



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA ELEKTROTECHNIKY

A KOMUNIKAČNÍCH TECHNOLOGIÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

ÚSTAV RADIOELEKTRONIKY

DEPARTMENT OF RADIO ELECTRONICS

MĚŘENÍ HLASOVÉ KVALITY TECHNOLOGIE

VOLTE/VOWIFI

TESTING OF VOICE QUALITY OF VOLTE/VOWIFI TECHNOLOGY

DIPLOMOVÁ PRÁCE

MASTER'S THESIS

AUTOR PRÁCE

AUTHOR

Bc. Martin Kráčala

VEDOUCÍ PRÁCE

SUPERVISOR

prof. Ing. Stanislav Hanus, CSc.

BRNO 2016

Diplomová práce

magisterský navazující studijní obor **Elektronika a sdělovací technika**
Ústav radioelektroniky

Student: Bc. Martin Kráčala

ID: 146868

Ročník: 2

Akademický rok: 2015/16

NÁZEV TÉMATU:

Měření hlasové kvality technologie VoLTE/VoWiFi

POKYNY PRO VYPRACOVÁNÍ:

Seznamte se s metodami měření hlasové kvality v telekomunikačních sítích, zejména pak s algoritmem ITU-T P.863 POLQA a metodami přenosu hlasu v sítích LTE a WiFi. Prostudujte možnosti využití E-modelu a R-faktoru dle ITU-T G.107 pro neintrusivní měření hlasové kvality hovoru na základě změřených parametrů v síti LTE s využitím technologie Voice over LTE (VoLTE).

Navrhněte algoritmus výpočtu R-faktoru na základě parametrů dostupných z měření běžným uživatelským zařízením s operačním systémem Android (např. Samsung Galaxy S5). Takto navržený algoritmus kalibrujte na co nejvyšší možnou korelaci výsledků R-faktoru a hlasové kvality MOS / POLQA. Na základě zjištěných poznatků navrhňte jednoduchý skript pro zařízení s OS Android pro provádění cyklického autonomního měření R-faktoru v sítích s technologiemi VoLTE a VoWiFi.

DOPORUČENÁ LITERATURA:

[1] Recommendation ITU-T P.863 (2011): Perceptual objective listening quality assessment. Switzerland: ITU, 2011

[2] Recommendation ITU-T G.107 (2012): Amendment 1: New Appendix IV – Use of the E-model in conjunction with noise reduction or echo canceller systems in the network or the terminal equipment. Switzerland: ITU, 2012

Termín zadání: 8.2.2016

Termín odevzdání: 19.5.2016

Vedoucí práce: prof. Ing. Stanislav Hanus, CSc.

Konzultant diplomové práce: Ing. Josef Fencel

doc. Ing. Tomáš Kratochvíl, Ph.D., předseda oborové rady

UPOZORNĚNÍ:

Autor diplomové práce nesmí při vytváření diplomové práce porušit autorská práva třetích osob, zejména nesmí zasahovat nedovoleným způsobem do cizích autorských práv osobnostních a musí si být plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č.40/2009 Sb.

ABSTRAKT

Tato diplomová práce se zabývá možnostmi použití ITU-T G.107 E-modelu pro objektivní neintrusivní měření kvality přenosu hlasu v sítích LTE a Wi-Fi. V první části práce jsou prezentovány používané metody měření hlasové kvality, zejména algoritmus POLQA, rozbor E-modelu a přehled technologií VoLTE a VoWiFi. Hlavní částí práce je návrh algoritmu výpočtu R-faktoru, lineárního ukazatele kvality přenosu hlasu, vycházejícího z parametrů měřených uživatelskými zařízeními s OS Android. Návrh vychází z rozsáhlých měření technologie VoLTE a jeho úspěšnost predikce hlasové kvality je vyhodnocována na základě srovnání s měřeními algoritmem POLQA. V práci jsou dále uvedeny možnosti implementace algoritmu pro měření hlasové kvality na zařízeních s OS Android.

KLÍČOVÁ SLOVA

Hlasová kvalita, POLQA, E-model, R-faktor, VoLTE, VoWiFi.

ABSTRACT

This master's thesis deals with application of the ITU-T G.107 E-model for objective non-intrusive voice transmission quality measurements in LTE and Wi-Fi networks. The first part presents techniques used for voice quality measurements, particularly algorithm POLQA, analysis of the E-model and overview of the VoLTE and VoWiFi technologies. The main part of this paper consists of design of the R-factor calculation formula using parameters measured by Android OS powered devices. The algorithm design is based on extensive VoLTE measurements and its voice quality prediction successfulness is evaluated by comparison with POLQA measurements. The paper also presents implementation possibilities of the proposed algorithm on devices with Android OS.

KEYWORDS

Speech quality, POLQA, E-model, R-factor, VoLTE, VoWiFi.

KRÁČALA, Martin *Měření hlasové kvality technologie VoLTE/VoWiFi: diplomová práce.* Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav radioelektroniky, 2016. 95 s. Vedoucí práce byl prof. Ing. Stanislav Hanus, CSc.

PROHLÁŠENÍ

Prohlašuji, že svou diplomovou práci na téma „Měření hlasové kvality technologie VoLTE/VoWiFi“ jsem vypracoval(a) samostatně pod vedením vedoucího diplomové práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

Jako autor(ka) uvedené diplomové práce dále prohlašuji, že v souvislosti s vytvořením této diplomové práce jsem neporušil(a) autorská práva třetích osob, zejména jsem nezasáhl(a) nedovoleným způsobem do cizích autorských práv osobnostních a/nebo majetkových a jsem si plně vědom(a) následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů, včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č. 40/2009 Sb.

Brno

.....

podpis autora

PODĚKOVÁNÍ

Rád bych poděkoval zejména konzultantovi diplomové práce panu Ing. Josefu Fenclovi za konzultace problémů a podnětné návrhy k vypracování práce. Poděkování patří také prof. Ing. Stanislavu Hanusovi, CSc. za pedagogické vedení práce, Ing. Michalu Vojtíškovi za konzultace týkající se OS Android a společnosti T-Mobile Czech Republic, a. s. za poskytnutí měřicího systému Accuver.

Brno

.....

podpis autora

OBSAH

Úvod	11
1 Metody měření hlasové kvality	12
1.1 Definice kvality přenosu hlasu	12
1.2 Poslechové a konverzační testy	12
1.3 Subjektivní metody	13
1.4 Způsoby hodnocení kvality vzorků	13
1.5 Objektivní metody	15
1.5.1 Intrusivní metody	15
1.5.2 Neintrusivní metody	21
2 E-model a jeho využití pro měření hlasové kvality hovoru	24
2.1 Vstupní parametry	24
2.2 R-faktor	25
2.3 Wideband E-model	26
2.4 Faktory degradace	28
2.4.1 Poměr signálu k šumu, $R_{0,WB}$	28
2.4.2 Zkreslení synchronní se signálem, $I_{s,WB}$	29
2.4.3 Degradace vlivem zpoždění, $I_{d,WB}$	29
2.4.4 Efektivní faktor degradace použitou technikou, $I_{e,eff,WB}$	30
2.4.5 Faktor zvýhodnění, A	32
2.5 R-faktor ve vztahu k hlasové kvalitě	32
3 Metody přenosu hlasu v sítích LTE a Wi-Fi	36
3.1 Hlasové služby v sítích LTE	36
3.1.1 Hlasové služby OTT	36
3.1.2 Circuit-switched fallback (CSFB)	37
3.1.3 Voice over LTE (VoLTE)	37
3.2 Hlasové služby v sítích Wi-Fi	39
3.3 IMS protokoly	40
3.3.1 Session Initiation Protocol (SIP)	41
3.3.2 User Datagram Protocol (UDP)	44
3.3.3 RTP/RTCP	44
3.4 Hlasové kodeky	46
3.4.1 Adaptive Multi-Rate (AMR)	46
3.4.2 Adaptive Multi-Rate Wideband (AMR-WB)	46
3.4.3 Enhanced Voice Services (EVS)	49

4	Návrh a optimalizace algoritmu výpočtu R-faktoru	51
4.1	Měření hlasové kvality v síti LTE	51
4.1.1	Měřicí systém Accuver XCAL	51
4.1.2	Měřicí scénář	53
4.1.3	Měřené parametry	53
4.1.4	Měření hlasové kvality algoritmem POLQA	59
4.2	Aplikace E-modelu	61
4.2.1	M2E zpoždění	62
4.2.2	AMR-WB kodek	63
4.2.3	Modelování ztrát paketů	64
4.3	Adaptace algoritmu na základě měření	66
4.3.1	Modelování vlivu jitteru	66
4.3.2	Optimalizace parametrů hlasového kodeku	69
4.3.3	Odhad M2E zpoždění	71
4.4	Analýza výsledků navrženého algoritmu	72
4.4.1	Výsledky poslechového R-faktoru	73
4.4.2	Výsledky hlasové kvality MOS	75
4.4.3	Výsledky odhadu M2E zpoždění	76
5	Implementace v OS Android	79
5.1	Omezení OS Android	79
5.2	Alternativní způsoby	79
6	Závěr	80
	Literatura	82
	Seznam symbolů, veličin a zkratk	87
	Obsah příloženého DVD	94

SEZNAM OBRÁZKŮ

1.1	Časté zkratky používané spolu s MOS	15
1.2	Zjednodušené blokové schéma algoritmu PESQ	16
1.3	Zjednodušené blokové schéma algoritmu POLQA	18
1.4	Ilustrace přiřazení úseků nahrávky a volba ideální konfigurace	20
1.5	Příklad zarovnání signálů algoritmem POLQA	20
1.6	Výpočet modifikovaných kritických pásem algoritmem POLQA	21
1.7	Blokové schéma algoritmu P.563 (3SQM)	23
2.1	Parametrický popis E-modelu	25
2.2	Lineární škála R-faktoru dle G.107.1 s příslušnými kategoriemi kvality	27
2.3	Závislost R-faktoru (NB) na velikosti úplného zpoždění T_a	30
2.4	Závislost $I_{e,eff,WB}$ na ztrátovosti paketů P_{pl} vybraných kodeků	31
2.5	Vztah E-modelu k dalším metodám testování kvality přenosu hlasu	32
2.6	Závislost MOS-CQEW na hodnotě R-faktoru	34
2.7	Statistické kvalitativní ukazatele <i>GoB</i> a <i>PoW</i> v závislosti na R-faktoru	35
3.1	Systémová architektura LTE definovaná 3GPP	36
3.2	Logická architektura jádra IMS	38
3.3	OSI model vyšších vrstev technologií VoLTE a VoWiFi	41
3.4	Minimální struktura VoIP paketu	41
3.5	Signalizace protokolu SIP při sestavování a ukončování hovoru	43
3.6	Hlavička protokolu UDP	44
3.7	Hlavička protokolu RTP	45
3.8	Používané frekvenční šířky pásma hlasových kodeků	47
3.9	Přesah okna LPC analýzy v algoritmu AMR-WB	48
3.10	Obecná struktura rámců AMR-WB IF1 a IF2	48
3.11	Generování AMR/AMR-WB rámců v závislosti na VAD	49
3.12	Subjektivní hodnoty MOS čistých hl. vzorků kodeků AMR a EVS	50
4.1	Zapojení měřicího systému Accuver XCAL	52
4.2	Grafické znázornění měřicí sekvence	53
4.3	Ilustrace zdrojového a paketového zpoždění	56
4.4	Ilustrace spojování paketů	58
4.5	Metoda měření jednosměrného zpoždění prostřednictvím RTCP zpráv	58
4.6	Hlasová nahrávka použitá při měření algoritmem POLQA	59
4.7	Srovnání originální a degradované nahrávky algoritmem POLQA	60
4.8	Analýza M2E zpoždění algoritmem POLQA	60
4.9	Zjednodušené znázornění vstupů a výstupů navrhovaného algoritmu	61
4.10	Vrstvový model VoLTE spojení s typickými hodnotami zpoždění	62
4.11	Závislost faktoru degradace I_d na hodnotě M2E zpoždění T_a	63

4.12	Diagram algoritmu pro určení charakteru ztracených paketů	64
4.13	Dvoustavový Markovův model ztrát paketů	65
4.14	Korelační diagram závislosti MOS na jitteru a M2E zpoždění	67
4.15	Průběh jitteru pro různé intervaly M2E zpoždění	68
4.16	Korelační diagram spojování paketů <i>Bundlr</i> a jitteru J_{p94}	69
4.17	Histogram hodnot R získaných z měření MOS s minimální degradací	70
4.18	Korelační diagramy R před a po optimalizaci parametrů	70
4.19	Histogram průměrného M2E zpoždění při měření alg. POLQA	71
4.20	Korelační diagram měření hl. kvality POLQA a výsledků alg. R_{LQ} . . .	74
4.21	Korelační diagram měření MOS-LQO a výsledků alg. MOS-LQE . . .	76
4.22	Korelační diagram měření M2E zpoždění alg. POLQA a odhadu T_a .	77

SEZNAM TABULEK

1.1	Hodnoty stupnice MOS pro poslechové ACR a DCR testy	14
1.2	Operační módy algoritmu ITU-T P.863 POLQA	18
1.3	Dosažitelné skóre MOS-LQO vybraných hlasových kodeků	19
2.1	Soubor parametrů E-modelu a jejich vlastností	27
2.2	Hodnoty $I_{e,WB}$ a B_{pl} pro vybrané kodeky	31
2.3	Příklady hodnot faktoru zvýhodnění A pro různé kom. systémy	32
2.4	Definice kategorií vnímané kvality přenosu	33
3.1	Standardizované vlastnosti QCI	39
3.2	Seznam často používaných SIP metod	42
3.3	Operační módy a odpovídající velikosti rámce kodeku AMR-WB . . .	47
4.1	Analýza korelace měření hl. kvality POLQA a výsledků alg. R_{LQ} . . .	74
4.2	Analýza korelace měření MOS-LQO a výsledků alg. MOS-LQE	75
4.3	Analýza korelace měření M2E zpoždění alg. POLQA a odhadu T_a . .	77

ÚVOD

Mobilní hlasová komunikace na dlouhé vzdálenosti prostřednictvím telekomunikačních sítí je v dnešní době již samozřejmou součástí každodenního života. Zatímco v předchozích dekadách mobilním komunikacím jasně dominoval hlasový provoz v okruhově spínaných sítích, v současnosti se již většina pozornosti věnuje ryze datovým službám postaveným na síťových protokolech TCP/IP. V bezdrátových sítích LTE a WLAN se jedná o technologie Voice over LTE (VoLTE) a Voice over Wi-Fi (VoWiFi).

Ze strany uživatelů, kteří jsou navyknuti na vysoký standard kvality tradičních hlasových služeb, je však na tyto nové technologie kladen tlak na zajištění vysoké kvality komunikace. To je v rozporu s principem fungování běžných IP sítí, které negarantují vysokou kvalitu spojení. Kvalita hlasových služeb je závislá především na ztrátovosti paketů v síti, transportnímu zpoždění a jitteru.

Proto hledají poskytovatelé těchto služeb způsoby, jakými by bylo možné vnímanou hlasovou kvalitu objektivně vyhodnotit. Hlavní náplní této práce je návrh výpočetního algoritmu, který by na základě parametrů měřených uživatelskými zařízeními dokázal neintrusivně odhadnout aktuální vnímanou kvalitu konverzace. Problematika parametrických modelů hlasové kvality je dlouhodobě vyšetřována ITU a je obsahem doporučení ITU-T G.107, ze kterého návrh algoritmu vychází. Optimalizace parametrů navrženého algoritmu proběhla s využitím rozsáhlých měření hlasové kvality technologie VoLTE, přičemž jako reference byly použity výsledky algoritmu POLQA.

Diplomová práce se skládá celkem z pěti částí. Kapitola 1 popisuje používané metody měření hlasové kvality, tj. konverzační, poslechové, subjektivní a objektivní, a způsoby, jakými se úroveň hlasové kvality vyjadřuje. Podrobně je popsán především algoritmus ITU-T P.863 POLQA. Druhá část prezentuje popis E-modelu a způsob výpočtu tzv. R-faktoru, skalárního ukazatele predikované kvality komunikační sítě. Třetí kapitola obsahuje přehled používaných metod přenosu hlasu v sítích LTE a Wi-Fi se zaměřením zejména na funkce zprostředkované IMS a protokoly, které k uskutečnění hovorů využívá. Hlavní část práce je soustředěna v kapitole 4, která obsahuje popis měření technologie VoLTE spolu s návrhem a optimalizací hledaného algoritmu výpočtu R-faktoru. Součástí kapitoly je také srovnání výsledků navrženého algoritmu s měřeními alg. POLQA. Poslední kapitola krátce představuje možnosti implementace navrženého algoritmu pro zařízení s OS Android a bezpečnostní omezení, které tuto implementaci výrazně znesnadňují.

1 METODY MĚŘENÍ HLASOVÉ KVALITY

Následující kapitola prezentuje přehled metod používaných k měření hlasové kvality v telekomunikačních systémech. Je definován pojem hlasová kvalita a parametry, které ji ovlivňují, a na základě standardů Mezinárodní telekomunikační unie ITU jsou popsány způsoby jejího měření se zaměřením na objektivní algoritmy ITU-T P.563 (3SQM), ITU-T P.862 (PESQ) a především nejnovější algoritmus ITU-T P.863 (POLQA).

1.1 Definice kvality přenosu hlasu

Pojmem kvalita přenosu hlasu nebo hlasová kvalita se v telekomunikačních sítích rozumí subjektivní hodnocení jakosti řečového signálu po průchodu komunikačním řetězcem. Kvalita nepatří mezi technické parametry, jedná se o uživatelské hodnocení jakosti poskytované služby. Měření kvality hlasového přenosu tak zohledňuje aspekty lidského vnímání řeči a vliv degradace hlasové kvality na celkový subjektivní dojem koncového uživatele.

Z hlediska lidského vnímání se tak měření hlasové kvality zaměřuje na kvalitativní parametry jako hlasitost, srozumitelnost, šum, echo, přeslechy, výpadky a další. Nejvýznamnější podíl na degradaci vnímané kvality mají dle testů srozumitelnost, zpoždění a echo [1].

1.2 Poslechové a konverzační testy

Metody měření hlasové kvality lze dle komplexity testů rozlišovat na konverzační a poslechové. Konverzační testy jsou více všestranné a na rozdíl od poslechových testů umožňují do celkového hodnocení kvality přenosu hlasu zahrnout i vliv proměnného zpoždění přenášeného hlasového signálu mezi účastníky, které má nezanedbatelný dopad na vnímanou kvalitu služby uživatelem. Běžný subjektivní konverzační test se sestává z rozhovoru dvou účastníků testu prostřednictvím telekomunikačního spojení na předem definované téma. Testující následně subjektivně hodnotí kvalitu hovoru podle sledovaných kvalitativních kritérií. Jednou z nevýhod konverzačních testů je jejich závislost na prováděné úloze. Vliv zpoždění na subjektivní kvalitativní dojem se totiž značně mění s mírou interaktivity rozhovoru. Mimo to, konverzační testy jsou zdlouhavé, více organizačně náročné a tedy i výrazně nákladnější. V praxi se proto provozují méně často. Metodikou konverzačních testů se zabývají doporučení Mezinárodní telekomunikační unie ITU-T P.800 [2] a P.805 [3].

Poslechové testy jsou založeny na kvalitativním hodnocení hlasových nahrávek. Tato metoda sice nedokáže plně simulovat reálné situace při použití komunikačních systémů, z principu je však jednodušší a v důsledku i častěji používaná. Poslechové testy proto analyzují především věrnost reprodukce přenášeného řečového signálu. Omezení ve vyhodnocení dynamiky rozhovoru jsou vykoupena snazší reprodukovatelností měření, která se následně statisticky vyhodnocují. V důsledku tohoto zjednodušení je ale nutné postupovat při testování podle jasně stanovených pravidel navržených v doporučení ITU-T P.800 [2], jejichž účelem je simulovat parametry reálné komunikace, které jsou jinak v případě konverzačních testů chápány jako přirozené.

1.3 Subjektivní metody

Metodika poslechových i konverzačních testů popsaná v předchozí kapitole zapadá do kategorie subjektivních metod měření. Tyto testy zahrnují skupinu lidských respondentů, kteří subjektivně hodnotí kvalitu simulovaných hovorů. Ačkoliv jsou účastníci testu vybíráni nahodile, měly by splňovat určité podmínky. Například by neměli být odborníky v oboru testování kvality telekomunikačních systémů ani zpracování signálů. Testy probíhají obvykle ve specializovaných laboratořích, které jsou schopny vyhovět náročným požadavkům pro zajištění kredibility měření. Tyto požadavky jsou standardizovány v doporučeních ITU-T P.800 [2], P.830 [4] a P.835 v případě telekomunikačních systémů s potlačením šumu.

Podle doporučení by se nahrávky používané při poslechových testech měly skládat z prostých vět, kterým je jednoduché porozumět, např. z novin či neodborné literatury. Jednotlivé nahrávky by měly obsahovat 2 až 3 věty náhodně seřazené za sebou, kdy každá by měla být v délce 2 až 3 sekund. Důraz na krátké nahrávky je kladen proto, aby se předešlo přirozenému jevu, kdy závěr dlouhé nahrávky hraje větší roli v hodnocení celkové kvality nahrávky. Podobně jsou v doporučeních popsány požadavky na výběr mluvčích a posluchačů, prostředí, systém i procedury pro nahrávání a poslech.

I přes striktní dodržování všech stanovených podmínek souvisejících se subjektivním testováním mohou být ale výsledky ovlivněny aspekty vyplývajícími ze subjektivity hodnocení. Aktuální nálada, věk či pohlaví respondentů patří mezi faktory, které by měly být při celkovém vyhodnocení výsledků brány v potaz.

1.4 Způsoby hodnocení kvality vzorků

Doporučení dále stanovují, jakým způsobem by měli respondenti hodnotit kvalitu vzorků a následnou metodu vyhodnocení výsledků. Nejčastěji používaným způso-

bem hodnocení je stupnice MOS (Mean Opinion Score). MOS je definován jako průměr číselných hodnot vyjadřujících hodnocení jednotlivých účastníků testu na předdefinované stupnici výkonosti přenosového řetězce [5]. Doporučení definuje více druhů stupnic MOS v závislosti na druhu prováděného měření. Tabulka 1.1 uvádí nejpoužívanější stupnici MOS určenou pro poslechové ACR (Absolute Category Rating) testy. Poslechová kvalita je vyjádřena diskrétní hodnotou v číselném rozsahu od 1 do 5, kde 5 odpovídá nejvyšší možné kvalitě.

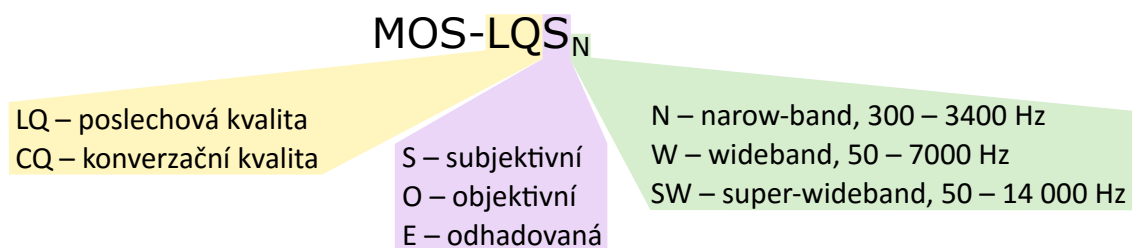
Tab. 1.1: Hodnoty stupnice MOS pro poslechové ACR a DCR testy [2]

Kvalita řeči (MOS)	Degradace (DMOS)	Hodnocení
Vynikající	Neznamenatelná	5
Dobrá	Znamenatelná, ale neobtěžující	4
Přijatelná	Mírně obtěžující	3
Špatná	Obtěžující	2
Velmi špatná	Velmi obtěžující	1

Mimo ACR testů, hodnotících celkovou kvalitu nahrávky, obsahuje tabulka 1.1 také definice hodnocení DCR (Degradation Category Rating) testů, u kterých účastníci namísto celkové kvality hodnotí míru degradace hlasové kvality v přímém srovnání s původní hlasovým záznamem. Tato metoda je speciálně určená pro komunikační systémy dosahující vysoké kvality. Subjektivní hodnocení jsou totiž při přímém srovnání více citlivé i na malé rušivé efekty. Opět se používá stupnice hodnocení 1 – 5, v tomto případě označovaná DMOS (Degradation MOS).

Pro dosažení statisticky významných výsledků je třeba pro každé testované podmínky zajistit dostatečně velký počet vzorků. Výsledná hodnota MOS odpovídající posuzovaným parametrům komunikačního řetězce je získána výpočtem průměru z jednotlivých diskrétních hodnocení. Při zpracování výsledků je třeba mít na paměti, že se jedná o subjektivní hodnocení, a proto se například hodnocení mužských a ženských mluvčích mohou významně lišit.

Mimo poslechových subjektivních testů se hodnocení MOS používá také pro konverzační testy a objektivní metody měření (popsány v kap. 1.5). Doporučení ITU-T P.800.1 [5] proto stanovuje jasnou terminologii v souvislosti s MOS, která umožňuje výsledky rozdílných druhů testů spolehlivě rozlišit. Na obrázku 1.1 je graficky znázorněn stanovený formát koncovky popisující druh testu spolu s používanou frekvenční šířkou pásma přenášeného hlasového signálu.



Obr. 1.1: Časté zkratky používané spolu s MOS

1.5 Objektivní metody

Značnou nevýhodou subjektivních metod měření hlasové kvality je jejich organizační, finanční a časová náročnost, která znemožňuje jejich nasazení pro monitorování telekomunikačních sítí. Proto byly vyvinuty výpočetní algoritmy, které typicky na základě analýzy přeneseného hlasového signálu umožňují odhadnout výsledný subjektivní dojem průměrného uživatele. Výstupní hodnocení kvality výpočetních modelů je stejně jako u subjektivních měření ve formě hodnoty na stupnici MOS, případně jiné stupnice, která má s MOS definovanou převodní charakteristiku.

Úspěšnost objektivních metod se obvykle posuzuje dle schopnosti aproximovat výsledky subjektivních testů, uvažovaných jako referenční, při co nejširším rozsahu testovaných situací. Nejnovější výpočetní algoritmy již dosahují velmi dobrých výsledků, tj. vysoké korelace se subjektivními poslechovými testy, při většině testovaných podmínek. Současné modely posuzují hlasovou kvalitu bez ohledu na parametry související se subjektivním vnímáním reálné řečové komunikace (dynamika rozhovoru, kontext, očekávání, osobní preference).

Objektivní metody lze rozdělit dle několika aspektů. Podle vstupních hodnot se dělí na signálově založené, analyzující přímo přenášené řečové signály a parametrické, zaměřující se na technické parametry přenosového systému. Signálově založené algoritmy se v závislosti na dostupných informacích na jejich vstupu člení na tzv. referenční (FR) a bez reference (NR). Prvně zmíněné mají k dispozici k vzájemnému srovnání řečové signály ze vstupu i výstupu komunikačního systému, zatímco druhá skupina nemá k informacím o referenčním signálu přístup. Tyto 2 kategorie ve většině případů korespondují s dělením na intrusivní a neintrusivní metody měření.

1.5.1 Intrusivní metody

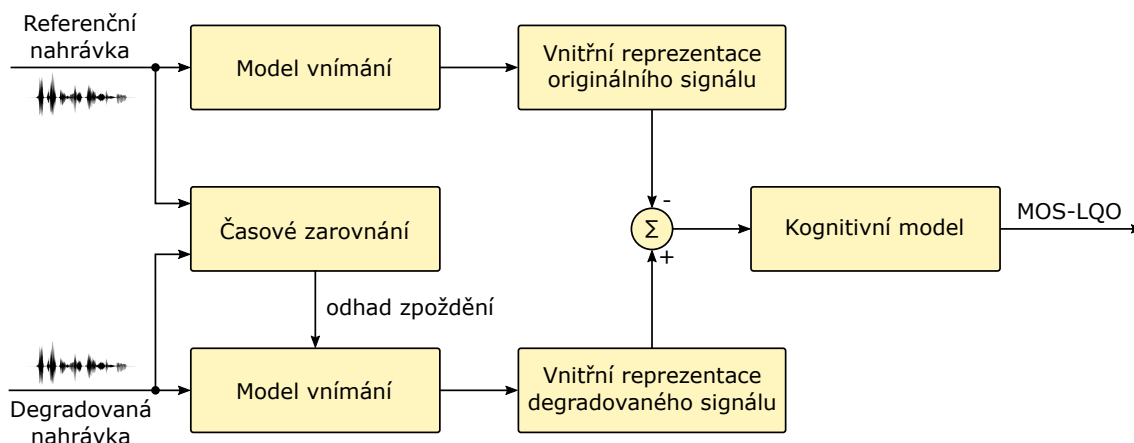
Intrusivní metody pro vyhodnocení kvality generují umělé telefonní hovory prostřednictvím oboustranného měřicího systému. Většinou při vyhodnocování hlasové kvality využívají FR přístup, vychází tedy z přímého srovnání původního a přeneseného hlasového signálu. V důsledku toho dosahují bližších výsledků vzhledem

k subjektivním testům než neintrusivní metody [6]. Pro matematický popis lidského vnímání zvuku využívají psychoakustických modelů.

