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 heterogenous 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


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

6.2.4 Videobewertungsmodell ........................................... 83
6.2.5 Grundszenario ......................................................... 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 Adoptionsverhaltens ....................... 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 Signalrückstandspegelwerte 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
2.8 DVB-IPTV-Protokollstapel nach [6] ................................................................... 18
2.9 Dienstgültigkeitsprüfung 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 zeittlichen Interpixelredundanzreduktion .......................................... 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 Netzgesteuerter Datenratenadaptation 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
<table>
<thead>
<tr>
<th>Sechstelnummer</th>
<th>Abbildung/Zeichnung</th>
<th>Seite</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.4</td>
<td>3inrow(d)</td>
<td>74</td>
</tr>
<tr>
<td>6.5</td>
<td>smity1(m)</td>
<td>74</td>
</tr>
<tr>
<td>6.6</td>
<td>washdc(c)</td>
<td>74</td>
</tr>
<tr>
<td>6.7</td>
<td>football(i)</td>
<td>74</td>
</tr>
<tr>
<td>6.8</td>
<td>Der vorausgewählte Videosequenzraum (zur weiteren Eingrenzung)</td>
<td>77</td>
</tr>
<tr>
<td>6.9</td>
<td>PSNR, 1DK</td>
<td>78</td>
</tr>
<tr>
<td>6.10</td>
<td>PSNR, 1DKf</td>
<td>78</td>
</tr>
<tr>
<td>6.11</td>
<td>PSNR, 2DK</td>
<td>78</td>
</tr>
<tr>
<td>6.12</td>
<td>Datenanteile, 1DK 0B</td>
<td>79</td>
</tr>
<tr>
<td>6.13</td>
<td>Datenanteile, 1DK 2B</td>
<td>79</td>
</tr>
<tr>
<td>6.14</td>
<td>Datenanteile, 1DK 11B</td>
<td>79</td>
</tr>
<tr>
<td>6.15</td>
<td>Datenanteile, 1DKf 0B</td>
<td>79</td>
</tr>
<tr>
<td>6.16</td>
<td>1DKf 2B</td>
<td>79</td>
</tr>
<tr>
<td>6.17</td>
<td>Datenanteile, 1DKf 11B</td>
<td>79</td>
</tr>
<tr>
<td>6.18</td>
<td>Datenanteile, 2DK 0B</td>
<td>79</td>
</tr>
<tr>
<td>6.19</td>
<td>Datenanteile, 2DK 2B</td>
<td>79</td>
</tr>
<tr>
<td>6.20</td>
<td>Datenanteile, 2DK 11B</td>
<td>79</td>
</tr>
<tr>
<td>6.21</td>
<td>Der ausgewählte Videosequenzraum für die Verfahrensuntersuchungen</td>
<td>80</td>
</tr>
<tr>
<td>6.22</td>
<td>Datenraten, Ausschnitt, 1DK</td>
<td>81</td>
</tr>
<tr>
<td>6.23</td>
<td>Datenraten, Ausschnitt, 1DKf</td>
<td>81</td>
</tr>
<tr>
<td>6.24</td>
<td>Datenraten, Ausschnitt, 2DK</td>
<td>81</td>
</tr>
<tr>
<td>6.25</td>
<td>Kanalverlaufshistogramm</td>
<td>83</td>
</tr>
<tr>
<td>6.26</td>
<td>Kanalverlauf</td>
<td>83</td>
</tr>
<tr>
<td>6.27</td>
<td>Protokolloverhead</td>
<td>85</td>
</tr>
<tr>
<td>6.28</td>
<td>Videobildverluste</td>
<td>85</td>
</tr>
<tr>
<td>6.29</td>
<td>Anteil der hohen Qualität</td>
<td>86</td>
</tr>
<tr>
<td>6.30</td>
<td>Verhältnis $D_{ef-eff}$</td>
<td>86</td>
</tr>
<tr>
<td>6.31</td>
<td>Verhältnis $D_{ef-all}$</td>
<td>86</td>
</tr>
<tr>
<td>6.32</td>
<td>Umschaltvorgänge</td>
<td>86</td>
</tr>
<tr>
<td>6.33</td>
<td>Adaptionsvorschläge</td>
<td>86</td>
</tr>
<tr>
<td>6.34</td>
<td>football 1DK 2B, Abspielpuffer</td>
<td>88</td>
</tr>
<tr>
<td>6.35</td>
<td>football 1DK 2B, empfangene Videobilder</td>
<td>88</td>
</tr>
<tr>
<td>6.36</td>
<td>football 1DK 2B, Datenmenge der fehlerfrei empfangenen Videobilder</td>
<td>88</td>
</tr>
<tr>
<td>6.37</td>
<td>football 1DK 2B, PSNR-Verlauf</td>
<td>88</td>
</tr>
<tr>
<td>6.38</td>
<td>intros 1DK 2B, Abspielpuffer</td>
<td>89</td>
</tr>
<tr>
<td>6.39</td>
<td>intros 1DK 2B, empfangene Videobilder</td>
<td>89</td>
</tr>
<tr>
<td>6.40</td>
<td>intros 1DK 2B, Datenmenge der fehlerfrei empfangenen Videobilder</td>
<td>89</td>
</tr>
<tr>
<td>6.41</td>
<td>intros 1DK 2B, PSNR-Verlauf</td>
<td>89</td>
</tr>
<tr>
<td>6.42</td>
<td>vtc1nw 1DK 2B, Abspielpuffer</td>
<td>90</td>
</tr>
<tr>
<td>6.43</td>
<td>vtc1nw 1DK 2B, empfangene Videobilder</td>
<td>90</td>
</tr>
<tr>
<td>6.44</td>
<td>vtc1nw 1DK 2B, Datenmenge der fehlerfrei empfangenen Videobilder</td>
<td>90</td>
</tr>
<tr>
<td>6.45</td>
<td>vtc1nw 1DK 2B, PSNR-Verlauf</td>
<td>90</td>
</tr>
<tr>
<td>6.46</td>
<td>football 1DK 2B, PVU</td>
<td>91</td>
</tr>
<tr>
<td>6.47</td>
<td>intros 1DK 2B, PVU</td>
<td>91</td>
</tr>
<tr>
<td>6.48</td>
<td>vtc1nw 1DK 2B, PVU</td>
<td>91</td>
</tr>
<tr>
<td>6.49</td>
<td>football 1DK 0B, PVU</td>
<td>91</td>
</tr>
<tr>
<td>6.50</td>
<td>intros 1DK 0B, PVU</td>
<td>91</td>
</tr>
</tbody>
</table>
ABBILDUNGSVERZEICHNIS

6.51 **vtc1nw 1DK 0B, PVU** ................................. 91
6.52 Wirkungsweise für **football 1DK 11B Pb** ................................. 99
6.53 Vergleich für verschiedene Abspielpuffergrößen für **intros 1DK 2B Pb** ................................. 99
6.54 **1DK 0B** ......................................................... 101
6.55 **1DK 0B P6s** ......................................................... 101
6.56 **1DK 2B** ......................................................... 101
6.57 **1DK 2B P6s** ......................................................... 101
6.58 **1DK f 2B** ......................................................... 101
6.59 **1DK f 2B P6s** ......................................................... 101
6.60 **2DK 2B** ......................................................... 101
6.61 **2DK 2B P6s** ......................................................... 101
6.62 Verfahrensvergleich für **intros 1DK 2B** ......................................................... 103
6.63 **1DK 0B Imp_{\text{min}} = 4** ......................................................... 105
6.64 **1DK 2B Imp_{\text{min}} = 4** ......................................................... 105
6.65 **1DK 11B Imp_{\text{min}} = 4** ......................................................... 105
6.66 **1DK f 0B Imp_{\text{min}} = 4** ......................................................... 105
6.67 **1DK f 2B Imp_{\text{min}} = 4** ......................................................... 105
6.68 **1DK f 11B Imp_{\text{min}} = 4** ......................................................... 105
6.69 **1DK 2B Imp_{\text{min}} = 0** ......................................................... 107
6.70 **1DK f 2B Imp_{\text{min}} = 0** ......................................................... 107
6.71 **1DK f 11B Imp_{\text{min}} = 4** ......................................................... 107
6.72 Histogramm der Unterbrechungsdauern für **football 1DK 2B BAb** ......................................................... 108

A.1 Datenraten, **1DK** ......................................................... 115
A.2 Datenraten, **1DK f** ......................................................... 115
A.3 Datenraten, **2DK** ......................................................... 115
A.4 Örtliche Auflösung ......................................................... 116
# Tabellenverzeichnis

