions	141
6.3	Voicing Controlled Packet Loss Concealment for CELP Encoded Speech Signals	141
6.3.1	Voicing Classification	142
6.3.2	Parameter Estimation Depending on Voicing Transition	143
6.3.2.1	LSF Interpolation	144
6.3.2.2	Parameter Estimation: Transition <i>voiced-voiced</i>	144
6.3.2.3	Parameter Estimation: Transition <i>voiced-unvoiced</i>	146
6.3.2.4	Parameter Estimation: Transition <i>unvoiced-voiced</i>	146

6.3.2.5	Parameter Estimation: Transition <i>unvoiced-unvoiced</i>	147
6.3.3	Performance Results	148
6.4	Improved Packet Loss Concealment by Transmission of Low Rate Side Information	148
6.4.1	Sender-assisted PLC Approach	149
6.4.2	Side Information and Concealment Methods for CELP	
Codec Parameters		149
6.4.2.1	Spectral Envelope	151
6.4.2.2	Pitch Lag	154
6.4.2.3	Adaptive and Fixed Codebook Gains	157
6.4.2.4	Fixed Codebook (Innovation)	159
6.4.3	Performance Results	159
6.4.4	Approaches for Side Information Transmission	162
6.5	Steganographic Transmission of Side Information for PLC	163
6.5.1	System Concept	163
6.5.2	Data Hiding Scheme for ACELP Codecs	164
6.5.2.1	Impact of Data Hiding on Speech Quality	165
6.5.3	Impact of Bit Errors and Channel Coding	166
6.5.4	Steganographic PLC in a Packet Network with Circuit-Switched GSM Access	166
6.5.4.1	GSM Bit Level Simulations	167
6.5.4.2	Steganographic VoIP simulations	168
6.5.4.3	Simulation Results	169
6.6	Conclusions	170
7	Summary	173
A	IP, UDP, and RTP Protocols	179
A.1	IP - Internet Protocol	179
A.2	UDP - User Datagram Protocol	180
A.3	RTP - Real-Time Transport Protocol	180
A.4	RTCP - RTP Control Protocol	181
B	RTP Payload Format for MP3 Music Signals	183
B.1	MP3 Bit Stream Format	183
B.2	MP3 Frame Packetization	184
C	Overview of Packet Sizes	185

D	Wireless Packet Transmission Standards: UMTS and WLAN	187
D.1	UMTS Packet-Switched (PS) Channels	187
D.1.1	Packet Data Convergence Protocol	187
D.1.2	Radio Link Control Protocol	188
D.1.3	Medium Access Control (MAC) Protocol	188
D.1.4	Physical Layer and Channel Coding	188
D.2	Wireless LAN (IEEE 802.11)	188
D.2.1	Medium Access Control (MAC) Protocol	189
D.2.2	Physical Layer Convergence Protocol (PLCP)	189
D.2.3	Physical Layer	190
E	Deriving Channel Models for UMTS & WLAN	191
E.1	Model Training and Assessment	191
E.1.1	Deriving Model Parameters from Channel Measurements and Simulations	191
E.1.2	Channel Model Adaptation Based on Feedback Reports . . .	192
E.1.3	Testing the Goodness of Fit	192
E.1.3.1	Graphical Evaluation of the Goodness of Fit	193
E.1.3.2	Chi-Square Goodness of Fit Test	193
E.2	Simulation and Modeling of UMTS and WLAN Channels	194
E.2.1	UMTS Channel Model	194
E.2.2	WLAN Channel Model	195
F	Probabilities of Specific Loss Patterns in the Extended Gilbert- Elliott Model	203
F.1	Probability of Compound Loss Patterns	204
F.2	Probability of a Repeated Occurrence of a Specific Pattern	205
G	Concatenation of Channel Models	207
G.1	Concatenation of Bernoulli Models	208
G.2	Concatenation of Simplified Gilbert Model and Bernoulli Model . . .	208
G.3	Concatenation of Gilbert(-Elliott) Model and Bernoulli Model	208
G.4	Concatenation of Simplified Gilbert Models	209
G.5	Concatenation of Gilbert-Elliott Models	209

H	Perceived Quality Assessment and Prediction	213
H.1	Means of Assessing and Predicting the Perceived Quality	214
H.2	Objective Speech Quality Evaluation with PESQ	215
H.3	Prediction of Quality with the ITU-T E-Model	216
H.3.1	Delay impairment factor I_d	216
H.3.2	Effective Equipment Impairment Factor $I_{e,eff}$	217
H.3.3	Categories of Speech Quality and According Rating Factors .	218
H.4	Deriving Equipment Impairment Factors from Instrumental Models .	219
H.5	Deriving Equipment Impairment Factors for the AMR Speech Codec from PESQ Measurements	220
I	Deutschsprachige Kurzfassung	221
	Bibliography	231