Algoritmů pro posuzování kvality přenosu hlasu intrusivní metodou se v průběhu vývoje objevila celá řada. V roce 1996 standardizovala ITU výpočetní model P.861 PSQM, určený pro úzkopásmové telefonní systémy (NB, 300 – 3400 Hz). PSQM se vyznačoval pokročilým modelem lidského vnímání hlasu. Zanedlouho ale vyšlo najevo, že PSQM má vážné nedostatky v oblasti časové synchronizace a při zpracování hlasových nahrávek s výrazným šumovým pozadím. Alternativou byl proprietární algoritmus PAMS, který nejprve referenční a přenesený signál podroboval amplitudové a časové korelaci po krátkých dílčích blocích.

ITU-T P.862 PESQ

Některé nové prvky využití v PAMS se staly základem doporučení P.862 [7], vydaném v roce 2001. Algoritmus Perceptual evaluation of speech quality, zkráceně PESQ, nahrazoval PSQM a řadu let byl nejpoužívanějším zástupcem intrusivních metod měření kvality přenosu hlasu. Jako ostatní intrusivní modely je založen na srovnání přeneseného hlasového signálu na výstupu komunikačního řetězce s referenčním signálem na vstupu. Výstupní hodnocení, ve formě MOS-LQOn, je predikcí vnímané kvality, kterou by respondenti ohodnotili přenesený hlasový signál při subjektivním poslechovém testu. Postup srovnání signálů je ilustrován na obrázku 1.2.



Obr. 1.2: Zjednodušené blokové schéma algoritmu PESQ [7]

Oba vstupní hlasové signály jsou nejprve podrobeny amplitudové ekvalizaci pro následné zpracování, přičemž chyby způsobené příliš vysokou, či nízkou hlasitostí zůstávají zachovány. Algoritmus PESQ je postaven na vzájemném srovnání spektrálních charakteristik referenčního a přeneseného signálu po velmi krátkých časových blocích, proto je vzájemné časové zarovnání odpovídajících si úseků obou signálů

kritickou součástí. Blok časového zarovnání využívá rekurentní korelace. Nejdříve se odhadne zpoždění celého vzorku. Ten se následně rozdělí na více kratších úseků, pro které se znovu pomocí korelace vypočítá zpoždění, což se opakuje, dokud nejsou jednotlivé segmenty příliš krátké nebo již nedochází ke zlepšení korelace.

Amplitudově a časově korelované signály jsou poté vstupem pro psychoakustický model lidského vnímání. Vzorek je rozdělen Hanningovým oknem na segmenty délky 32 ms s 50% překryvem (tzn. úseky 16 ms při úvaze překryvu) a pro každý segment je vypočítáno frekvenční spektrum prostřednictvím 256 bodové FFT. Reprezentace signálů v transformované doméně je analogická způsobu vnímání zvuku lidských sluchovým ústrojím. Výsledky FFT se rozdělí do 17 kritických frekvenčních pásem slyšení pro posouzení frekvenčního maskování a pro každé z nich se vypočte suma energie. Tato suma je následně převedena zpět na akustickou hlasitost a vážena křivkou hlasitosti popisující frekvenční závislost citlivosti lidského sluchu.

Na výstupu modelu vnímání jsou oba signály reprezentovány pro každý úsek 16 ms vektorem 17 hodnot, tvořící dvojrozměrnou matici. Matice originálního a přeneseného signálu se vzájemně srovnávají. Jednotlivé rozdíly se hodnotí podle způsobu lidského vnímání, lidský sluch je totiž více citlivý na přidání zkreslení než na chybějící části signálu [6]). Celková suma rozdílů určuje interní hodnocení MOS v rozsahu $-0,5$ až $4,5$, které se přepočítává dle mapovací funkce aproximující hodnocení kvality lidských respondentů stanovené v doporučení P.862.1 [8] na běžnější formu MOS-LQO_N v rozsahu od 1 do 4,5.

V roce 2005 bylo vydáno rozšíření WB-PESQ [9] přidávající podporu pro systémy využívající širokopásmovou (WB) telefonii ve frekvenčním rozsahu 50 – 7000 Hz.

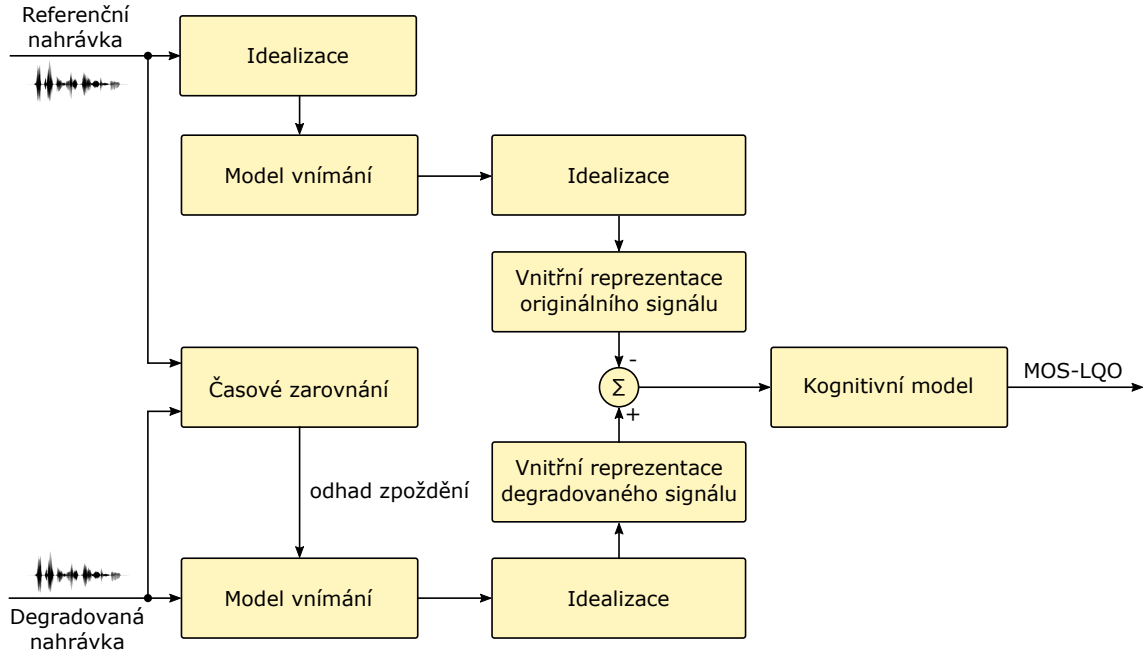
ITU-T P.863 POLQA

S rozvojem pokročilejších technik kódování hlasového signálu a očekávaným rozšířením použití technologie HD Voice ale vyvstala potřeba pro zásadnější modernizaci modelu. Po více než 4 letech výzkumu, na počátku roku 2011, byla oficiálně vydána třetí generace objektivního měřicího algoritmu pod označením Perceptual objective listening quality assessment, zkráceně POLQA [10, 11, 12, 13].

Ve srovnání s předchozí generací se nejednalo o revoluci. Algoritmus POLQA staví na základech, které již byly definovány v PESQ, o čemž vypovídá i vzájemná podoba blokových schémat obou algoritmů na obrázcích 1.2 a 1.3. Hlavními cíli bylo rozšířit sféru použití a adresovat oblasti, na které nebyl zastarávající algoritmus PESQ připraven nebo v nich nepodával spolehlivé výsledky:

- nové hlasové kodeky,
- vyhodnocení úrovně hlasitosti,
- rozdíly vzorkovacího kmitočtu na přijímací a vysílací straně,

- Dynamic Time Warping (DTW),
- Voice Quality Enhancement (VQE) systémy pro vylepšení kvality hlasu, tj. odšumění, odstranění echa, zvýraznění řeči, umělé rozšíření spektra;
- úpravy přenášené šířky pásma hlasového signálu, podpora SWB.



Obr. 1.3: Zjednodušené blokové schéma algoritmu POLQA [13]

POLQA zachová přístup, podle kterého se v algoritmu zpracovává pouze referenční a přenesený hlasový signál. Žádné vedlejší informace o přenosovém systému, jako používaný hlasový kodek, nejsou využity. V závislosti na tom, jaký typ referenčního signálu je použit, rozlišuje POLQA dva operační módy uvedené v tabulce 1.2.

Tab. 1.2: Operační módy algoritmu ITU-T P.863 POLQA [11]

Zkratka	Operační mód	Šířka pásma [Hz]	Popis
NB	Narrow-band	300 – 3400	Zpětně kompatibilní s PESQ
SWB	Super-wideband	50 – 14 000	Ref. řečový signál frekv. neomezen

V případě úzkopásmového (NB) módu by měla do přenosového řetězce být vložena úzkopásmová reference, pro testování sítí s podporou WB a SWB kodeků musí být referenční hlasový signál SWB frekvenčního rozsahu. Volba operačního módu přímo ovlivňuje dosažitelné výsledky predikce subjektivní kvality přenosu hlasu MOS-LQO. Tabulka 1.3 obsahuje průměry dosažitelných hodnot MOS v obou módech pro vybrané techniky kódování hlasu při přenosu získaných na základě testů

popsaných v [11, 12]. Je evidentní, že výsledky MOS-LQO_N a MOS-LQO_{SW} není možné vzhledem k odlišné referenci vzájemně srovnávat.

Tab. 1.3: Dosažitelné skóre MOS-LQO vybraných hlasových kodeků [11, 12]

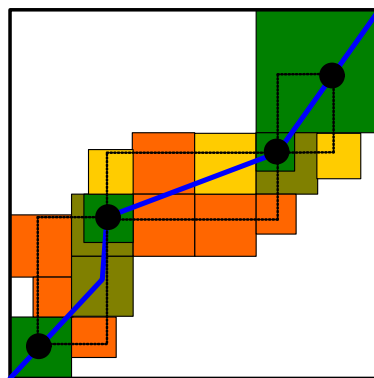
Technika kódování hlasu při přenosu	POLQA NB	POLQA SWB
Transparentní přenos 50 – 14 000 Hz	-	4,75
Transparentní přenos 50 – 7000 Hz	-	4,3
AMR-WB 23.85 kbit/s	-	4,0
AMR-WB 12.65 kbit/s	-	3,9
Transparentní přenos 50 – 3800 Hz	4,5	3,7
AMR 12.2 kbit/s	4,2	3,3
AMR 7.95 kbit/s	3,9	3,1

Zásadním přepracováním ve srovnání s algoritmem PESQ prošel blok časového zarovnání. Mezi problematické oblasti časové synchronizace patří systémy s asynchronními zařízeními (pracující s rozdílnými hodinovými kmitočty), u kterých se vzájemné zpoždění signálů neustále mění. Ještě náročnější v tomto ohledu jsou sítě založené na bázi protokolu IP, vyznačující se proměnlivým zpožděním jednotlivých bloků, způsobující problémy na přijímací straně. Hlasové kodeky podporující DTW dokáží tento efekt kompenzovat adaptivním zpomalováním a zrychlováním řeči i pauz v promluvě, někdy dokonce přesouváním celých bloků v čase, aniž by to příjemce upozoroval. DTW ale znesnadňuje vzájemné srovnání přeneseného signálu s referenčním.

Při časovém zarovnání algoritmus nejdříve rozpoznává části přeneseného signálu s konstantním zpožděním a označí je jako tzv. mezníky (angl. landmarks). Mezníky slouží při následném zpracování oblastí mezi nimi jako výchozí body. Jednotlivé oblasti se postupně dělí na menší a menší intervaly a pro každý z nich je stanovena odpovídající část referenčního signálu a míra spolehlivosti jejich přiřazení. Pro méně spolehlivá přiřazení se analyzuje širší úsek signálu a naopak. Na principu Viterbiho algoritmu se pak prostřednictvím informací o pozici mezníků a spolehlivosti intervalů mezi nimi určí nejpravděpodobnější konfigurace přiřazení jednotlivých částí k referenčnímu signálu. Tento proces je ilustrován na obrázku 1.4.

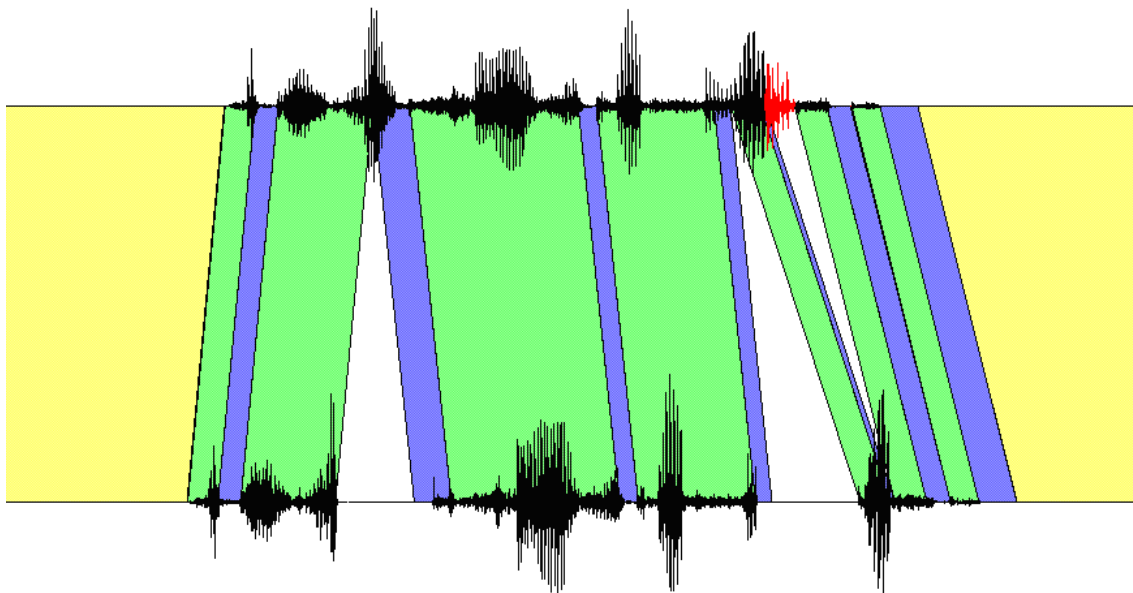
Výstupem je tabulka obsahující časy začátku a konce jednotlivých odpovídajících si úseků v referenčním a přeneseném signálu. Označeny jsou také části, které nemají svou obdobu v referenčním signálu i úseky, které v přeneseném signálu chybí.

Grafický příklad výstupu bloku časového zarovnání je na obrázku 1.5. V horní části je zobrazen referenční signál a v dolní části je přenesený signál. Zelené oblasti spojují úseky určené s vysokou mírou spolehlivosti, modře spojené úseky byly určeny



Obr. 1.4: Ilustrace přiřazení úseků nahrávky a volba ideální konfigurace [12]

se spolehlivostí nižší. Červená část referenčního signálu byla ztracena při přenosu. Nepřiřazené bílé úseky se využívají pro posouzení nepřijemnosti šumového pozadí v přeneseném signálu.



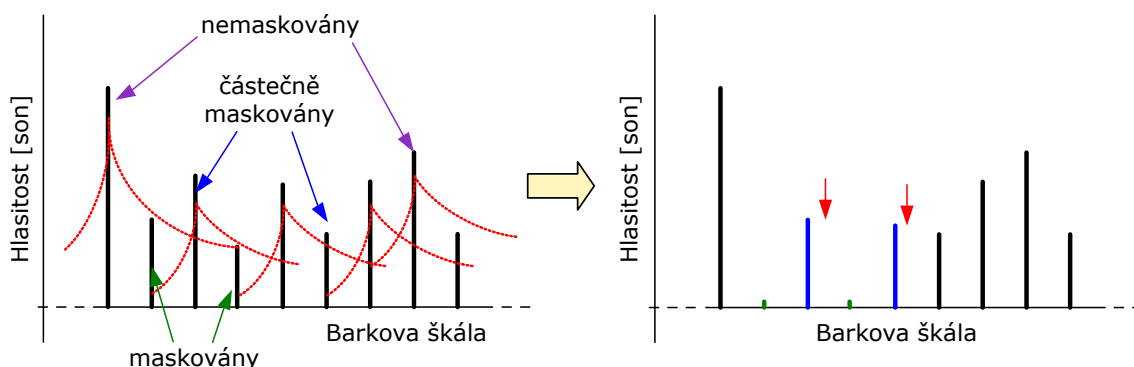
Obr. 1.5: Příklad zarovnání signálů algoritmem POLQA [12]

Vylepšeními prošel i model lidského vnímání. Běžně používané psychoakustické modely v multimediálních kodecích nebo algoritmu PESQ nerozlišují mezi zdroji rušení, což ale dle nových výzkumů v případě některých typů zkreslení není v souladu s lidským vnímáním řeči [12]. Mezi problematické zdroje degradace signálu patří výrazný šum na pozadí, dozvuk, echo a spektrální zkreslení. POLQA tato rušení na vstupu modelu vnímání kompenzuje, přičemž jejich vliv je zohledněn separátně až při závěrečném výpočtu MOS.

Naprosto novým prvkem algoritmu POLQA je koncept idealizace referenčního

signálu. Cílem je korekce mírných zkreslení, rušivých vlivů a vylepšení zabarvení hlasu k zdokonalení referenčního signálu. Důvodem je přístup hodnocení posluchačů u subjektivních ACR testů. Ti totiž nesrovnávají přímo referenční a degradovaný řečový signál, ale degradovaný signál s jejich představou toho, jak referenční signál má správně znít. Idealizace tak při testech, ve kterých je referenční signál mírně zašuměn a je transparentně přenesen na přijímací stranu, způsobí nižší výsledné hodnocení MOS, než by tomu bylo v případě přímého srovnání referenčního a přeneseného signálu.

Změny v modelu lidského vnímání proběhly i ve způsobu vyhodnocení rušení ve spektrální oblasti. Algoritmus PESQ využívá tzv. rozmazané spektrum, které bere v úvahu pouze rozdíly ve spektru zahrnujícím i maskovací úrovně. POLQA se naopak zaměřuje na vnímání nemaskovaných rozdílů ve spektru, maskovací úrovně slouží pouze k vážení jednotlivých frekvenčních složek. Způsob zpracování je znázorněn na obrázku 1.6. Plně maskované složky jsou odebrány, částečně maskované složky jsou redukovány a nemaskované části spektra zůstávají zachovány. Výsledkem je tzv. zostřené spektrum, jehož výhodou je vyšší spektrální rozlišení pro následnou analýzu jemných rozdílů.



Obr. 1.6: Výpočet modifikovaných kritických pásem algoritmem POLQA [12]

Výše popsané spektrální srovnání probíhá v krátkých časových oknech. Pro výpočet MOS se nejprve vliv spektrálních rozdílů obou signálů integruje, podobně jako v PESQ, a následně se provede vážení výsledné hodnoty údaji o echu, dozvucích, spektrálním tvarování a šumu odebraných před zpracováním signálů.

1.5.2 Neintrusivní metody

Neintrusivní metody jsou založeny na jednostranném pasivním monitorování probíhajícího spojení a na rozdíl od intrusivních metod obvykle při vyhodnocení nevyužívají srovnání s referenčním hlasovým signálem ze vstupu komunikačního řetězce. Proto

také neintrusivní metody obecně dosahují nižšího korelačního koeficientu se subjektivními testy, neboť analýza přeneseného signálu je založena na předpokladech o signálu referenčním. Přesto jsou neintrusivní metody zajímavou možností pasivního monitorování hlasové kvality v komunikačních sítích bez nutnosti generování umělých hovorů speciálně určených pro účely měření.

Neintrusivní metody se dělí dle druhu použitých parametrů pro vyhodnocení kvality do dvou kategorií na signálově založené, podobně jako algoritmy PESQ a POLQA, a parametrické. Parametrické metody vycházejí z plánovacích modelů, jejichž cílem je predikce poskytované hlasové kvality na základě technických parametrů sítě. Výhodou těchto modelů je relativně snadná možnost zahrnutí efektů zpoždění a echa pro predikci konverzační kvality namísto obvykle používané poslechové kvality.

Nejvýraznějším zástupcem parametrických neintrusivních metod je E-model, definovaný doporučením ITU-T G.107 [14], který je podrobně rozebrán v kapitole 2. Dalšími příklady jsou standardy P.561 (INMD) [15] a P.562 (CCI) [16]. INMD stanovuje soubor parametrů, které je třeba pro predikci kvality vyhodnocovat. Patří sem odstup signálu od šumu, úroveň hlasitosti, řečová aktivita, echo a další. CCI tento koncept rozšiřuje návodem, jak soubor těchto parametrů analyzovat za účelem získání jednoduchého ukazatele kvality.

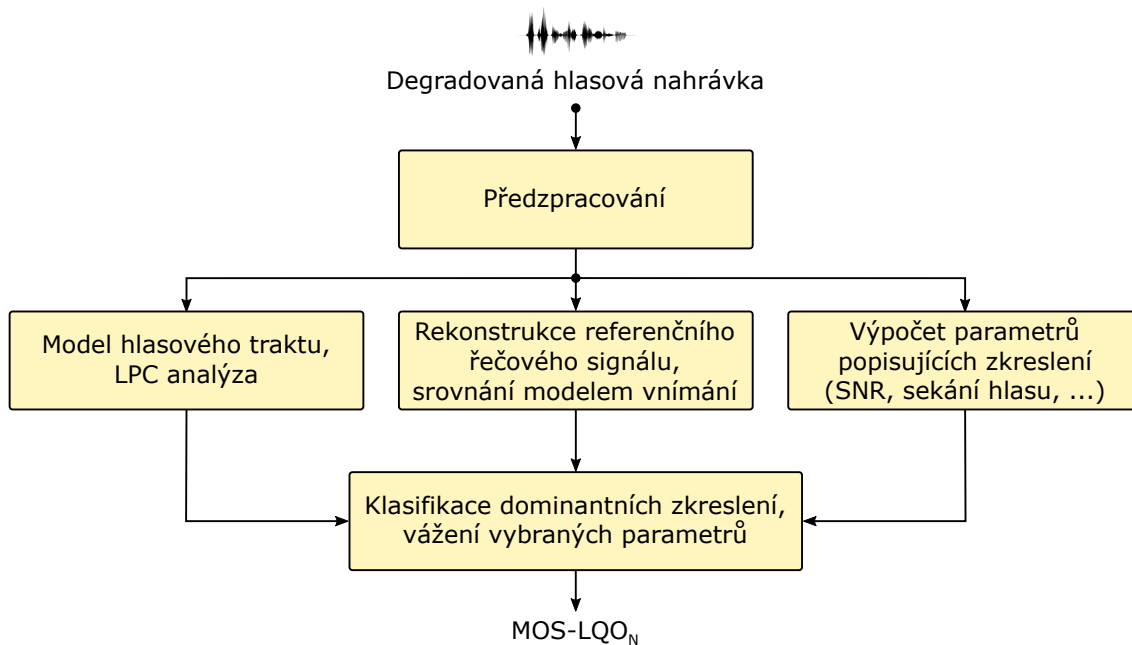
ITU-T P.563 3SQM

V roce 2004 byl ITU-T definován standard P.563 (3SQM [17]), zapadající do kategorie signálově založených objektivních metod. Je určen pro jednostrannou predikci subjektivní hlasové kvality v úzkopásmových telefonních sítích, které jsou ovlivněny šumovým pozadím, filtrací, proměnným zpožděním a zkresleními vnesenými hlasovými kodeky a chybami kanálu [18].

Na obrázku 1.7 je znázorněna koncepce algoritmu. 3SQM se sestává z 3 samostatných částí, přičemž každá používá jinou metodu výpočtu MOS. Nejprve je však degradovaný hlasový signál amplitudově normalizován a dvakrát zkopírován s využitím rozdílných frekvenčních filtrů pro adaptaci na různorodé možnosti použití. Dále jsou pomocí VAD na základě výkonu hlasového signálu určeny pasáže promluvy.

Po předzpracování je přenesený hlasový signál analyzován 3 různými technikami. První metoda spočívá v použití precizního LPC modelu lidského hlasového traktu jako syntezátoru pro reprodukci degradovaného hlasového signálu. Tato reprodukce je srovnána s přijatým degradovaným signálem a vzájemné rozdíly, identifikované jako nepřirozené lidské řeči, jsou interpretovány jako zkreslení způsobená přenosem v komunikačním řetězci. Míra těchto zkreslení určuje predikci MOS.

Druhý přístup se snaží pomocí opravných funkcí rekonstruovat z degradovaného signálu signál referenční. Oba dva signály jsou pak vstupem zjednodušeného algo-



Obr. 1.7: Blokové schéma algoritmu P.563 (3SQM) [18]

ritmu na bázi PESQ (bez časového zarovnání), který oba signály podrobí modelu lidského vnímání a na základě rozdílů vypočítá predikci MOS kvality. Algoritmus je blíže popsán v kapitole 1.5.1 na straně 16.

Třetí technikou vyhodnocení kvality je identifikace charakteristických druhů rušení při hlasovém přenosu, jako jsou krátká přerušení, ořezání řeči, robotizace hlasu a šum. Hodnoty parametrů získaných analýzou těchto rušení jsou následně využity k výpočtu MOS.

Na závěr zpracování jsou všechny kritéria vzájemně srovnány, nejvzdálenější odhady se zahodí a výsledné MOS-LQO_N algoritmu se získá průměrováním zbylých dílčích výsledků [6].

2 E-MODEL A JEHO VYUŽITÍ PRO MĚŘENÍ HLASOVÉ KVALITY HOVORU

E-model se řadí mezi parametrické modely, které vnímanou kvalitu stanovují na základě definovaného matematického vztahu mezi charakteristikami jednotlivých elementů sítě a celkovou kvalitou přenosové cesty. Právě parametrických přístupů umožňuje jeho primární užití jako nástroje pro plánování telekomunikačních sítí.

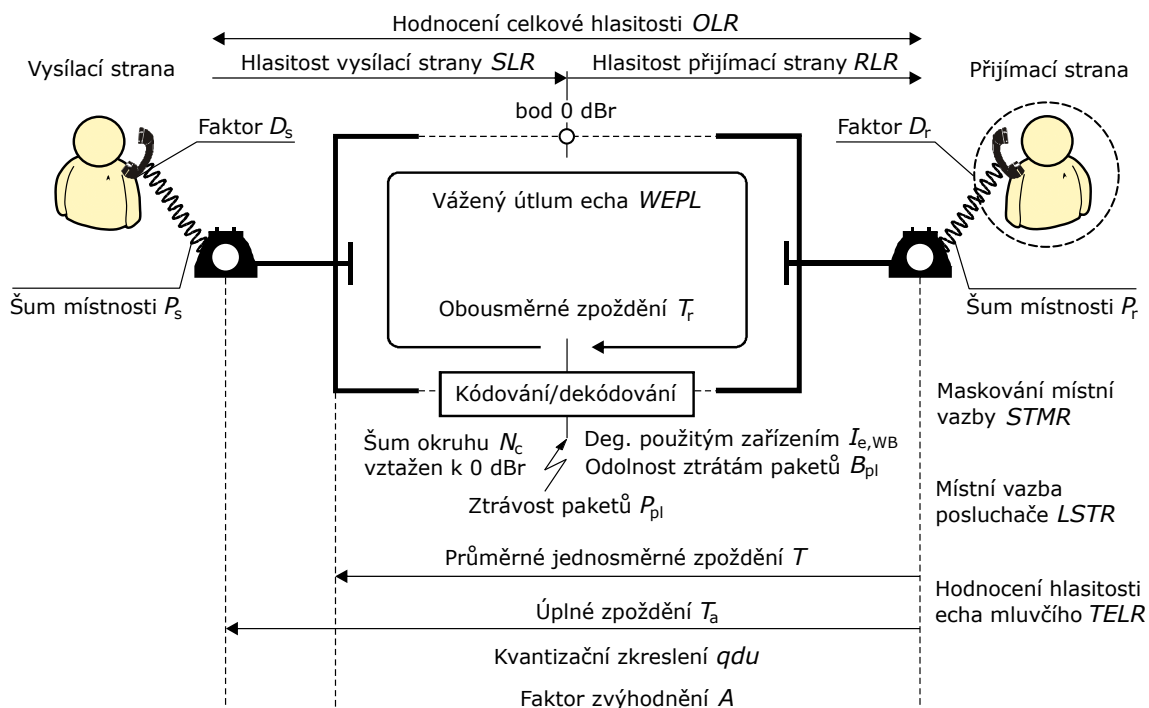
Návrh E-modelu byl výsledkem výzkumného týmu švédského experta Nils-Olof Johannessona v 90. letech. Prvně byl publikován v ETSI TR 250 [19, 1996], z čehož pramení také označení E-model (ETSI model). Nedlouho nato byl model přijat Mezinárodní telekomunikační unií a standardizován v doporučení ITU-T G.107 [14, 1998].

Johannessonova idea, na které se E-model zakládá, vychází do značné míry z parametrických modelů vyvinutých v 70. a 80. letech pro tehdejší analogové komunikační sítě, zejména modelů BcTR a OPINE [20]. Základem OPINE modelu je předpoklad, že „psychologické faktory jsou na psychologické stupnici aditivní“ a jednotlivé rušivé zdroje při přenosu mohou být do těchto psychologických faktorů transformovány. Kombinací jednotlivých skalárních přínosů psychologických faktorů vzniká výsledná predikce kvality komunikační sítě na skalární R-škále s jasně definovaným minimem a maximem.

Původní verze E-modelu se později ukázala jako jednoduchá a nepřesná, což vyústilo v řadu vylepšení a rozšíření umožňujících jeho využití pro plánování moderních komunikačních systémů. V roce 2011 bylo standardizováno samostatné doporučení ITU-T G.107.1 [21], které zavádí analogickou verzi E-modelu pro širokopásmové aplikace (WB, 50 – 7000 Hz).

