

Zur „Geschichte“ des E-Modells

Sebastian Möller, Marcel Wältermann

Quality and Usability Lab, Telekom Innovation Labs, TU Berlin, E-Mail: sebastian.moeller@telekom.de

Einleitung

Das sogenannte E-Modell in ITU-T-Empfehlung G.107 [1] ist das aktuelle Planungswerkzeug zur Vorhersage der Sprachqualität von Telefonnetzen. Es schätzt auf der Basis einer parametrischen Beschreibung der gesamte Übertragungsstrecke Mund-zu-Ohr die Gesamtqualität, wie sie ein „mittlerer“ Nutzer der betrachteten Telefonverbindung bewerten würde. Die Vorhersage fußt auf 19 Eingangsparametern, die mittels komplexer nichtlinearer Zusammenhänge über eine sog. psychologische Skala auf einen skalaren Qualitätsschätzwert transformiert werden. Aufgrund der Entwicklung des Modells in verschiedenen Standardisierungsgremien über einen Zeitraum von über 20 Jahren bleibt der Ursprung vieler Zusammenhänge aber leider im Dunkeln.

In diesem Beitrag soll deshalb ein komprimierter Überblick über die historischen Grundlagen des Modells gegeben werden. Dazu werden zunächst zwei grundlegende Konzepte des Modells beschrieben: Das Konzept der *Loudness Ratings* sowie die Additivität unterschiedlicher Störungsarten. Anschließend wird ein geschichtlicher Abriss der verschiedenen Vorgängermodelle und Entwicklungsstufen des E-Modells gegeben, und es wird gezeigt, durch welche Veränderungen das Modell trotz seiner über 80-jährigen Geschichte auch bei modernsten paketorientierten und mobilen Vermittlungstechniken weiterhin höchste Akzeptanz genießt.

Loudness Ratings

Die Eingangsgrößen des E-Modells sind Parameter, die Planern von Telefonnetzen typischerweise aus den Spezifikationen der beteiligten Elemente zur Verfügung stehen. Für die einzelnen Übertragungswege werden daher keine Übertragungsfunktionen verwendet, sondern Einzahlwerte zur Beschreibung der Abschwächung, der Laufzeit, sowie der in diesem Teil des Übertragungswegs akkumulierten Störungen.

Abschwächungen werden durch sog. *Loudness Ratings* beschrieben [2]. Diese bezeichnen die Einfügungsdämpfung, die man einem Referenzpfad geben muss, um subjektiv die gleiche Lautheit wie der betrachtete Übertragungspfad zu erlangen. Bei Definition eines geeigneten Referenzpfades (hierzu wird heute das *Intermediate Reference System, IRS* nach ITU-T Rec. P.48 verwendet) können *Loudness Ratings* aber nicht nur subjektiv, sondern auch algorithmisch aus gemessenen Sensitivitäten des Übertragungspfad und angenommenen Sprach- und Rauschspektren berechnet werden. Hierzu wird das Lautheits-Modell nach Fletcher & Munson [3] verwendet, welches auf dem *Sensation Level Z* als Abstand zwischen der um etwaiges Rauschen

angehobenen Ruheshwelle (Kurve c in Abb. 1) und dem übertragenen Sprachspektrum (b) berechnet wird. Aus Z wird über eine asymptotische Wachstumsfunktion Q und über Frequenz-Gewichtungsfaktoren B' die Lautheit λ durch Integration über den betrachteten Frequenzbereich gewonnen:

$$\lambda = \text{const} \cdot \int Q(Z)B' df \quad (1)$$

Bei Richards [2] sind unterschiedliche Gewichtungsfaktoren B' für Gesamtqualität, Höranstrengung und Artikulation angegeben.

