

VERGLEICH DER TESTSEQUENZ FÜR INTRUSIVE MESSUNG VTQoS MIT DEN SPEECHSEQUENZEN IN DER UMGEBUNG DES FESTNETZES

Peter Počta, Martin Vaculík

Universität Žilina, Univerzitná strasse 1, 010 26 Žilina, Slowakische Republik

Abstract: In diesem Artikel werden die Simulationen der Testsequenzübertragung und der Speechsequenzenübertragung für die intrusive Messung VTQoS durch die Umgebung des Festnetzes beschrieben. Das Ziel der Simulationen war die Detektion des Einflusses SNR und BER auf die Qualität der übertragenen Sequenzen. Die Bewertung der Simulationstestsequenzen und der Simulationsspeechsequenzen werden aufgrund der Ausrechnung mean square measure bewertet. Die Aufgabe dieses Artikels war die Untersuchung der Angemessenheit der Benutzung der Testsequenz auf der intrusive Messung VTQoS in der Umgebung des Festnetzes. Die Testsequenz besteht aus den simplen Signale.

1. Der Prolog

Die Qualitätsbewertung der Speechübertragung ist der wichtigen Qualitätskomponente der geleisteten Dienste. VTQoS ist sowie für den Kunden, als auch für den Betreiber sehr wichtig. Die heutige Kommunikationsnetze enthalten grosse Menge der ganz verschiedenen Übertragungstechnologien. Die Qualitätsmessung der Speechübertragung ist eine der wenigen Plattformen, die den Vergleich der ganz verschiedenen Übertragungstechnologien aus dem Blick der Speechqualität ermöglicht. Diese Plattform ist sehr nahe zum Blick der einzelnen Zielteilnehmer.

Die Übertragungsparameter der Netze kann man natürlich messen und auswerten. In Bezug auf die Komplexitäten der Netztechnologien bietet die optimalen Ergebnisse die Qualitätsbewertung an. Die Qualitätsbewertung „end-to-end“ ist ähnlich wie die Bewertung der Teilnehmer.

Der gebräuchteste Dienst ist der Lautdienst. Der empfangende Teilnehmer benutzt bei diesem Dienst auch die Prediktioneigenschaften und die Filtereigenschaften des Gehirnes, deshalb ist die optimale Qualitätsbewertung gerade bei diesem Dienst sehr wichtig. Für die Qualitätsbewertung der Speechübertragung benutzt man die intrusive und non-intrusive Methode, objektiv oder subjektiv. Bei der non-intrusive Methode kontrolliert man nur den bestehenden Dialog. Der Nachteil dieser Methode ist, dass die Quellenspeechprobe dem bewertenden Algorithmus nicht verfügbar ist, deshalb ist es schwere, gewissen Typen der Störung bei der Übertragung zu detektieren. Bei den intrusive Methoden entstehen Verbindungen, die ausschliesslich für die Übertragung der Speechproben dienen. Die heutige technische und programmatische Ausstattung ermöglicht die Objektivierung dieser Messverfahren. Die Objektivierung wird auf der Sendung der Speechprobe, auf dem Empfang dieser Speechprobe an der Zielseite und auf dem Vergleich des Quellenspeechprobe und des Zielspeechprobe mit Hilfe des passenden Algorithmus gegründet. Dieser Algorithmus bildet die Wahrnehmung und die Bewertung der Meinung über die

Übertragungsqualität des durchschnittlichen Teilnehmers nach. Es geht z.B. um diese Algorithmen:

E-model definiert in ETR-250, oder Algorithmus PSQM (Perceptual Speech Quality Measurement) definiert in der Referenz P.861 ITU-T und Algorithmus PESQ (Perceptual Evaluation of Speech Quality) definiert in der Referenz P.862 ITU-T.

Die Wahl der optimalen Testsequenz ist bei allen diesen Methoden sehr wichtig. Die Testsequenz sollte aus den non-speech-like (fully artificial) Signale bestehen. Diese Signale sind in der Referenz P.501 ITU-T näher definiert. Diese Referenz verteilt die Signale an die deterministischen und zufälligen Signale. Ein grosser Vorteil dieser Signale ist die Simplizität und die grosse Sensibilität auf die Störung und auch die Möglichkeit die Ergebnisse aus den verschiedenen Sprachregionen zu vergleichen. Diese Testsequenz ermöglicht also den Vergleich der Netze, die liegen in den einzelnen Landschaften im Rahmen einer Korporation (zum Beispiel Deutsche Telekom, Orange, Vodafone, usw.) liegen, aus dem Blick VTQoS.

Zur Zeit wird die Messung VTQoS mit Hilfe der Speechproben realisiert. In diesem Fall ist aber der Vergleich nur im Rahmen einer Sprachregion möglich.