2.1 Vstupní parametry

E-model předpovídá vnímanou konverzační kvalitu na základě sady parametrů popisujících komunikační systém od úst mluvčího po ucho posluchače (Mouth-to-Ear, M2E). Ve svém plném rozsahu se jedná o velmi komplexní nástroj pro široké spektrum komunikačních přístupů. Standard definuje celkem 21 vstupních parametrů reprezentujících vlastnosti komunikačního řetězce na transportní trase i na přijímací a vysílací straně (tj. komunikační zařízení, vlastnosti sítě, komunikačního kanálu a prostředí). Soubor parametrů je schématicky znázorněn na obrázku 2.1. Tabulka 2.1 na straně 27 obsahuje jejich celkový seznam s příslušnými jednotkami, výchozími hodnotami a povolenými rozsahy.



Obr. 2.1: Parametrický popis E-modelu [21]

Tento komplexní přístup umožňuje použití teoreticky pro jakýkoliv telefonní systém, tj. se spínanými okruhy, s přepínáním paketů; kabelové i mobilní. Při predikci vnímané kvality ve VoIP sítích se však často používá zjednodušená verze, která neanalyzuje celý přenosový kanál M2E, ale pro většinu vstupních proměnných uvažuje výchozí hodnoty a vyhodnocení probíhá jen na základě několika nejvýznamnějších parametrů jako komunikační zpoždění, ztrátovost paketů a typ použitého hlasového kodeku [22]. Stejný přístup je zvolen i při návrhu algoritmu pro měření hlasové kvality technologií VoLTE a VoWiFi v kapitole 4.

2.2 R-faktor

Výstupem E-modelu je tzv. R-faktor. Jak již bylo popsáno v úvodu této kapitoly, E-model se zakládá na aditivním vyjádření jednotlivých rušivých zdrojů formou skalárních parametrů, které dohromady dávají výsledné hodnocení konverzační kvality sítě na lineární R-škále. R-faktor může nabývat hodnot od 0 do 100, přičemž s rostoucí hodnotou R monotónně roste i odhadovaná vnímaná kvalita průměrným uživatelem.

Výpočet R-faktoru je definován základním vztahem [14]:

$$R = R_0 - I_s - I_d - I_{e,eff} + A, \quad (2.1)$$

kde $R \in \langle 0; R_{0,max} + A \rangle$. R_0 představuje poměr signálu k šumu, zahrnuje SNR v komunikačním kanále i šum okolního prostředí. I_s reprezentuje rušení související s hlasovým signálem, tj. neoptimální hlasitost, místní vazba, příp. kvantizační zkreslení. I_d seskupuje faktory související s přenosovým zpožděním, I_e reprezentuje zkreslení plynoucí z používaného hlasového kodeku a je zahrnuto do rozšířeného parametru $I_{e,eff}$ zahrnující i vliv ztrát paketů ve VoIP. Faktor zvýhodnění A představuje, o kolik bude průměrný uživatel očekávat zhoršení konverzační kvality hlasové služby na základě určitého zvýhodnění, např. při použití mobilních sítí v obtížně přístupných lokalitách. Jednotlivé faktory jsou podrobněji představeny v kapitole 2.4.

Faktor zvýhodnění A je závislý na kontextu využití E-modelu, při měření hlasové kvality se obvykle nepoužívá, a proto je možné jej v těchto případech zcela vypustit. Rovnici 2.1 lze následně pro větší názornost uvádět ve formě

$$R = R_0 - I_{tot}, \quad (2.2)$$

kde I_{tot} představuje souhrn všech zdrojů rušení, které výslednou hodnotu R-faktoru snižují, tedy

$$I_{tot} = I_s + I_d + I_{e,eff} = R_0 - R. \quad (2.3)$$

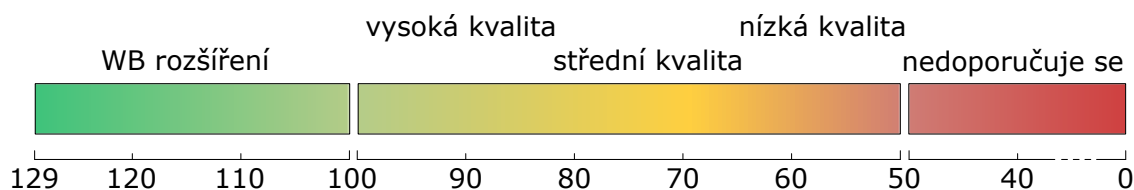
Hodnota R_0 omezuje nejvyšší dosažitelnou hodnotu R shora. Právě tento parametr zahrnuje i vliv využití šířky pásma při přenosu hlasového signálu. Pro NB platí, že $R_{0,max,NB} = 93,2$.

2.3 Wideband E-model

Paralelně s doporučením G.107 [14], které upravuje pravidla pro využití E-modelu pro plánování NB telefonních systémů s vzorkovací frekvencí 8 kHz, platí od roku 2011 samostatné doporučení G.107.1 [21] pro WB systémy s vzorkovací frekvencí 16 kHz.

Využití větší přenosové šířky pásma hlasového signálu má při transparentním spojení za následek nárůst vnímané uživatelské kvality. Proto musela být pro WB kontext lineární R-škála vnímané kvality rozšířena nad dosavadní teoretický limit $R \leq 100$. Na základě testů popsanych v [21] bylo zjištěno, že přechodem od NB k WB referenci je lineární škála kvality rozšířena o přibližně 29 %. Doporučení G.107.1 tak zavádí $R_{0,max,WB} = 129$, přičemž NB část R-stupnice zůstala zachována v nedotčené podobě a Wideband E-model je možno využít i ve smíšeném NB/WB kontextu. Jednotlivé části stupnice v závislosti na odhadované kvalitě sítě jsou graficky znázorněny na obrázku 2.2.

Na základě subjektivních testů provedených v laboratořích Deutsche Telekom [23]



Obr. 2.2: Lineární škála R-faktoru dle G.107.1 s příslušnými kategoriemi kvality

bylo již také navrženo rozšíření škály R-faktoru pro SWB (50 – 14 000 Hz) přenosy na rozsah od 0 do 179 analogicky s WB rozšířením. Tato škála by umožňovala vzájemné srovnání přenosové kvality a zkreslení v NB, WB i SWB systémech. Pro adaptaci ITU-T je však zapotřebí dalších konfirmačních testů.

Tab. 2.1: Soubor parametrů E-modelu a jejich vlastností [21]

Parametr	Zkratka	Jednotka	Hodnota	Rozsah
Hlasitost vysílací strany	SLR	dB	+8	$\langle 0; +18 \rangle$
Hlasitost přijímací strany	RLR	dB	+2	$\langle -5; +14 \rangle$
Celková hlasitost	OLR	dB	+10	$\langle -5; +32 \rangle$
Maskování místní vazby	$STMR$	dB	15	$\langle 10; 20 \rangle$
Místní vazba posluchače	$LSTR$	dB	18	$\langle 13; 23 \rangle$
Hodnota D telefonu, vysílací strana	D_s	-	3	$\langle -3; +3 \rangle$
Hodnota D telefonu, přijímací strana	D_r	-	3	$\langle -3; +3 \rangle$
Hodnocení hlasitosti echa mluvčího	$TELR$	dB	65	$\langle 5; 65 \rangle$
Vážený útlum echa	$WEPL$	ms	110	$\langle 5; 110 \rangle$
Průměrné jednosměrné zpoždění echa	T	ms	0	$\langle 0; 500 \rangle$
Obousměrné zpoždění ve smyčce	T_r	ms	0	$\langle 0; 1000 \rangle$
Úplné zpoždění	T_a	-	0	$\langle 0; 500 \rangle$
Počet „jednotek kvant. zkraslení“	qdu	-	1	$\langle 1; 14 \rangle$
Faktor degradace použitým zařízením	$I_{e,WB}$	-	0	$\langle 0; 56 \rangle$
Odolnost proti ztrátovosti paketů	B_{pl}	-	4,3	$\langle 4,3; 7,3 \rangle$
Ztrátovost paketů	P_{pl}	%	0	$\langle 0; 20 \rangle$
Šum v okruhu vztažený k bodu 0 dB _r	N_c	dBm _{0p}	-70	$\langle -80; -40 \rangle$
Šumové pozadí na přijímací straně	N_{for}	dB _m p	-96	-
Šum místnosti na vysílací straně	P_s	dB(A)	35	$\langle 35; 85 \rangle$
Šum místnosti na přijímací straně	P_r	dB(A)	35	$\langle 35; 85 \rangle$
Faktor zvýhodnění	A	-	0	$\langle 0; 20 \rangle$

2.4 Faktory degradace

V případě Wideband E-modelu uvádí doporučení G.107.1 [21] rovnici 2.1 ve formě

$$R = R_{0,WB} - I_{s,WB} - I_{d,WB} - I_{e,eff,WB} + A. \quad (2.4)$$

2.4.1 Poměr signálu k šumu, $R_{0,WB}$

Poměr signálu k šumu zahrnuje šum v komunikačním obvodu i šum místnosti na přijímací a vysílací straně. Je funkcí SLR , RLR , $LSTR$, D_s , N_c , N_{for} , P_s a P_r . Základní vztah pro výpočet $R_{0,WB}$ je definován [21]:

$$R_{0,WB} = 20 - 1,5(N_{0,WB} + SLR), \quad (2.5)$$

kde $N_{0,WB}$ [dBm0p] představuje součet příspěvků jednotlivých zdrojů šumu vtažených k bodu 0 dBr:

$$N_{0,WB} = 10 \log \left(10^{\frac{N_c}{10}} + 10^{\frac{N_{os,WB}}{10}} + 10^{\frac{N_{or}}{10}} + 10^{\frac{N_{fo,WB}}{10}} \right). \quad (2.6)$$

N_c je sumou výkonů všech šumů v okruhu. $N_{os,WB}$ a N_{or} jsou ekvivalenty šumu místnosti na vysílací a přijímací straně a $N_{fo,WB}$ reprezentuje šumové pozadí na přijímací straně. Zbylé výpočty jsou uvedeny v rovnicích 2.7–2.10.

$$N_{os,WB} = P_s - SLR - D_s - 97, \quad (2.7)$$

$$N_{or} = RLR - 121 + P_{re} + 0,008(P_{re} - 35)^2, \quad (2.8)$$

$$P_{re} = P_r + 10 \log \left[1 + 10^{\frac{10-LSTR}{10}} \right], \quad (2.9)$$

$$N_{fo,WB} = N_{for,WB} + RLR. \quad (2.10)$$

V případě transparentního přenosu nebo při zanedbání vlivu šumu v obvodu a v okolním prostředí lze za všechny výše zmíněné parametry dosadit jejich výchozí hodnoty dle tabulky 2.1, potom

$$R_0 = R_{0,max}, \quad (2.11)$$

kde $R_{0,max,WB} = 129$, případně $R_{0,max,NB} = 93,2$ pro NB E-model. Této vlastnosti bylo využito i při návrhu algoritmu měření hlasové kvality v kapitole 4. Dosazením do rovnice 2.2 lze zároveň také odvodit rozdílovou hodnotu I_{tot} mezi NB a WB kontextem:

$$I_{tot} = R_{0,max,WB} - R_{0,max,NB} = 129 - 93,2 = 35,8. \quad (2.12)$$

2.4.2 Zkreslení synchronní se signálem, $I_{s,WB}$

I_s představuje aditivní vliv zkreslení, které nastávají synchronně s přenosem hlasového signálu. Je definován jako součet příspěvků příliš vysoké, či nízké úrovně hlasitosti *OLR*, neoptimální místní vazby (vliv *STMR* a *TELR*) a degradace způsobené kvantizačním zkreslením *qdu*. Nezahrnuje ale vliv hlasových kodeků s nízkou přenosovou rychlostí, který je vyjádřen I_e .

Vliv $I_{s,WB}$ není dle [21] doposud dostatečně prozkoumán, je proto doporučeno jej při výpočtu R-faktoru vypustit a do rovnice 2.4 dosadit $I_{s,WB} = 0$. Mimo to, zkreslení popsaná parametrem I_s nejsou pro mobilní VoIP služby typická a v rámci simplifikace měřících metod hlasové kvality se jejich vliv zanedbává [22].

2.4.3 Degradace vlivem zpoždění, $I_{d,WB}$

I_d reprezentuje vliv zpoždění hlasových signálů na vnímané zhoršení konverzační kvality spojení. Konkrétně se jedná o dopad celkového zpoždění na interaktivitu komunikace a vnímaného echa na straně mluvčího i posluchače. Je vyjádřeno jako součet 3 složek [21]:

$$I_{d,WB} = I_{dte,WB} + I_{dle,WB} + I_{dd}, \quad (2.13)$$

kde $I_{dte,WB}$ jako funkce T , RLR , $TELR$ a $N_{0,WB}$ je odhadem vlivu echa na straně mluvčího. Faktor $I_{dle,WB}$ představuje echo na straně posluchače a je závislé na T_r , $R_{0,WB}$ a $WEPL$.

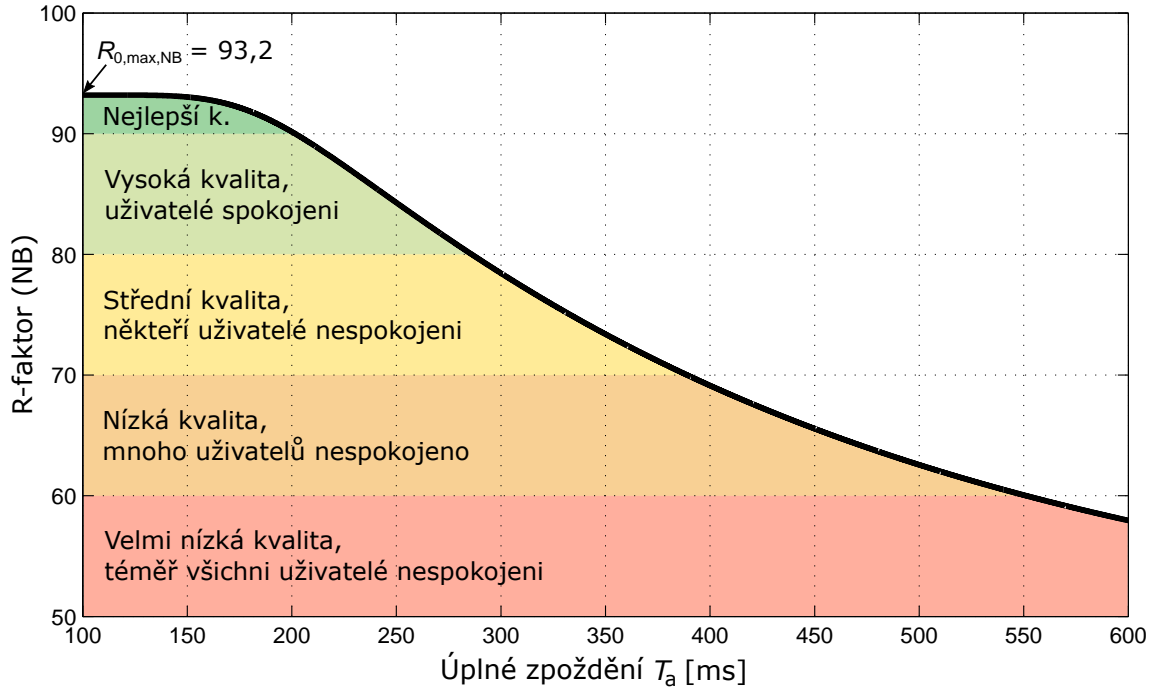
Složka I_{dd} v rovnici 2.13 vyjadřuje vliv úplného (absolutního, M2E) zpoždění T_a na vnímanou kvalitu konverzace. Na základě subjektivních testů bylo zjištěno, že celkové zpoždění menší než 150 ms má minimální vliv na vnímanou kvalitu hovoru [24], proto pro $T_a \leq 100$ ms je $I_{dd} = 0$. Pro $T_a > 100$ ms je faktor I_{dd} definován dle [21] následovně:

$$I_{dd} = 25 \left\{ \left(1 + X^6\right)^{\frac{1}{6}} - 3 \left[1 + \left(\frac{X}{3}\right)^6\right]^{\frac{1}{6}} + 2 \right\}, \quad (2.14)$$

kde

$$X = \frac{\log\left(\frac{T_a}{100}\right)}{\log 2}. \quad (2.15)$$

Vliv úplného zpoždění T_a na výslednou hodnotu R-faktoru (NB) reprezentující predikovanou konverzační kvalitu je znázorněn na obrázku 2.3. Pro zajištění uživatelské spokojenosti a vysoké kvality přenosu hlasu definované doporučením ITU-T G.109 (tabulka 2.4, [25]) nesmí být úplné zpoždění větší než 280 ms.



Obr. 2.3: Závislost R-faktoru (NB) na velikosti úplného zpoždění T_a

2.4.4 Efektivní faktor degradace použitou technikou, $I_{e,eff,WB}$

Faktor degradace použitou technikou $I_{e,WB}$ vyjadřuje vliv hlasového kodeku na vnímanou kvalitu přenosu hlasu. Rozšířená verze faktoru, $I_{e,eff,WB}$, zahrnuje navíc také efekt ztrátovosti paketů při přenosu. Je vyjádřena prostřednictvím $I_{e,WB}$, odolnosti proti ztrátám paketů B_{pl} a ztrátovosti paketů P_{pl} rovnicí [21]:

$$I_{e,eff,WB} = I_{e,WB} + (95 - I_{e,WB}) \cdot \frac{P_{pl}}{P_{pl} + B_{pl}}. \quad (2.16)$$

Při nulové ztrátovosti paketů (tj. $P_{pl} = 0$) je po dosazení do rovnice 2.16 zřejmé, že $I_{e,eff,WB} = I_{e,WB}$. Parametr $I_{e,WB}$ je nezávislý na ostatních vstupních parametrech E-modelu, jeho provizorní hodnoty jsou k dispozici v doporučení ITU-T G.113 [26] a stanovují se na základě subjektivních nebo objektivních (POLQA) testů. Tabulka 2.2 uvádí doporučené hodnoty $I_{e,WB}$ pro několik vybraných hlasových kodeků. Hodnoty $I_{e,WB}$ pro NB kodeky lze odvodit z I_e vztahem [26]:

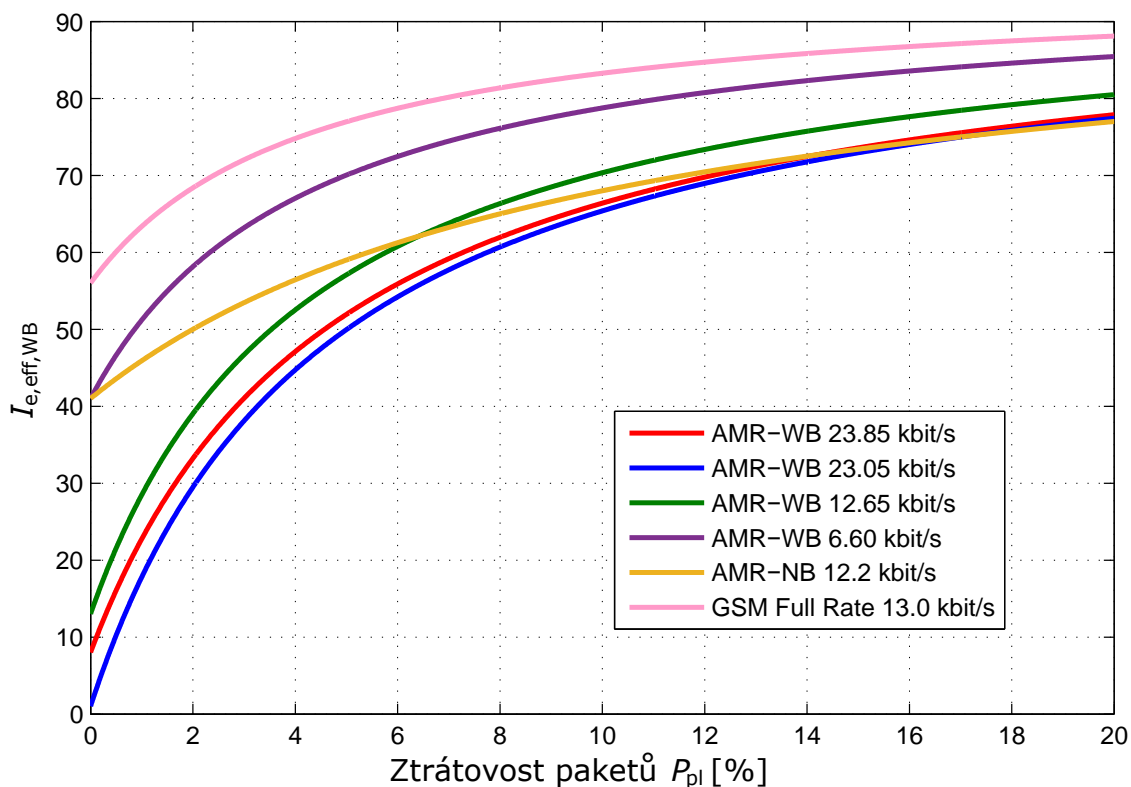
$$I_{e,WB} = I_e + (R_{0,max,WB} - R_{0,max,NB}) = I_e + 35,8. \quad (2.17)$$

Dosazením rovnice 2.17 do rovnice 2.12 lze zjistit, že degradace hlasové kvality v důsledku lineárního zkreslení použitím menší přenosové šířky pásma je vyjádřena právě faktorem $I_{e,WB}$.

Tab. 2.2: Hodnoty $I_{e,WB}$ a B_{pl} pro vybrané kodeky [26]

Kodek	Bitová rychlost [kbit/s]	Monotické		Diotické	
		$I_{e,WB}$	B_{pl}	$I_{e,WB}$	B_{pl}
CELP(AMR-WB)/G.722.2	23,85	8	-	10	4,9
	23,05	1	-	8	4,6
	12,65	13	-	20	4,3
	6,60	41	-	56	-
GSM Full-rate/GSM 06.10	13,0	56	-	56	-
GSM Enhanced-FR/AMR-NB	12,2	41	10,0	41	10,0

Na obrázku 2.4 je znázorněn průběh závislosti efektivního faktoru degradace použitou technikou $I_{e,eff,WB}$ na ztrátovosti paketů P_{pl} dle vztahu 2.16 vycházejí z provizorních hodnot $I_{e,WB}$ a B_{pl} v tabulce 2.2.



Obr. 2.4: Závislost $I_{e,eff,WB}$ na ztrátovosti paketů P_{pl} vybraných kodeků

2.4.5 Faktor zvýhodnění, A

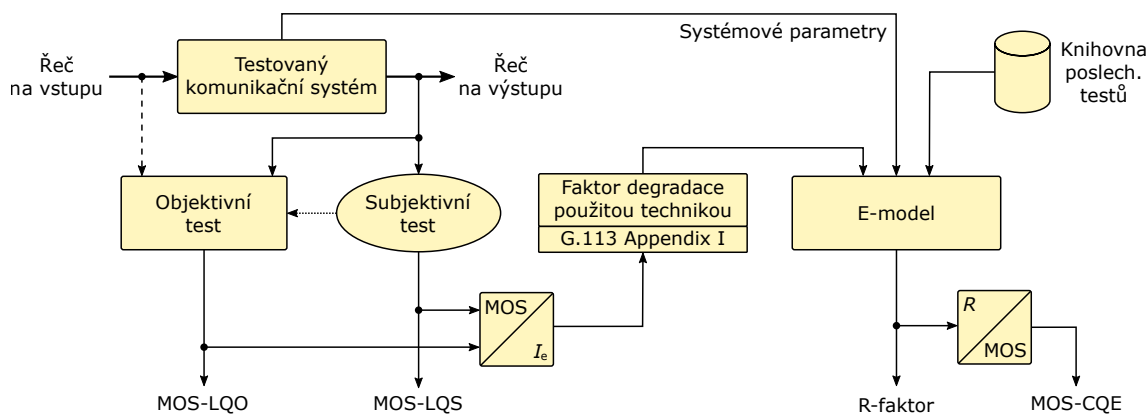
Faktor zvýhodnění A představuje potenciální zvýhodnění vnímané konverzační kvality uživatelem na základě mobility a dostupnosti v obtížně přístupných lokalitách [26]. Tabulka 2.3 uvádí příklady provizorních hodnot A pro nejčastější situace. Efekt faktoru zvýhodnění však dosud není potvrzen prostřednictvím subjektivní testů a jeho vliv ve WB kontextu také nebyl doposud dostatečně prozkoumán.

Tab. 2.3: Příklady hodnot faktoru zvýhodnění A pro různé kom. systémy [14]

Příklad komunikačního systému	Maximální hodnota A
Konvenční (drátové)	0
Mobilní zařízení v budově	5
Mobilní zařízení v terénu nebo jedoucím vozidle	10
Velmi obtížně přístupné lokality	20

2.5 R-faktor ve vztahu k hlasové kvalitě

Vzájemný vztah R-faktoru, resp. E-modelu, a hodnot MOS (kapitola 1.4) získaných prostřednictvím subjektivních (kapitola 1.3) nebo signálově založených objektivních (kapitola 1.5) testů je ilustrován na obr. 2.5.



Obr. 2.5: Vztah E-modelu k dalším metodám testování kvality přenosu hlasu [24]

Evidentním rozdílem ve vyhodnocení kvality přenosu hlasu jsou vstupní parametry jednotlivých metod. Zatímco objektivní a subjektivní testy hodnotí kvalitu řečových signálů na výstupu komunikačního systému, případně i ve srovnání se vstupním signálem, E-model k vyhodnocení využívá technické parametry komunikačního spojení. Vstupem E-modelu je také parametr I_e , který se odvozuje z výsledků

subjektivních a objektivních testů a další definované parametry stanovené na základě poslechových testů. Standardní výstup E-modelu ve formě R-faktoru lze převést na MOS-CQE nebo MOS-LQE, predikci kvality na stupnici MOS, která pak může být srovnávána s výsledky subjektivních a objektivních testů.

Tabulka 2.4 obsahuje přehled hodnocení kvality a předpokládané uživatelské spokojenosti v závislosti na hodnotě R-faktoru. Dále je doplněna odpovídajícími intervaly hodnot predikce kvality MOS-CQE v těchto kategoriích.

Tab. 2.4: Definice kategorií vnímané kvality přenosu [25]

Kvalita řeči	R-faktor	MOS-CQE	Uživatelská spokojenost
Nejlepší	90 – 100	4,3 – 4,5	Velmi spokojeni
Vysoká	80 – 90	4,0 – 4,3	Spokojeni
Střední	70 – 80	3,6 – 4,0	Někteří nespokojeni
Nízká	60 – 70	3,1 – 3,6	Mnozí nespokojeni
Velmi nízká	50 – 60	2,6 – 3,1	Téměř všichni nespokojeni
Nedoporučeno	0 – 50	0 – 2,6	Všichni nespokojeni

Hodnoty MOS-CQE byly odvozeny prostřednictvím převodní funkce [21]:

$$MOS = 1 + 0,035 R_x + R_x (R_x - 60) (100 - R_x) \cdot 7 \cdot 10^{-6}, \quad (2.18)$$

kde R_x představuje hodnotu R-faktoru transformovanou na NB škálu. V případě měření kvality v NB kontextu je $R_x = R$, kde $R \in \langle 0; 100 \rangle$. Hlasová kvality se potom označuje jako MOS-CQE_N. Pro NB/WB případy je

$$R_x = \frac{R}{1,29}, \quad (2.19)$$

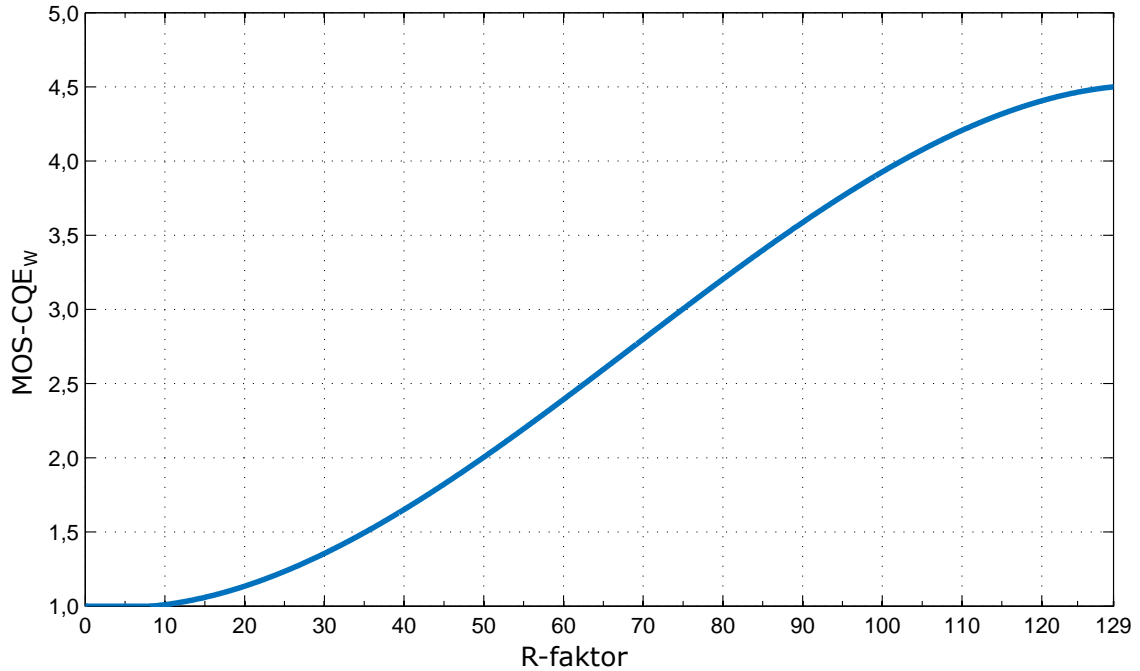
kde $R \in \langle 0; 129 \rangle$. Výsledek ve formě MOS-CQE_W následně odpovídá WB referenci. Převodní S funkce získaná dosazením do rovnic 2.18 a 2.19 je znázorněna na 2.6.