<table>
<thead>
<tr>
<th>Kapitel</th>
<th>Beschreibung</th>
<th>Seite</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Five-Grade-Scale-Notenaufstellung</td>
<td>42</td>
</tr>
<tr>
<td>3.2</td>
<td>Abbildung von PSNR auf MOS [8], [9]</td>
<td>44</td>
</tr>
<tr>
<td>3.3</td>
<td>Zuordnung der Videosequenzen zu Videokategorien (aus [7])</td>
<td>45</td>
</tr>
<tr>
<td>6.1</td>
<td>Beschreibung des vorausgewählten unkomprimierten Videosequenzsatzes</td>
<td>75</td>
</tr>
<tr>
<td>6.2</td>
<td>Unterschiedliche Videostrukturen</td>
<td>76</td>
</tr>
<tr>
<td>6.3</td>
<td>Statistikwerte der Transportbitraten (kbit/s)</td>
<td>84</td>
</tr>
<tr>
<td>6.4</td>
<td>Statistikwerte (Quantile) der Transportbitraten (kbit/s)</td>
<td>84</td>
</tr>
<tr>
<td>6.5</td>
<td>Anzahl der Ereignisse im Adapptionskontext</td>
<td>93</td>
</tr>
<tr>
<td>6.6</td>
<td>Phasen der hohen Qualität (Statistik der Dauern in Sekunden)</td>
<td>95</td>
</tr>
<tr>
<td>6.7</td>
<td>Adaptionseffizienz</td>
<td>96</td>
</tr>
<tr>
<td>6.8</td>
<td>Videobildwichtigkeiten</td>
<td>100</td>
</tr>
<tr>
<td>A.1</td>
<td>Verwendete Komprimierungsparameter für 1DKf</td>
<td>113</td>
</tr>
<tr>
<td>A.2</td>
<td>Protokolloverhead, 308 kbit/s</td>
<td>114</td>
</tr>
</tbody>
</table>
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)
<table>
<thead>
<tr>
<th>Abkürzung</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>COFDM</td>
<td>Coded Orthogonal Frequency Division Multiplexing</td>
</tr>
<tr>
<td>CQI</td>
<td>Channel Quality Indicator</td>
</tr>
<tr>
<td>DAB</td>
<td>Digital Audio Broadcasting</td>
</tr>
<tr>
<td>DASH</td>
<td>Dynamic Adaptive Streaming over HTTP</td>
</tr>
<tr>
<td>DCH</td>
<td>Dedicated Transport Channel</td>
</tr>
<tr>
<td>DiffServ</td>
<td>Differentiated Services</td>
</tr>
<tr>
<td>DMB</td>
<td>Digital Multimedia Broadcasting</td>
</tr>
<tr>
<td>DMB-T</td>
<td>Digital Multimedia Broadcasting – Terrestrial</td>
</tr>
<tr>
<td>DQPSK</td>
<td>Differential Quadrature Phase-Shift Keying</td>
</tr>
<tr>
<td>DSCP</td>
<td>Differentiated Services Code Point</td>
</tr>
<tr>
<td>DVB</td>
<td>Digital Video Broadcasting</td>
</tr>
<tr>
<td>DVB-H</td>
<td>Digital Video Broadcasting – Handhelds</td>
</tr>
<tr>
<td>DVB-IPDC</td>
<td>DVB IP-Datacast</td>
</tr>
<tr>
<td>DVB-T</td>
<td>Digital Video Broadcasting – Terrestrial</td>
</tr>
<tr>
<td>ECN</td>
<td>Explicit Congestion Notification</td>
</tr>
<tr>
<td>EDGE</td>
<td>Enhanced Data Rates for GSM Evolution</td>
</tr>
<tr>
<td>ETSI</td>
<td>European Telecommunications Standards Institute</td>
</tr>
<tr>
<td>FDMA</td>
<td>Frequency Division Multiple Access</td>
</tr>
<tr>
<td>FEC</td>
<td>Forward Error Correction</td>
</tr>
<tr>
<td>GPRS</td>
<td>General Packet Radio Service</td>
</tr>
<tr>
<td>GSM</td>
<td>Global System for Mobile Communications</td>
</tr>
<tr>
<td>HARQ</td>
<td>Hybrid Automatic Repeat Request</td>
</tr>
<tr>
<td>HbbTV</td>
<td>Hybrid Broadcast Broadband TV</td>
</tr>
<tr>
<td>HDS</td>
<td>HTTP Dynamic Streaming</td>
</tr>
<tr>
<td>HEVC</td>
<td>High Efficiency Video Coding</td>
</tr>
<tr>
<td>HLS</td>
<td>HTTP Live Streaming</td>
</tr>
<tr>
<td>HS-DSCH</td>
<td>High Speed Downlink Shared Channel</td>
</tr>
<tr>
<td>HSDPA</td>
<td>High Speed Downlink Packet Access</td>
</tr>
<tr>
<td>Abkürzung</td>
<td>Deutscher Begriff</td>
</tr>
<tr>
<td>-----------</td>
<td>------------------</td>
</tr>
<tr>
<td>HSPA</td>
<td>High Speed Packet Access</td>
</tr>
<tr>
<td>HTTP</td>
<td>Hypertext Transfer Protocol</td>
</tr>
<tr>
<td>HVS</td>
<td>Human Visual System</td>
</tr>
<tr>
<td>IBN</td>
<td>Intent Based Networking</td>
</tr>
<tr>
<td>IETF</td>
<td>Internet Engineering Task Force</td>
</tr>
<tr>
<td>iMB</td>
<td>Integrated Mobile Broadcast</td>
</tr>
<tr>
<td>IMT-2000</td>
<td>International Mobile Telecommunications at 2000 MHz</td>
</tr>
<tr>
<td>IP</td>
<td>Internet Protocol</td>
</tr>
<tr>
<td>IPDC</td>
<td>IP-Datacast</td>
</tr>
<tr>
<td>IPTV</td>
<td>Internet Protocol Television</td>
</tr>
<tr>
<td>IPv4</td>
<td>Internet Protocol Version 4</td>
</tr>
<tr>
<td>IPv6</td>
<td>Internet Protocol Version 6</td>
</tr>
<tr>
<td>ISDB</td>
<td>Integrated Services Digital Broadcasting</td>
</tr>
<tr>
<td>ISMA</td>
<td>Internet Streaming Media Alliance</td>
</tr>
<tr>
<td>ISO</td>
<td>International Organization for Standardization</td>
</tr>
<tr>
<td>ITU</td>
<td>International Telecommunication Union</td>
</tr>
<tr>
<td>ITU-R</td>
<td>International Telecommunication Union - Radiocommunication Standardization Sector</td>
</tr>
<tr>
<td>ITU-T</td>
<td>International Telecommunication Union - Telecommunication Standardization Sector</td>
</tr>
<tr>
<td>LAN</td>
<td>Local Area Network</td>
</tr>
<tr>
<td>LOS</td>
<td>Line Of Sight</td>
</tr>
<tr>
<td>LTE</td>
<td>Long Term Evolution</td>
</tr>
<tr>
<td>MAC-d</td>
<td>Medium Access Control For Dedicated Transport Channel</td>
</tr>
<tr>
<td>MAC-hs</td>
<td>Medium Access Control For High Speed Downlink Shared Channel</td>
</tr>
<tr>
<td>MANET</td>
<td>Mobile Ad Hoc Network</td>
</tr>
<tr>
<td>MBMS</td>
<td>Multimedia Broadcast Multicast Service</td>
</tr>
<tr>
<td>MDC</td>
<td>Multiple Description Coding</td>
</tr>
<tr>
<td>MECN</td>
<td>Multilevel Explicit Congestion Notification</td>
</tr>
<tr>
<td>Abkürzung</td>
<td>Beschreibung</td>
</tr>
<tr>
<td>-----------</td>
<td>--------------------------------------------------------</td>
</tr>
<tr>
<td>Media-FLO</td>
<td>Media Forward Link Only</td>
</tr>
<tr>
<td>MIMO</td>
<td>Multiple Input Multiple Output</td>
</tr>
<tr>
<td>MP4</td>
<td>MPEG-4 Dateiformat</td>
</tr>
<tr>
<td>MPE</td>
<td>Multiprotocol Encapsulation</td>
</tr>
<tr>
<td>MPE-FEC</td>
<td>Forward Error Correction in Multiprotocol Encapsulation</td>
</tr>
<tr>
<td>MPEG</td>
<td>Moving Picture Expert Group</td>
</tr>
<tr>
<td>MPEG-2</td>
<td>Standard von Moving Picture Expert Group</td>
</tr>
<tr>
<td>MPEG-2-TS</td>
<td>MPEG-2-Transportstrom</td>
</tr>
<tr>
<td>MPEG-4</td>
<td>Standard von Moving Picture Expert Group</td>
</tr>
<tr>
<td>MPEG-DASH</td>
<td>MPEG Dynamic Adaptive Streaming over HTTP</td>
</tr>
<tr>
<td>NATO</td>
<td>North Atlantic Treaty Organization</td>
</tr>
<tr>
<td>NLOS</td>
<td>Non Line Of Sight</td>
</tr>
<tr>
<td>OFDM</td>
<td>Orthogonal Frequency Division Multiplexing</td>
</tr>
<tr>
<td>OFDMA</td>
<td>Orthogonal Frequency Division Multiple Access</td>
</tr>
<tr>
<td>OIPF</td>
<td>Open IPTV Forum</td>
</tr>
<tr>
<td>OMA</td>
<td>Open Mobile Alliance</td>
</tr>
<tr>
<td>OSI</td>
<td>Open Systems Interconnection</td>
</tr>
<tr>
<td>PAL</td>
<td>Phase Alternating Line</td>
</tr>
<tr>
<td>PC</td>
<td>Personal Computer</td>
</tr>
<tr>
<td>PSK</td>
<td>Phase-Shift Keying</td>
</tr>
<tr>
<td>PSNR</td>
<td>Peak Signal-To-Noise Ratio</td>
</tr>
<tr>
<td>PSS</td>
<td>Packet-Switched Streaming Service</td>
</tr>
<tr>
<td>QAM</td>
<td>Quadrature Amplitude Modulation</td>
</tr>
<tr>
<td>QoS</td>
<td>Quality of Service</td>
</tr>
<tr>
<td>QPSK</td>
<td>Quadrature Phase-Shift Keying</td>
</tr>
<tr>
<td>RAN</td>
<td>Radio Access Network</td>
</tr>
<tr>
<td>RLC</td>
<td>Radio Link Control</td>
</tr>
<tr>
<td>RNC</td>
<td>Radio Network Controller</td>
</tr>
<tr>
<td>Abkürzungen</td>
<td>Definition</td>
</tr>
<tr>
<td>-------------</td>
<td>------------</td>
</tr>
<tr>
<td>RNS</td>
<td>Radio Network Subsystem</td>
</tr>
<tr>
<td>RTCP</td>
<td>RTP Control Protocol</td>
</tr>
<tr>
<td>RTMFP</td>
<td>Secure Real-Time Media Flow Protocol</td>
</tr>
<tr>
<td>RTMP</td>
<td>Real-Time Messaging Protocol</td>
</tr>
<tr>
<td>RTMPT</td>
<td>Real-Time Messaging Protocol Tunneled</td>
</tr>
<tr>
<td>RTP</td>
<td>Real-Time Transport Protocol</td>
</tr>
<tr>
<td>RTSP</td>
<td>Real-Time Streaming Protocol</td>
</tr>
<tr>
<td>SCTE</td>
<td>Society of Cable Telecommunications Engineers</td>
</tr>
<tr>
<td>SDN</td>
<td>Software Defined Networking</td>
</tr>
<tr>
<td>SDP</td>
<td>Session Description Protocol</td>
</tr>
<tr>
<td>SDTV</td>
<td>Standard Definition Television</td>
</tr>
<tr>
<td>SMIL</td>
<td>Synchronized Multimedia Integration Language</td>
</tr>
<tr>
<td>SMPTE</td>
<td>Society of Motion Picture and Television Engineers</td>
</tr>
<tr>
<td>SR</td>
<td>Segment Routing</td>
</tr>
<tr>
<td>SRv6</td>
<td>Segment Routing Version 6</td>
</tr>
<tr>
<td>SSIM</td>
<td>Structural Similarity</td>
</tr>
<tr>
<td>STANAG 4609</td>
<td>NATO Standardisation Agreement 4609 on Air Reconnaissance Primary Imagery Data Standard</td>
</tr>
<tr>
<td>SVC</td>
<td>Scalable Video Coding</td>
</tr>
<tr>
<td>T-DMB</td>
<td>Terrestrial Digital Multimedia Broadcasting</td>
</tr>
<tr>
<td>TCP</td>
<td>Transmission Control Protocol</td>
</tr>
<tr>
<td>TDMA</td>
<td>Time Division Multiple Access</td>
</tr>
<tr>
<td>TFRC</td>
<td>TCP-Friendly Rate Control</td>
</tr>
<tr>
<td>TISPAN</td>
<td>Telecommunications and Internet converged Services and Protocols for Advanced Networking</td>
</tr>
<tr>
<td>TISPAN IPTV</td>
<td>TISPAN Standard for IPTV</td>
</tr>
<tr>
<td>TV</td>
<td>Television</td>
</tr>
<tr>
<td>UDP</td>
<td>User Datagram Protocol</td>
</tr>
<tr>
<td>UE</td>
<td>User Equipment</td>
</tr>
<tr>
<td>Abkürzungen</td>
<td>Definition</td>
</tr>
<tr>
<td>---------------------</td>
<td>------------------------------------------------------</td>
</tr>
<tr>
<td>UMTS</td>
<td>Universal Mobile Telecommunications System</td>
</tr>
<tr>
<td>US</td>
<td>United States (of America)</td>
</tr>
<tr>
<td>UTRAN</td>
<td>UMTS Terrestrial Radio Access Network</td>
</tr>
<tr>
<td>VSNC</td>
<td>Videostreaming-Netz-Controller</td>
</tr>
<tr>
<td>WCDMA</td>
<td>Wideband Code Division Multiple Access</td>
</tr>
<tr>
<td>WLAN</td>
<td>Wireless LAN</td>
</tr>
</tbody>
</table>
### Symbole

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A$</td>
<td>Parameter der exponentiellen Glättung</td>
</tr>
<tr>
<td>$BuffOccup$</td>
<td>Pufferfüllstand</td>
</tr>
<tr>
<td>$DR_{ef\rightarrow all}$</td>
<td>Verhältnis der Datenmenge von fehlerfrei empfangenen Videobildern im adaptiven Fall zur Datenmenge aller empfangenen Videobilddaten im adaptionslosen Fall</td>
</tr>
<tr>
<td>$DR_{ef\rightarrow ef}$</td>
<td>Datenmengenverhältnis von fehlerfrei empfangenen Videobildern im adaptiven und im adaptionslosen Fall</td>
</tr>
<tr>
<td>$Imp_i$</td>
<td>Videobildwichtigkeit</td>
</tr>
<tr>
<td>$Imp_{\min}$</td>
<td>Schwellwert für die minderwertigen Wichtigkeitswerte</td>
</tr>
<tr>
<td>$k_{hys}$</td>
<td>Hysterese-Faktor</td>
</tr>
<tr>
<td>$e_i$</td>
<td>Aktueller Schätzwert</td>
</tr>
<tr>
<td>$e_{i-1}$</td>
<td>Vorangegangener Schätzwert</td>
</tr>
<tr>
<td>$m_i$</td>
<td>Aktueller Messwert</td>
</tr>
<tr>
<td>$PSNR$</td>
<td>Peak Signal-To-Noise Ratio</td>
</tr>
<tr>
<td>$r_e$</td>
<td>Prognostizierte Übertragungsdatenrate</td>
</tr>
<tr>
<td>$vr_{\text{avail},i}$</td>
<td>Verfügbare diskrete Videodatenrate</td>
</tr>
<tr>
<td>$vr_{\text{current}}$</td>
<td>Aktuell verwendete Videodatenrate</td>
</tr>
<tr>
<td>$vr_{\text{maxDec}}$</td>
<td>Maximale vom Decoder dekodierbare Videodatenrate</td>
</tr>
<tr>
<td>$\delta$</td>
<td>Schwellwert für den Pufferfüllstand</td>
</tr>
</tbody>
</table>
Symbole
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 Datenkompresseionsverfahren und die Entwicklung neuer Kommunikationstechnologien - insbesondere die Erfolgsgeschichte der paketbasierten Kommunikation auf der Basis des Internet-Protokolls (IP) - ermöglichten neuartige Kommunikationssysteme, 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.