In meinem Artikel beschäftige ich mich mit dem Einfluss BER und SNR auf die Qualität der übertragenen Sequenzen durch die Umgebung des Festnetzes.

2. Die Beschreibung der Testsequenz

Die Dauer der Testsequenz ist an 90 Sekunden definiert. Diese Zeit entspricht der Telefongesprächdauer des durchschnittlichen Teilnehmers. Diese Testsequenz besteht aus den simplen Signale. Ihre Schicklichkeit wurde in [1] nachgeprüft. In der Testsequenz werden diese Signale mit den zugehörigen Parameter benutzt:

- Das Signal des sinusförmigen Verlaufs mit den Frequenzen 300, 800, 1000, 1700, 2400, 3000 Hz.

- Das Signal des bipolarförmigen Verlaufs mit den Frequenzen 300, 400, 500, 600, 635, 670 Hz
- Gaussian white noise mit diesen Parameter: $\mu = 0$ a $\delta = 0,0001; 0,001; 0,01; 0,1; 0,5; 1$.

Die Konfiguration auf dem Bild 1 wird in der zugehörigen Testsequenz sechs mal benutzt. Die Testsequenz besteht aus sechs Teilen. Die Signale in einzelnen Teilen der Testsequenz erlangen sukzessive die Valoren, die bei der Beschreibung der benutzten Signalen für die Gestaltung der Testsequenzen eingeführt haben. Aus voriger Beschreibung ergibt sich, dass die Signale im nächsten Teil der Testsequenz (von 15 Sekunde bis 30 Sekunde) diese Valoren haben werden: das Signal des bipolarförmigen Verlaufs $f = 400$ Hz, Gaussian white noise $\delta = 0,001$ und das Signal des sinusförmigen Verlaufs $f = 800$ Hz. Die Valoren der Signale im ersten Teil der Testsequenz (von 0 Sekunde bis 15 Sekunde) werden auf dem Bild 1 abgebildet.

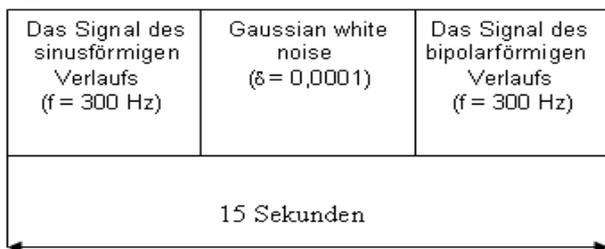


Bild 1 Der Prologteil der Testsequenz

Die Wahl der passenden Testsequenz wurde in [5] realisiert.

3. Die Beschreibung der Speechsequenzen

Die Referenz P.830 ITU-T empfiehlt für die Testung der Qualität des Telefongesprächs im Netz mindestens zwei weibliche und zwei männliche Stimmen zu benutzen. Die Benutzung acht weiblicher, acht männlicher und acht kindlicher Stimmen ist aber am besten. Die zwei weiblichen und männlichen Stimmen werden für die Gestaltung der Speechsequenzen benutzt. Die zwei weiblichen und männlichen Speechsequenzen mit der Dauer von 90 Sekunden werden für die Simulationen benutzt. Diese Speechsequenzen werden aus den Speacheinspielungen gestaltet, die aus der slowakischen Dateibase Speechdat stammen. Die detailliertere Beschreibung der Speechsequenzen:

- Die Speechsequenz „FrauSK1“ wurde mit einer Frau aufgenommen. Diese Frau stammt aus der Westslowakei. Sie ist 57 Jahre alt.
- Die Speechsequenz „FrauSK2“ wurde mit einer Frau aufgenommen. Diese Frau stammt aus der Westslowakei. Sie ist 52 Jahre alt.
- Die Speechsequenz „MannSK1“ wurde mit einem Mann aufgenommen. Dieser Mann

stammt aus der Westslowakei. Er ist 38 Jahre alt.

- Die Speechsequenz „MannSK2“ wurde mit einem Mann aufgenommen. Dieser Mann stammt aus der Mittelslowakei. Er ist 39 Jahre alt.

4. Die Beschreibung der Simulationsmodelle

Die Sequenzen werden in der Umgebung Matlab als die Sequenz der Digitalproben modelliert. Diese Digitalproben werden durch die kompetenten Typen der Kommunikationskanäle transportiert (Der AWGN (Add white Gaussian noise) Kanal und BSC (Binary Symmetric Channel)).