Pro opačný převod, z R-faktoru na MOS-CQE, definuje doporučení G.107 [14] pro R_x v intervalu $\langle 6,5; 100 \rangle$ vztah

$$R_x = \frac{20}{3} \left(8 - \sqrt{226} \cos \left(h + \frac{\pi}{3} \right) \right), \quad (2.20)$$

kde

$$h = \frac{1}{3} \arctan 2 \left(18566 - 6750 MOS, 15 \sqrt{-903522 + 1113960 MOS - 202500 MOS^2} \right) \quad (2.21)$$



Obr. 2.6: Závislost MOS-CQEw na hodnotě R-faktoru

a

$$\arctan2(x, y) = \begin{cases} \arctan\left(\frac{y}{x}\right) & \text{pro } x \geq 0, \\ \pi - \arctan\left(\frac{y}{-x}\right) & \text{pro } x < 0. \end{cases} \quad (2.22)$$

Alternativním statistickým vyhodnocením predikce vnímané kvality přenosu jsou funkce vyjadřující podíl osob, kteří přenosovou kvalitu ohodnotí jako dobrou či lepší (*GoB*) a podíl osob, kteří kvalitu ohodnotí jako špatnou nebo horší (*PoW*). Jsou definovány prostřednictvím Gaussovy chybové funkce [14]:

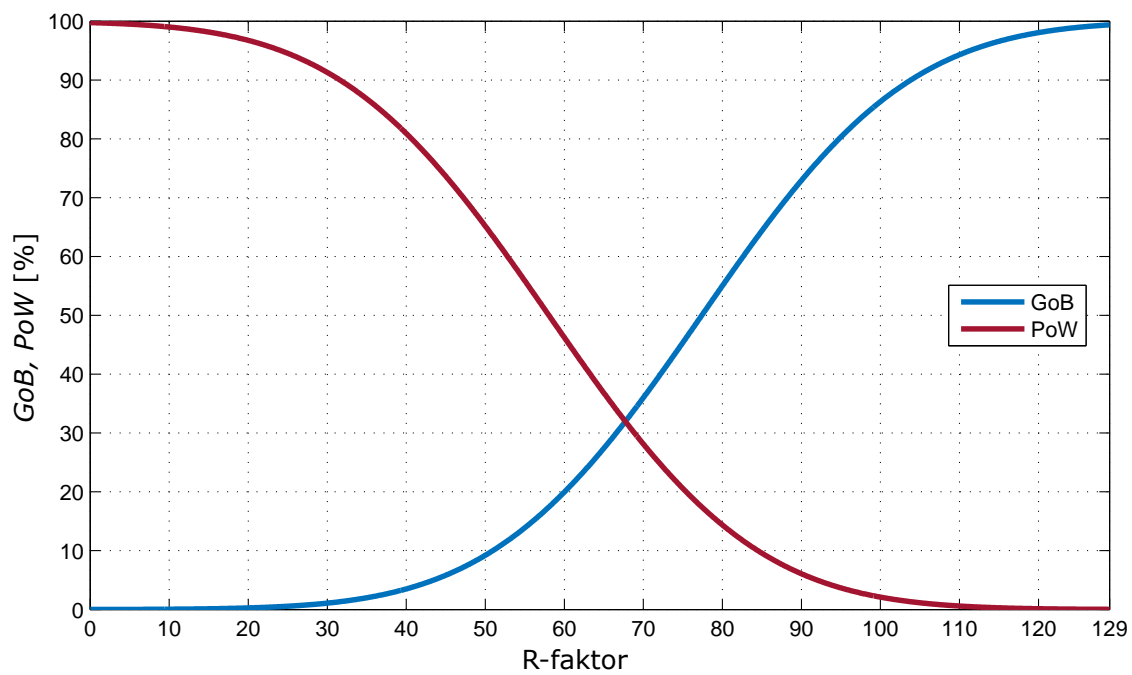
$$E(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt. \quad (2.23)$$

Pro funkce *GoB* a *PoW* v závislosti na hodnotě R-faktoru platí:

$$GoB = 100 E\left(\frac{R_x - 60}{16}\right) \%, \quad (2.24)$$

$$PoW = 100 E\left(\frac{45 - R_x}{16}\right) \%, \quad (2.25)$$

kde pro R_x je použita stejná úprava jako v rovnici 2.19. Průběh funkcí *GoB* a *PoW* je ilustrován na obrázku 2.7.



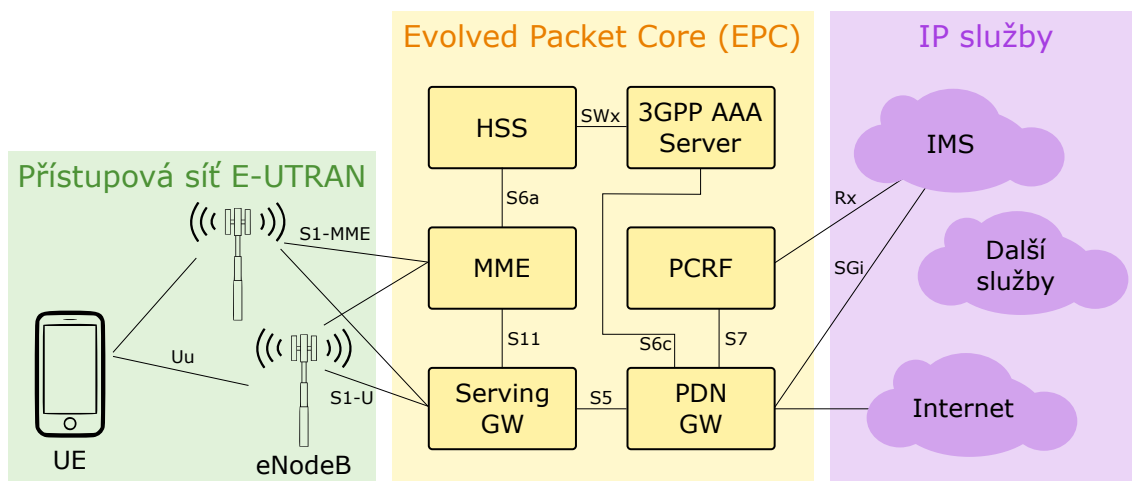
Obr. 2.7: Statistické kvalitativní ukazatele *GoB* a *PoW* v závislosti na R-faktoru

3 METODY PŘENOSU HLASU V SÍTÍCH LTE A WI-FI

Následující kapitola prezentuje používané metody přenosu hlasu v sítích LTE a Wi-Fi se zaměřením na funkci jádra sítě IMS. V druhé části jsou popsány nejdůležitější protokoly pro sestavení a přenos hlasového hovoru v technologiích VoLTE a VoWiFi, zejména signalizační protokol SIP, transportní protokoly RTP/RTCP a hlasový kodek AMR-WB.

3.1 Hlasové služby v sítích LTE

Na rozdíl od předchozích sítí druhé a třetí generace, LTE, založené na architektuře all-IP jádra Evolved Packet Core (EPC, obr. 3.1), nepodporuje okruhově spínanou (CS) komunikaci. Proto bylo zapotřebí pro přenos hlasových služeb v sítích LTE přejít k technologii VoIP, umožňující hlasové hovory v paketových sítích.



Obr. 3.1: Systémová architektura LTE definovaná 3GPP

3.1.1 Hlasové služby OTT

Tzv. over-the-top komunikační služby, např. známý Skype, jsou jedním ze způsobů využití datových sítí LTE pro hlasovou nebo multimediální komunikaci prostřednictvím síťového protokolu IP. OTT služby ke komunikaci využívají externí VoIP server umístěný mimo síť poskytovatele LTE. Jejich nevýhodou je však nemožnost zajištění QoS v přístupových sítích, omezená interoperabilita s okruhově spínanými sítěmi

a jinými poskytovateli hlasových služeb a také problematické zajištění nouzových volání.

3.1.2 Circuit-switched fallback (CSFB)

V první fázi rozvoje sítí LTE, bez podpory IMS jádra, je obvykle všechen hlasový provoz obsluhován prostřednictvím starších CS sítí 2G a 3G, zatímco LTE zajišťuje pro podporovaná uživatelská zařízení výhradně datový provoz. Tento přístup se označuje zkratkou CSFB.

Implementace spočívá v propojení MME v síti LTE a MSC serveru obsluhujícího CS komunikaci v sítích 2G a 3G signalizačním rozhraním SGs. Důsledkem toho je uživatelskému zařízení v síti LTE umožněna současná registrace v PS i CS síti. V případě příchozích či odchozích hovorů jsou pak zařízení připojená v síti LTE přeměrována do sítí 2G/3G, ve kterých je hovor obslužen.

3.1.3 Voice over LTE (VoLTE)

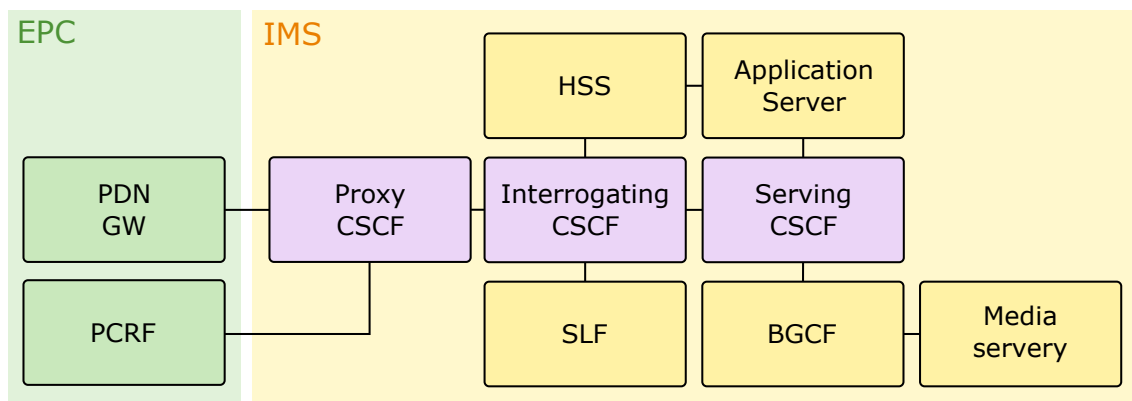
Skutečnou hlasovou podporu v sítích LTE (VoLTE) umožňuje až nasazení IP Multimedia Core Network Subsystem, zkráceně IMS, dle standardu MMTel.

IMS

IMS je samostatná řídicí síť definovaná 3GPP již v Release 5 [27]. Asociace GSM tento koncept pro multimediální služby MMTel adoptovala a definuje jej v GSMA IR.92 [28].

IMS je přístupově nezávislé jádro, které vzájemně propojuje existující hlasové a datové sítě, tj. PS s CS. Umožňuje sestavení peer-to-peer spojení mezi podporovanými zařízeními prostřednictvím síťového protokolu IP a zajišťuje QoS v přístupových sítích. Hlavním signalizačním protokolem IMS je Session Initiation Protocol (SIP, kap. 3.3.1), široce používaný pro VoIP aplikace. Přenos hlasu je zajištěn multimediálním transportním protokolem RTP, který je blíže popsán v kapitole 3.3.3.

Na obrázku 3.2 je znázorněna architektura jádra IMS. Základním prvkem je Call Session Control Function (CSCF) dělicí se na 3 logické části: Proxy, Interrogating a Serving. P-CSCF obstarává spojení IMS s vnějšími elementy sítě, je tak jediným prvkem viditelným pro UE [29]. P-CSCF ve spojení s PCRF spravuje QoS, zajišťuje kompresi signalizačních zpráv a udržuje šifrované spojení IPsec tunelem s UE. S-CSCF je spojovacím bodem s IMS multimediálními servery, registruje všechny UE pod jeho správou a sestavuje mezi nimi vzájemná spojení. I-CSCF figuruje jako proxy mezi uživatelem (UE) a S-CSCF. Zároveň je rozhraním pro signalizační zprávy z jiných IMS [30].



Obr. 3.2: Logická architektura jádra IMS

Dalšími prvky jsou HSS, databáze informací o užívatelích zodpovědná také za jejich autentizaci a autorizaci, správce HSS v síti SLF a multimediální servery propojené k S-CSCF skrze BGCF. Pro hlasové služby je důležitým multimediálním serverem Media Gateway Control Function (MGCF), který má na starosti překódování kodeků a konverzi RTP paketového toku při spojení s CS sítěmi.

SRVCC

V období stále se rozvíjejících sítí LTE je důležitou součástí VoLTE infrastruktury systém označovaný SRVCC, který umožňuje plynulý handover (předání hovoru) mezi PS sítěmi LTE a staršími systémy CS sítí 2G (GSM) a 3G (UMTS). Běžně by všechna aktivní spojení v momentě opuštění oblasti stabilního LTE pokrytí selhaly. SRVCC umožňuje sítěmi druhé a třetí generace tato slepá místa vykrýt. Průběh handoveru i celý hovor před i po handoveru je řízen prostřednictvím IMS, které současně informuje MME a MSC Server o vzájemném předání.

Quality of Service

Vymežujícím faktorem technologie VoLTE oproti jiným OTT VoIP aplikacím je komplexní systém QoS, který umožňuje v síti LTE přiřadit datovým tokům hovorů VoLTE zvýšenou prioritu a v důsledku tak zaručit i vyšší kvalitu hovoru.

Zatímco běžný datový provoz v síti LTE využívá výchozí sdílený EPS bearer¹, fungující na tzv. best-effort principu, nezaručují QoS, pro speciální druhy datového provozu, jako je technologie VoLTE, umožňuje LTE vytvoření vyhrazeného EPS beareru. Vyhrazený EPS bearer izoluje na základě zdrojových a cílových IP adres a portů konkrétní typ provozu od ostatních, předchází tak do určité míry negativním dopadům vysokého transportního zpoždění, jitteru a ztrát paketů na kvalitu

¹logické spojení mezi UE a PDN-GW

hovoru. Své QoS vlastnosti má bearer definován přiřazením určitého QCI (QoS Class Identifier). Tabulka 3.1 obsahuje všechny třídy QCI definované 3GPP.

Tab. 3.1: Standardizované vlastnosti QCI [31]

QCI	Typ zdrojů	Priorita	Horní limity		Příklady služeb
			Zpoždění	Chybovost	
1	GBR	2	100 ms	10^{-2}	Konverzační hlas (VoLTE)
2		4	150 ms	10^{-3}	Konverzační video (ViLTE)
3		3	50 ms	10^{-3}	Hry v reálném čase
4		5	300 ms	10^{-6}	Bufferované video
65		0,7	75 ms	10^{-2}	MCPTT
66		2	100 ms	10^{-2}	Non-MCPTT
5	non-GBR	1	100 ms	10^{-6}	IMS signalizace
6		6	300 ms	10^{-6}	Video, TCP služby (MPS)
7		7	100 ms	10^{-3}	Hlas, video, on-line hry
8		8	300 ms	10^{-6}	Video, TCP služby (VIP)
9		9	300 ms	10^{-6}	Video, TCP služby
69		0,5	60 ms	10^{-6}	MCPTT signalizace
70		5,5	200 ms	10^{-6}	Kriticky důležitá data

QCI je charakterizován typem přiřazení zdrojů², prioritou ve frontách, maximálním přípustným E2E zpožděním mezi UE a PDN-GW a maximální přípustnou paketovou chybovostí při příjmu na PDCP vrstvě. Při pokusu o VoLTE hovor PCRF vždy inicializuje sestavení vyhrazeného beareru s QCI = 1 [31].

3.2 Hlasové služby v sítích Wi-Fi

Další inovací v oblasti mobilních služeb, která využívá IMS, je podpora volání v sítích WLAN standardu IEEE 802.11, tzv. Wi-Fi Calling. Podobně jako v sítích LTE bylo již dříve možné využívat pro volání skrze přístupové sítě Wi-Fi OTT aplikací, nejnovější mobilní zařízení s operačními systémy Android a iOS už ale podporují Wi-Fi Calling nativně bez nutnosti přímé interakce s uživatelem. Výhodou této technologie je snížení datové zátěže a snadné rozšíření pokrytí sítí LTE o vnitřní prostory budov prostřednictvím domácích a podnikových sítí WLAN.

²GBR — garantovaná bitová rychlost, nebo non-GBR — negarantovaná bitová rychlost

Technologie Wi-Fi Calling je specifikována Asociací GSM v dokumentu GSMA IR.51 [32]. Technologicky se jedná o rozšíření služeb operátora poskytovaných prostřednictvím jádra IMS, umožňující snadnou koexistenci s technologií VoLTE včetně obousměrných handoverů.

Služby VoWiFi mohou být dle standardů 3GPP komunikovat s IMS přes EPC prostřednictvím důvěryhodného (Trusted) nebo nedůvěryhodného (Untrusted) přístupu. Důvěryhodnost lokální Wi-Fi sítě posuzuje operátor a závisí na jejím zabezpečení. V případě důvěryhodné přístupu jsou Wi-Fi přístupové body (AP) připojeny skrze tzv. Trusted WLAN Access Gateway (TWAG), která je s EPC spojena standardním S2a rozhraním. V případě nedůvěryhodného přístupu koncová zařízení sestaví šifrovaný IPsec tunel s blokem evolved Packet Data Gateway (ePDG) na vstupu EPC. Veškerá komunikace je potom šifrována. VoWiFi, stejně jako VoLTE, využívá internetový signalizační protokol SIP a multimediální transportní protokol RTP. Autentizace je zajištěna protokolem EAP-AKA.

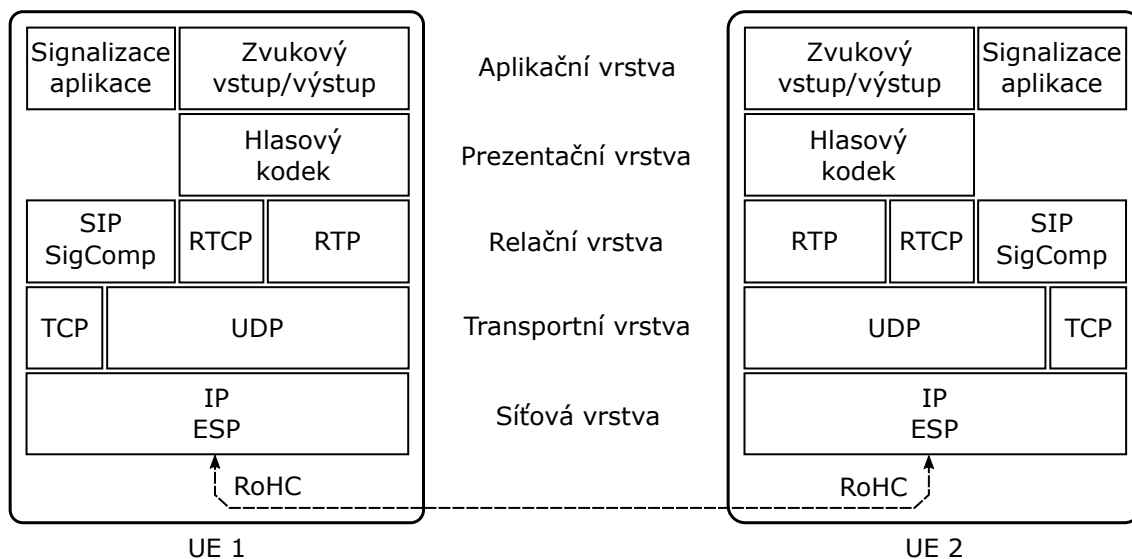
Připojení přes EPC umožňuje snadnou mobilitu mezi přístupovými body Wi-Fi a LTE sítí, zachovávající identitu koncového zařízení v síti. Alternativní VoWiFi architekturou je přímý přístup k IMS bez zapojení EPC. Zařízením připojícím se přímo prostřednictvím šifrované komunikace v síti Internet však není umožněno předávání hovorů mezi přístupovými body nebo sítí LTE.

3.3 IMS protokoly

Své uplatnění má ve VoIP systémech VoLTE a VoWiFi celá řada komunikačních protokolů. Následující kapitola s ohledem na jejich důležitost při měření a vyhodnocení hlasové kvality popisuje několik z nich.

Z pohledu UE jsou hlavními komunikačními protokoly signalizační protokol SIP (kap. 3.3.1) a transportní protokoly RTP/RTCP (kap. 3.3.3). Pro přenos signalizace se obvykle využívá transportního protokolu TCP a pro přenos hlasu výhradně protokol UDP (kap. 3.3.2). Na síťové vrstvě se pro směrování paketů používá protokolu IP rozšířeného o zabezpečovací principy sady protokolů IPsec, zejména ESP. Na aplikační vrstvě se pro zabezpečení komunikace používá šifrování TLS. Obrázek 3.3 zobrazuje logicky uspořádané VoIP protokoly ve vztahu k referenčnímu modelu OSI.

Důležitou součástí technologií VoLTE a VoWiFi je také metoda komprese hlaviček RoHC. Hlasové kodeky používané v těchto technologiích generují malé pakety, které jsou přenášeny v ideálním případě každých 20 ms. Nejnížší možná souhrnná velikost hlaviček takového paketu, zahrnující IPv4, UDP a RTP, je 40 B, viz obrázek 3.4. Samotný náklad paketu, přenášející užitečná data, tak může mít i menší velikost než hlavičky protokolů. RoHC umožňuje velikost hlaviček během spojení mnohonásobně



Obr. 3.3: OSI model vyšších vrstev technologií VoLTE a VoWiFi

redukovat, a to až na velikost 3 B [33].

20 bytů	8 bytů	12 bytů	
IPv4 hlavička	UDP hlavička	RTP hlavička	Náklad (závislý na kodeku)

Obr. 3.4: Minimální struktura VoIP paketu

3.3.1 Session Initiation Protocol (SIP)

Signalizační protokol SIP se používá v internetové telefonii. Jedná se textově založený TCP/IP protokol aplikační vrstvy. Je definován IETF v RFC 3261 [34], v kombinaci s IMS má ale řadu specifických rozšíření. Výchozím portem protokolu je TCP/UDP port 5060, případně port 5061 pro šifrovanou komunikaci. Každý účastník sítě SIP, tzv. User Agent, je identifikován unikátním URI (Uniform Resource Identifier) ve formátu `sip:username:password@host:port`. Protokol umožňuje lokalizaci účastníka, jeho autentizaci, zjištění jeho aktuálního stavu a možnosti, navázání, řízení a ukončení spojení.

SIP zprávy

Základem funkce protokolu je výměna zpráv mezi účastníky sestavovaného spojení. Klient zasílá žádost, která vyvolá určitou metodu na serveru a ten následně odešle

alespoň jednu odpověď zpět [34]. Tabulka 3.2 obsahuje seznam často používaných SIP metod.

Tab. 3.2: Seznam často používaných SIP metod

SIP metoda	Popis	Definice
REGISTER	Registrace klienta na SIP serveru	RFC 3261
INVITE	Žádost o sestavení nové relace	RFC 3261
ACK	Potvrzení příjmu konečné odpovědi na INVITE žádost	RFC 3261
BYE	Ukončení relace	RFC 3261
CANCEL	Zrušení čekajících žádostí	RFC 3261
OPTIONS	Dotaz na možnosti serveru	RFC 3261
PRACK	Průběžné potvrzení	RFC 3262
SUBSCRIBE	Přihlásit se k odběru upozornění na události	RFC 3265
NOTIFY	Upozornění na událost	RFC 3265
UPDATE	Aktualizace parametrů relace	RFC 3311

SIP odpovědi jsou vždy charakterizovány standardizovaným trojmístným kódem. První číslice dělí odpovědi do 6 skupin: 1xx označuje průběžné odpovědi (100 Trying, 180 Ringing, 183 Session Progress), 2xx informuje o úspěšné operaci (200 OK), kategorie 3xx upozorňuje na nutnost provedení další akce, 4xx varuje klienta, že požadavek je chybný nebo nemůže být zpracován (400 Bad Request, 401 Unauthorized, 486 Busy Here), 5xx značí chybu na straně serveru (500 Server Internal Error, 503 Service Unavailable) a 6xx signalizuje globální chybu informující o nemožnosti vykonat žádost (604 Does Not Exist Anywhere) [34].

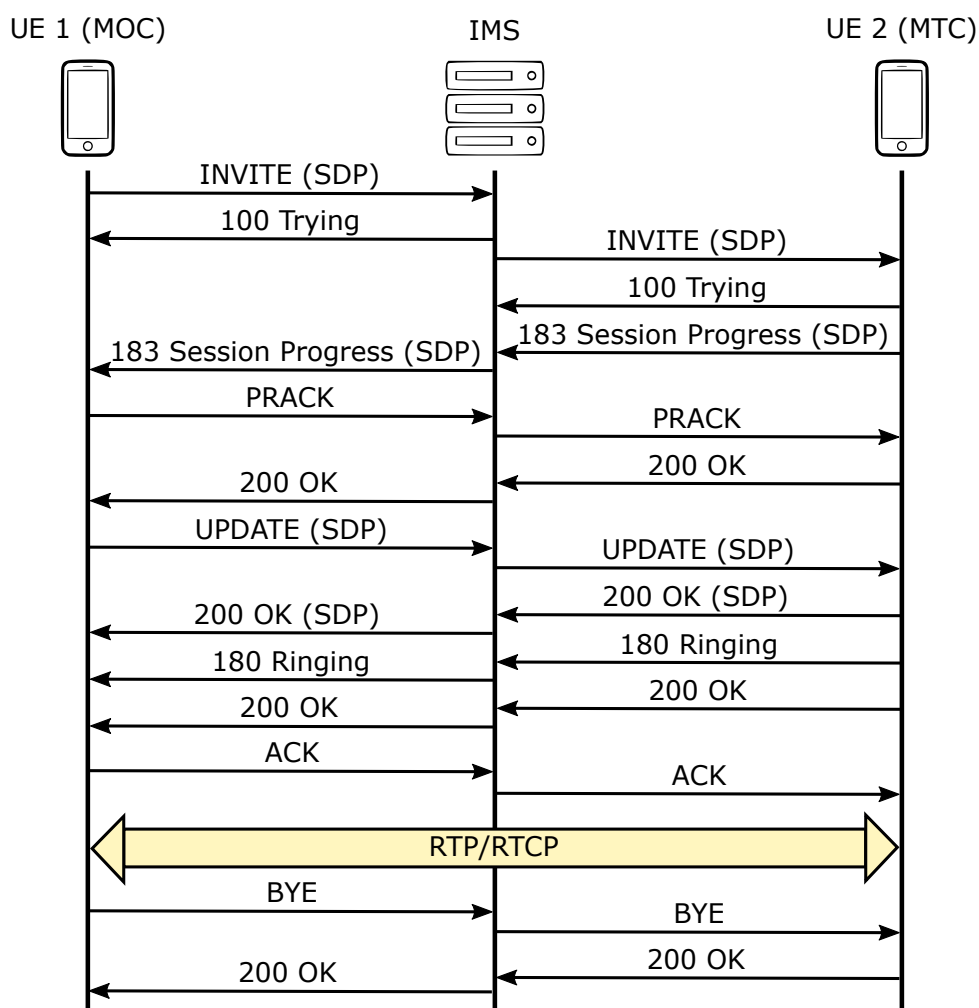
Session Description Protocol (SDP)

Detaily sestavovaného spojení, např. čísla portů, protokoly, kodeky, se sjednávají na základě SDP zpráv transportovaných uvnitř SIP paketů. Protokol SDP je definován IETF v RFC 4566 [35]. Stejně jako v případě SIP se jedná o textově založený protokol. Vzájemná domluva účastníků spojení probíhá na principu nabídka-odpověď, kdy jedna strana navrhne parametry spojení a druhá strana odpoví protinávrhem dle svých vlastních možností.

Sestavení spojení

Po úspěšné registraci klienta (metodou REGISTER) v IMS je UE připraveno k sestavení hovoru. Sekvence signalizačních zpráv vyměňovaných mezi oběma UE a IMS při sestavování spojení je na obrázku 3.5. Spojení je iniciováno volajícím UE, v tomto

případě UE 1, které zasílá SIP INVITE žádost, obsahující mimo volaného telefonního čísla také SDP zprávu s nabídkou podporovaných formátů pro přenos hlasu. Poté následuje výměna SIP zpráv, která se může mírně lišit v závislosti na konkrétní implementaci. UE 2 odpovídá UE 1 protinávrhem zvoleného hlasového kodeku v další SDP zprávě. V případě signalizace na obr. 3.5 je toto nastavení oběma UE potvrzeno ještě jednou výměnou SDP zpráv.



Obr. 3.5: Signalizace protokolu SIP při sestavování a ukončování hovoru

Po přijetí hovoru příjemcem (UE 2), je zaslána odpověď 200 OK, načež UE 1 potvrzuje inicializaci spojení metodou ACK. IMS následně přidělí oběma UE vyhrazený bearer s QCI 1 pro přenos hlasové komunikace protokolem RTP. Ukončení hovoru je opět signalizováno pomocí protokolu SIP, konkrétně metodou BYE a následným potvrzením protistrany.