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.
Kapitel 1. Einführung und Überblick

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


Während die in einem Rechnerverbund wie dem Internet eingesetzten Datentransportprotokolle für die Rechner-Rechner-Kommunikation paketbasiert (UDP\(^1\)) bzw. bytestrombasiert (TCP\(^2\)) 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**

- **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**

Betrachtet man dedizierte digitale Audio- und Videoverteilkommunikationsnetze\(^3\), 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*

---

\(^1\)UDP: User Datagram Protocol
\(^2\)TCP: Transmission Control Protocol
\(^3\)Broadcastnetze
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 Videoströmung steht für die Übertragung und das gleichzeitige Abspielen eines Streamingvideos im Empfänger.


Während die Videoinhalte in ihren Datenvolumen dank der fortschrittlichen Komprimierungsverfahren reduziert werden, steigen die Datenraten in den Internetzugangsnetzen. 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-


\[ \frac{352 \times 288 \times 12 \text{bit}}{s} \times \frac{30 \text{Videobilder}}{s} = 36,49536 \frac{\text{Mbit}}{s} \]

⁶Ein Videokomprimierungsstandard legt die Encoder-Implementierung nicht fest.
Videostreaming

Abbildung 1.1: Die drei treibenden Kräfte für die Videostreaminganwendungen

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

1.2 Überblick über diesen Beitrag


Das Kapitel 5 beschreibt die eigens für diese Arbeit entwickelte experimentelle Untersuchungs umgebung. 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 Videostreamingadaption 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 über einen terrestrischen Kanal (s. Abschnitt 2.2) betrachtet.


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: DVB\(^2\), ATSC\(^3\), ISDB\(^4\). Eine kompakte Parameterübersicht der digitalen Fernsehstandards findet sich in [17]. Für den mobilen Fernsehempfang (auf einem Smartphone)

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


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


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.


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 Verteilkommunikationsnetzten 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.

2.2 Terrestrischer Funkkanal

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

2.2 Terrestrischer Funkkanal

Der Begriff terrestrischer Funkkanal beschreibt einen erdbundenen 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/Geomorphologie (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 – 40 dB 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 Teilnehmerge- schwindigkeit abhängt und bis zu mehreren 100 Hz betragen kann [19], was zu Pegelleinbrüchen 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 langsamer verlaufen. Deshalb nennt man dieses Phänomen *langsamer Schwund*. Neben der Freiraumdämpfung\(^{14}\) spielt hier die Funkfelddämpfung durch den materiellen Einfluss\(^{15}\) eine große Rolle. Die folgenden Umstände tragen zum langsamen Schwund bei:

---

\(^{13}\)Da er durch die Überlagerung mehrerer Signalpfade zustande kommt, wird er auch als *Interferenzschwund* bezeichnet.

\(^{14}\)Die Empfangsleistung nimmt im materiellen Raum quadratisch mit dem Abstand zum Sender ab.

\(^{15}\)Die 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 steeperen 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.


- **Vegetationsdämpfung**

- **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**

- **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**
- **Impulsförmige Störungen**

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


So wird der langsame Schwund aufgrund der Empfängerbewegung in GSM \(^{16}\) 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 OFDM \(^{17}\), MIMO \(^{18}\) und den höheren Modulationsarten \(^{19}\) 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 Schwundes bei der Netzdimensionierung berücksichtigt [15].

---

\(^{16}\) GSM: Global System for Mobile Communications ist ein Mobilfunksystem der zweiten Generation.

\(^{17}\) OFDM: Orthogonal Frequency Division Multiplexing ist ein Mehrträger-Modulationsverfahren.

\(^{18}\) MIMO: Multiple Input Multiple Output bedeutet eine gleichzeitige Übertragung mehrerer unabhängiger Datenströme über denselben Kanal [27].

\(^{19}\) Z. B. 256-QAM

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 Funkkanalausprägungen\(^{20}\) unterscheiden:

- **Rayleigh-Kanal**


- **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.


- **Gauß-Kanal**

\(^{20}\)Es 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.
2.2 Terrestrischer Funkkanal

Diese Kanalmodelle beschreiben den schnellen Schwund. Ihre Gültigkeit ist auf kleine Gebiete begrenzt\(^{21}\). 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**


In modernen Übertragungssystemen (basierend auf WCDMA\(^{22}\), OFDM bzw. OFDMA\(^{23}\) und COFDM\(^{24}\)) wurde die mittlere Übertragungsraten gegenüber den älteren FDMA\(^{25}/TDMA\(^{26}\)-Systemen (GSM, GPRS\(^{27}\) mit EDGE\(^{28}\)) erheblich erhöht. Der Einfluss des schnellen Schwundes auf die Kanalqualität wurde insbesondere bei OFDM-Techniken (LTE) durch das Verwenden mehrerer zueinander orthogonalen Unterträger stark reduziert. Die Zeitvarianz des langsamten 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.

---

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

\(^{22}\) WCDMA: Wideband Code Division Multiple Access ist ein codebasiertes Vielfachzugriffsverfahren in den Mobilfunknetzen der dritten Generation.

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

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

\(^{25}\) FDMA: Frequency Division Multiple Access ist ein frequenzbasiertes Vielfachzugriffsverfahren.

\(^{26}\) TDMA: Time Division Multiple Access ist ein zeitbasiertes Vielfachzugriffsverfahren.

\(^{27}\) GPRS: General Packet Radio Service ist eine GSM-Erweiterung.

\(^{28}\) EDGE: Enhanced Data Rates for GSM Evolution ist eine GSM-Erweiterung.
### 2.3 Datenstreaming


![Datenstreaming Diagram](image)

**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) zureordnen.

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 Datenstreaming können im weiten Sinne die HTTP\textsuperscript{31}/TCP-basierten Methoden wie MPEG-DASH\textsuperscript{32}, Adobe RTMPT\textsuperscript{33}, Adobe HDS\textsuperscript{34}, Apple HLS\textsuperscript{35}, Microsoft Smooth Streaming\textsuperscript{36}, das TCP-basierte Adobe RTMP\textsuperscript{37}, das UDP-basierte RTP\textsuperscript{38}, Verfahren (mit oder ohne RTSP\textsuperscript{39}) sowie das UDP-basierte Adobe RTMP\textsuperscript{40} gezählt werden.

Ein klarer Fall des synchronisierten Datenstreaming sind die IP-basierten Fernsehproduktionsverfahren, wie sie in der SMPTE\textsuperscript{41}-ST-2110-Standardfamilie\textsuperscript{42} 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 ISO\textsuperscript{43}-Referenzmodell als Schichtenarchitektur wurde von der ISO eingeführt mit dem Ziel, Kommunikationsvorgänge zwischen offenen Systemen zu standardisieren\textsuperscript{44}\textsuperscript{45}. Das Videostreaming lässt sich darauf abbilden.


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-


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 Videostreams. 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 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-TS\(^{46}\)).

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.

\(^{46}\)MPEG-2-TS: MPEG-2 Transport Stream
2.5 Videostreaming im Schichtenmodell

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

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

Abbildung 2.5: ISMA-Protokollstapel nach [2]


---

\textsuperscript{47}IETF: Internet Engineering Task Force.
\textsuperscript{48}SDP: Session Description Protocol.
werden (siehe den Protokollsstapel in der Abbildung 2.5). Der ISMA-Protokollsstapel 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 (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 Protokollsstapel ist in der Abbildung 2.7 zu sehen. Er gilt sowohl bei MBMS (UMTS/HSPA) als auch bei eMBMS (LTE). Das Videostreaming erfolgt

\(^{49}\)3GPP: 3rd Generation Partnership Project ist ein Zusammenschluss mehrerer Standardisierungsorganisationen im Mobilfunkbereich.

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]


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.
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 Diensterbringung 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 Adaptionsmöglichkeit 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 Vollständigkeitshalber werden sie an dieser Stelle behandelt. Das Ziel der Vielfachzugriffsverfahren ist die bestmögliche Aufteilung der gemeinsamen Funkressourcen unter den vorhandenen Kanälen.
  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 Schwundes erreicht.

- **Modulation**
  Es gibt zwei adaptive Modulationsarten:
  - **Adaptive Modulation**
adaptive Modulation darin, dass die momentan verfügbare Übertragungsdatenrate je nach Kanalsituation schwankt.

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

- **Fehlerkorrektur**
  
  - **Adaptive Kanalkodierung**
  
  - **Hybrides ARQ**
    ARQ ist ein Mechanismus der erneuten Übertragungen 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 Videostreaminganwendung wird aus der kommunikationstechnischen Sicht im Wesentlichen von den folgenden Faktoren beeinflusst:

---

54 ARQ: Automatic Repeat Request

55 Sendewiederholungen
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 Kompressionsrate\(^{56}\) 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.

\(^{56}\)Es 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].
2.5 Videostreaming im Schichtenmodell

Abbildung 2.9: Dienstgüteabschätzung einer Videostreaminganwendung


2.5.4 Zusammenfassende Diskussion


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

Die TCP-basierten Varianten sind aufgrund der Protokollnatür 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
Über IP

Über MPEG–2–TS

Über TCP

Über UDP

Adaptiv über HTTP

Ggf. adaptiv über RTP/RTCP (ggf. mit RTSP)

Progressiver Videodownload

Andere Verfahren wie z. B. RTMP

Andere Verfahren wie z. B. RTMFP


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-

Abbildung 2.10: Videostreamingarten


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 Adaptionseinscheidungen 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.

Das Mobilfunkzugangsnetz 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 Basistation 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 entsprechen-
der Wert bei UMTS/HSPA lag bei 42 ms. Hat man nur die Downlink-Richtung vermisser, 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 Datenratenunter-
schiede zwischen UMTS/HSPA und LTE irrelevant. Es werden hier Videos mit den Videodaten-
raten 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 unterschieden, 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 Übertra-
gung zu trennen. Im Einzelnen gibt es:
  - **Logische Kanäle**
    Ein logischer Kanal beschreibt einen Übertragungskanal für die Nutz- oder Steuer-
daten innerhalb der Sicherungsschicht des ISO-OSI-Schichtenmodells und dient der

- **Transportkanäle**

  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.

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 Videostreaminganwendungen in den Mobilfunknetzen identifizieren:

- **Grundsätzliche Zeitvarianz**
  - *Zeitvariable Wellenausbreitungssituation durch die Teilnehmermobilität*
    Mit der adaptiven Modulation und der adaptiven Kanalcodierung reagiert ein Mobilfunkzugangsnetz auf die zeitvariblen 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 Störrauschen aufgrund der asynchronen Übertragung der teilnehmerspezifischen orthogonalen Codes merklich, was die Dekodierbarkeit beeinflusst.

- **Zugangsnetzheterogenität**
  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.
Kapitel 2. Transport von Streamingvideo über terrestrische Funknetze
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 Abrufvideodienstes\textsuperscript{1} 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 Videodatenspeicherformat 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

\textsuperscript{1}Engl: 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 Übertragungressourcen über die Veränderung des zu übertragenden Datenvolumens anwendungsgerecht angepasst\(^2\).

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 Wahrnehmungs-system 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.

---

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

\(^3\)RGB: Rot, Grün, Blau
3.2 Videoerzeugung

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.