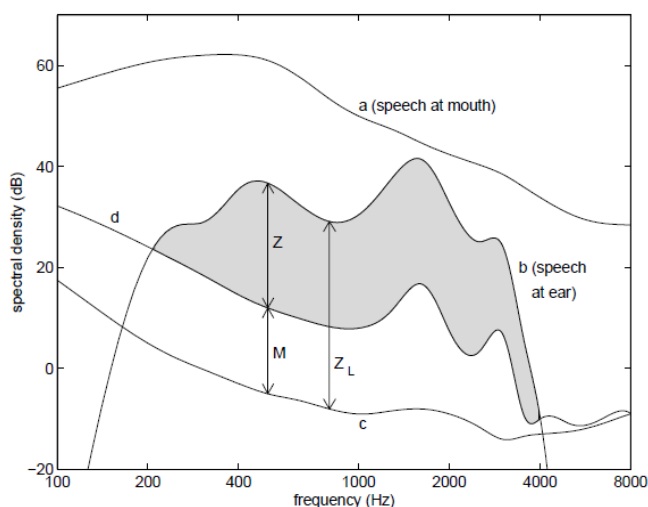


Abbildung 1: Berechnung der Lautheit aus dem Sensation Level Z durch Integration über die graue Fläche (Bild aus [4], Werte aus [2]).

Ein wichtiges Charakteristikum dieser Methode ist, dass *Loudness Ratings* für einzelne Teile des Übertragungspfad additiv sind, d.h. dass sich die Gesamt-Einfügungsdämpfung annähernd aus den Einfügungsdämpfungen der Teilpfade ergibt.

Additivität von Störungen

Ein weiteres Charakteristikum des E-Modells ist, dass es unterschiedlichste Arten von Störungen zu einem Gesamt-Qualitätswert integriert. Der dabei verwendete Weg beruht auf Beobachtungen von Allnatt [5], der für durch Rauschen und Echos gestörte Videobilder eine Additivität der Störfähigkeit auf einer sog. psychologischen Skala feststellte. Auch wenn die Beschreibungen dieser Skala nicht ganz einheitlich sind kann man davon ausgehen, dass hierbei von einer Ratioskala ausgegangen wird. Bereits das Bellcore-Modell als Vorgängermodell des E-Modells verwendete hierzu die sog. *Transmission Rating Scale R*, welche dem Prinzip nach vom E-Modell übernommen wurde. Auf dieser Skala addieren sich die Störungen durch Abschwächung, Rauschen, nicht-optimales Rückhören, Echos und nichtlineare Verzerrungen sowie zeitvariantes Verhalten des Übertragungskanal zu einem Gesamtwert R ,

der dann mittels einer S-förmigen Beziehung auf einen *Mean Opinion Score*, MOS, umgerechnet werden kann. Letzterer entspricht dem Urteil einer mittleren Versuchsperson auf einer 5-stufigen Kategorienskala.

Entstehung des E-Modells

Seit Ende der 1960er Jahren arbeiteten die größeren Netzbetreiber an eigenständigen Modellen zur Vorhersage der Sprachqualität bei der Netzplanung. Drei Hauptmodelle sind dabei entstanden, die als direkte Vorgänger des E-Modells angesehen werden können und maßgeblichen Einfluss auf seine aktuelle Form haben [6]:

1. Das Bellcore-Transmission-Rating-Modell. Es basiert auf Testergebnissen von 1965-1972, welche noch mittels des Referenzsystems EARS erfasst und später umgerechnet wurden. Es umfasste Leitungsrauschen, *Loudness Ratings*, Sprecher- und Hörerechos (1977-78), Rückhören (1980), lineare Verzerrungen, Quantisierungsrauschen, sowie sende- und empfangsseitiges Hintergrundgeräusch. Die Ergebnisse wurden auf einer dem E-Modell vergleichbaren *Transmission Rating Scale* angegeben, allerdings war noch keine Transformation auf MOS definiert, stattdessen aber auf den Anteil an Urteilen im Bereich gut oder besser (GoB) bzw. dürftig oder schlechter (PoW).
2. Das CATNAP-Modell der British Telecom (*Computer-Aided Telephone Network Assessment Program*). Dieses Modell verfolgte die Idee der Sichtbarmachung der Ursache-Wirkungs-Prinzipien der menschlichen Wahrnehmung und arbeitete mit Sensitivitäten und Rauschspektren anstelle der *Loudness Ratings*. Diese wurden ähnlich wie in Gl. (1) zunächst in einen Wert für die Höranstrengung (dann mit anderen Werten für B') umgerechnet und erst anschließend in einen MOS für die Konversationsituation transformiert.
3. Das OPINE-Modell der NTT (*Overall Performance Index model for Network Evaluation*). Es wurde im Zeitraum 1983-1986 entwickelt und umfasste lineare Verzerrungen, Leitungsrauschen, Hintergrundgeräusche, Quantisierungsrauschen, Sprecherecho und Rückhören. Dieses Modell basierte bereits auf dem Allnatt'schen Additivitätsprinzip.