3.3.2 User Datagram Protocol (UDP)

UDP je protokolem transportní vrstvy patřící do rodiny protokolů TCP/IP. Je charakteristický minimálními režijními požadavky, absencí řízení toku nebo potvrzování přijatých dat [36]. Právě proto je tento tzv. nespolehlivý protokol používán namísto spolehlivého protokolu TCP pro přenos multimediálních služeb v reálném čase. Služby jako VoIP jsou totiž vysoce závislé na nízkém a neměnném transportním zpoždění. Pro VoIP je obvykle akceptovatelnější nenulová ztrátovost paketů nad vysokým síťovým zpožděním, které retransmise ztracených paketů protokolu TCP způsobují.

Na obrázku 3.6 je zobrazena jednoduchá struktura hlavičky datagramů protokolu. UDP zprostředkovává komunikaci mezi dvěma procesy na obou koncích přenosu na základě tzv. portů. Hlavička obsahuje 16 bitová čísla portů přiřazená odesílající a přijímající aplikaci. Další 16 bitová pole obsahují velikost datagramu v bytech a kontrolní součet, umožňující kontrolu chyb při přenosu.

1	16 17	32
Číslo zdrojového portu	Číslo cílového portu	
Celková délka	Kontrolní součet	

Obr. 3.6: Hlavička protokolu UDP

3.3.3 RTP/RTCP

Real-time Transport Protocol (RTP) je nástavbou protokolu UDP určenou speciálně pro přenos multimediálních služeb, např. VoIP. Použití protokolu však není omezeno jen na spojení s UDP. Inicializace spojení mezi účastníky není obsahem standardu RTP a v případě VoIP aplikací je řešena prostřednictvím protokolů SIP a SDP.

Protokol RTP je specifikován IETF v RFC 3550 [37]. Tento standard definuje 2 protokoly: RTP pro transport multimediálních dat a RTCP pro řízení a periodické monitorování kvalitativních parametrů spojení. RTP a RTCP používají ve spojení s transportním protokolem UDP pro každý separátní tok dat pár po sobě následujících portů (RTP lichý, RTCP sudý).

Real-time Transport Protocol (RTP)

RTP zajišťuje multimediálním aplikacím podporu prostřednictvím identifikace pořadí, časování a obsahu jednotlivých rámců, určení transportních ztrát nebo zabezpečení nákladu. Obecný formát hlavičky protokolu RTP je uveden na obrázku 3.7.

1	2	3	4	5	6	7	8	9	10	16	17	32
V	P	X	CC		M	Typ nákladu (PT)				Pořadové číslo (SN)		
Časová značka (TS)												
Identifikátor synchronizačního zdroje (SSRC)												
Identifikátor přispívajících zdrojů (CSRC)												
Rozšiřující hlavička												

Obr. 3.7: Hlavička protokolu RTP

Prvních 12 bytů je povinných, pole CSRC a rozšířené hlavičky jsou volitelná v závislosti na obsahu kontrolních polí X a CC. Kontrolní pole V, P a M nesou informaci o verzi protokolu a bytovém zarovnání. Pole PT určuje typ nákladu, v případě VoIP používaný hlasový kodek. Pořadové číslo je identifikátor, který je inkrementován s každým odeslaným paketem a umožňuje jejich správné seřazení na přijímací straně. TS obsahuje informaci o přesném čase prvního vzorku přenášeného nákladu umožňující synchronizaci přehrávacího zařízení anebo výpočet transportního zpoždění. CSRC umožňuje rozlišit jednotlivé zdroje multimediálních dat v případě použití více oddělených proudů. Počáteční hodnoty SN, TS a CSRC jsou voleny náhodně.

Každý RTP paket pak za hlavičkou obsahuje multimediální náklad, jehož obsah je určen konkrétní aplikací. Formát nákladu pro specifickou aplikaci je definován tzv. RTP profilem. Profil je identifikován hodnotou PT v hlavičce RTP paketu [36].

RTP Control Protocol (RTCP)

RTCP zajišťuje přenos statistických a kontrolních informací protokolu RTP skrze oddělený UDP port. RTCP zprávy jsou zasílány prostřednictvím IP uni- nebo multicast ostatním členům aktivního RTP spojení v kvaziperiodických časových intervalech. Zprávy protokolu RTCP se využívají také pro synchronizaci několika souběžných RTP proudů dat.

Mezi základní typy RTCP paketů dle RFC 3550 [37] patří zpráva příjemce RR, nesoucí informaci o kvalitě spojení z pohledu příjemce obsahu (nejvyšší přijaté sekvenční číslo, počet ztracených paketů, jitter, časová značka pro výpočet obousměrného zpoždění), zpráva odesílatele SR, obsahující navíc stavové informace o odeslaných paketech, a tzv. SDES (Source Description Items) zprávy. RTCP SDES pakety přenáší textové informace identifikující zdroj přijímaného RTP proudu, např. uživatelské jméno, telefonní číslo nebo e-mailovou adresu. Narozdíl od identifikátoru CSRC protokolu RTP se obsah SDES paketů v průběhu spojení nemění. Dalším typem RTCP paketu je BYE, který oznamuje ukončení spojení.

3.4 Hlasové kodeky

Všechny VoIP systémy využívají pro zakódování analogového hlasového signálu na vysílací straně a dekodování přeneseného signálu na přijímací straně služeb tzv. hlasových kodeků. Své uplatnění má v telekomunikacích velké množství audio kodeků v závislosti na konkrétní aplikaci. Část z nich je standardizována ITU-T v doporučeních G.711 až G.729, jiné jsou proprietární. Mezi vlastnosti jednotlivých kodeků patří stupeň komprese a s tím související bitová rychlost a vzorkovací frekvence, odolnost proti chybovosti a ztrátám paketů aj. Použitý typ hlasového kodeku je jedním z nejdůležitějších faktorů výsledné hlasové kvality sítě.

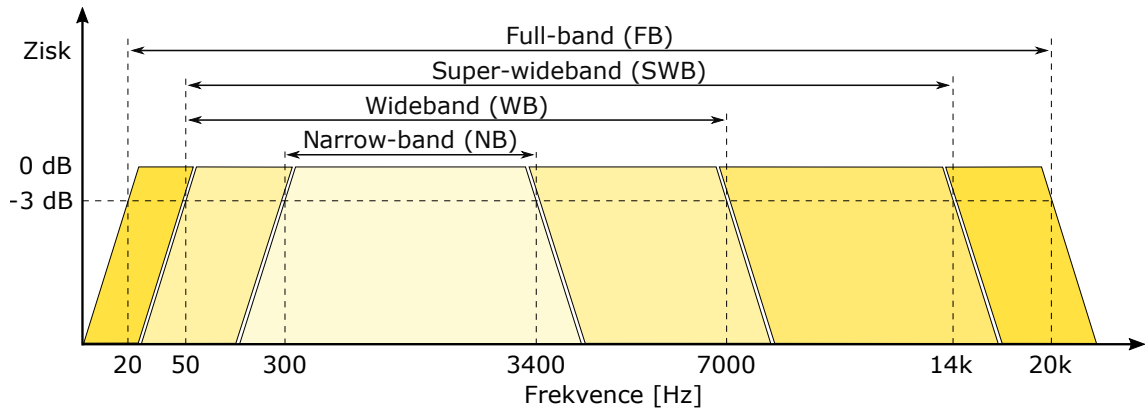
3.4.1 Adaptive Multi-Rate (AMR)

V mobilních sítích GSM a UMTS se v současné době dle norem 3GPP používají hlasové kodeky AMR (označován také AMR-NB) a AMR-WB. Přestože původně byly navrženy pro okruhově spínané mobilní systémy, díky jejich flexibilitě a odolnosti proti chybám se nyní využívají i v sítích VoIP.

AMR-NB je kodek, který podporuje 8 kódovacích módů s bitovými rychlostmi od 4,75 do 12,2 kbit/s. Kodek je schopný mezi těmito módy přepínat v závislosti na rychle se měnících rádiových podmínkách spojení a kompenzovat tak nárůst chybovosti zvýšením zabezpečení na úkor kvality přenášeného hlasového signálu. Ve VoIP aplikacích se ale adaptivních vlastností kodeku běžně nevyužívá a nejčastěji je volen fixní mód s nejvyšší bitovou rychlostí [22]. Kodek AMR používá vzorkovací frekvenci 8 kHz, což odpovídá NB frekvenčnímu kanálu (viz obr. 3.8). Kódování hlasu probíhá po rámcích délky 20 ms, tzn. 160 vzorcích původního analogového hlasového signálu [38]. Technologie VoLTE a VoWiFi používají kodek AMR-NB jen pro zpětnou kompatibilitu se sítěmi a zařízeními nepodporujícími technologii HD Voice.

3.4.2 Adaptive Multi-Rate Wideband (AMR-WB)

HD Voice kodek AMR-WB má podobně jako AMR-NB definovaných 9 operačních módů, lišících se svou bitovou rychlostí a využitou šířkou pásma, viz tabulka 3.3. Algoritmus využívá metodu kódování hlasu označovanou ACELP (Algebraic Code Excited Linear Prediction). WB frekvenčnímu rozsahu odpovídá použitá vzorkovací frekvence 16 kHz [38]. Použitý rozšířený frekvenční rozsah hlasového signálu přináší výrazné zlepšení vnímané hlasové kvality, hlas je čistší, více přirozený a snáze srozumitelný [39]. Na rozdíl od své NB předlohy je kodek AMR-WB standardizován ITU-T v doporučení G.722.2 [40].



Obr. 3.8: Používané frekvenční šířky pásma hlasových kodeků

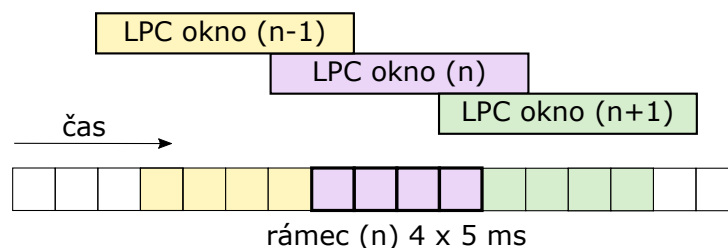
Tab. 3.3: Operační módy a odpovídající velikosti rámce kodeku AMR-WB [41]

Mód	Typ rámce	Počet bitů nákladu	Počet bytů rámce IF2
AMR-WB 6.60 kbit/s	0	132	18
AMR-WB 8.85 kbit/s	1	177	23
AMR-WB 12.65 kbit/s	2	253	33
AMR-WB 14.25 kbit/s	3	285	37
AMR-WB 15.85 kbit/s	4	317	41
AMR-WB 18.25 kbit/s	5	365	47
AMR-WB 19.85 kbit/s	6	397	51
AMR-WB 23.05 kbit/s	7	461	59
AMR-WB 23.85 kbit/s	8	477	61
AMR-WB SID	9	40	6

Zpoždění

Stejně jako v případě kodeku AMR-NB má každý aktivní AMR-WB rámec délku 20 ms, skládající se ze 4 subrámců délky 5 ms (80 vzorků). Samotná LPC analýza algoritmu ACELP však po předzpracování vstupního navzorkovaného signálu probíhá přes okno délky 30 ms, tj. s 5 ms symetrickým přesahem, viz obrázek 3.9. Jedná se o tzv. look ahead zpoždění. Při zanedbání zpoždění vloženého filtrací a signálovým zpracováním, které se v závislosti na implementaci pohybuje okolo 1 ms [42], je celkové zpoždění vložené AMR-WB kódérem

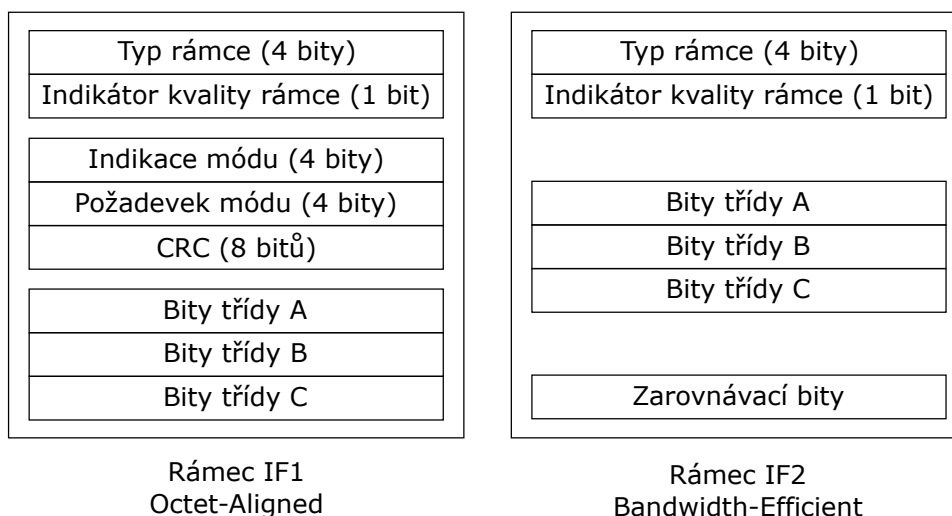
$$T_{\text{AMR-WB}} = T_{\text{frame}} + T_{\text{LA}} = 20 \text{ ms} + 5 \text{ ms} = 25 \text{ ms} \quad (3.1)$$



Obr. 3.9: Přesah okna LPC analýzy v algoritmu AMR-WB

Formát rámce

Obsah AMR-WB kodeků může být formátován dvěma způsoby: vychází bytově zarovnaném (Octet-Aligned, IF1), nebo úsporném (Bandwidth-Efficient, IF2). V prvně jmenovaném případě jsou jednotlivá pole rámce včetně hlavičky zarovnána do oktětů pro maximální výkonnost implementace kodeku. Při úsporném přístupu je bytově zarovnan jen celý obsah rámce pro snížení počtu redundantních bitů na minimum. Pouze bytově zarovnaný mód má však možnost využít robustního třídění, prokládání a CRC pro zvýšení odolnosti proti chybovosti a ztrátám paketů. Struktura rámců IF1 a IF2 je znázorněna na obrázku 3.10.



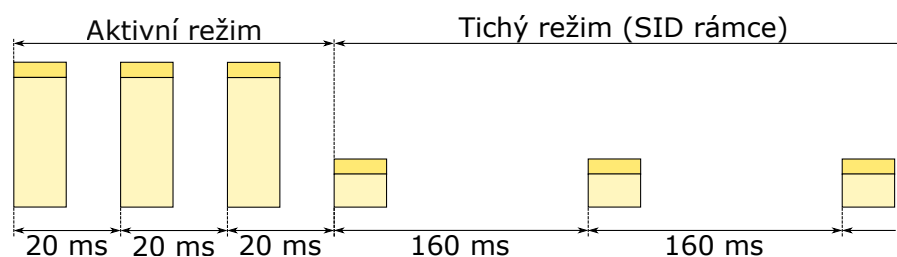
Obr. 3.10: Obecná struktura rámců AMR-WB IF1 a IF2 [41]

Protože při kódování řeči kodekem nemají všechny bity nákladu stejnou důležitost pro kvalitu hlasu, jsou definovány třídy A, B a C se sestupnou citlivostí na chyby přenosu a tomu odpovídající úrovní zabezpečení [38].

DTX/VAD

Pro redukci spotřebované energie a přenášených dat má kodek integrovaný detektor aktivity hlasu (VAD). V případě, že právě zpracováváný rámeček neobsahuje řeč, přenáší

se v omezených intervalech (DTX) pouze informace o úrovni šumu na pozadí, díky níž je účastník na druhé straně spojení ujistěn, že nedošlo k výpadku hovoru. Rámec obsahující informaci o úrovni šumu se nazývá Silence Insertion Descriptor (SID) a je kódován jako typ rámce 9. Obrázek 3.11 znázorňuje činnost AMR/AMR-WB kodéru při aktivní řeči a v tichém režimu, výška bloků symbolizuje velikost hlavičky a nákladu rámce.



Obr. 3.11: Generování AMR/AMR-WB rámců v závislosti na VAD

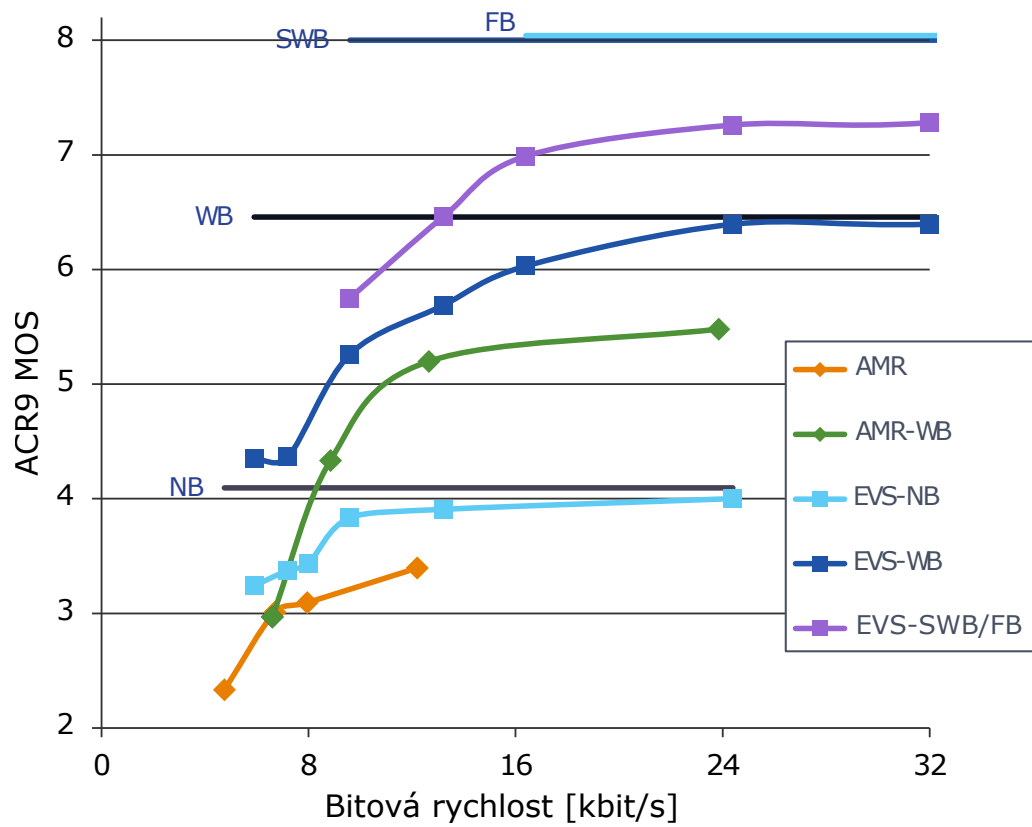
Formát RTP nákladu

IETF RFC 4867 [38] definuje formát RTP nákladu při použití kodeků AMR-NB a AMR-WB. Časová značka TS nese informaci pořadovém čísle prvního přenášeného audio vzorku v daném RTP paketu. RTP marker M je využit k indikaci úseků řeči, dle specifikace má být nastaven vždy u prvního rámce aktivní řeči následujícího SID rámec. Mimo to, jeden paket může pro redukci režijních dat obsahovat více po sobě jdoucích AMR-WB rámců. RFC 4867 podrobně uvádí systém jejich řazení.

3.4.3 Enhanced Voice Services (EVS)

V roce 2014 byl organizací 3GPP standardizován nový Evolved HD Voice audio kodek pod označením Enhanced Voice Services (EVS). Prvního nasazení ve veřejné síti se dočkal v dubnu 2016 operátorem T-Mobile US [43]. Tento kodek přináší ve srovnání s ostatními 3GPP kodeky zlepšení hlasové kvality, zvýšení účinnosti komprese a odolnosti proti chybám [39]. Byl navržen speciálně pro použití s technologií VoLTE. Jeho aplikace není omezena jen na přenos hlasu, díky podpoře až FB (20 – 20 000 Hz) frekvenčního pásma dokáže účinně přenášet i jiné audio signály, např. hudbu. Při vývoji byl kladen důraz také na zpětnou kompatibilitu s kodeky AMR-WB a AMR-NB a mobilními sítěmi 2. a 3. generace.

Výsledky subjektivních testů vybraných hlasových kodeků na obrázku 3.12 ukazují, že kodek EVS kvalitativně překonává AMR kodeky i při výrazně vyšší kompresi hlasových dat. Například EVS-SWB 9,6 kbit/s dosahuje vyšších hodnocení kvality než AMR-WB 23,85 kbit/s.



Obr. 3.12: Subjektivní hodnoty MOS čistých hl. vzorků kodeků AMR a EVS [44]

4 NÁVRH A OPTIMALIZACE ALGORITMU VÝPOČTU R-FAKTORU

Následující kapitola popisuje návrh výpočetního algoritmu R-faktoru na základě E-modelu a měření hlasové kvality technologie VoLTE. Podkapitola 4.1 uvádí popis měřených podmínek, použitého měřicího systému a podrobný rozbor jednotlivých měřených parametrů. Část 4.2 obsahuje teoretický návrh výpočetního algoritmu R-faktoru vycházejícího z doporučení ITU-T G.107.1 [21]. Podkapitola 4.3 rozšiřuje navržený algoritmus na základě poznatků z měření hlasové kvality a předkládá návrh korekce parametrů navrhovaného algoritmu pro maximální korelaci jeho výsledků s měřeními algoritmem POLQA. V poslední části, podkapitole 4.4, jsou představeny výsledky měření hlasové kvality navrženým algoritmem a je analyzována jejich korelace s měřeními algoritmem POLQA.

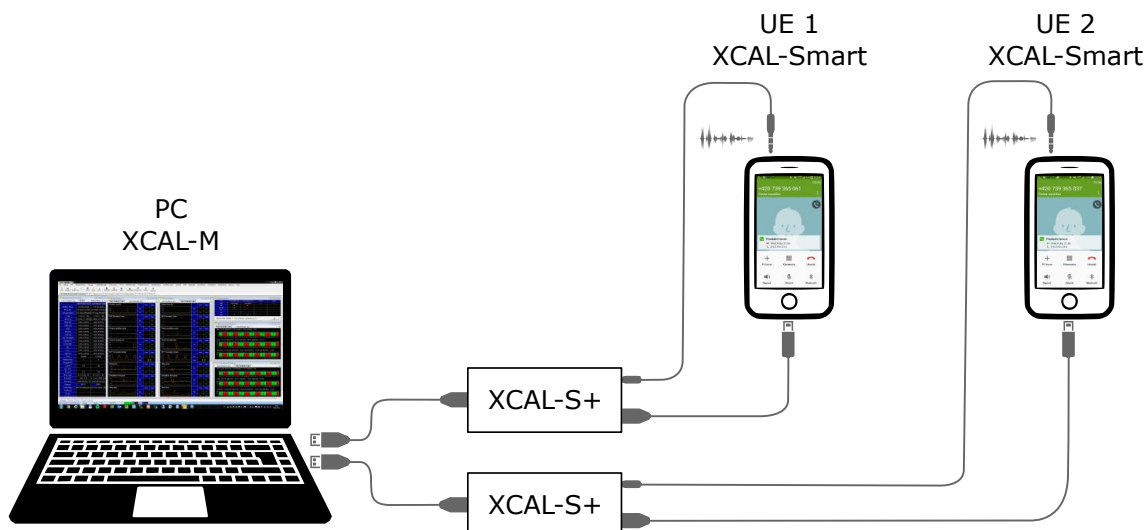
4.1 Měření hlasové kvality v síti LTE

Za účelem návrhu a optimalizace výpočetního algoritmu R-faktoru byla provedena rozsáhlá měření hlasové kvality technologie VoLTE telekomunikačního operátora T-Mobile Czech Republic, a. s. Měření byla koncipována tak, aby pokrývala široké množství reálných podmínek, tj. měření uvnitř budov, venkovní měření a měření za jízdy (autem, hromadnou dopravou). Zároveň byly pro rozmanitost měřených hodnot hlasové kvality voleny měřené indoor lokality s ohledem na místní rádiové podmínky sítě LTE.

4.1.1 Měřicí systém Accuver XCAL

Nástrojem pro měření byl systém Accuver XCAL, určený pro monitorování a analýzu činnosti bezdrátových komunikačních sítí. Systém umožňuje prostřednictvím měřicích uživatelských zařízení v reálném čase zobrazit a zaznamenat rádiovou i síťovou komunikaci a na základě těchto hodnot vyhodnotit parametry QoS.

Zapojení měřicího systému je znázorněno na obrázku 4.1. Řídící funkci obstarává počítač s běžící aplikací XCAL-M, ve které probíhá nastavení parametrů měření, měřicího scénáře a zpracování měřených dat. Samotné měření probíhá na připojených uživatelských zařízeních, konkrétně se jedná o mobilní telefony Samsung Galaxy S5 s operačním systémem Android.



Obr. 4.1: Zapojení měřicího systému Accuver XCAL

Specifikace měřicího systému:

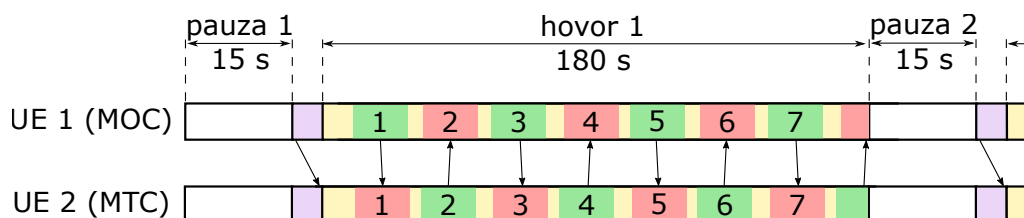
- PC
 - Program XCAL-M O3.3.6.17N (24. 3. 2016)
 - Program XCAP-U D5.11.120.0 (9. 3. 2016)
 - Algoritmus P.863 POLQA v2.4, knihovna 1.22
- UE 1, UE 2
 - Aplikace XCAL-Smart O4.10.01
 - Aplikace XCAL MDN Info O1.2.07
 - Model telefonu SM-G900F (Samsung Galaxy S5)
 - Chipset model: MSM8974PRO
 - OS: Android 5.0
 - Verze jádra 3.4.0-5952091
 - Číslo sestavení: LRX21T.G900FXXS1PPC7
 - Verze základního pásma: G900FXXU1BOJ1
- XCAL-S+
 - Model XDM-100S
- GPS přijímač
 - NaviLock NL-402U

Pro měření technologie VoLTE musí mít telefony firmware a UICC s ISIM operátora umožňujícího VoLTE volání, nainstalovanou aplikaci XCAL-Smart, povolen USB debugging pro spojení s PC a v závislosti na typu připojeného telefonu případně další úpravy nastavení nebo firmware umožňující zachytávání veškeré TCP/IP komunikace. Pro měření hlasové kvality jsou telefony připojeny k PC skrze USB Audio

Codec zařízení XCAL-S⁺. Audio záznamy jsou z telefonu do XCAL-S⁺ přenášeny ze sluchátkového výstupu, zatímco ostatní data a řídicí příkazy se posílají skrze USB sériový port a ADB rozhraní. V případě venkovních měření je možné pro záznam polohy k počítači připojit také GPS přijímač.

4.1.2 Měřicí scénář

Na obrázku 4.2 je znázorněno nastavení měřicí sekvence, která se neustále opakuje až do manuálního ukončení měření. Před každým vytáčením předchází 15 s přestávka sloužící k uvedení obou UE do nečinného stavu po ukončení předchozího hovoru. UE 1 následně automaticky vytočí telefonní číslo UE 2, které očekává příchozí spojení. Pro sestavení hovoru je vyhrazen maximální čas 30 s, v technologii VoLTE však tato procedura obvykle trvá méně než 3 s.



Obr. 4.2: Grafické znázornění měřicí sekvence

Doba trvání hovoru je omezena na 180 s, během kterých probíhá měření hlasové kvality algoritmem POLQA. Před zahájením přehrávání hlasové nahrávky je však na začátku hovoru držena ještě ochranná doba 5 s. Algoritmus POLQA následně v duplexním režimu vyhodnocuje kvalitu spojení mezi oběma telefony, dokud nedojde k ukončení hovoru. Činnost algoritmu POLQA je teoreticky popsána v kapitole 1.5.1, praktické měření je blíže specifikováno v kap. 4.1.4.

4.1.3 Měřené parametry

Přestože jsou v programu XCAL-M měřená data vyhodnocována v reálném čase a prezentována prostřednictvím nastavitelného GUI, pro potřeby pozdější podrobné analýzy byl využit nástroj Accuver XCAP-U. Jeho hlavní funkcí pro návrh zmiňovaného algoritmu vyhodnocení hlasové kvality byla extrakce informací z hlaviček jednotlivých RTP paketů a ze zpráv protokolu RTCP. Kromě údajů síťové komunikace byly extrahovány také výsledky měření algoritmu POLQA, které byly použity jako referenční měření hlasové kvality pro návrh algoritmu. Pro vysokou přesnost a plnou kontrolu nad činností navrhovaného algoritmu byly extrahované údaje z jednotlivých paketů podrobeny dalšímu zpracování již mimo prostředí měřicího systému Accuver.

Před samotnou analýzou datového toku RTP je však pro systém Accuver nezbytné nejdříve dekodovat a analyzovat SIP komunikaci při sestavování spojení (popsáno v kap. 3.3.1). Výpis 4.1 obsahuje zprávu SDP protokolu, která byla součástí odpovědi 200 OK odeslané z UE 2. Význam jednotlivých polí je definován RFC 4566 [35]. Pro měření je důležité zejména pole `m=`, obsahující typ přenášeného média (audio), UDP port, transportní protokol (RTP) a seřazený list kódů příslušících formátům podporovaných kodeků, a dále také pole atributů (`a=`).