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

- **Redundanzreduktion**
  

- **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


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
Kapitel 3. Grundlagen des adaptiven Videostreaming und der Auswahl von Videosequenzen

Abbildung 3.1: Klassifizierung der Redundanzarten

Kodierungsredundanz  Interpixelredundanz  Spektrale Redundanz

(Statistische) Redundanz

Räumliche Redundanz  Zeitliche Redundanz

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.


Abbildung 3.2: Redundanzreduktion

Die Grenze der Kompressionsrate der Redundanzreduktionsverfahren wird durch die Entropie 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 den Einsatz der zeitlichen Interpixelredundanzreduktion werden erhebliche Steigerungen der Kompressionsrate [72] erreicht. Diese Verfahren führen aber zur Entstehung von

---

4Dort hat das Signal eine redundantärere 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 erkauft:

- **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**

![Abbildung 3.3: Einfluss der zeitlichen Interpixelredundanzreduktion](image)

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

### 3.2.2.2 Irrelevanzreduktion

Die Methode der Irrelevanzreduktion ist eine wirksame Methode zur Reduzierung des Videodatenumfanges 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

---

7 fixed interval synchronization
- nicht wahrnehmen kann oder
- sich dafür nicht interessiert.

Die Irrelevanzreduktion ist per Definition eine verlustbehafte 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 größeren bzw. ungleichen Quantisierungstufen im Frequenzbereich.


### 3.2.2.3 Komprimierungsstandards

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

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


Die bekanntesten modernen Komprimierungsstandards für den Endkonsum sind MPEG-4, MPEG-4 AVC\(^8\) (H.264) und MPEG-H Part 2 (HEVC\(^9\) bzw. H.265). MPEG-4 AVC soll im

---

\(^8\)MPEG-4 AVC: MPEG-4 Advanced Video Coding  
\(^9\)HEVC: 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 Videoqualität\(^{10}\) 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**

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

![Abbildung 3.4: MPEG-4-Videostruktur](image)


\(^{10}\)Für 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.


### 3.3 Videotransport


#### 3.3.1 Arten der Videodatenverluste


---

11 Unterschiedlich 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**

- **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)**

- **Fehlerkorrigierende Maßnahmen**

- **Disjunkte Wegeführung**
  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.


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 einewendungsgerechte Auslieferung von Videodaten für die beste erreichbare Videoqualität zu gewährleisten.

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].


### 3.4 Bewertung der Videoqualität

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


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 Übertragungsstrecke 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-
Kapitel 3. Grundlagen des adaptiven Videostreaming und der Auswahl von Videosequenzen

**Tabelle 3.1: Five-Grade-Scale-Notenauflistung**

<table>
<thead>
<tr>
<th>Quality</th>
<th>Impairment</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>Excellent</td>
</tr>
<tr>
<td>4</td>
<td>Good</td>
</tr>
<tr>
<td>3</td>
<td>Fair</td>
</tr>
<tr>
<td>2</td>
<td>Poor</td>
</tr>
<tr>
<td>1</td>
<td>Bad</td>
</tr>
</tbody>
</table>

| 5     | Imperceptible                     |
| 4     | Perceptible, but not annoying     |
| 3     | Slightly annoying                 |
| 2     | Annoying                          |
| 1     | Very annoying                     |


### 3.4.1 Subjektive Messverfahren


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 pixelbasiert Vergleich des zu beurteilenden Videos durchgeführt wird. Dafür werden Pixelmerkmale wie Helligkeits- und Chrominanzwerte verwendet.
3.4 Bewertung der Videoqualität

- **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**

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


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 quadratischer Fehler (MSE\(^{12}\)) 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].

\(^{12}\text{MSE: Mean Square Error}\)
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].

<table>
<thead>
<tr>
<th>PSNR, dB</th>
<th>MOS</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;37</td>
<td>5 (Excellent)</td>
</tr>
<tr>
<td>31 - 37</td>
<td>4 (Good)</td>
</tr>
<tr>
<td>25 - 31</td>
<td>3 (Fair)</td>
</tr>
<tr>
<td>20 - 25</td>
<td>2 (Poor)</td>
</tr>
<tr>
<td>&lt;20</td>
<td>1 (Bad)</td>
</tr>
</tbody>
</table>

#### 3.4.3 Auswahl von Videosequenzen


- **Spatial Perceptual Information (SI)**

- **Temporal Perceptual Information (TI)**


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.

---

13 Innerhalb eines Videobildes
14 Änderungen vom Videobild zum Videobild
15 In der Helligkeitsebene
3.4 Bewertung der Videoqualität

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

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

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


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).
Kapitel 3. Grundlagen des adaptiven Videostreaming und der Auswahl von Videosequenzen
4 Eingebettete netzgesteuerte Videostreamingadaption


4.1 Konzept


Der VSNC befindet sich in einem Mobilfunkzugangsnetz (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 Videostreamingservers und idealerweise über die Verarbeitungsfähigkeiten des Videoclients. Aufgrund ihrer genauen Kenntnisse über die momentane Ressourcensituation und ggf. unter Einhaltung von netzspezifischen Optimierungskriterien und anderweitigen...
Randbedingungen kann sie Trends präzise und schnell erkennen und eine Videoadaptionsentscheidung in Bezug auf die Anpassung der Videodatenrate treffen. Dem Videostreamingserver wird dies in Form eines Videoadaptionsvorschlags mitgeteilt.

- **Active-Queue-Management (AQM)**

**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 Videodatenratenadaptation* (Im Bild *netzgesteuerte Adaption*) steuert die Anpassung der
4.2 Netzgesteuerte Datenratenadaption

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.


\(^{1}\)LAN: Local Area Network
Kapitel 4. Eingebettete netzgesteuerte Videostreamingadaption


4.2 Netzgesteuerte Datenratenadaption


In [105] und in [106] wird vorgeschlagen, in einem Mobilfunkzugangssnetz (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 momentane Pufferfüllstände). Es ist keine videostreamingspezifische Lösung.


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.


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

\(^2\)Gemeint ist ein RAN von IMT-2000.

\(^3\)MECN: Multilevel Explicit Congestion Notification


4.2.2 Diskussion und Einordnung

In der Literaturübersicht wurde deutlich, dass alle vorgeschlagenen Lösungansätze 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.


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-

\[\text{SSIM: Structural Similarity} \]
\[\text{WLAN: Wireless LAN} \]
\[\text{SDN: Software Defined Networking} \]
4.2 Netzgesteuerte Datenratenadaption

Es sei 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. \textit{RLC}\textsuperscript{7} und \textit{MAC-hs}\textsuperscript{8} sind Komponenten der Sicherungsschicht.


Die Datenratenadaptionsschicht weist vier Funktionsblöcke auf:

- \textit{Beobachtung}
  

\footnotesize\textsuperscript{7}RLC: Radio Link Control
\footnotesize\textsuperscript{8}MAC-hs: Medium Access Control for High Speed Downlink Shared Channel
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**
### 4.2 Netzgesteuerte Datenratenadaption

- **Kommunikation**


  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 Sicherungsschicht\(^9\) ab, und entnimmt diesem die augenblickliche Übertragungsdatenrate\(^10\).

  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 Adaptionsempfindlichkeit 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 Videoadoptionsvorschlags für den Videostreamingserver. Die Entscheidung über einen Videoadoptionsvorschlag hat die folgenden Randbedingungen:

\[^9\text{Genau gesagt in der MAC-hs-Schicht.}\]

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

\[ r_u = k_{\text{pow}} \cdot r_s \]

Alledings 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 Videostreamingservers ($v_{avail_i}$)

- Die aktuell verwendete Videodatenrate $v_{current}$

- Die maximale vom Videoclient dekodierbare Videodatenrate ($v_{maxDec}$)

- 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 Videoadaption entscheidung wird wie folgt getroffen:

- **Heraufschalten**
  Es wird auf die nächsthöhere verfügbare Videodatenrate $v_{avail_i}$ heraufgeschaltet, wenn die folgende Bedingung wahr ist:

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

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

- **Herunter schalten**
  Es wird auf die nächstkleinere Videodatenrate $v_{avail_i}$ heruntergeschaltet, wenn die folgende Bedingung wahr ist:

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

Die Fairness der Verkehrsströme untereinander muss nicht betrachtet werden, da das Mobilfunkzugangsnetz 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 Videostreamingserver über den ermittelten Videoadaptionsvorschlag.

### 4.3 AQM

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


\(^{11}\)CBR: Constant Bit Rate
Kapitel 4. Eingebettete netzgesteuerte Videostreamingadaption

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


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-

\textsuperscript{12}DiffServ: Differentiated Services

\textsuperscript{13}QoS: Quality of Service


14TCP-Friendly Rate Control


$^{15}$FER: Frame Error Rate


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 entsprechenden Abschnitten ausführlich beschrieben.


### 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


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.


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 $Buf_{Occu}$ überwacht. Beim Überschreiten eines Schwellwertes $\delta$ 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 videobildtypbasierten 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 \((Imp_{\text{min}})\) 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:

\[
(Buf\text{Occu}p > 5) \land (Imp_i \leq Imp_{\text{min}})
\]

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

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 eingeführt. 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-Mobilfunkzugangsnetz (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 Maßnahmearten 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].


Im Folgenden wird die Untersuchungsumgebung näher vorgestellt.

5.2 Systemübersicht

Die Untersuchungsumgebung wurde mit C++ entwickelt und läuft unter Linux\(^1\). 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 Bildeauflosung dem Komprimierungsverfahren unterzogen. Beim Komprimierungsverfahren 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].

\(^1\) Ein Unix-basiertes Betriebssystem
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 Video- bild, 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.


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 16 byte) des IPv4-Protokollheaders versehen. Das RTP wird nicht verwendet. Dies stellt für die durchgeführten Untersuchungen keine Einschränkung dar, da die Funktionalität des Protokolls nicht verwendet wird. Der GesamtprotokollobjektiveHead beträgt 44 byte4. Zum Vergleich

Abbildung 5.1: Experimentelle Untersuchungsumgebung
beträgt der minimale Gesamtprotokollobjekt des realen RTP/UDP-basierten Videostrams 40\,byte\textsuperscript{5}.

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.


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.


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

\textsuperscript{5}20(\textit{IP – Paketheader}) + 8(\textit{UDP – Datagrammheader}) + 12(\textit{RTP – Sequenzheader}) = 40
5.2.2 PVU-Diagramm

Bei der Verwendung von PSNR als Videoqualitätsmetrik für die Bewertung einer Video-
streamingsitzung 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 demselben 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 Video-streamingsitzungen unter dem
Aspekt der Unterbrechungsdauern durchgeführt. Es nahmen 15 Personen an der Bewertung
teil. Zusammenfassend lässt sich feststellen, dass Unterbrechungen, die eine Sekunde über-
schreiten, als störend empfunden werden.

In dieser Arbeit wird vorgeschlagen, die PSNR-Verteilungen und Unterbrechungsdauer-
Verteilungen in Form von abgewandelten Box-Plot-Darstellungen in ein und demselben Dia-
gramm darzustellen. Ein Box-Plot-Diagramm zeigt das 1.Quartil (0,25-Quantil, im Folgen-
den \( Q_1 \) 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 einheit-
llich. Es wird hier vorgeschlagen, für die Ausreißer den Minimum- und den Maximumwert zu
verwenden.

Die Abbildung 5.4 zeigt das neuartige \textit{PVU-Diagramm} (\textit{PSNR-VideoUnterbrechungs-
Diagramm}) am Beispiel einer bewerteten Video-streamingsitzung. Die Unterbrechungsdauer-

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure52.png}
\caption{Messaufbau}
\end{figure}

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure53.png}
\caption{Konnektivität im Messaufbau}
\end{figure}
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.
Kapitel 5. Experimentelle Untersuchungsumgebung
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.


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

\footnote{Im Wesentlichen 4:2:2 und 4:2:0}
Abbildung 6.1: Vorauswahl der Videosequenzen (vgl. Tabelle 3.3)


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.