Die Simulationen der Signalübertragung durch diese Übertragungskette stellen das Ziel dieser Modelle vor. Die Simulationen werden aus dem Standpunkt des Einflusses BER und SNR auf die Qualität der übertragenen Sequenzen durch die Umgebung des Festnetzes realisiert. Auf dem Bild 2 ist ein prinzipielle Schema der Simulationsmodelle. Zwei Modelle wurden für die Simulationen gabaut. Das erste Modell ist für die Übertragung durch den AWGN Kanal bestimmt. Der Einfluss SNR auf die Qualität der übertragenen Sequenzen wird mit Hilfe des AWGN Kanals modelliert. Das zweite ist für die Übertragung durch BSC bestimmt. Der Einfluss BER auf die Qualität der übertragenen Sequenzen wird mit Hilfe BSC modelliert. Diese zwei Modelle sind sehr ähnlich, deshalb wird nur ein Simulationsmodell präsentiert. Auf dem Bild 3 ist ein Simulationsmodell für den AWGN Kanal. Das Simulationsmodell für BSC beinhaltet gegenüber dem Simulationsmodell für AWGN Kanal diese Blöcke nicht: Zero-Order Hold 2, Saturation 2, Quantizer 2. Im Simulationsmodell für BSC wird das AWGN Kanal gegen BSC ausgetauscht.

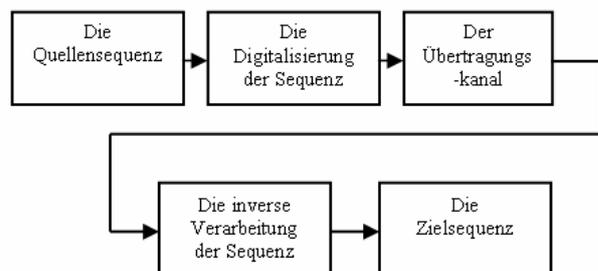


Bild 2 Das prinzipielle Schema der Simulationsmodelle

4.1 Die detailliertere Beschreibung der Simulationsmodells

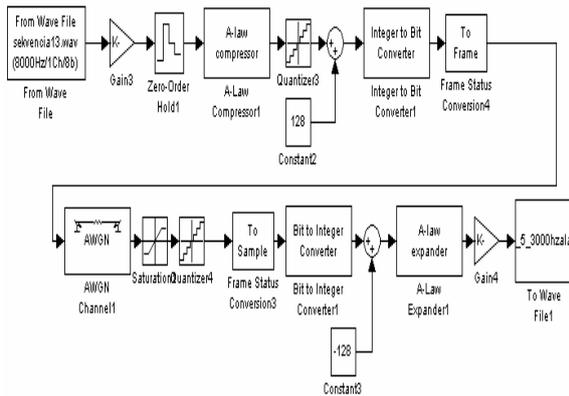


Bild 3 Das Simulationsmodell für den AWGN Kanal

Dieses Modell wird in [1] beschrieben.

4.2 Das Prinzip der Simulation

Die Quellensequenzen werden mit Hilfe des Programms Soundforge gestaltet. Das Prinzip der Gestaltung der Quellensequenzen wird in den Kapitel 2 und 3 beschrieben. Die Zielsequenzen werden mit Hilfe der Simulationen gestaltet. Die Dauer der Simulation wird an 90 Sekunden definiert. Aus der Dauer der Simulation ergibt sich, dass die Dauer der Zielsequenz 90 Sekunden beträgt. Diese Dauer ist dieselbe wie die Dauer der Quellensequenz. Nach der Beendigung der Simulation werden die Quellensequenz und Zielsequenz verglichen. Das Prinzip des Vergleiches wird in dem Algorithmus des Vergleiches beschrieben.

Der Algorithmus des Vergleiches:

- 1) Die Einlesung der Quellensequenz und Zielsequenz
- 2) Die Segmentierung der Sequenzen auf n Intervalle mit dem Umfang 8000 Proben
Auf jedem Intervall wird mean square measure gerechnet.
- 3) Die Rechnung des Mittelwerts mean square measure

$$\bar{d} = \frac{\sum_{i=1}^n d_i}{n}, \quad (1)$$

wo n die Anzahl des Intervalls ist, d_i ist mean square measure des i -ten Intervalls.

Genauere Ergebnisse werden mit der Segmentierung der Sequenz auf n Intervalle und mit der nachfolgenden Rechnung mean square measure auf diesen Intervalle versichert.

Mean square measure wird auf dem Vergleich des Quellenmikrosegmentsspektrums und Zielmikrosegmentsspektrums gegründet. Die am

meistens benutzte L_2 Norm, wird mit Hilfe der folgenden Formel definiert:

$$d_2(t, r) = \left[\sum_{j=1}^N (y_{tj} - y_{rj})^2 \right]^{1/2}, \quad (2)$$

wo N ist die Anzahl des FFT Punktes in kompetentem Mikrosegment, y_{tj} ist der absolute Valor des j -ten FFT Koeffizienten des Zielmikrosegments, y_{rj} ist der absolute Valor des j -ten FFT Koeffizienten des Quellenmikrosegments.