Výpis 4.1: SDP zpráva obsažená v odpovědi 200 OK odeslaná z UE 2

```
1 v=0
2 o=SAMSUNG-IMS-UE 1234562 1 IN IP4 100.77.188.30
3 s=SS VOIP
4 c=IN IP4 100.77.188.30
5 b=AS:41
6 b=RS:612
7 b=RR:1837
8 t=0 0
9 m=audio 1234 RTP/AVP 116 111
10 b=AS:41
11 b=RS:612
12 b=RR:1837
13 a=rtpmap:116 AMR-WB/16000/1
14 a=fmtp:116 mode-change-capability=2;max-red=0
15 a=rtpmap:111 telephone-event/16000
16 a=fmtp:111 0-15
17 a=curr:qos local sendrecv
18 a=curr:qos remote sendrecv
19 a=des:qos mandatory local sendrecv
20 a=des:qos mandatory remote sendrecv
21 a=sendrecv
22 a=ptime:20
23 a=maxptime:240
```

Atributy blíže upřesňují podporované mediální formáty. Atribut `rtpmap` definuje pro dynamicky číslované kodeky jejich přesný formát, vzorkovací frekvenci a specifickou verzi. `116 AMR-WB/16000/1` jednoznačně určuje kodek AMR-WB v úsporném (Bandwidth-Efficient) formátu s nulovou redundancí (`max-red=0`). Atribut `ptime:20` nese informaci o délce média přenášené jedním RTP paketem v ms.

Po navázání spojení je zahájena vzájemná výměna RTP a RTCP paketů mezi oběma telefony. Výpis 4.2 obsahuje informace extrahované programem XCAL-M z hlaviček protokolů IP (neuveden plný výpis), UDP, RTP a AMR-WB z jednoho audio rámce přijatého UE 2.

Výpis 4.2: Audio rámeček přijatý UE 2

```
1 Frame 690
2   Arrival Time: Apr 24, 2016 11:13:14.250603000 CEDT
3   Epoch Time: 1461489194.250603000 seconds
4   Protocols in frame: raw:ip:udp:rtp:amr
5 Raw packet data
6   No link information available
7 Internet Protocol Version 4,
8   Src Addr: 10.225.139.196, Dst Addr: 100.66.88.191
9   ...
10 User Datagram Protocol, Src Port: 10932, Dst Port: 1234
11   Source port: 10932 (0x2AB4)
12   Destination port: 1234 (0x04D2)
13   Length: 81
14   Checksum: 0x79cd [validation disabled]
15 Real-Time Transport Protocol
16   Stream setup by SDP (frame 15)
17   10.. .... = Version: RFC 1889 Version (2)
18   ..0. .... = Padding: False
19   ...0 .... = Extension: False
20   .... 0000 = Contributing source identifiers count: 0
21   0... .... = Marker: False
22   Payload type: AMR-WB (116)
23   Sequence number: 349
24   Extended sequence number: 65885
25   Timestamp: 6246208
26   Synchronization Source identifier: 0x0005dd8d (384397)
27 Adaptive Multi-Rate
28   Payload decoded as RFC 3267 bandwidth-efficient mode
29   1111 .... = CMR: No mode request (15)
30   .... 0... = F bit: Last frame in this payload
31   .... .100 0... .... = FT bits: AMR-WB 23.85 kbit/s (8)
32   .1.. .... = Q bit: 0k
33 Frame Data (60 Bytes)
```

Ztráty paketů

Program Accuver XCAP-U vyhodnocuje pořadí přijatého paketu v RTP toku dat na základě jeho sekvenčního čísla. Proto jsou sekvenční čísla přijatých paketů zároveň indikátorem ztrát při přenosu. Detekce ztrát paketů funguje dle nastavení tolerance jejich zpřeházení. Přestože výchozím nastavením povoleného zpřeházení je hodnota 1, pro účely vyhodnocení hlasové kvality byla zvolena hodnota 4. Příkladem může být příjem paketů v pořadí 45-47-49-48-50-51-46-52. Pakety 46 a 48 jsou přijaty v nesprávném pořadí, pouze paket 46 je však vyhodnocen jako ztracený. Důvodem

je zohlednění funkce adaptivního jitter bufferu, kdy opožděně přijaté pakety s přípustným zpožděním nejsou zahozeny, namísto toho je přehrávání hlasového signálu pozdrženo o omezenou dobu s ohledem na plynulost komunikace, tzv. funkce DTW.

Zdrojové zpoždění

Nástroj XCAP-U zároveň vyhodnocuje časové značky protokolů RTP a RTCP v kombinaci s časy odeslání a příjmu rámců na jednotlivých UE (tzv. chipset time) pro určení několika druhů zpoždění popisujících latenci a plynulost spojení. Zdrojové zpoždění, nebo také zpoždění kodeku, představuje interval vytváření jednotlivých RTP rámců ve zdroji a indikuje správný interval příjmu a přehrávání paketů. Rovnice 4.1 uvádí výpočet zdrojového zpoždění jako rozdílu časových značek právě přijatého a posledního přijatého RTP paketu poděleného vzorkovací frekvencí kodeku:

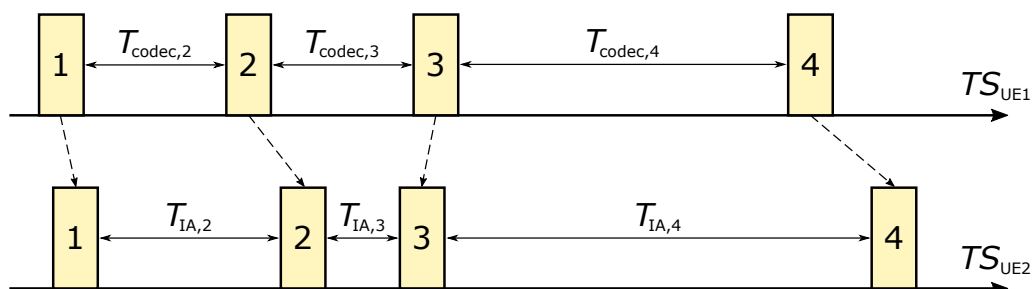
$$T_{\text{codec},i} = 10^3 \cdot \frac{TS_{\text{RTP},i} - TS_{\text{RTP},i-1}}{f_s}. \quad [\text{ms}] \quad (4.1)$$

Paketové zpoždění

Paketové zpoždění (angl. interarrival delay) reprezentuje čas mezi příchodem aktuálního a posledního přijatého paketu. V ideálním případě je paketové zpoždění rovno zdrojovému zpoždění. Počítá se dle rovnice 4.2 na základě rozdílu chipset časových značek příchozí posledních dvou přijatých paketů:

$$T_{\text{IA},i} = TS_{\text{UE},i} - TS_{\text{UE},i-1}. \quad [\text{ms}] \quad (4.2)$$

Na obrázku 4.3 je ilustrováno zdrojové a paketové zpoždění paketů přijímaných UE 2 při proměnném síťovém zpoždění.



Obr. 4.3: Ilustrace zdrojového a paketového zpoždění

Sítové zpoždění

Bez synchronizovaných hodin mezi vysílací a přijímací stranou (na obr. 4.3 TS_{UE1} a TS_{UE2}) není možné z RTP paketů získat přesnou informaci o čase, za který jednotlivé pakety dorazí od odesílatele k příjemci. Proto se používá tzv. sítové zpoždění, které je definováno RFC 3550 [37] jako rozdíl paketového a zdrojové zpoždění:

$$T_{\text{network},i} = T_{\text{IA},i} - T_{\text{codec},i}. \quad [\text{ms}] \quad (4.3)$$

Sítové zpoždění nenesení informaci o fixních složkách transportního zpoždění, namísto toho je ukazatelem proměnných složek v síti, např. čekání ve frontách. Při plynulém přenosu paketů sítí se sítové zpoždění téměř nemění a je blízké nule.

Jitter

Jitter je váženým průměrem absolutní hodnoty sítového zpoždění T_{network} a používá se jako numerický ukazatel proměnlivosti transportního zpoždění. RFC 3550 [37] jitter definuje následovně:

$$J_i = J_{i-1} + \frac{|T_{\text{network}}| - J_{i-1}}{16}. \quad (4.4)$$

Parametr $1/16$ je nastaven optimálně tak, aby se omezil efekt šumu v podobě ojedinělých náhodných skoků T_{network} a zároveň aby hodnota jitteru J_i měla dostatečnou míru konvergence.

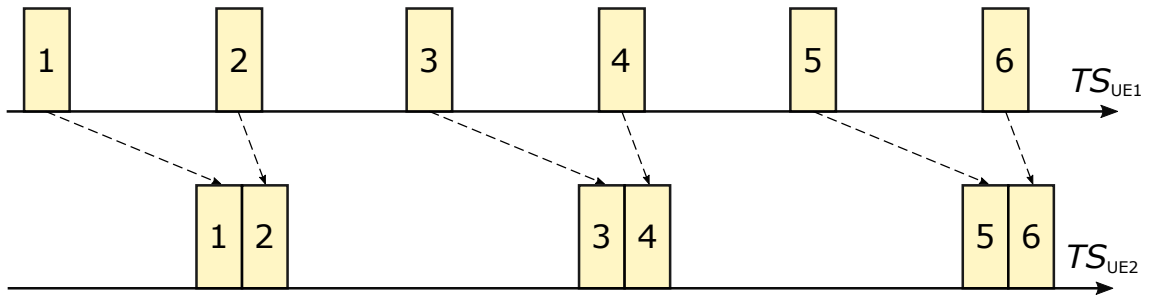
Spojování paketů

Program XCAP-U dále uvádí v sekundových intervalech počet a celkový poměr paketů, které byly přijaty v krátkém sledu za sebou, tj. s $T_{\text{IA}} = 0 \text{ ms}^1$, viz obr. 4.4.

RTCP transportní zpoždění

Přestože měřicí systém Accuver umožňuje při propojení záznamů měření z obou měřicích telefonů vyhodnocení transportního jednosměrného zpoždění RTP paketů, analýza výsledků testovacích měření ukázala, že synchronizace hodin připojených telefonů není dostatečně přesná. Výsledkem jsou „plovoucí“ hodnoty zpoždění, v některých případech dokonce až záporné. Proto byla pro odhad jednosměrného zpoždění zvolena metoda měření prostřednictvím vyměňovaných RTCP zpráv, kterou stanovuje RFC 3550 [37].

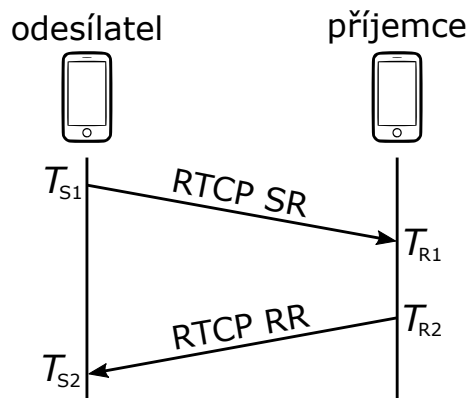
¹XCAP-U dle analýzy výsledků uvažuje jako připojené také pakety s $T_{\text{IA}} \leq 1 \text{ ms}$. Hlubší rozbor však ukázal, že v některých případech, obzvláště při dočasně vyšším vytížení procesoru UE, dochází ke spojování paketů i při vyšších indikovaných hodnotách T_{IA} .



Obr. 4.4: Ilustrace spojování paketů

Princip měření je znázorněn na obrázku 4.5. RTCP SR zprávy jsou průběžně zasílány protistraně v intervalu několika sekund. Pro vyhodnocení obousměrného zpoždění RTT se využívá vztahu 4.5. Odesílatel, který chce vyhodnotit zpoždění, zašle druhé straně RTCP SR. Příjemce² při odesílání další RTCP zprávy zahrne tzv. $DLSR$, tj. rozdíl T_{R2} a T_{R1} , který odesílatel následně využije k vyjádření obousměrného zpoždění RTCP paketů:

$$RTT = T_{S2} - T_{S1} - (DLSR) = T_{S2} - T_{S1} - (T_{R2} - T_{R1}). \quad [\text{ms}] \quad (4.5)$$



Obr. 4.5: Metoda měření jednosměrného zpoždění prostřednictvím RTCP zpráv

Pro odhad jednosměrného transportního zpoždění v síti se využívá předpokladu, že doba transportu paketu z UE 1 k UE 2 je přibližně stejná, jako doba transportu ve směru opačném, tedy:

$$T_{\text{one-way}} = \frac{RTT}{2}. \quad [\text{ms}] \quad (4.6)$$

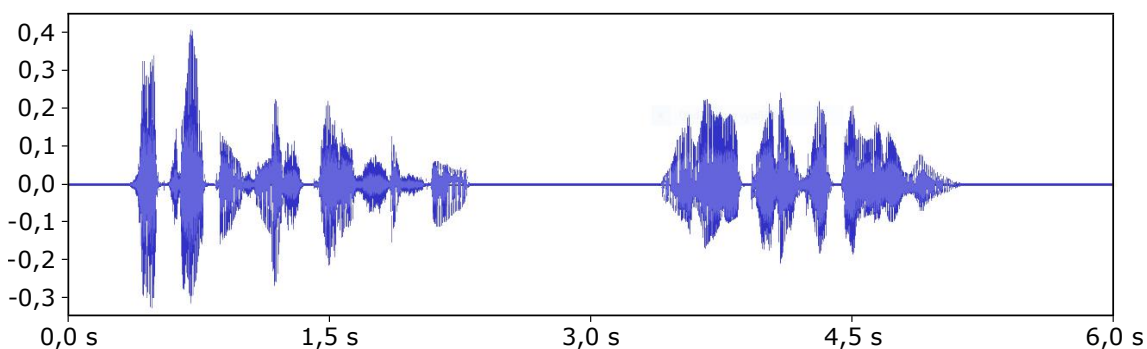
²V praxi jsou obvykle aktivní oba účastníci spojení, jsou tedy zároveň oba odesílatelé a vyměňují si jen RTCP SR zprávy. Metoda měření se však nemění.

Typ AMR-WB rámce

Mimo zpracování informací z hlaviček RTP/RTCP paketů program XCAP-U umožňuje i čtení jejich obsahu, tj. AMR-WB rámců, které se opět skládají z hlavičky a datového nákladu. Protože technologie VoLTE operátora T-Mobile v době měření využívá pouze kódování AMR-WB 23.85 kbit/s (typ 8) bez markeru M, pro vyhodnocení hlasové kvality je jedinou užitečnou, avšak zásadní, informací zda daný rámeček je typu 8, nesoucí hlasová data, nebo typu 9, obsahující tzv. Silence Insertion Descriptor, případně také zda tento rámeček žádná data neobsahuje (typ 15).

4.1.4 Měření hlasové kvality algoritmem POLQA

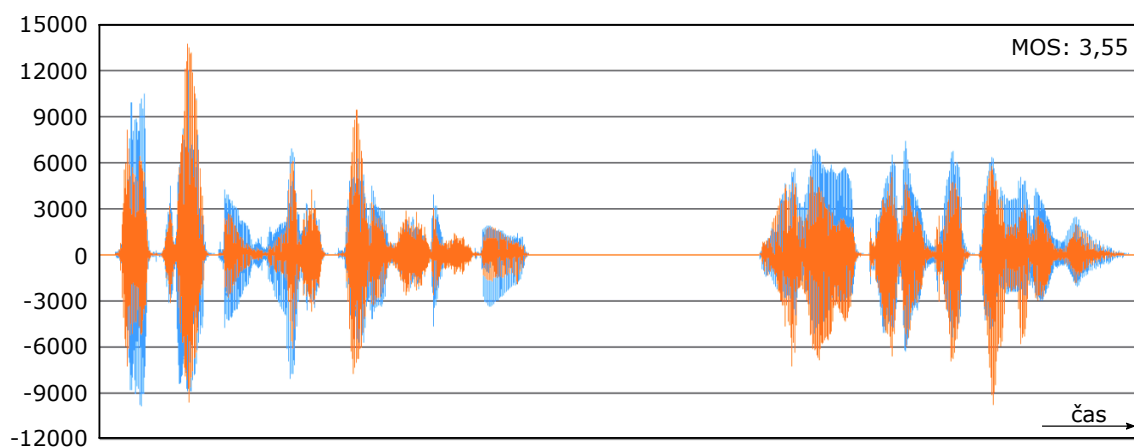
Referenční měření hlasové kvality byly provedeny využitím algoritmu ITU-T P.863 POLQA. Jak již bylo uvedeno při popisu měřicího scénáře v kapitole 4.1.2, měření hlasové kvality probíhala v duplexním režimu. Pro měření byla využita 6 sekundová SWB nahrávka v českém jazyce poskytovaná společností OPTICOM GmbH, spoluautorem algoritmu POLQA. Nahrávka se v souladu s doporučeními ITU-T pro subjektivní a objektivní testování poslechové kvality sestává ze dvou krátkých vět muže a ženy, viz obrázek 4.6.



Obr. 4.6: Hlasová nahrávka použitá při měření algoritmem POLQA

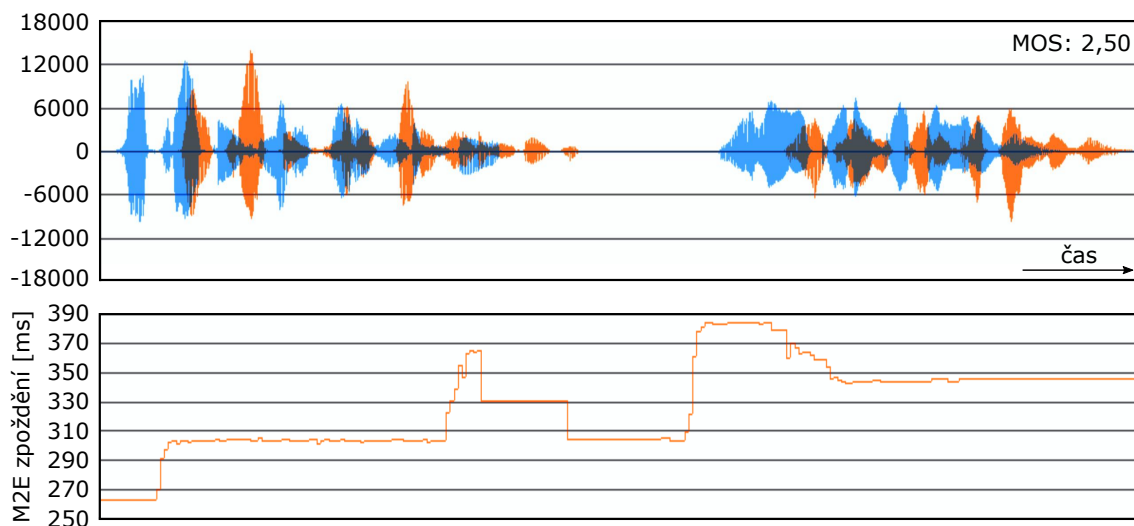
Délka použité nahrávky hrála důležitou roli v rozhodnutí o časové délce analyzujícího okna navrhovaného algoritmu výpočtu R-faktoru. Aby byla zajištěna nejvyšší možná korelace obou typů měření a výsledky nebyly negativně ovlivněny nedokonalou implementací tzv. průběžného R-faktoru, navrhovaný algoritmus operuje ve stejném časovém intervalu 6 s. Na obrázku 4.7 je k vidění příklad srovnání originální (modré) a degradované (oranžové) nahrávky po přenosu sítě LTE. Tato konkrétní nahrávka obdržela výstupní objektivní hodnocení poslechové kvality MOS 3,55. Hodnocení je sníženo oproti teoretické maximální hodnotě 4,75 především z důvodu komprese

a omezené frekvenční šířky pásma kodeku AMR-WB 23.85 kbit/s (přibližně do 7000 kHz) a také kvůli ztrátě paketů na začátku nahrávky.



Obr. 4.7: Srovnání originální a degradované nahrávky algoritmu POLQA

Kromě hodnoty hodnoty MOS bylo z měření algoritmem POLQA využito také měření M2E zpoždění, znázorněné na obrázku 4.8. Měření je výstupem bloku časového zarovnání, který páruje korespondující úseky originální a degradované nahrávky získávané ze sluchátkového výstupu mobilního telefonu. Program XCAP-U reportuje pod pojmem M2E zpoždění počáteční zpoždění nahrávky a dále průměrné, maximální a minimální zpoždění v nahrávce.



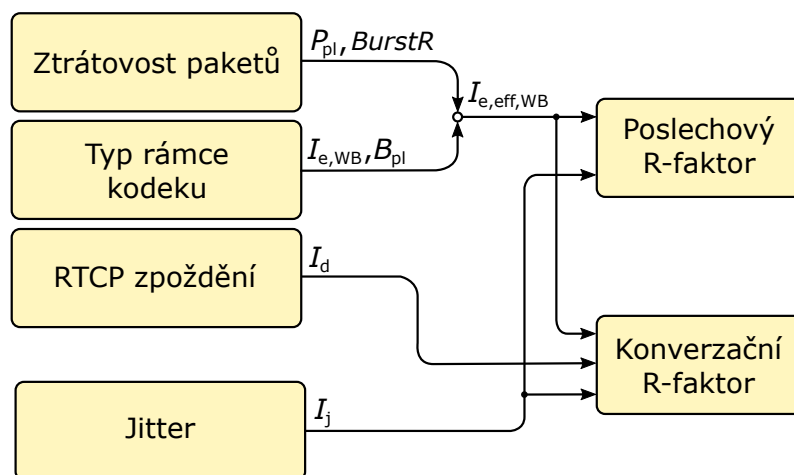
Obr. 4.8: Analýza M2E zpoždění algoritmem POLQA

Celkem bylo při zahrnutí i prvotních testů při sestavování měřicího systému a hledání optimální konfigurace naměřeno celkem 27 835 vzorků hlasové kvality. Pro

finální návrh a optimalizaci algoritmu výpočtu R-faktoru bylo však využito jen 17 006 z nich. Podstatnou část nepoužitých vzorků tvoří měření hlasové kvality ve starší verzi programu Accuver XCAL-M O3.3.4. Obě verze používají nejnovější verzi algoritmu POLQA 2.4, ale zatímco aktuální verze O3.3.6 pro měření algoritmu POLQA využívá knihovny kalibračních měření verze 1.22, starší verze pracovala s knihovnou verze 1.18, která u některých hlasových vzorků nejvyšší kvality produkovala výrazně vyšší hodnocení MOS-LQO_{sw}³. Zbylou část nepoužitých vzorků potom tvoří měření, která v některém ohledu selhala, tj. buď systém Accuver pro daný hovor nedokázal zachytit RTP komunikaci nebo byla hodnota MOS rovna 0, indikující neúspěšné měření, případně měření obsahovala jiné nesmyslné hodnoty.

4.2 Aplikace E-modelu

Návrh algoritmu výpočtu hlasové kvality prostřednictvím R-faktoru vychází z matematických vztahů definovaných v parametrickém modelu pro predikci hlasové kvality telefonního spojení, tzv. E-modelu, podrobně popsáném v kapitole 2. Protože je E-model objektivním modelem konverzační kvality a algoritmus POLQA, používaný pro referenční měření, je modelem poslechové kvality, výstupem navrhovaného algoritmu jsou dvě hodnoty reprezentující zvláště poslechovou a konverzační kvalitu spojení, viz obrázek 4.9. Konverzační R-faktor na rozdíl od poslechového zahrnuje navíc vliv M2E zpoždění na kvalitu konverzace.



Obr. 4.9: Zjednodušené znázornění vstupů a výstupů navrhovaného algoritmu

³Při použití knihovny verze 1.18 bylo přibližně 11,6% vzorků MOS > 4 s tím, že nejvyšší hodnoty byly rovny dokonce 4,39. Měření s knihovnou verze 1.22 přesahují MOS = 4 jen v 0,08% případů a nejvyšší naměřená hodnota je 4,04.

Navrhovaný algoritmus definuje R-faktor na základě rovnice

$$R = 129 - I_{d,WB} - I_{e,eff,WB} - I_j. \quad (4.7)$$

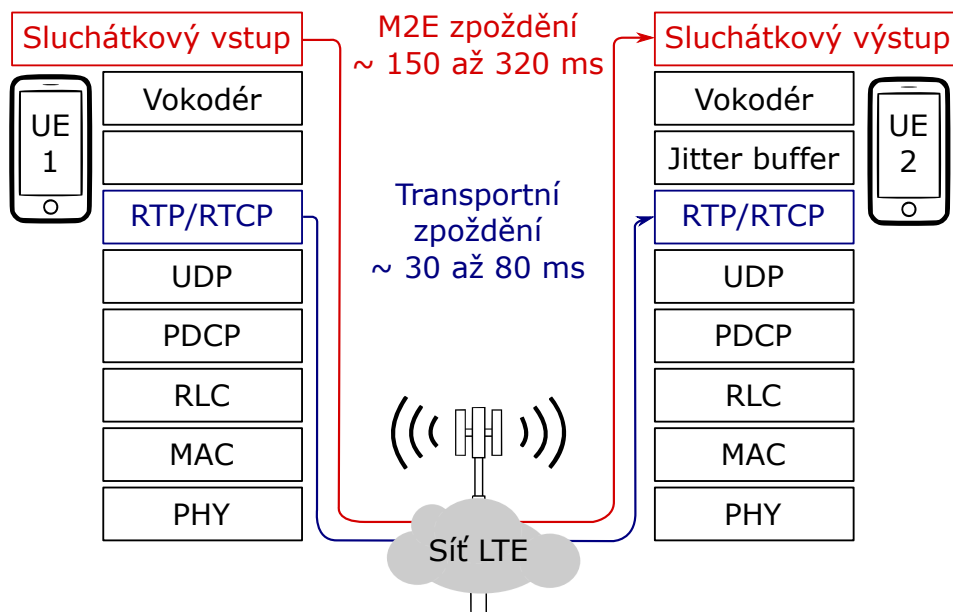
Jedná se o adaptaci rovnice 2.4, ve které bylo využito vztahu 2.11 pro dosažení $R_{0,max} = 129$ za $R_{0,WB}$ a přidán faktor degradace jitterem I_j , jehož opodstatnění je uvedeno v kap. 4.3.1. Rovnice 4.7 předpokládá dokonalé potlačení echa a místní vazby, proto faktor $I_{d,WB}$ reprezentuje pouze vliv M2E zpoždění. Faktor $I_{e,eff,WB}$ modeluje vliv hlasového kodeku a ztrát paketů. Pro jasné rozlišení budou v této práci hodnoty R vypočítané na základě rovnice 4.7 uváděny jako tzv. konverzační R-faktor R_{CQ} a v případě dosažení $I_{d,WB} = 0$ jako poslechový R-faktor R_{LQ} .

4.2.1 M2E zpoždění

Transportní zpoždění RTP/RTCP paketů $T_{one-way}$, které je možné analýzou síťové komunikace měřit, je jen jednou ze složek M2E (absolutního) zpoždění. Kromě něj se absolutní zpoždění skládá také z fixního zpoždění hlasového kodeku (viz rov. 3.1) a zpoždění T_{buffer} vkládaného jitter bufferem na přijímací straně:

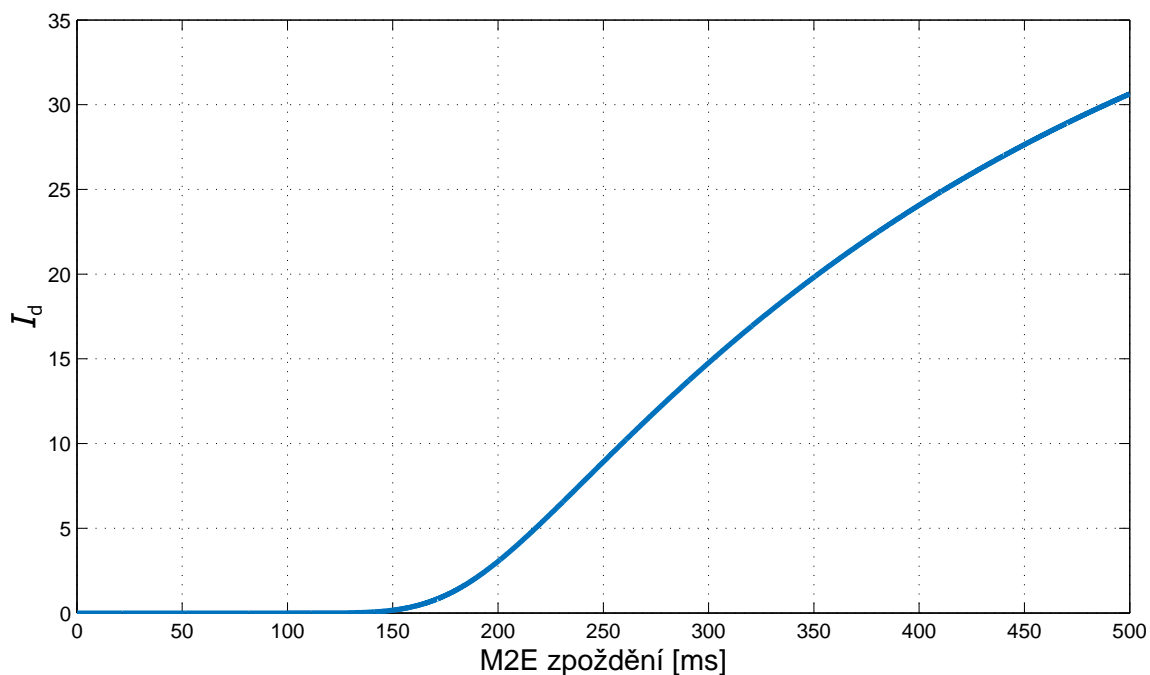
$$T_a = T_{one-way} + 2T_{frame} + T_{LA} + T_{buffer}. \quad [\text{ms}] \quad (4.8)$$

Obrázek 4.10 ilustruje pomocí vrstevného modelu technologie VoLTE transportní cestu hlasového signálu a uvádí typické hodnoty zpoždění.