<table>
<thead>
<tr>
<th>Eigenschaften</th>
<th>vtctnw(f)</th>
<th>introso(o)</th>
<th>3inrow(d)</th>
<th>smity1(m)</th>
<th>washdc(c)</th>
<th>football(i)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inhalt</td>
<td>Teilnehmerin einer Videokonferenz</td>
<td>Mehrere Personen stellen sich vor, Kamerabewegungen</td>
<td>Drei Personen am Tisch, Kamerabewegungen</td>
<td>Ein Händler im TV-Shop</td>
<td>Stadtplan, Hand- und Stiftbewegungen</td>
<td>American-Football-Spiel</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Originale Videotestsequenzen in den Datenbanken</th>
</tr>
</thead>
<tbody>
<tr>
<td>Länge, Videobilder</td>
</tr>
<tr>
<td>Zeitliche Auflösung, Videobilder/s</td>
</tr>
<tr>
<td>Räumliche Auflösung, Pixel</td>
</tr>
<tr>
<td>Farbunterabtastungs- und Speicherformat</td>
</tr>
<tr>
<td>Bezugsquelle</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Für die Untersuchungen vorbereitete Videosequenzen (Grundsequenzen)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Länge, Videobilder</td>
</tr>
<tr>
<td>Räumliche Auflösung, Pixel</td>
</tr>
<tr>
<td>Farbunterabtastungs- und Speicherformat</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Zu verwendende Videosequenzen (Aneinanderhängen der Grundsequenzen)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Länge, Videobilder</td>
</tr>
<tr>
<td>Länge, s</td>
</tr>
</tbody>
</table>

a Alternative Quelle  
b Alternative Quelle  
c Alternative Quelle mit einem anderen, aber mit [156] überlappenden Ausschnitt.  
d Bei der Konvertierung zur Bildrate von 25 Videobildern/s ergab sich eine Länge von 217 Videobildern. In dieser Ausprägung wies das Video einen enormen T1-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].
6.1.2 Komprimierung


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.
6.1 Videosequenzen

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 darstellen. 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


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: PNSR, 2DK
6.1.5 Datenraten


In den beiden 1DK-Fällen sind die Datenratenunterschiede zwischen den unabhängigen und abhängigen Videobildern weniger stark als bei 2DK. Der 1DK-Fall weist breitere Verteilungen die mittleren 50% der Daten als der 1DKf-Fall auf. Das bedeutet, dass bei 1DKf 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 0B (u. a. Überwachungsnetzwerkkameras), 2B (Broadcastsbereich), 1IB (Reduzierung der Videobildabhängigkeiten und somit der Fehlerfortpflanzung) wurden als die wichtigen zu untersuchenden Strukturen ausgewählt.

Die 1DK-, 1DKf- und 2DK-Vorgehensweisen sollen als weitere Dimension untersucht werden. 1DK und 2DK können bei Live-Streaming bzw. Abrufvideodiensten vorliegen. 1DKf reduziert die Variabilität der Videodatenraten innerhalb einer Videobildgruppe.

<table>
<thead>
<tr>
<th>Videoinhalt</th>
<th>Vorgehensweise</th>
</tr>
</thead>
<tbody>
<tr>
<td>vtcn1nw(f)</td>
<td>2DK</td>
</tr>
<tr>
<td>intros(o)</td>
<td>1DKf</td>
</tr>
<tr>
<td>football(i)</td>
<td>1DK</td>
</tr>
</tbody>
</table>

Abbildung 6.21: Der ausgewählte Videosequenzraum für die Verfahrensuntersuchungen
6.1 Videosequenzen

Abbildung 6.22: Datenraten, Ausschnitt, \( IDK \)

Abbildung 6.23: Datenraten, Ausschnitt, \( IDK_f \)

Abbildung 6.24: Datenraten, Ausschnitt, \( 2DK \)
6.2 Systemmodell

6.2.1 Systembeschreibung


6.2.2 Emulationsmodell


6.2.3 Videostreamingmodell


---

\(^{3}\)HS-DSCH: High Speed Downlink Shared Channel  
\(^{4}\)DCH: Dedicated Transport Channel  
\(^{5}\)MAC-d: Medium Access Control For Dedicated Transport Channel  
\(^{6}\)BLER: Block Error Rate  
\(^{7}\)CQI: Channel Quality Indicator
6.2 Systemmodell

6.2.4 Videobewertungsmodell


6.2.5 Grundscenario

Der mobile Teilnehmer, für den der emulierte Videostrom bestimmt is, bewegt sich mit 30 km/h. Je nach Untersuchung werden weitere CBR\(^8\)-erzeugte Videoströme 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.

\(^8\)CBR: Constant Bit Rate
Der Mobilfunkanal-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 *vtc1nw* jeweils 110 Mal übertragen.


### Tabelle 6.3: Statistikkennze der Transportbitraten (kbit/s)

<table>
<thead>
<tr>
<th>Minimum</th>
<th>Maximum</th>
<th>Mittelwert</th>
<th>Standardabweichung</th>
<th>Varianz</th>
<th>Variationskoeffizient</th>
<th>quadrierter Variationskoeffizient</th>
</tr>
</thead>
<tbody>
<tr>
<td>343</td>
<td>2996</td>
<td>867,6</td>
<td>601,5</td>
<td>3,618 · 10^5</td>
<td>0,6933</td>
<td>0,4806</td>
</tr>
</tbody>
</table>

### Tabelle 6.4: Statistikkennze (Quantile) der Transportbitraten (kbit/s)

<table>
<thead>
<tr>
<th>(Q_1) (Median)</th>
<th>(Q_2)</th>
<th>(Q_3)</th>
<th>(Q_{0.05})</th>
<th>(Q_{0.95})</th>
<th>(Q_{0.99})</th>
<th>(Q_{0.999})</th>
</tr>
</thead>
<tbody>
<tr>
<td>343</td>
<td>1005</td>
<td>1005</td>
<td>343</td>
<td>1985</td>
<td>2643</td>
<td>2996</td>
</tr>
</tbody>
</table>

### 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.

---

\(^{9}\) TTI: Transmission Time Interval
6.3 Verfahrensergebnisse

6.3.1 Netzgesteuerte Datenratenadaptation

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

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.


\( ^{10} \) Ein dynamisches Zuschalten von konkurrierenden mobilen Teilnehmern wird vom Emulator nicht unterstützt.
Abbildung 6.29: Anteil der hohen Qualität

Abbildung 6.30: Verhältnis $D_{ref} - ef$

Abbildung 6.31: Verhältnis $D_{ref} - all$

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 Adapptionsfall zur Datenmenge der fehlerfrei empfangenen Videobilder im adaptionslosen Fall $DR_{ef-f}$. Im adaptiven Fall treten im Allgemeinen bis zu $k_{hys} = 30$ keine Videobildverluste 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 Adapptionsfall zur Datenmenge aller im adaptionslosen Fall empfangenen Videobilddaten $DR_{ef-all}$. Im adaptionslosen Fall gehen ca. 25 % der Videobildbytes im Netz verloren. Im Adapptionsfall 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 Adapptionsvorschläge 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 Adapptionsvorschläge 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 Adapptionsvorschläge 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 1DK-2B-Ausprägung beurteilt. Die anderen Ausprägungen weisen das gleiche Verhalten auf.

Bei 900 kbit/s ohne Berücksichtigung des Protokolloverheads ergibt sich eine Videobildanzahl = ((212000 byte·8 bit) : 900000 byte/s) · 25 Videobilder/s = 47,11 Videobilder
Qualität zu. Im adaptiven Fall treten keine Pufferleerläufe auf. Der adaptive Fall überdeckt in den Diagrammen den adaptionslosen 308 kbit/s-Fall (die blaue Kurve wird überdeckt).

Abbildung 6.42: \textit{vtc1nw 1DK 2B}, Abspiel-puffer

Abbildung 6.43: \textit{vtc1nw 1DK 2B}, empfangene Videobilder

Abbildung 6.44: \textit{vtc1nw 1DK 2B}, Datenmenge der fehlerfrei empfangenen Videobilder

Abbildung 6.45: \textit{vtc1nw 1DK 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 308kbit/s-Fall werden alle Videobilder fehlerfrei empfangen. Im adaptionslosen 900kbit/s-Fall treten Videobildverluste auf. Im Adaptionsfall verläuft die Übertragung fehlerfrei. Das Verhältnis $D_{ref-ef}$ zwischen der 900kbit/s-Empfangskurve und der Empfangskurve des adaptiven Falls liegt über 90%. Im Abschnitt 6.3.1.4 wird das $D_{ref-ef}$ gesondert betrachtet.

Abbildung 6.46: football 1DK 2B, PVU
Abbildung 6.47: intros 1DK 2B, PVU
Abbildung 6.48: vtc1nw 1DK 2B, PVU
Abbildung 6.49: football 1DK 0B, PVU
Abbildung 6.50: intros 1DK 0B, PVU
Abbildung 6.51: vtc1nw 1DK 0B, PVU

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 Adaptionsverhaltens

Der Videocontainerserver 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 Vertrauensintervallen. Für die anderen Spalten sind aus den Übersichtsgründen keine Vertrauensintervalle angegeben.