Die Integration dieser drei bestehenden Modelle, weiterer Modellansätze sowie neuer Faktoren und Testergebnisse in ein einheitliches Modell wurde schließlich Anfang der 1990er Jahre innerhalb einer Arbeitsgruppe der ETSI (BTC2), maßgeblich durch Johannesson [7] vorangetrieben. Durch Kombination des Additivitäts-Prinzip des OPINE-Modells, des *Transmission Ratings R* sowie der *Loudness Ratings* und der Echo-Störungen des Bellcore-Modells, sowie des Quantisierungsrauschens des CATNAP-Modells wurde ein einheitliches Werkzeug geschaffen, mit dessen Hilfe Netzbetreiber weltweit nach dem gleichen Standard Sprachqualität planen können. Neue, erstmals im E-Modell betrachtete Aspekte waren die Verzerrungen durch nichtlineare Kodierer (insbes. GSM) sowie die Betrachtung des psychologischen „Qualitätszuschlags“, welcher bei

mobiler Telefonie oder bei Satellitenverbindungen häufig beobachtet wird. Gemäß der für seine Entwicklung verantwortliche Organisation wurde das ETSI-Modell schließlich als „E-Modell“ bezeichnet und erstmalig 1996 in ETR 250 [8] vollständig beschrieben, 1998 dann auch bei der ITU-T als Rec. G.107 [1] standardisiert.

Aktuelle Entwicklungen und Ausblick

Seit seiner ersten Standardisierung bei ETSI und ITU wurden insgesamt 7 wesentliche Veränderungen am Modell vorgenommen, welche in Neufassungen der ITU-Empfehlung resultierten. Diese umfassen die Berücksichtigung des Lombard-Effekts bei Hintergrundgeräuschen auf Sendeseite sowie Verbesserung bei Quantisierungsrauschen (2000), die Ableitung von Faktoren für Kodierverzerrungen aus subjektiven Tests (neue ITU-T Rec. P.833, 2001) bzw. mittels instrumenteller Schätzer (neue ITU-T Rec. P.834, 2002), die Berücksichtigung von unabhängigen Paketverlusten (2002), eine verbesserte Modellierung bei leisem Rückhören (2003), die Berücksichtigung von abhängigen Paketverlusten (2005), sowie die Definition einer Transmission-Rating-Skala und von Kodierverzerrungs-Faktoren für Breitband-Codex (2008). Ende 2011 wurde erstmalig eine komplette Breitband-Version des E-Modells vorgelegt, welche fast alle auch im Schmalbandfall betrachteten Störungsklassen berücksichtigt. Diese Version soll jetzt innerhalb der ITU-T Studiengruppe 12 validiert und verbessert werden, um auch bei der Planung zukünftiger Breitband-Sprachübertragung nützlich zu sein.

Literatur

- [1] ITU-T Rec. G.107. The E-Model, a Computational Model for Use in Transmission Planning. Int. Telecomm. Union, Genf, 2011.
- [2] Richards, D. L. *Telecommunication by Speech*. Butterworths, London, 1973.
- [3] Fletcher, H., Munson, W. Relation Between Loudness and Masking. *J. Acoust. Soc. Am.*, 9(1), 1–10, 1937.
- [4] Möller, S. *Assessment and Prediction of Speech Quality in Telecommunications*. Kluwer Academic Publ., Boston MA.
- [5] Allnatt, J. W. Subjective Rating and Apparent Magnitude. *Int. J. Man Machine Studies*, 7, 801–816, 1975.
- [6] ITU-T Suppl. 3 to P-Series Rec. Models for Predicting Transmission Quality from Objective Measurements. Int. Telecomm. Union, Genf, 1993.
- [7] Johannesson, N. O. The ETSI Computation Model: A Tool for Transmission Planning of Telephone Networks. *IEEE Communications Magazine*, Jan. 1997, 70–79.
- [8] ETSI Technical Report ETR 250. *Transmission and Multiplexing (TM); Speech Communication Quality from Mouth to Ear for 3,1 kHz Handset Telephony across Networks*. European Telecommunications Standards Institute, Sophia Antipolis, 1996.