Obr. 4.10: Vrstvový model VoLTE spojení s typickými hodnotami zpoždění

Pro aplikaci E-modelu je nutné navrhnout odhad M2E zpoždění T_a na základě proměnných složek v rov. 4.8. Kapitola 4.3.3 představuje návrh odhadu na základě měřeného zpoždění RTCP paketů korigovaném hodnotami jitteru. Odhadovaná hodnota T_a po dosazení do rovnice 2.14 udává faktor degradace kvality vlivem zpoždění I_d s průběhem zobrazeným na obrázku 4.11. Pro výpočet poslechového R-faktoru se I_d považuje nulové, což odpovídá minimálnímu zpoždění technologie VoLTE pohybujícímu se kolem 150 ms.



Obr. 4.11: Závislost faktoru degradace I_d na hodnotě M2E zpoždění T_a

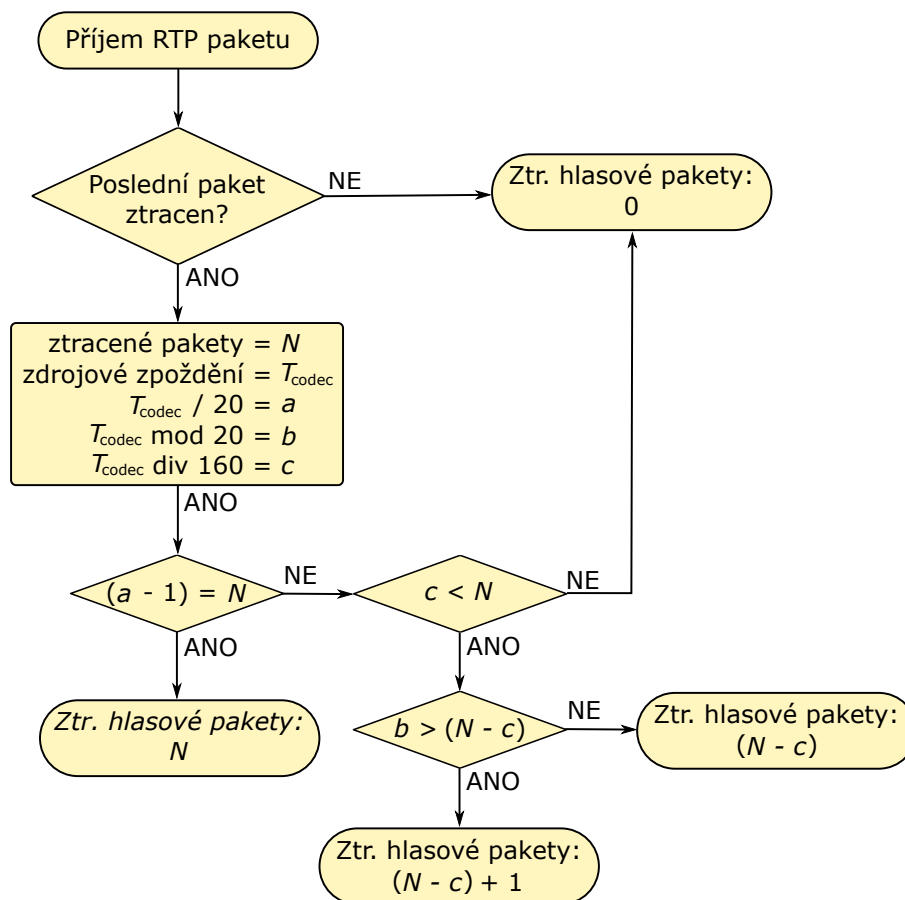
4.2.2 AMR-WB kodek

Operátor T-Mobile využívá pro technologii VoLTE výhradně kodek AMR-WB 23.85 kbit/s. E-model pro popis vlivu kodeku na hlasovou kvalitu využívá tzv. faktor degradace použitou technikou $I_{e,WB}$ spolu s koeficientem odolnosti proti ztrátám paketů B_{pl} . Tabulka 2.2 obsahuje provizorní hodnoty obou parametrů stanovené v dop. ITU-T G.113 [26].

V souvislosti s funkcí kodeku je při vyhodnocení hlasové kvality důležité zohlednit také činnost VAD, zejména při ztrátách paketů. Nejenže je distribuce ztrát paketů v úseku, ve kterém se zasílá pouze informace o šumu na pozadí, ovlivněna nižší přenosovou rychlostí, ale také má ztrátovost v těchto úsecích minimální vliv na vnímanou kvalitu účastníka na přijímací straně [45]. Dekodér AMR-WB totiž v tomto případě nahradí ztracené části odhadem přijatelné úrovně šumu. Při analýze hlasové

kvality je proto nutné úsekům aktivní řeči přiřadit větší váhu ve srovnání s úseky přijatelného šumu.

Komplikací tohoto přístupu je rozlišení aktivní a neaktivní řeči na přijímací straně, zejména v případě pro mobilní komunikace typických úseků ztrát paketů s nenáhodným rozložením, tj. se shluky ztracených paketů. Na přijímací straně není totiž možné ve všech případech s jistotou rozhodnout, zda ztracené pakety přenášely hlas, či nikoliv. Proto byl navržen algoritmus (obr. 4.12), který na základě znalostí o rozdílných přenosových rychlostech aktivní a neaktivní řeči minimalizuje možnou chybu určení charakteru ztracených paketů.



Obr. 4.12: Diagram algoritmu pro určení charakteru ztracených paketů

4.2.3 Modelování ztrát paketů

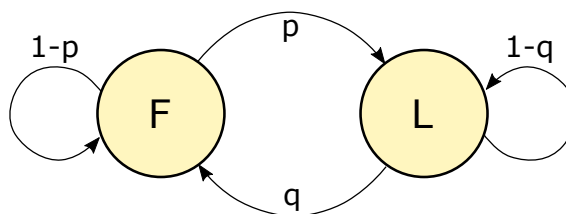
Modelování vlivu ztrát paketů na hlasovou kvalitu úzce souvisí s druhem použitého hlasového kodeku. E-model pro popis efektu ztrátovosti paketů zavádí tzv. efektivní faktor degradace použitou technikou, $I_{e,eff,WB}$, který je popsán v kap. 2.4.4. Vztah 2.16, vyjadřující $I_{e,eff,WB}$, vychází z předpokladu, že ztráty paketů při přenosu

jsou distribuovány náhodně. Pro mobilní sítě jsou však typické shluky ztrát paketů v krátkém časovém intervalu, např. při handoveru.

Pro popis tohoto chování je možno využít tzv. dvoustavový Markovův model, speciální druh Gilbert-Elliottova modelu [46]. Tento model popisuje ztráty paketů prostřednictvím dvou přenosových stavů: přenášejícího (Found) a ztrácejícího, (Lost) a pravděpodobností přechodu mezi nimi, viz obr. 4.13. Poměr shluků ztrát, angl. burst ratio, je definován následovně [46]:

$$BurstR = \frac{\text{Průměrná délka shluku ztrát}}{\text{Očekávaná průměrná délka shluku při náhodné distribuci ztrát}}, \quad (4.9)$$

kde shluk ztrát je definován jako počet ztracených paketů v řadě za sebou. Pokud je $BurstR = 1$, ztráty paketů jsou náhodné. $BurstR > 1$ znamená, že ztráty paketů jsou ve shlucích.



Obr. 4.13: Dvoustavový Markovův model ztrát paketů

Poměr shluků ztrát $BurstR$ je možné zároveň vyjádřit i přímo na základě Markovova modelu jako [47]:

$$BurstR = \frac{1}{p + q}, \quad (4.10)$$

kde p a q jsou pravděpodobnosti z obr. 4.13. Ty lze vypočítat prostřednictvím ztrátovosti paketů mlr a průměrné délky shluku ztrát mbl :

$$p = \frac{mlr}{mbl(1 - mlr)}, \quad q = \frac{1}{mbl}. \quad (4.11)$$

Kombinací rovnic 4.10, 4.11 a dosazením $P_{pl} = 100 \cdot mlr$ vznikne vztah [47]:

$$BurstR = mbl \left(1 - \frac{P_{pl}}{100} \right), \quad (4.12)$$

do kterého je dále možné za mbl dosadit:

$$mbl = \sum_{k=0}^{\infty} k P(Y = k), \quad (4.13)$$

kde $P(Y = k)$ je pravděpodobnost ztráty k paketů za sebou.

Výsledná hodnota $BurstR$ se dosazuje do rovnice 2.16 v následující podobě [14]:

$$I_{e,eff,WB} = I_{e,WB} + (95 - I_{e,WB}) \cdot \frac{P_{pl}}{\frac{P_{pl}}{BurstR} + B_{pl}}. \quad (4.14)$$

Z rovnice je evidentní, že s rostoucí hodnotou $BurstR$ roste efektivní faktor degradace použitou technikou $I_{e,eff,WB}$ a zároveň tedy klesá odhadovaná hlasová kvalita.

4.3 Adaptace algoritmu na základě měření

Algoritmus výpočtu R-faktoru, předložený v kapitole 4.2, je nutné pro dosažení přijatelné spolehlivosti predikce hlasové kvality kalibrovat referenčními měřeními pro danou technologii.

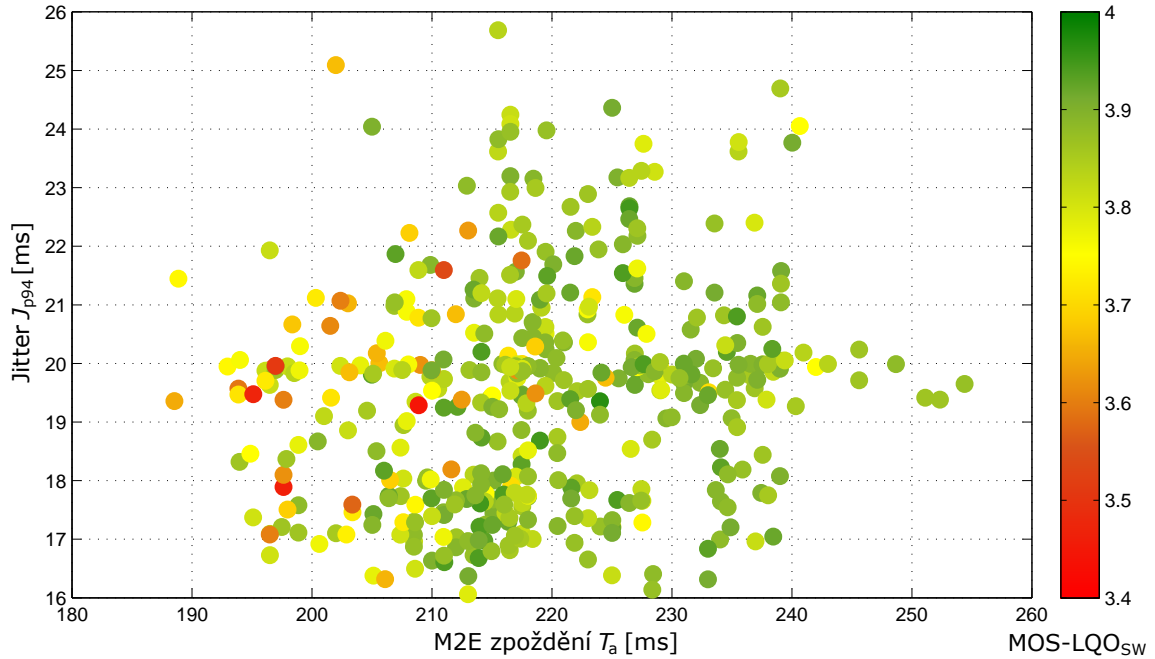
4.3.1 Modelování vlivu jitteru

Pro návrh výpočtu faktoru degradace jitterem J_j bylo nutné nejdříve stanovit parametry, na základě kterých bude úroveň jitteru posuzována. Rovnice 4.4, kterou používá systém Accuver pro výpočet jitteru, udává v čase váženou hodnotu jitteru takovou, aby dostatečně rychle konvergovala při změně charakteru komunikačního kanálu, avšak nebyla příliš ovlivněna náhodnými skoky síťového zpoždění. Z pohledu hlasové kvality hrají ale i tyto krátké výkyvy podstatnou roli v adaptivní funkci jitteru bufferu a eventuálně i kvalitě přenášeného hlasu. Proto byl faktor $1/16$ z rovnice 4.4 při určování velikosti jitteru nahrazen faktorem $1/8$ s rychlejší konvergencí:

$$J_i = J_{i-1} + \frac{|T_{network}| - J_{i-1}}{8}. \quad (4.15)$$

Systém Accuver zároveň nezohledňuje proměnnou dobu zdrojového zpoždění T_{codec} mezi úseky aktivní a neaktivní řeči. Z tohoto důvodu se průběžná hodnota jitteru RTP paketů počítá dle rovnice 4.15 jen v případě příjmu AMR-WB rámce typu 0 až 8, jinak je $J_i = J_{i-1}$. Z 6 sekundového analyzujícího okna se pak skalární velikost jitteru vyjádří jako 94. percentil J_{p94} ze všech rámců typu 0 až 8.

Na obrázku 4.14 je korelační diagram závislosti hodnocení MOS-LQO_{SW} na velikosti jitteru J_{p94} a M2E zpoždění na začátku nahrávky T_a , měřeném alg. POLQA, pro měření bez ztrát paketů. Z rozložení bodů reprezentujících jednotlivá 6 s měření lze vyčíst, že vysoké hodnoty jitteru nemají negativní dopad na hlasovou kvalitu v případě vyššího M2E zpoždění. Většina vzorků s MOS < 3,75 má M2E zpoždění menší než 210 ms.



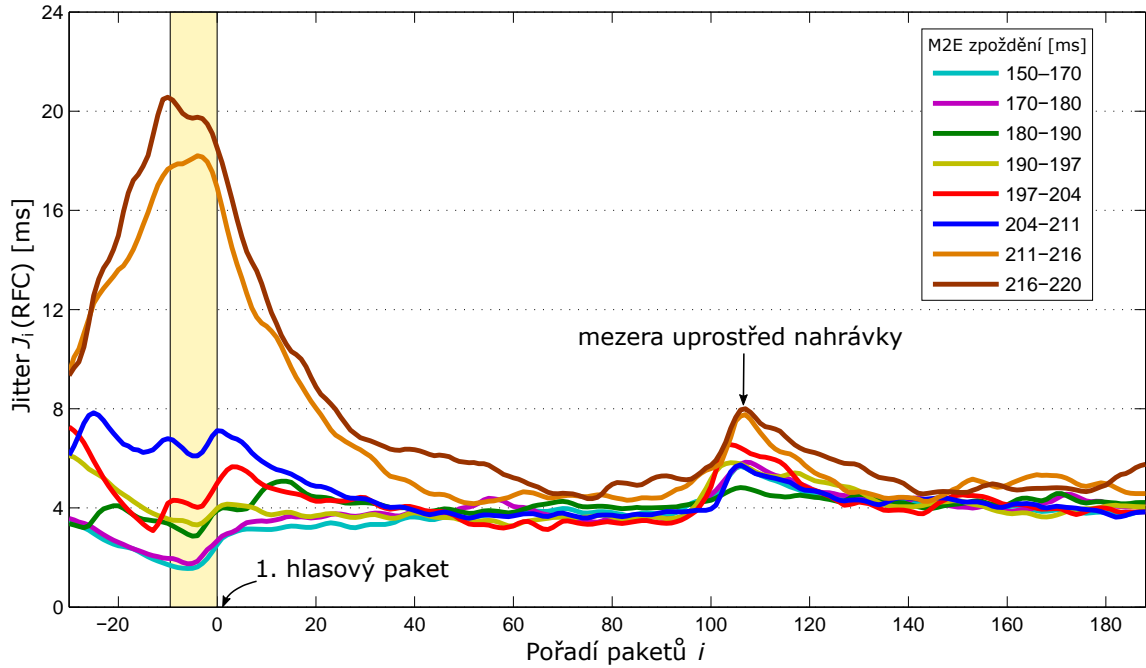
Obr. 4.14: Korelační diagram závislosti MOS na jitteru a M2E zpoždění

Pro zjištění charakteristických vlastností určujících velikost M2E zpoždění jednotlivých vzorků bylo analyzováno síťové zpoždění paketů v celém průběhu měřicího okna. Obr. 4.15 obsahuje vyhlazené křivky mediánů průběžných hodnot jitteru J_i (dle rov. 4.4) pro korespondující sekvenční čísla RTP paketů a kategorie M2E zpoždění. První hlasový paket v nahrávce je na vodorovné ose reprezentován hodnotou 1 a ostatní pakety jsou číslovány dle sekvenčního čísla relativně k němu. Jak je patrné z barevně vyznačené oblasti -7 . až 0 . paketu (úsek délky 1280 ms), s rostoucím jitterem v této části roste i velikost M2E zpoždění, zatímco v ostatní části sekvence paketů je korelace jitteru a M2E velmi nízká. Parametr J_{M2E} je definován:

$$J_{M2E} = \max_{-7 \leq i \leq 0} (J_i). \quad (4.16)$$

Velikost jitteru je do značné míry ovlivněna spojováním paketů při přenosu a je nutné jeho vliv kompenzovat. Např. chování ilustrované na obrázku 4.4 by vyústilo při zdrojovém zpoždění 20 ms na přijímací straně v jitter roven přibližně 20, přestože kromě mírně zvýšeného M2E zpoždění nebude pravděpodobně hlasová kvalita nijak postižena.

Systém Accuver reportuje poměr spojování paketů se sekundovým rozlišením, které není pro přesnou analýzu RTP paketů přípustné. Proto byla hodnota *BundLR* počítána dodatečně z jednotlivých paketů. Zároveň bylo analýzou záznamů měření zjištěno, že při vyšším zatížení procesoru telefonu je mnohdy $T_{IA} > 0$ ms, přestože dochází ke spojování paketů. *BundLR* tak bylo počítáno jako podíl paketů typu 0 až 8



Obr. 4.15: Průběh jitteru pro různé intervaly M2E zpoždění

s $T_{IA} \leq 5$ ms a všech přijatých hlasových paketů v daném časovém intervalu. Na korelačním diagramu 4.16 je vidět vztah mezi spojováním paketů $BundlR$ a jitterem J_{p94} . Z něj byl odvozen vztah vyjadřující velikost jitteru v měřeném úseku nezávisle poměru $BundlR$:

$$J_X = J_{p94} - X(BundlR), \quad (4.17)$$

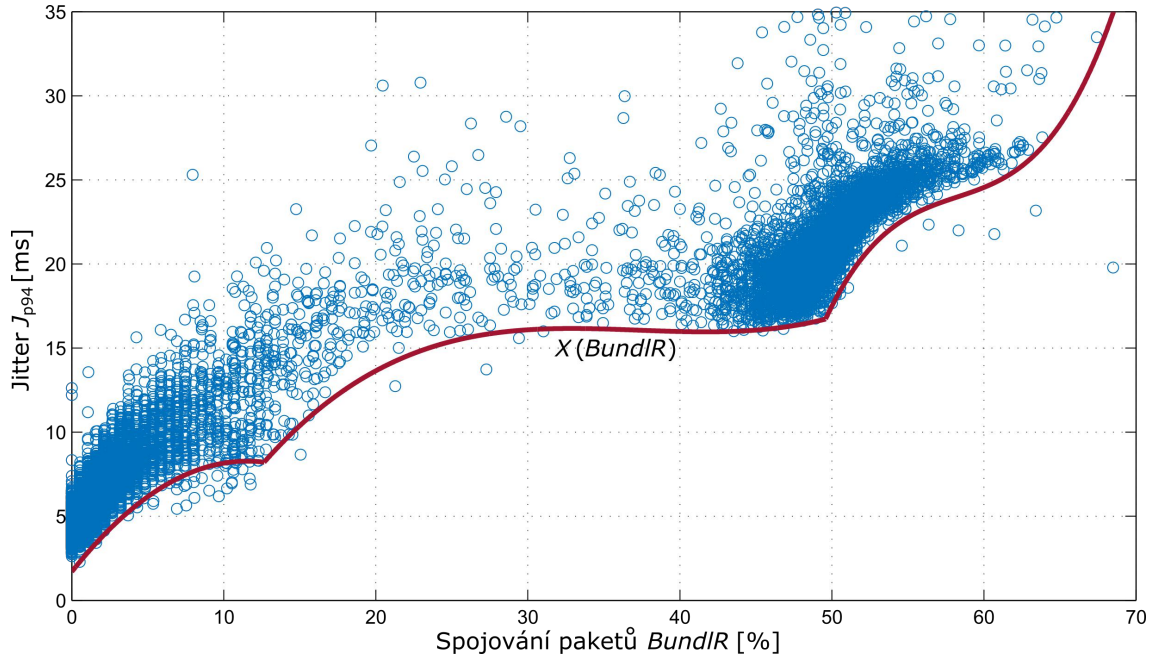
kde platí, že $J_X \geq 0$ a zároveň jsou pro intervaly v pořadí $BundlR \in \langle 0; 12,66\% \rangle$, $BundlR \in \langle 12,66\%; 49,58\% \rangle$ a $BundlR \geq 49,58\%$ definovány polynomické funkce

$$X(BundlR) = \begin{cases} -494 BundlR^2 + 114 BundlR + 1,7, \\ 581 BundlR^3 - 650 BundlR^2 + 239 BundlR - 12,8, \\ 7242 BundlR^3 - 12647 BundlR^2 + 7392 BundlR - 1422. \end{cases} \quad (4.18)$$

Výsledný faktor degradace jitterem I_j se pak spočítá na základě vztahu

$$I_j = \left(6 - \frac{\log(J_{M2E} + 1)}{\log(2)} \right) \left(\frac{\log(J_X + 4)}{\log(2)} - 2 \right) - 1, \quad (4.19)$$

přičemž opět platí, že faktor degradace jitter I_j je nezáporný.



Obr. 4.16: Korelační diagram spojujání paketů *BundlR* a jitteru J_{p94}

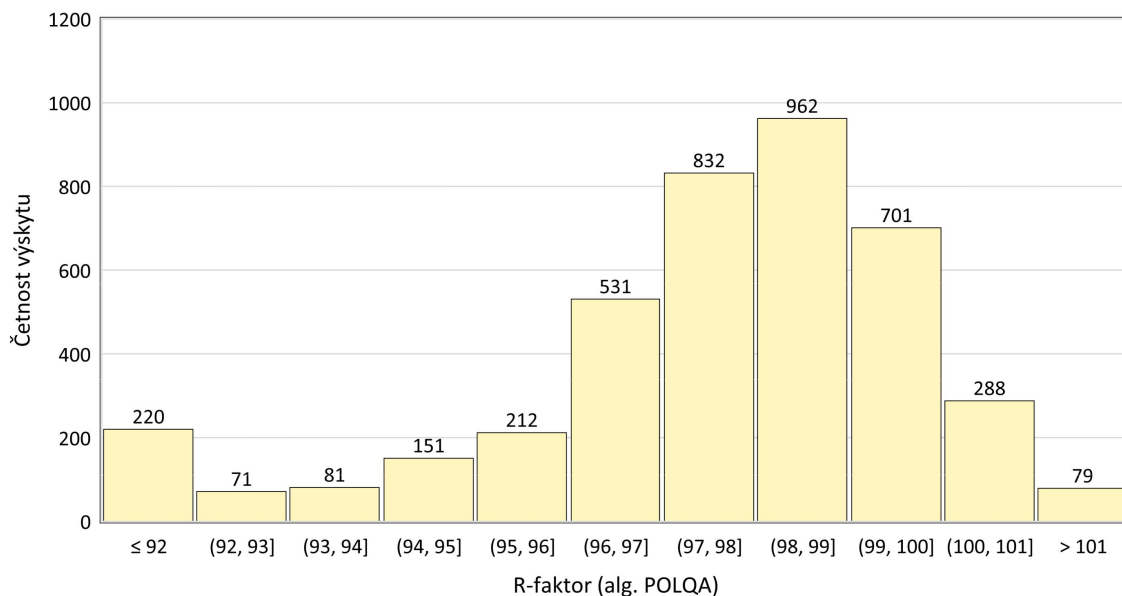
4.3.2 Optimalizace parametrů hlasového kodeku

Provizorní hodnoty koeficientů $I_{e,WB}$ a B_{pl} , obsažené v tabulce 2.2, se do značné míry rozcházejí s výsledky měření hlasové kvality algoritmem POLQA. Na obrázku 4.17 je zobrazen histogram hodnot R-faktoru, které byly převedeny pomocí rovnice 2.20 z hodnot $MOS-LQO_{SW}$ naměřených algoritmem POLQA, přičemž byly zahrnuty pouze měření bez výrazné degradace, tj. bez ztrát paketů ($P_{pl} = 0$) a s $I_j < 1$.

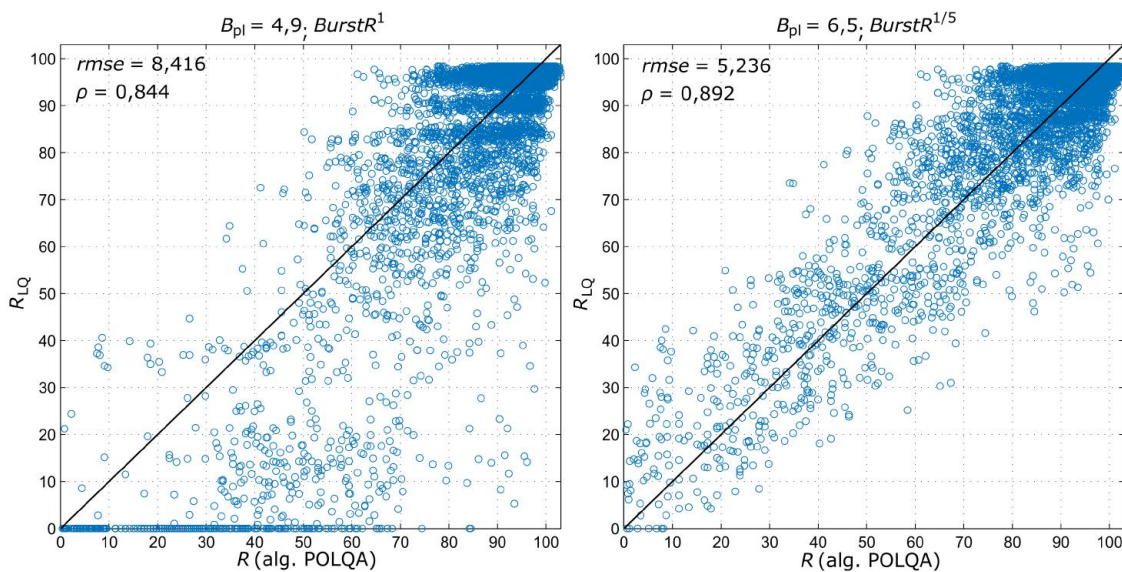
Doporučení ITU-T G.113 [26] uvádí pro monotický poslech (tj. jedním uchem) kodeku AMR-WB 23.85 kbit/s provizorní hodnotu $I_{e,WB} = 8$, což po dosažení do rovnice 4.14 dává $R = 121$. Medián naměřených hodnot v histogramu 4.17 je však roven 97,9. Protože je do této hodnoty zahrnuto i I_j , bylo na základě měření stanoveno

$$I_{e,WB} = R_{0,max} - R - \bar{I}_j = 129 - 97,9 - 0,5 = 30,6. \quad (4.20)$$

Faktor $I_{e,eff,WB}$ při ztrátách paketů je dle rovnice 4.14 závislý mimo měřené hodnoty ztrátovosti paketů P_{pl} na parametrech odolnosti kodeku proti ztrátám paketů B_{pl} a měřítku náhodnosti rozložení ztrát paketů $BurstR$. Ztrátovost paketů je měřena na základě algoritmu 4.12, navrženém v kapitole 4.2.2. Pro odolnost ztrátám paketů kodeku AMR-WB 23.85 kbit/s uvádí dop. G.113 [26] při diotickém poslechu provizorní hodnotu 4,9. Na obrázku 4.18 je uvedeno srovnání korelačních diagramů hodnot R-faktoru, získaných přepočtem z měření MOS algoritmem POLQA, a hodnot R_{LQ} , vypočítaných na základě navrženého algoritmu analýzy RTP paketů.



Obr. 4.17: Histogram hodnot R získaných z měření MOS s minimální degradací



Obr. 4.18: Korelační diagramy R před a po optimalizaci parametrů

V levé části je korelace po dosažení provizorní hodnoty B_{pl} a $BurstR$ ve formě uvedené v rov. 4.14, v pravé části po optimalizaci obou parametrů pro maximální korelaci s měřeními hlasové kvality algoritmem POLQA. Ideální optimalizace byla hledána na základě hodnot střední kvadratické chyby $rmse$ a Pearsonova korelačního koeficientu ρ všech měřených bodů i v jednotlivých intervalech R . Oba parametry jsou popsány v kap. 4.4.