---

1295%-Vertrauensintervall
Tabelle 6.5: Anzahl der Ereignisse im Adoptionskontext

<table>
<thead>
<tr>
<th>Video</th>
<th>IP-Pakete</th>
<th>TTIs</th>
<th>Datenratenänderungen</th>
<th>Vorschläge o. Glättung</th>
<th>Vorschläge m. Glättung</th>
<th>Vorschläge m. Glättung m. Hysterese</th>
<th>Umschaltvorgänge</th>
</tr>
</thead>
<tbody>
<tr>
<td>1DK</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>football 0B</td>
<td>74630</td>
<td>532123</td>
<td>252171</td>
<td>119086</td>
<td>4128</td>
<td>889 ±27</td>
<td>114 ±4</td>
</tr>
<tr>
<td>football 2B</td>
<td>76006</td>
<td>534403</td>
<td>253992</td>
<td>119269</td>
<td>4348</td>
<td>950 ±48</td>
<td>116 ±8</td>
</tr>
<tr>
<td>football 11B</td>
<td>76379</td>
<td>535212</td>
<td>254226</td>
<td>120301</td>
<td>4342</td>
<td>904 ±44</td>
<td>122 ±7</td>
</tr>
<tr>
<td>intros 0B</td>
<td>78301</td>
<td>545337</td>
<td>259296</td>
<td>118011</td>
<td>4887</td>
<td>1145 ±26</td>
<td>180 ±5</td>
</tr>
<tr>
<td>intros 2B</td>
<td>81676</td>
<td>547942</td>
<td>261764</td>
<td>119451</td>
<td>4983</td>
<td>1204 ±29</td>
<td>199 ±6</td>
</tr>
<tr>
<td>intros 11B</td>
<td>76739</td>
<td>531760</td>
<td>252200</td>
<td>117698</td>
<td>4386</td>
<td>997 ±27</td>
<td>126 ±8</td>
</tr>
<tr>
<td>vtc1nw 0B</td>
<td>79036</td>
<td>532615</td>
<td>252803</td>
<td>117709</td>
<td>4449</td>
<td>955 ±39</td>
<td>115 ±8</td>
</tr>
<tr>
<td>vtc1nw 2B</td>
<td>82931</td>
<td>534404</td>
<td>254083</td>
<td>116890</td>
<td>4727</td>
<td>1096 ±34</td>
<td>125 ±8</td>
</tr>
<tr>
<td>vtc1nw 11B</td>
<td>76246</td>
<td>535323</td>
<td>253935</td>
<td>118165</td>
<td>4341</td>
<td>993 ±35</td>
<td>125 ±11</td>
</tr>
<tr>
<td>1DKf</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>football 0B</td>
<td>75775</td>
<td>533004</td>
<td>252311</td>
<td>119708</td>
<td>4048</td>
<td>811 ±50</td>
<td>108 ±12</td>
</tr>
<tr>
<td>football 2B</td>
<td>77575</td>
<td>538377</td>
<td>255139</td>
<td>120636</td>
<td>4214</td>
<td>884 ±29</td>
<td>115 ±6</td>
</tr>
<tr>
<td>football 11B</td>
<td>80837</td>
<td>557258</td>
<td>263074</td>
<td>125561</td>
<td>4493</td>
<td>914 ±36</td>
<td>123 ±9</td>
</tr>
<tr>
<td>intros 0B</td>
<td>76485</td>
<td>534681</td>
<td>253145</td>
<td>116276</td>
<td>4469</td>
<td>999 ±27</td>
<td>118 ±8</td>
</tr>
<tr>
<td>intros 2B</td>
<td>82524</td>
<td>545688</td>
<td>259412</td>
<td>119852</td>
<td>4591</td>
<td>1031 ±42</td>
<td>169 ±12</td>
</tr>
<tr>
<td>intros 11B</td>
<td>79236</td>
<td>543243</td>
<td>252282</td>
<td>119594</td>
<td>3986</td>
<td>770 ±47</td>
<td>109 ±6</td>
</tr>
<tr>
<td>vtc1nw 0B</td>
<td>79048</td>
<td>532438</td>
<td>252710</td>
<td>117809</td>
<td>4417</td>
<td>954 ±51</td>
<td>117 ±9</td>
</tr>
<tr>
<td>vtc1nw 2B</td>
<td>79337</td>
<td>531702</td>
<td>251189</td>
<td>117684</td>
<td>4031</td>
<td>814 ±41</td>
<td>102 ±9</td>
</tr>
<tr>
<td>vtc1nw 11B</td>
<td>76023</td>
<td>534564</td>
<td>252195</td>
<td>118894</td>
<td>3996</td>
<td>784 ±48</td>
<td>114 ±8</td>
</tr>
<tr>
<td>2DK</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>football 0B</td>
<td>75286</td>
<td>528487</td>
<td>249120</td>
<td>115835</td>
<td>4337</td>
<td>910 ±31</td>
<td>114 ±10</td>
</tr>
<tr>
<td>football 2B</td>
<td>75807</td>
<td>526958</td>
<td>247656</td>
<td>110437</td>
<td>4364</td>
<td>984 ±25</td>
<td>123 ±8</td>
</tr>
<tr>
<td>football 11B</td>
<td>73552</td>
<td>520318</td>
<td>242901</td>
<td>109780</td>
<td>4455</td>
<td>872 ±42</td>
<td>136 ±9</td>
</tr>
<tr>
<td>intros 0B</td>
<td>85880</td>
<td>555212</td>
<td>264044</td>
<td>113882</td>
<td>6194</td>
<td>1583 ±19</td>
<td>276 ±9</td>
</tr>
<tr>
<td>intros 2B</td>
<td>86921</td>
<td>547655</td>
<td>260317</td>
<td>112297</td>
<td>5862</td>
<td>1500 ±29</td>
<td>246 ±5</td>
</tr>
<tr>
<td>intros 11B</td>
<td>82013</td>
<td>509589</td>
<td>237227</td>
<td>103478</td>
<td>5166</td>
<td>1379 ±32</td>
<td>163 ±5</td>
</tr>
<tr>
<td>HQ nicht konform</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>vtc1nw 0B</td>
<td>71887</td>
<td>459811</td>
<td>206453</td>
<td>101232</td>
<td>4824</td>
<td>1247 ±15</td>
<td>155 ±7</td>
</tr>
<tr>
<td>vtc1nw 2B</td>
<td>67448</td>
<td>419418</td>
<td>177041</td>
<td>91698</td>
<td>4772</td>
<td>1312 ±17</td>
<td>165 ±6</td>
</tr>
<tr>
<td>vtc1nw 11B</td>
<td>66040</td>
<td>403860</td>
<td>164744</td>
<td>84114</td>
<td>4689</td>
<td>1506 ±19</td>
<td>145 ±12</td>
</tr>
<tr>
<td>vtc1nw*0B</td>
<td>82460</td>
<td>550762</td>
<td>255939</td>
<td>125638</td>
<td>7201</td>
<td>1778 ±15</td>
<td>225 ±6</td>
</tr>
<tr>
<td>vtc1nw*2B</td>
<td>76475</td>
<td>518750</td>
<td>226383</td>
<td>123312</td>
<td>7226</td>
<td>1729 ±18</td>
<td>221 ±6</td>
</tr>
<tr>
<td>vtc1nw*11B</td>
<td>67718</td>
<td>492786</td>
<td>210664</td>
<td>117714</td>
<td>6591</td>
<td>1554 ±14</td>
<td>243 ±6</td>
</tr>
</tbody>
</table>

Die Tabelle ist in drei Bereiche 1DK, 1DKf und 2DK aufgeteilt. Der 2DK-Bereich besteht aus zwei Teilen. Der untere Teil des 2DK-Bereiches fasst Videos (vtc1nw) zusammen, deren Videodatenrate der höheren Qualitätsstufe aufgrund der Implementierung des verwendeten Videoco-
dec unter der einheitlichen Videodatenrate von 900 kbit/s liegt. D. h., die Ausprägungen der hohen Qualität sind videodatenratenbezogen nicht konform zu den anderen Videosequenzen.

Es werden zuerst die Bereiche 1DK, 1DKf und der 900 kbit/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 Adoptionsvorschlägen führen. Denn die zur Schätzung der Übertragungsdatenrate verwendeten Werte der augenblicklichen Übertragungsdatenrate werden pro TTI bestimmt.

Im nicht 900 kbit/s-konformen 2DK-Bereich sind auffallend niedrige TTI-Zahlen gelb hinterlegt. Das ist wie folgt begründet. Die vtc1nw-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 900 kbit/s verwendet (bei vtc1nw 0B 692 kbit/s, bei vtc1nw 2B 511 kbit/s und bei vtc1nw 11B 459 kbit/s). In diesen Fällen hat der mit 30 km/h fahrende Teilnehmer andere Mobilfunkanalbedingungen erfahren als die Videostreamingsituationen aus dem oberen Teil der Tabelle. Die mit einem Stern (*) gekennzeichneten vtc1nw-Zeilen zeigen die Ergebnisse für die Adaption an die tatsächliche Videodatenrate der Stufe der hohen Qualität.


13 Die 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)

<table>
<thead>
<tr>
<th>Video</th>
<th>Dauer insgesamt</th>
<th>Minimum</th>
<th>Maximum</th>
<th>(p(D &lt; 0.48s))</th>
<th>Q1</th>
<th>Q2</th>
<th>Q3</th>
<th>(Q_{0.95})</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1DK</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>football 0B</td>
<td>490.70 ±6.31</td>
<td>0.48</td>
<td>38.4</td>
<td>0</td>
<td>0.48</td>
<td>1.92</td>
<td>20.16</td>
<td>31.68</td>
</tr>
<tr>
<td>football 2B</td>
<td>502.37 ±7.67</td>
<td>0.48</td>
<td>67.68</td>
<td>0</td>
<td>0.96</td>
<td>1.92</td>
<td>7.68</td>
<td>31.2</td>
</tr>
<tr>
<td>football 11B</td>
<td>495.55 ±8.51</td>
<td>0.48</td>
<td>36.96</td>
<td>0</td>
<td>0.96</td>
<td>1.92</td>
<td>11.04</td>
<td>31.68</td>
</tr>
<tr>
<td>intros 0B</td>
<td>545.5 ±5.94</td>
<td>0.04</td>
<td>43.48</td>
<td>0.31</td>
<td>0.4</td>
<td>1.12</td>
<td>2.84</td>
<td>32.04</td>
</tr>
<tr>
<td>intros 2B</td>
<td>554.17 ±5.11</td>
<td>0.12</td>
<td>43.32</td>
<td>0.29</td>
<td>0.36</td>
<td>0.96</td>
<td>3.72</td>
<td>29.64</td>
</tr>
<tr>
<td>intros 11B</td>
<td>497.86 ±5.00</td>
<td>0.48</td>
<td>36.48</td>
<td>0</td>
<td>0.96</td>
<td>1.92</td>
<td>8.64</td>
<td>31.68</td>
</tr>
<tr>
<td>vtc1nw 0B</td>
<td>501.17 ±7.51</td>
<td>0.48</td>
<td>40.32</td>
<td>0</td>
<td>0.48</td>
<td>2.4</td>
<td>20.16</td>
<td>32.16</td>
</tr>
<tr>
<td>vtc1nw 2B</td>
<td>512.78 ±4.62</td>
<td>0.48</td>
<td>66.24</td>
<td>0</td>
<td>0.96</td>
<td>1.92</td>
<td>7.20</td>
<td>31.20</td>
</tr>
<tr>
<td>vtc1nw 11B</td>
<td>508.51 ±4.07</td>
<td>0.48</td>
<td>37.92</td>
<td>0</td>
<td>0.96</td>
<td>1.92</td>
<td>10.56</td>
<td>31.68</td>
</tr>
<tr>
<td><strong>1DKf</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>football 0B</td>
<td>487.92 ±9.53</td>
<td>0.48</td>
<td>38.4</td>
<td>0</td>
<td>0.96</td>
<td>1.92</td>
<td>20.64</td>
<td>33.60</td>
</tr>
<tr>
<td>football 2B</td>
<td>499.87 ±5.91</td>
<td>0.48</td>
<td>67.68</td>
<td>0</td>
<td>0.96</td>
<td>1.92</td>
<td>11.04</td>
<td>31.2</td>
</tr>
<tr>
<td>football 11B</td>
<td>502.17 ±5.46</td>
<td>0.48</td>
<td>36.36</td>
<td>0</td>
<td>0.96</td>
<td>2.4</td>
<td>11.04</td>
<td>31.20</td>
</tr>
<tr>
<td>intros 0B</td>
<td>512.84 ±6.09</td>
<td>0.04</td>
<td>73.20</td>
<td>0.003</td>
<td>0.48</td>
<td>1.92</td>
<td>17.36</td>
<td>31.32</td>
</tr>
<tr>
<td>intros 2B</td>
<td>535.48 ±6.04</td>
<td>0.12</td>
<td>35.64</td>
<td>0.25</td>
<td>0.36</td>
<td>1.32</td>
<td>4.92</td>
<td>29.64</td>
</tr>
<tr>
<td>intros 11B</td>
<td>487.25 ±8.31</td>
<td>0.48</td>
<td>58.08</td>
<td>0</td>
<td>0.96</td>
<td>2.4</td>
<td>18.24</td>
<td>31.68</td>
</tr>
<tr>
<td>vtc1nw 0B</td>
<td>499.49 ±9.32</td>
<td>0.48</td>
<td>37.92</td>
<td>0</td>
<td>0.96</td>
<td>1.92</td>
<td>17.76</td>
<td>31.68</td>
</tr>
<tr>
<td>vtc1nw 2B</td>
<td>490.94 ±7.87</td>
<td>0.48</td>
<td>65.76</td>
<td>0</td>
<td>0.96</td>
<td>1.92</td>
<td>21.6</td>
<td>31.20</td>
</tr>
<tr>
<td>vtc1nw 11B</td>
<td>490.75 ±7.56</td>
<td>0.48</td>
<td>37.92</td>
<td>0</td>
<td>0.96</td>
<td>2.4</td>
<td>16.8</td>
<td>31.68</td>
</tr>
<tr>
<td><strong>2DK</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>football 0B</td>
<td>489.07 ±5.01</td>
<td>0.48</td>
<td>37.92</td>
<td>0</td>
<td>0.96</td>
<td>1.92</td>
<td>20.16</td>
<td>31.68</td>
</tr>
<tr>
<td>football 2B</td>
<td>507.79 ±3.98</td>
<td>0.48</td>
<td>71.04</td>
<td>0</td>
<td>0.96</td>
<td>1.44</td>
<td>6.24</td>
<td>31.20</td>
</tr>
<tr>
<td>football 11B</td>
<td>482.77 ±6.46</td>
<td>0.48</td>
<td>58.56</td>
<td>0</td>
<td>0.96</td>
<td>1.92</td>
<td>7.68</td>
<td>29.28</td>
</tr>
<tr>
<td>intros 0B</td>
<td>579.86 ±6.18</td>
<td>0.04</td>
<td>68.60</td>
<td>0.38</td>
<td>0.32</td>
<td>0.76</td>
<td>2.44</td>
<td>29.08</td>
</tr>
<tr>
<td>intros 2B</td>
<td>575.54 ±4.39</td>
<td>0.12</td>
<td>36.96</td>
<td>0.31</td>
<td>0.36</td>
<td>0.84</td>
<td>2.88</td>
<td>27.96</td>
</tr>
<tr>
<td>intros 11B</td>
<td>481.20 ±7.67</td>
<td>0.48</td>
<td>36.48</td>
<td>0</td>
<td>0.48</td>
<td>1.44</td>
<td>5.76</td>
<td>26.88</td>
</tr>
</tbody>
</table>