5. Die Präsentation der Ergebnisse

Die Simulationen wurden mit Hilfe der beschriebenen Sequenzen für die A-law Kompresscharakteristik realisiert. Die Übertragung wurde durch den AWGN Kanal und durch BSC realisiert. Die Simulationen für den AWGN Kanal wurden für diese Valoren des SNR Parameters realisiert: 5, 10, 15, 25, 50 dB. Die Simulationen für BSC wurden für diese Valoren der Fehlerwahrscheinlichkeit realisiert: 0,0001; 0,001; 0,01; 0,1; 0,5; 1. Das Ziel dieser Simulationen war die Feststellung des Einflusses BER und SNR auf die Qualität der übertragenen Sequenzen. Diese Ergebnisse wurden mit Hilfe der Simulationen bekommen.

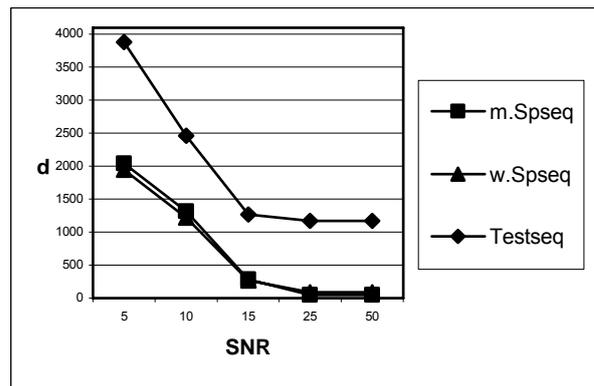


Bild 4 Die graphische Repräsentation der Simulationsergebnisse für den AWGN Kanal

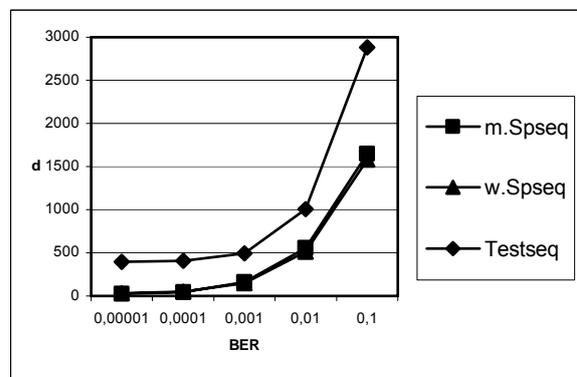


Bild 5 Die graphische Repräsentation der Simulationsergebnisse für BSC

Die Anmerkungen der Kurzformen aus den Graphen:

- Die Kurzform „w.Spseq“ stellt die weiblichen Speechsequenzen dar.
- Die Kurzform „m.Spseq“ stellt die männlichen Speechsequenzen dar.
- Die Kurzform „Testseq“ stellt die Testsequenz dar.

In den Graphen präsentieren wir nur die Mittelwerte der weiblichen und männlichen Speechsequenzensimulationsergebnisse.

6. Das Ende

Aus den Graphen können wir sehen, dass die Testsequenz mehr empfindlich als die Speechsequenzen auf die Störungseinflüsse ist, die bei der Übertragung durch die Umgebung des Festnetzes entstehen. Die Testsequenz besteht aus den simplen Signale. Die simplen Signale haben die höhere Sensibilität auf die Störung als die Speechsequenzen. Das ist die Ursache der höheren Sensibilität der Testsequenz auf die Störungseinflüsse. Die höhere Sensibilität der Testsequenz ist für die intrusive Messung VTQoS sehr günstige und ermöglicht uns genauere Messung der Störungseinflüsse, die in den Übertragungsketten entstehen. Die hohe Sensibilität ermöglicht uns die qualitativen Änderungen zu prognostizieren, die im Netz entstehen. Die Schicklichkeit dieser Testsequenz für intrusive Messung VTQoS wird mit Hilfe der Messungen im konvergenten Netz der Universität Žilina testen.

Die Danksagungen

Dieser Artikel wurde von dem Staatsprogramm Nummer 2003 SP 51/028 09 00/028 09 10 unterstützt.

Dieser Artikel wurde teils von dem Staatsprogramm der Forschung und der Entwicklung „Intelligent Speech Communications Interface“ unterstützt. Dieses Programm werde mit dem Ministerium für Ausbildung der Slowakei unterstützt.

Das Literaturverzeichnis

- [1] Počta, P., Vaculík, M. : Method of choice of test signals for automatic intrusive measurement VTQoS, MESAQIN 2005, Praha.
- [2] Matlab Help.
- [3] Psutka, J. : Komunikace s počítačem mluvenou rečí, ACADEMIA, 1995.
- [4] Franeková, M. : Modelovanie komunikačných systémov v prostredí Matlab, Simulink a Communications Toolbox, EDIS, 2003.
- [5] Počta, P., Vaculík, M. : Determination of optimal test sequence for automatic intrusive measurement VTQoS on environment of fixed telecommunication network, RTT 2005, Ostrava.