HQ nicht konform

<table>
<thead>
<tr>
<th>video</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>vtc1nw 0B</td>
<td>424.80 ±4.41</td>
<td>0.48</td>
<td>30.24</td>
<td>0</td>
<td>0.96</td>
<td>1.92</td>
<td>6.24</td>
<td>24.48</td>
</tr>
<tr>
<td>vtc1nw 2B</td>
<td>390.24 ±2.58</td>
<td>0.48</td>
<td>31.68</td>
<td>0</td>
<td>0.48</td>
<td>1.44</td>
<td>5.28</td>
<td>23.04</td>
</tr>
<tr>
<td>vtc1nw 11B</td>
<td>344.06 ±2.23</td>
<td>0.48</td>
<td>27.84</td>
<td>0</td>
<td>0.96</td>
<td>1.92</td>
<td>5.28</td>
<td>21.60</td>
</tr>
<tr>
<td>vtc1nw*0B</td>
<td>790.99 ±4.58</td>
<td>0.48</td>
<td>84.96</td>
<td>0</td>
<td>0.48</td>
<td>0.96</td>
<td>2.40</td>
<td>45.12</td>
</tr>
<tr>
<td>vtc1nw*2B</td>
<td>1064.22 ±1.23</td>
<td>0.48</td>
<td>255.36</td>
<td>0</td>
<td>0.48</td>
<td>0.96</td>
<td>2.4</td>
<td>58.56</td>
</tr>
<tr>
<td>vtc1nw*11B</td>
<td>1143.64 ±1.30</td>
<td>0.48</td>
<td>261.6</td>
<td>0</td>
<td>0.48</td>
<td>0.96</td>
<td>1.92</td>
<td>56.64</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Video</th>
<th>Zeitlicher Anteil der Phasen der hohen Qualität</th>
<th>DR_{ef-all}</th>
<th>DR_{ef-ef}</th>
</tr>
</thead>
<tbody>
<tr>
<td>1DK</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>football 0B</td>
<td>37,01 % ±0,48</td>
<td>78,35 % ±0,43</td>
<td>91,37 % ±0,80</td>
</tr>
<tr>
<td>football 2B</td>
<td>37,88 % ±0,58</td>
<td>79,11 % ±0,51</td>
<td>93,96 % ±0,93</td>
</tr>
<tr>
<td>football 11B</td>
<td>37,37 % ±0,64</td>
<td>78,91 % ±0,56</td>
<td>91,96 % ±0,97</td>
</tr>
<tr>
<td>intros 0B</td>
<td>41,19 % ±0,44</td>
<td>82,00 % ±0,41</td>
<td>96,37 % ±0,80</td>
</tr>
<tr>
<td>intros 2B</td>
<td>41,84 % ±0,39</td>
<td>82,78 % ±0,35</td>
<td>97,28 % ±0,74</td>
</tr>
<tr>
<td>intros 11B</td>
<td>37,59 % ±0,37</td>
<td>78,82 % ±0,34</td>
<td>92,32 % ±0,71</td>
</tr>
<tr>
<td>vtc1nw 0B</td>
<td>37,84 % ±0,57</td>
<td>79,13 % ±0,84</td>
<td>92,13 % ±0,90</td>
</tr>
<tr>
<td>vtc1nw 2B</td>
<td>38,72 % ±0,35</td>
<td>79,96 % ±0,32</td>
<td>94,82 % ±0,69</td>
</tr>
<tr>
<td>vtc1nw 11B</td>
<td>38,39 % ±0,31</td>
<td>79,45 % ±0,28</td>
<td>93,48 % ±0,64</td>
</tr>
<tr>
<td>1DKf</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>football 0B</td>
<td>36,80 % ±0,71</td>
<td>78,19 % ±0,64</td>
<td>90,96 % ±1,05</td>
</tr>
<tr>
<td>football 2B</td>
<td>37,70 % ±0,44</td>
<td>79,16 % ±0,40</td>
<td>92,57 % ±0,78</td>
</tr>
<tr>
<td>football 11B</td>
<td>37,87 % ±0,41</td>
<td>81,70 % ±0,35</td>
<td>95,15 % ±0,73</td>
</tr>
<tr>
<td>intros 0B</td>
<td>38,72 % ±0,45</td>
<td>79,78 % ±0,41</td>
<td>93,86 % ±0,80</td>
</tr>
<tr>
<td>intros 2B</td>
<td>40,43 % ±0,45</td>
<td>81,30 % ±0,41</td>
<td>95,93 % ±0,81</td>
</tr>
<tr>
<td>intros 11B</td>
<td>36,79 % ±0,61</td>
<td>78,09 % ±0,55</td>
<td>91,63 % ±0,95</td>
</tr>
<tr>
<td>vtc1nw 0B</td>
<td>37,71 % ±0,70</td>
<td>79,03 % ±0,63</td>
<td>92,40 % ±1,05</td>
</tr>
<tr>
<td>vtc1nw 2B</td>
<td>37,07 % ±0,59</td>
<td>78,32 % ±0,53</td>
<td>91,82 % ±0,93</td>
</tr>
<tr>
<td>vtc1nw 11B</td>
<td>37,05 % ±0,57</td>
<td>78,31 % ±0,51</td>
<td>91,89 % ±0,91</td>
</tr>
<tr>
<td>2DK</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>football 0B</td>
<td>36,88 % ±0,37</td>
<td>78,22 % ±0,35</td>
<td>91,35 % ±0,71</td>
</tr>
<tr>
<td>football 2B</td>
<td>38,29 % ±0,30</td>
<td>79,53 % ±0,28</td>
<td>102,67 % ±0,70</td>
</tr>
<tr>
<td>football 11B</td>
<td>36,41 % ±0,49</td>
<td>78,32 % ±0,49</td>
<td>93,94 % ±0,90</td>
</tr>
<tr>
<td>intros 0B</td>
<td>43,78 % ±0,47</td>
<td>86,03 % ±0,49</td>
<td>102,09 % ±0,93</td>
</tr>
<tr>
<td>intros 2B</td>
<td>43,46 % ±0,33</td>
<td>84,78 % ±0,27</td>
<td>107,16 % ±0,70</td>
</tr>
<tr>
<td>intros 11B</td>
<td>36,33 % ±0,58</td>
<td>77,66 % ±0,54</td>
<td>90,77 % ±0,94</td>
</tr>
</tbody>
</table>

HQ nicht konform

| vtc1nw 0B  | 32,07 % ±0,33                               | 72,90 % ±0,26 | 79,73 % ±0,35 |
| vtc1nw 2B  | 29,46 % ±0,19                               | 76,10 % ±0,08 | 80,53 % ±0,36 |
| vtc1nw 11B | 25,98 % ±0,17                               | 78,27 % ±0,05 | 81,28 % ±0,33 |
| vtc1nw*0B  | 59,72 % ±0,34                               | 90,85 % ±0,27 | 99,36 % ±0,62 |
| vtc1nw*2B  | 80,35 % ±0,09                               | 97,02 % ±0,06 | 102,66 % ±0,41 |
| vtc1nw*11B | 86,35 % ±0,10                               | 97,75 % ±0,08 | 101,52 % ±0,42 |

Te beinhaltet das bereits eingeführte Verhältnis der Datenmengen der *fehlerfrei* empfangenen Videobilder im Adoptions- und adaptionlosen 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
6.3 Verfahrensergebnisse

Folge des bereits besprochenen agilen Adoptionsverhaltens 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 900 kbit/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 Adaptionssfall 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 900 kbit/s-konformen Tabellenbereich liegen um die 80%. Die besonders adaptionsagile *intros-OB*-Videosequenz weist einen Wert von 86% auf.

Die nicht 900 kbit/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.


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 200 000 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


Um die Grundknappheit der Ressourcen zu erzeugen, werden bei jeder Untersuchung drei weitere CBR-erzeugte Videostreams 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 212 000 Byte groß. Die in den Puffer passende Datennenge entspricht somit einem Abspielintervall zwischen 5,25 s und 5,3 s (je nach sequenzspezifischem Overhead) für eine 308 kbit/s-Videosequenz.

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

6.3 Verfahrensergebnisse

Abbildung 6.52: Wirkungsweise für football 1DK 11B Pb

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

Schwellwert des RNC-Pufferfüllstandes \( \delta \), bei dessen Überschreitung das proaktive Verwerfen von Videobildaten durchgeführt wird.

Die blaue Kurve zeigt die im Netz verworfenen Videobilder. Beim Schwellwert \( \delta = 1 \) greift das proaktive Verwerfen nicht. Die Gesamtverluste 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 \( \delta = 0,97 \) komplett\(^\text{14} \). Durch die Veränderung der Videobildverluststruktur sinkt somit die Gesamtverluste.

Ein Videostreamingclient verwendet einen Abspelpuffer, 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 \( \delta \) kann abhängig von der Größe des Abspelpuffers 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 \( \delta \) für unterschiedlich große Abspelpuffer im Videoclient. Die jeweiligen Minimumwerte der Videobildverluste entsprechen der Grenze für den Schwellwert, unter der keine Verluste aufgrund einer zu späten Ankunft auftreten. Liegt z. B. das Minimum der Videobildverluste bei \( \delta = 0,6 \), so gibt es für \( \delta < 0,6 \) keine Verluste aufgrund einer zu späten Ankunft.

\(^{14}\)Der im Puffer frei gewordene Raum entspricht \( 0,03 \cdot 212000 = 6360 \) Byte. Die mittlere \( I \)-Videobildgröße liegt für diese Videosequenz bei 2880 Byte\(\pm\)32. Da finden im Mittel zwei \( I \)-Videobilder Platz.
Tabelle 6.8: Videobildwichtigkeiten

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
</tr>
</thead>
<tbody>
<tr>
<td>0B</td>
<td>I/11</td>
<td>P/10</td>
<td>P/9</td>
<td>P/8</td>
<td>P/7</td>
<td>P/6</td>
<td>P/5</td>
<td>P/4</td>
<td>P/3</td>
<td>P/2</td>
<td>P/1</td>
<td>P/0</td>
<td>I/11</td>
</tr>
<tr>
<td>2B</td>
<td>I/13</td>
<td>B/0</td>
<td>B/0</td>
<td>B/0</td>
<td>B/0</td>
<td>B/0</td>
<td>B/0</td>
<td>B/0</td>
<td>B/0</td>
<td>B/0</td>
<td>B/0</td>
<td>B/0</td>
<td>I/13</td>
</tr>
<tr>
<td>11B</td>
<td>I/22</td>
<td>B/0</td>
<td>B/0</td>
<td>B/0</td>
<td>B/0</td>
<td>B/0</td>
<td>B/0</td>
<td>B/0</td>
<td>B/0</td>
<td>B/0</td>
<td>B/0</td>
<td>B/0</td>
<td>I/22</td>
</tr>
</tbody>
</table>

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 besitzt.15

Die hier vorgeschlagenen Verfahren führen das proaktive Verwerfen anhand der Videobildwichtigkeit $I_{mp}$ als Hauptkriterium (vgl. Abschnitt 4.3.3) durch. Ein Video-IP-Paket wird verworfen, wenn $I_{mp}$ der darin enthaltenen Videobilddaten dem Schwellwert $I_{mp,min}$ gleich ist oder darunter liegt. Ist der Schwellwert beispielsweise $I_{mp,min} = 4$, so können bei 0B ggf. alle P-Videobilder mit $I_{mp} = 4$ und kleiner verworfen werden. Bei 2B können ggf. alle B-Videobilder und alle P-Videobilder mit $I_{mp} = 4$ 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 verwertenden 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 $I_{mp,min}$ 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 $\delta$ von 0,35 bis 1,0 mit einer in den relevanten Bereichen geringen Schrittweite ermittelt wurde.


Über alle Videosequenzen betrachtet kann festgestellt werden, dass der optimale Schwellwert (für den die Videobildverluste minimal wird) je nach Sequenz $I_{mp,min} = 4$ bzw. $I_{mp,min} = 0$ ist. Der Fall $I_{mp,min} = 4$ ist optimal im Allgemeinen für die gängigsten Ausprägungen 1DK 0B, 1DK 2B und ggf. 2DK 2B. Der Fall $I_{mp,min} = 0$ ist optimal für die Ausprägungen, die besonders viele Daten in den B-Videobildern aufweisen, wie 1DK 2B.

Die aus Platzgründen nicht dargestellte Ausprägung 1DKf 0B zeigt ein ähnliches Verhalten wie 1DK 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 $I_{mp,min} = 0$ 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.
6.3 Verfahrensergebnisse

Abbildung 6.54: 1DK 0B

Abbildung 6.55: 1DK 2B

Abbildung 6.56: 1DKf 2B

Abbildung 6.57: 1DK 2B P

Abbildung 6.58: 1DKf 2B P

Abbildung 6.59: 2DK 2B

Abbildung 6.60: 2DK 2B P

Abbildung 6.61: 2DKf 2B

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 1DK 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 *vtc1nw* bessere Ergebnisse als *intros* erzielt, da die erste Sequenz etwas mehr Daten in den P-Videobildern aufweist. Betrachtet man aber die zu verwertenden P-Videobilder für Imp<sub>min</sub> = 4, so stellt man fest, dass die mittlere Videobildgröße und das 1. und 2. Quartil der Videobildgrößenverteilung bei *vtc1nw* am kleinsten unter den drei Sequenzen sind. Insofern für Imp<sub>min</sub> = 4 hat das Verfahren im Mittel bei *vtc1nw* weniger Daten zum proaktiven Verwerfen.

Bei 1DK 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 *vtc1nw* 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 *vtc1nw*. Die Sequenz *intros* liegt dazwischen. Das Verhalten von *football* bei Imp<sub>min</sub> = 0 wird in Abschnitt 6.3.2.4 behandelt.

Bei 1DKf 2B fallen die Kurven für *intros* und *vtc1nw* 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.

6.3 Verfahrensergebnisse

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 \( Im_{\text{min}} = 4 \) bzw. \( Im_{\text{min}} = 0 \) entsprechend verwendet. Die 2DK-Ausprägungen werden nicht weiter untersucht. Die optimalen Einstellungen für den Schwellwert \( \delta \) liegen\(^{16}\) für die \( 0B \)- 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 \( \delta \) 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 \( Im_{\text{min}} = 4 \) im \( 0B \)-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 \( \delta \) würde dies in einer andauernden Überlastsituation zu größeren Werten des RNC-Pufferfüllstandes bei \( 0B \)- 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 6 s aufweist.

Die Abbildung 6.62 zeigt die Videobildverluste in Abhängigkeit vom Schwellwert \( \delta \) für drei Verfahren \( Pb \), \( Bb \) und \( BAb \) bei \( Im_{\text{min}} = 4 \) im Falle einer gängigen Ausprägung \( \text{intros 1DK 2B} \). 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

---

\(^{16}\)Für die zu verwendenden Werte von \( Im_{\text{min}} = 4 \) bzw. \( Im_{\text{min}} = 0 \)
Video-IP-Paket\textsuperscript{17} 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 \(vtc1nw-2DK-11B\)-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 \(1DK\ 0B\) (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 schmalsten\textsuperscript{18}. Bei \(vtc1nw\) ist sie am breitesten\textsuperscript{19}. 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 \(vtc1nw\) 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 \(1DK\ 0B\) (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 \(vtc1nw\) verlagert, was die Absenkung der \(intros\)-Kurve erklärt. Die Minderung der Videobildverluste durch den Einsatz von \(Pb\), \(Bb\) bzw. \(BAb\) gegenüber \(OA\) beträgt bei \(vtc1nw\) ca. 40\%, 41\% bzw. 48\%, bei \(intros\) ca. 43\%, 45\% bzw. 53\% und bei \(football\) ca. 49\%, 50\% bzw. 59\%.

Bei \(1DK\ 2B\) (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 \(vtc1nw\) ca. 26\%, 31\% bzw. 41\%, bei \(intros\) ca. 34\%, 37\% bzw. 47\% und bei \(football\) ca. 42\%, 45\% bzw. 55\%.

\textsuperscript{17}Die verwendete maximale IP-Paketgröße beträgt hier 1500 Byte.

\textsuperscript{18}Der Variationskoeffizient beträgt 0.21

\textsuperscript{19}Der Variationskoeffizient liegt bei 0.62
Abbildung 6.63: $\text{IDK} \ 1B \ \text{Im}_{\min} = 4$

Abbildung 6.64: $\text{IDK} \ 2B \ \text{Im}_{\min} = 4$

Abbildung 6.65: $\text{IDK} \ 0B \ \text{Im}_{\min} = 4$

Abbildung 6.66: $\text{IDK} \ 0B \ \text{Im}_{\min} = 4$

Abbildung 6.67: $\text{IDK} \ 1B \ \text{Im}_{\min} = 4$

Abbildung 6.68: $\text{IDK} \ 1B \ \text{Im}_{\min} = 4$
Bei 1DKf2B hat die Verlagerung der Daten in die B- und P-Videobilder insbesondere bei vtc1nw 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 vtc1nw ca. 48%, 63% bzw. 66%, bei intros ca. 46%, 64% bzw. 66% und bei football ca. 43%, 52% bzw. 60%.


Die Abbildung 6.69 zeigt football 1DKf2B für Imp\textsubscript{min} = 0. Pb erzielt hier bessere Ergebnisse als bei Imp\textsubscript{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\textsubscript{min} = 4 die bessere Wahl (vgl. Abbildung 6.64).

Die Abbildung 6.70 zeigt 1DKf2B für Imp\textsubscript{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\textsubscript{min} = 4. Bb erzielt sogar dieselben Ergebnisse wie BAb bei Imp\textsubscript{min} = 4. Die Minderung der Videobildverluste durch den Einsatz von Pb, Bb bzw. BAb gegenüber OA beträgt bei vtc1nw ca. 56%, 67% bzw. 67%, bei intros ca. 54%, 67% bzw. 67% und bei football ca. 50%, 59% bzw. 59%. Dieser Fall hat gegenüber Imp\textsubscript{min} = 4 Vorteile in der Aufteilung der Videoverluste über die Zeit. Die Unterbrechungsduern bzw. die Dauern des Einfrierens des Videobildes werden kürzer. Denn es werden keine Videobilder mit den Videobildabhängigkeiten verworfen.
6.3 Verfahrensergebnisse

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 Videobildeinfriemens 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-Hauptsäulen 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).
6.3.2.6 Zusammenfassende Diskussion


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.


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


# Anhang

## Tabelle A.1: Verwendete Komprimierungsparameter für 1DKf

<table>
<thead>
<tr>
<th>Video</th>
<th>1DKf-Parameter</th>
<th>Bemerkung</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>bq</td>
<td>iq</td>
</tr>
<tr>
<td>vtcInw OB</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>vtcInw nB</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>intros OB</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>intros nB</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>3inrow OB</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>3inrow nB</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>smityl OB</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>smityl nB</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>washdc OB</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>washdc nB</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>football 0B</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>football 5B</td>
<td>0.4</td>
<td>nicht verwendet</td>
</tr>
<tr>
<td>football 8B</td>
<td>0.4</td>
<td>nicht verwendet</td>
</tr>
<tr>
<td>football 11B</td>
<td>0.4</td>
<td>nicht verwendet</td>
</tr>
<tr>
<td>football nB</td>
<td>0.4</td>
<td>nicht verwendet</td>
</tr>
</tbody>
</table>
## Tabelle A.2: Protokolloverhead, 308 kbit/s

<table>
<thead>
<tr>
<th>Video</th>
<th>512 / 40</th>
<th>512 / 44</th>
<th>1024 / 40</th>
<th>1024 / 44</th>
<th>1456 / 40</th>
<th>1456 / 44</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1DK</td>
<td>2DK</td>
<td>1DKf</td>
<td>1DK</td>
<td>2DK</td>
<td>1DKf</td>
</tr>
<tr>
<td>vtctnw 0B</td>
<td>9.1</td>
<td>9.3</td>
<td>9.1</td>
<td>10</td>
<td>10.3</td>
<td>10</td>
</tr>
<tr>
<td>vtctnw 2B</td>
<td>9</td>
<td>9.6</td>
<td>9.1</td>
<td>9.9</td>
<td>10.6</td>
<td>10</td>
</tr>
<tr>
<td>vtctnw 11B</td>
<td>9</td>
<td>9.8</td>
<td>9.1</td>
<td>9.9</td>
<td>10.8</td>
<td>10.1</td>
</tr>
<tr>
<td>introa 0B</td>
<td>9.2</td>
<td>9.3</td>
<td>9.1</td>
<td>10</td>
<td>10.2</td>
<td>9.9</td>
</tr>
<tr>
<td>introa 2B</td>
<td>9</td>
<td>9.4</td>
<td>9.1</td>
<td>9.9</td>
<td>10.4</td>
<td>10</td>
</tr>
<tr>
<td>introa 11B</td>
<td>9.1</td>
<td>9.4</td>
<td>9</td>
<td>10</td>
<td>10.3</td>
<td>9.9</td>
</tr>
<tr>
<td>football 0B</td>
<td>9.1</td>
<td>9.1</td>
<td>8.9</td>
<td>10</td>
<td>10.8</td>
<td>9.8</td>
</tr>
<tr>
<td>football 2B</td>
<td>9</td>
<td>9.2</td>
<td>9</td>
<td>9.9</td>
<td>10.1</td>
<td>9.9</td>
</tr>
<tr>
<td>football 11B</td>
<td>9.1</td>
<td>9.1</td>
<td>8.9</td>
<td>10</td>
<td>10.8</td>
<td>9.8</td>
</tr>
</tbody>
</table>
Längen von B-Serien

Bilddatenraten, kbit/s

Abbildung A.1: Datenraten, 1DK

Abbildung A.2: Datenraten, 1DKf

Abbildung A.3: Datenraten, 2DK
Abbildung A.4: Orliche Auflösung

<table>
<thead>
<tr>
<th>Bildformat</th>
<th>Auflösung</th>
</tr>
</thead>
<tbody>
<tr>
<td>4k Film</td>
<td>4096 x 3112</td>
</tr>
<tr>
<td>2k Film</td>
<td>2048 x 1536</td>
</tr>
<tr>
<td>HDTV</td>
<td>1920 x 1080</td>
</tr>
<tr>
<td>SIF (NTSC)</td>
<td>352 x 240</td>
</tr>
<tr>
<td>SDTV PAL</td>
<td>720 x 480</td>
</tr>
<tr>
<td>SDTV NTSC</td>
<td>720 x 576</td>
</tr>
<tr>
<td>CIF</td>
<td>352 x 288</td>
</tr>
<tr>
<td>CIF (NTSC)</td>
<td>352 x 240</td>
</tr>
<tr>
<td>UHDTV</td>
<td>7680 x 4320</td>
</tr>
</tbody>
</table>
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.


[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.


[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.


[55] ETSI TS 102 470 v1.2.1, "Digital Video Broadcasting (DVB); IP Datacast: Program Speci-

fic 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:


[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, Fehlerbeseiti-

[59] Roland Schmitz, Roland Kiefer, Johannes Maucher, Jan Schulze, and Thomas Suchy. Kom-


[61] BITMOVIN. Optimal Adaptive Streaming Formats MPEG-DASH & HLS Segment

Length, 2019.


ing within the 3GPP Packet-Switched Streaming Service. IEEE Network, pages 34–40,

March/April 2006.


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


[65] H. Wu and K. Rao, editors. Digital Video Image Quality and Perceptual Coding. CRC


[66] Y. Q. Shi and H. Sun. Image and Video Compression for Multimedia Engineering: Funda-


[70] Q. Cai, L. Song, G. Li, and N. Ling. Lossy and lossless intra coding performance eva-

luation: 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.


[141] https://www.tcpdump.org/.


[150] http://media.xiph.org/video/derf/y4m/vtc1nw_422_cif.y4m.