Nejdříve byla hledána optimalizace parametru $BurstR$, který v mnoha případech

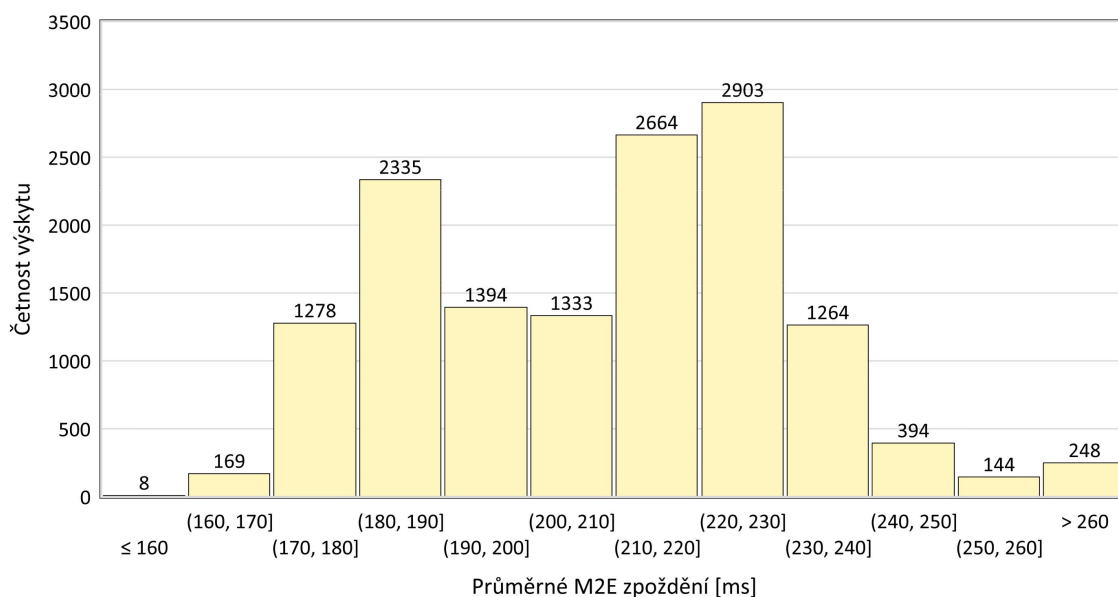
s vysokým počtem ztracených hlasových paketů způsoboval v důsledku krátkodobého charakteru měření R-faktoru přílišný nárůst faktoru $I_{e,\text{eff},\text{WB}}$. Na obr. 4.18 je tento efekt patrný z rozložení měřených bodů pod korelační křivkou v levé polovině korelačního diagramu, často s $R_{\text{LQ}} = 0$. Optimální řešením se ukázalo dosazení $BurstR$ do rov. 4.14 pod pátou odmocninou:

$$I_{e,\text{eff},\text{WB}} = I_{e,\text{WB}} + (95 - I_{e,\text{WB}}) \cdot \frac{P_{\text{pl}}}{\frac{P_{\text{pl}}}{\sqrt[5]{BurstR}} + B_{\text{pl}}}. \quad (4.21)$$

Následně byla na základě minimální hodnoty střední kvadratické chyby $rmse$ v jednotlivých intervalech R upravena velikost parametru B_{pl} z 4,9 na 6,5. Výsledkem bylo snížení celkové střední kvadratické chyby $rmse$ z 8,416 na 5,331 a zvýšení korelačního koeficientu ρ z 0,844 na 0,907.

4.3.3 Odhad M2E zpoždění

Posledním vstupním parametrem pro výpočet konverzačního R-faktoru R_{CQ} je faktor degradace vlivem zpoždění $I_{d,\text{WB}}$, jehož velikost závisí na hodnotě M2E zpoždění T_a . Při měřeních technologie VoLTE se průměrné M2E zpoždění T_a pohybovalo ve většině případů mezi 160 a 260 ms, jak je patrné z obrázku 4.19. Rovnice 4.8 uvádí jednotlivé složky zpoždění, které se na celkové hodnotě T_a podílí. Jedná se o fixní komponenty zpracování hlasovým kodekem $2T_{\text{frame}} + T_{\text{LA}}$ spolu s proměnnými složkami vložného zpoždění jitter bufferu T_{buffer} a transportního zpoždění RTP paketů $T_{\text{one-way}}$.



Obr. 4.19: Histogram průměrného M2E zpoždění při měření alg. POLQA

Protože jedním z cílů navrhovaného algoritmu je, aby byl neintrusivní, namísto přesného měření transportního zpoždění jednotlivých RTP paketů, které vyžaduje synchronizaci obou měřících stran, je voleno měření zpoždění na základě výměny RTCP zpráv, popsané v kap. 4.1.3. Analýza vztahu mezi M2E zpožděním měřeným alg. POLQA a RTCP zpožděním $T_{\text{one-way}}$ ukázala, že fixní složky zpoždění dávají dohromady přibližně 140 ms. Proto je výpočet T_a definován jako

$$T_a = 140 + T_{\text{one-way}} + 0,5 J_{\text{M2E}} + 2 J_{\text{p94}}. \quad [\text{ms}] \quad (4.22)$$

Přídavné složky $0,5 J_{\text{M2E}}$ a $2 J_{\text{p94}}$ slouží ke kompenzaci adaptivního zpoždění jitter bufferu přijímače. Následně se pro vyjádření dopadu zpoždění na vnímanou kvalitu konverzace dosazuje T_a do rovnice 2.14 definované E-modelem.

4.4 Analýza výsledků navrženého algoritmu

Pro srovnání výsledků referenčních měření hlasové kvality algoritmem POLQA s výsledky navrženého neintrusivního algoritmu výpočtu R-faktoru bylo využito Pearsonova korelačního koeficientu ρ , střední kvadratické chyby $rmse$ a průměrné chyby me .

Pearsonův korelační koeficient

Pearsonův korelační koeficient ρ je statistický parametr umožňující určit míru lineární korelace mezi dvěma proměnnými x a y . Může nabývat hodnot v intervalu $\langle -1; 1 \rangle$, přičemž -1 vyjadřuje úplnou zápornou korelaci, 0 žádnou korelaci a 1 úplnou kladnou korelaci. Je definován jako poměr kovariancí souborů dat dvou proměnných normalizovaný jejich směrodatnými odchylkami [48]:

$$\rho = \frac{\text{cov}(x, y)}{\sigma_x \sigma_y}. \quad (4.23)$$

Ze souboru dat proměnných x a y o n bodech se ρ vyjádří dosazením do rovnice

$$\rho = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}. \quad (4.24)$$

Střední kvadratická chyba

Pro kvantifikování souhrnné chyby měření hlasové kvality byla využita střední kvadratická chyba $rmse$. Máme-li dva srovnávané soubory proměnných x a y s n měřeními

body, $rmse$ se vypočítá na základě rovnice

$$rmse = \sqrt{\frac{\sum_{i=1}^n (x_i - y_i)^2}{n}}. \quad (4.25)$$

V důsledku umocnění jednotlivých chyb je hodnota $rmse$ značně závislá na velikosti a četnosti velkých chyb v souboru měření. Chyba referenčního měření hlasové kvality algoritmem POLQA byla pro zjednodušení zanedbána.

Průměrná chyba

Střední kvadratická chyba je výborným ukazatelem souhrnné chyby měření, ale v důsledku použití druhé mocniny neudává informaci o znaménku chyby měření. Proto byla zavedena ještě průměrná chyba

$$me = \frac{\sum_{i=1}^n (x_i - y_i)}{n}, \quad (4.26)$$

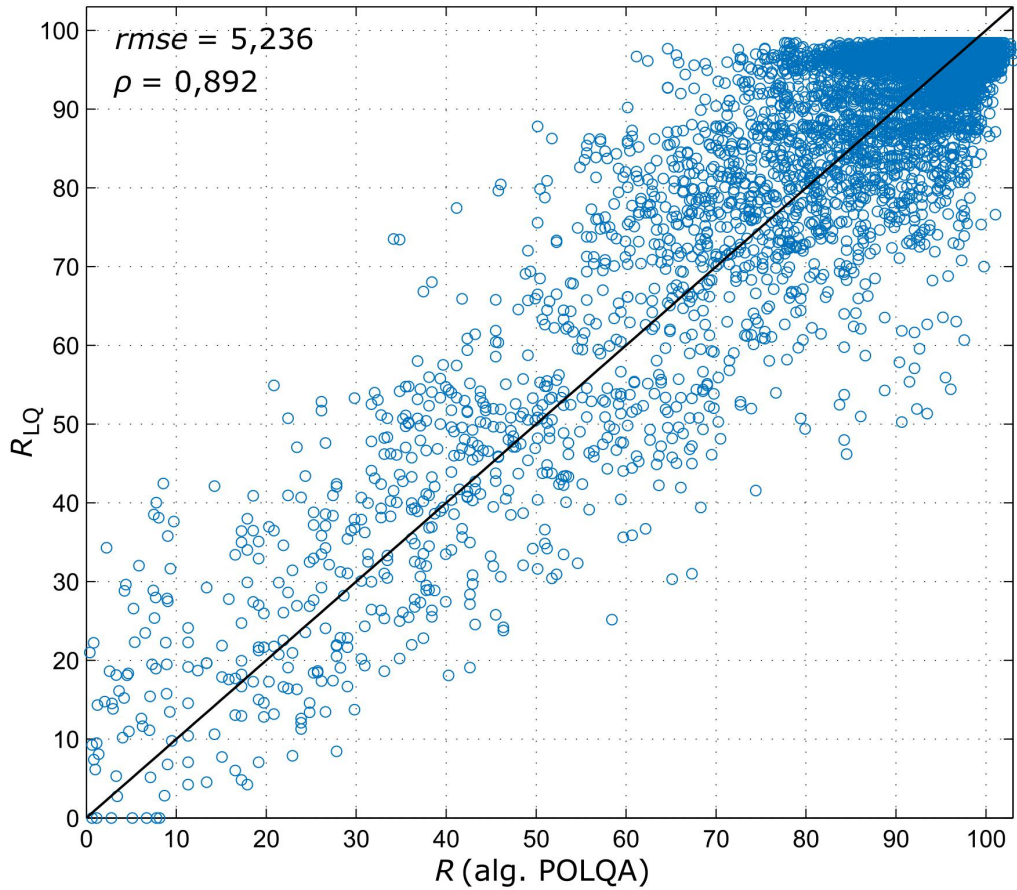
kde stejně jako v případě $rmse$ jsou $\{x_1, \dots, x_n\}$ a $\{y_1, \dots, y_n\}$ soubory srovnávaných proměnných.

4.4.1 Výsledky poslechového R-faktoru

Obrázek 4.20 zachycuje korelační diagram měření hlasové kvality MOS-LQO_{SW} algoritmem POLQA, které byly převedeny pomocí rovnice 2.20 na ekvivalentní hodnotu WB R-faktoru, a výsledků navrženého algoritmu výpočtu poslechového R-faktoru R_{LQ} na základě analýzy RTP/RTCP paketů. Tabulka 4.1 obsahuje statistickou analýzu této závislosti pro jednotlivé intervaly měření R algoritmem POLQA.

Celkový korelační koeficient $\rho = 0,892$ a střední kvadratická chyba $rmse = 5,23$ jsou na parametrický model měření hlasové kvality velmi kvalitní hodnoty, obzvláště při uvážení určité chyby měření algoritmu POLQA. Protože měření nejsou rozdělena rovnoměrně, je důležité ale brát ohled i na výsledky algoritmu v jednotlivých intervalech měřicí stupnice. V intervalu $R \in \langle 10; 80 \rangle$ je velikost $rmse$ podstatně vyšší než při souhrnu všech měření. Možným důvodem většího rozptylu měření je přirozená nejistota závislosti hlasové kvality na ztrátách paketů, které na ni mají v tomto intervalu nejvýraznější dopad. Protože navržený algoritmus na základě statistické analýzy staví všechny pakety přenášející hlasová data na stejnou úroveň, nedokáže rozpoznat rozdíly ve vnímané kvalitě se stejnou přesností jako signálově založené modely.

Kromě toho je z průběhu průměrné chyby me patrné, že navržený algoritmus ve stejném intervalu v průměru mírně nadhodnocuje vnímanou hlasovou kvalitu, což je v celkovém souhrnu kompenzováno naopak nižší predikcí pro vysoké hodnoty R .



Obr. 4.20: Korelační diagram měření hl. kvality POLQA a výsledků alg. R_{LQ}

Tab. 4.1: Analýza korelace měření hl. kvality POLQA a výsledků alg. R_{LQ}

R (POLQA)	Počet měření	ρ	me	$rmse$
$\langle 100; 104 \rangle$	1129	0,003	-3,74	4,00
$\langle 90; 100 \rangle$	13403	0,258	-0,84	3,16
$\langle 80; 90 \rangle$	1212	0,292	4,71	8,98
$\langle 70; 80 \rangle$	445	0,207	6,42	12,60
$\langle 60; 70 \rangle$	251	0,064	5,01	15,11
$\langle 50; 60 \rangle$	155	0,230	4,64	15,13
$\langle 40; 50 \rangle$	112	0,203	3,01	12,11
$\langle 30; 40 \rangle$	112	0,161	5,22	12,66
$\langle 20; 30 \rangle$	76	0,093	3,20	12,23
$\langle 10; 20 \rangle$	50	0,273	4,56	10,90
Celkem	16945	0,892	-0,21	5,23

Možnou korekcí by bylo snížení odpovídajícího parametru odolnosti proti ztrátám paketů B_{pl} , které by ale zvýšilo chybu měření pro hodnoty $R \geq 80$.

Principiálně bude ale vždy při srovnání signálově založeného objektivního algoritmu s parametrickým modelem ve výsledcích určitý rozpor z důvodu odlišných filozofií určení hlasové kvality. Zatímco POLQA je konstruována pro měření hlasové kvality izolovaně pro jednotlivé krátké vzorky hlasového signálu, u kterých se může kvalita rychle měnit, parametrické modely vychází z plánovacích principů využívajících statisticky významné charakteristiky neměnné v delším časovém horizontu. Aplikace těchto charakteristik na krátkodobá měření není samozřejmá.

V tabulce 4.1 chybí interval $R \in (0; 10)$. Z definice vztahu 2.20 mezi MOS a R-faktorem totiž vyplývá, že $\text{MOS} = 1,01$ odpovídá $R = 10,03$ a nižší hodnoty R nejsou tak na základě převodního vztahu definovány. Na obr. 4.20 byla proto pro hodnoty R v tomto intervalu vypočítána náhodná hodnota s rovnoměrným rozdělením pravděpodobnosti.

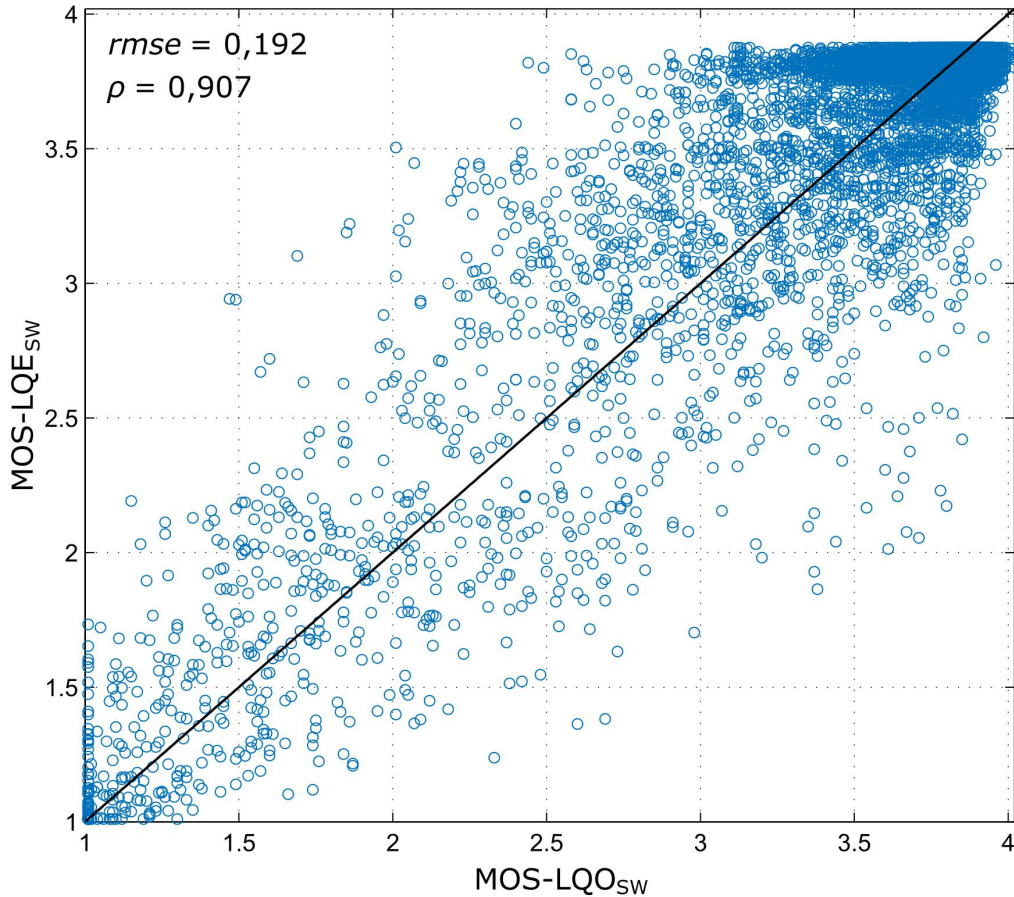
4.4.2 Výsledky hlasové kvality MOS

Kromě převodu měřené hlasové kvality $\text{MOS-LQO}_{\text{SW}}$ na hodnotu R je možné srovnání výsledků měření i z opačného pohledu, kdy budou výsledky navrženého algoritmu R_{LQ} převedeny na hlasovou kvalitu stupnice $\text{MOS-LQE}_{\text{SW}}$. Taková závislost je znázorněna na obrázku 4.21. Stejně jako v předchozím případě, tabulka 4.2 obsahuje analýzu měření pro jednotlivé intervaly měřené hlasové kvality $\text{MOS-LQO}_{\text{SW}}$.

Tab. 4.2: Analýza korelace měření MOS-LQO a výsledků alg. MOS-LQE

$\text{MOS-LQO}_{\text{SW}}$	Počet měření	ρ	me	$rmse$
$\langle 3,5; 4,1 \rangle$	15015	0,293	-0,029	0,122
$\langle 3,0; 3,5 \rangle$	1008	0,297	0,191	0,402
$\langle 2,5; 3,0 \rangle$	371	0,299	0,208	0,541
$\langle 2,0; 2,5 \rangle$	204	0,332	0,237	0,617
$\langle 1,5; 2,0 \rangle$	176	0,279	0,148	0,426
$\langle 1,0; 1,5 \rangle$	232	0,551	0,153	0,323
Celkem	17006	0,907	-0,004	0,192

Oproti závislosti z obr. 4.20 má chyba měření relativně k měřenému intervalu hlasové kvality větší rozptyl. Tento efekt je obzvlášť patrný ve strmé části převodní funkce znázorněné na obr. 2.6. Přesto je výsledný Pearsonův korelační koeficient ρ roven 0,907, tedy o 0,015 více než v předchozím případě. Celková střední kvadratická chyba $rmse$ je rovna 0,192. Při srovnání s průměrnou hodnotou $rmse^* 0,167$, uváděnou pro algoritmus POLQA verze 2.4 [49], působí výsledky velmi pozitivně, opět je však třeba brát ohled i na úseky s chybou výrazně vyšší. Průběh průměrné



Obr. 4.21: Korelační diagram měření MOS-LQO a výsledků alg. MOS-LQE

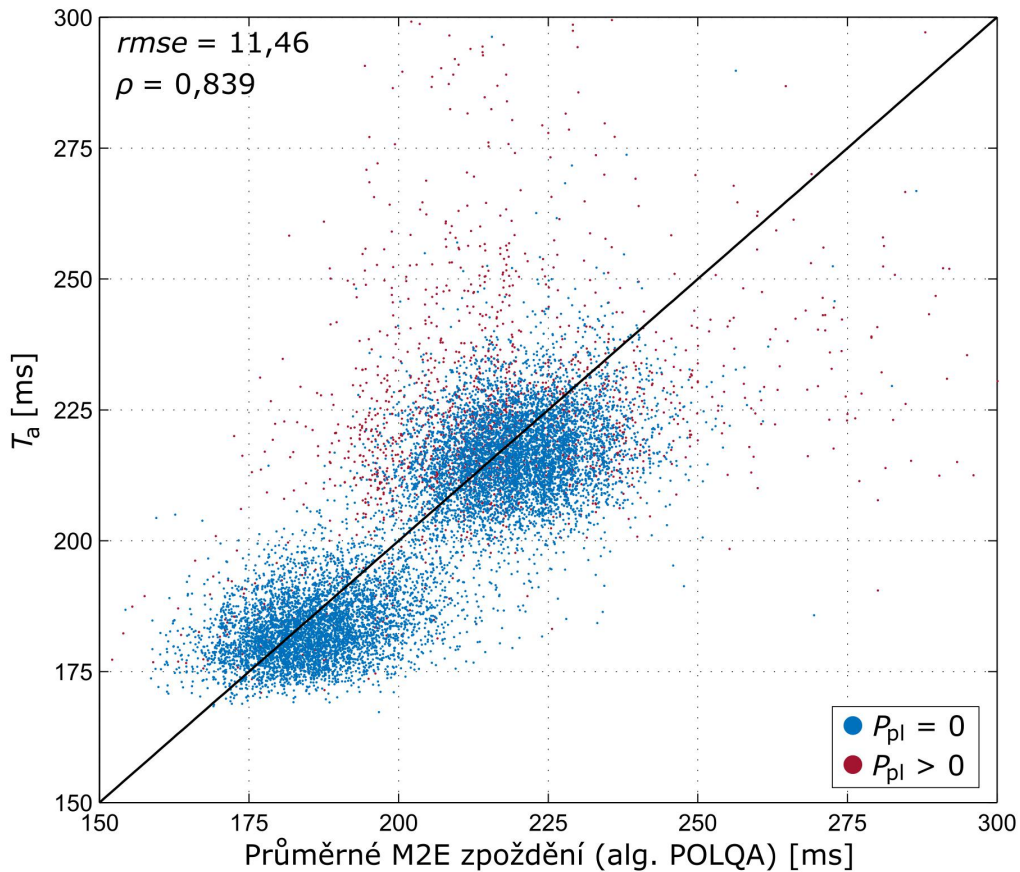
chyby me a střední kvadratické chyby $rmse$ je s ohledem na rozdílné měřítko hlasové kvality podobný jako v tabulce 4.1.

4.4.3 Výsledky odhadu M2E zpoždění

Protože algoritmus POLQA při měření hlasové kvality nezohledňuje efekt M2E zpoždění, nelze výsledky výpočtu konverzačního R-faktoru R_{CQ} přímo srovnávat s měřeními algoritmu POLQA. Součástí algoritmu POLQA je však nezbytný blok časového zarovnání referenční a degradované nahrávky, z jehož výstupu je možné velikost M2E zpoždění vyhodnotit. Na obrázku 4.22 je znázorněna závislost tohoto zpoždění na odhadu M2E zpoždění T_a , navrženém v kapitole 4.3.3.

Modrou barvou jsou zobrazena měření, při kterých nedošlo ke ztrátám paketů. Tmavě červená měření se ztrátami paketů jsou vedena odděleně kvůli extrémní proměnlivosti reportovaného M2E zpoždění algoritmem POLQA, pohybujícím se mezi -1 až $2,4$ s. Ze stejného důvodu nejsou zahrnuty ani do analýzy výsledků v tab. 4.3.

Přestože je z obrázku 4.22 evidentní vztah mezi měřeným M2E zpožděním a jeho



Obr. 4.22: Korelační diagram měření M2E zpoždění alg. POLQA a odhadu T_a

Tab. 4.3: Analýza korelace měření M2E zpoždění alg. POLQA a odhadu T_a

Prům. zpoždění	Počet měření	ρ	me	$rmse$
$\langle 150; 175 \rangle$	615	0,008	9,50	11,61
$\langle 175; 200 \rangle$	4447	0,334	-2,33	8,70
$\langle 200; 225 \rangle$	4711	0,384	-1,16	11,05
$\langle 225; 250 \rangle$	1780	0,135	-12,02	15,78
$\langle 250; 300 \rangle$	26	0,361	-30,15	35,21
Celkem	11579	0,839	-2,78	11,15

odhadem, v krajních intervalech měřených hodnot je střední kvadratická chyba výrazně vyšší. Zároveň pro určité intervaly průměrného zpoždění je odhad znatelně podhodnocován, přičemž pro nízké měřené hodnoty je odhad naopak soustavně vyšší. Tyto nedostatky plynou jednak z principu měření vstupních parametrů odhadu T_a , svou roli hraje ale i nezanedbatelná nejistota měření zpoždění algoritmem POLQA, která exponenciálně vzrůstá s nedokonalostí analyzované degradované nahrávky.

Odhad T_a se zakládá na měřené hodnotě RTT zpoždění RTCP paketů. V prů-

běhu analyzované 6 sekundové nahrávky je tato hodnota měřena nejvýše jednou, proto srovnání s průměrným zpožděním celé nahrávky není v mnoha případech přesné, obzvláště pokud je transportní doba v obou směrech komunikace proměnlivá. Přesnějším řešením by byl odhad T_a na základě měření transportního zpoždění jednotlivých RTP paketů, nezbytnou podmínkou by však byla komplikovaná synchronizace obou zařízení. Přesto, jak je prokázáno hodnotami celkového korelačního koeficientu $\rho = 0,839$ a střední kvadratické chyby $rmse = 11,49$, algoritmus udává ve většině případů dobrý odhad M2E zpoždění pro přibližné určení faktoru degradace vlivem zpoždění I_d , resp. konverzačního R-faktoru.

5 IMPLEMENTACE V OS ANDROID

Součástí zadání práce je návrh jednoduchého měřicího skriptu určeného pro zařízení s OS Android, který by byl schopen provádět autonomní cyklická měření hlasové kvality technologií VoLTE a VoWiFi na základě navrženého algoritmu výpočtu R-faktoru.

5.1 Omezení OS Android

Původní idea při tvorbě této části zadání vycházela z předpokladu, že mobilní telefony s OS Android podporující technologii VoLTE pro správu kvality hovoru potřebují během hovoru měřit parametry RTP komunikace a že je možné tyto parametry využít pro zamýšlené měření hlasové kvality přímo v telefonu. Po seznámení se s vývojářským prostředím Android [50] a analýze jiných řešení na trhu se ale ukázalo, že možnosti měření hlasové kvality na základě analýzy síťové komunikace jsou v OS Android značně limitovány. Ve verzích Android 5.0 a novějších není z bezpečnostních důvodů možné k VoLTE komunikaci běžně přistupovat. UDP metody třídy `android.net.TrafficStats`, umožňující přístup ke statistickým údajům provozu na transportní vrstvě, nejsou dále podporovány již od verze Android 4.3 [51].

5.2 Alternativní způsoby

Existují způsoby, jakými je možné tato bezpečnostní opatření obejít a získat přístup k IP IMS Stack. Možnosti se obvykle liší v závislosti na typu telefonu, ale pro zachytávání veškeré TCP/IP komunikace přímo na telefonu je nezbytná úprava firmware zařízení odemykající běžně nepřístupné vývojářské funkce.

Systém Accuver využívá pro zachytávání datového provozu USB připojení k PC v ladícím režimu. Prerekvizitami měření je povolení CP logging v režimu obnovy telefonu, uvedení telefonu do diagnostického módu (DM) a nastavení USB ovladačů pro ADB/DM/modem připojení. Pro většinu typů telefonů je nezbytný také tzv. root telefonu, tj. získání privilegovaných oprávnění. Síťový provoz je ve výsledku směrován skrze USB připojení na klienta v PC prostřednictvím Android Debug Bridge (ADB) [52] rozhraní.

6 ZÁVĚR

Diplomová práce se zabývala metodami měření hlasové kvality se zaměřením na technologie VoLTE a VoWiFi. Cílem byl návrh algoritmu výpočtu R-faktoru na základě parametrů měřených uživatelskými zařízeními s operačním systémem Android.

V první části byly nejprve definovány pojmy hlasová kvalita a konverzační a poslechové testování, pomocí nichž se úroveň hlasové kvality posuzuje. Na základě doporučení ITU-T byly charakterizovány metody subjektivního a objektivního testování, přičemž popis byl zaměřen především na funkci algoritmu POLQA, který byl později při návrhu výpočtu R-faktoru využit pro referenční měření hlasové kvality.

Práce se dále zabývá možnostmi využití E-modelu, plánovacího nástroje definovaného doporučením ITU-T G.107, pro neintrusivní měření hlasové kvality. Je představena obecná definice výpočtu tzv. R-faktoru, skalárního ukazatele predikované kvality komunikační sítě, a parametry, na jejichž základě se hodnota R počítá.

Třetí kapitola uvádí používané metody přenosu hlasu v sítích LTE a Wi-Fi a zaměřuje se především na funkce zprostředkovávané IMS. Při měření hlasové kvality mají zásadní roli zejména protokoly SIP, SDP, RTP/RTCP a hlasové kodeky, proto se popis technologií VoLTE a VoWiFi soustředí především na ně.

Pro návrh výpočetního algoritmu R-faktoru bylo provedeno měřicím systémem Accuver XCAL přibližně 80 hodin měření hlasové kvality technologie VoLTE. Navržený algoritmus vychází z analýzy příchozích RTP/RTCP paketů, měřeními vstupními parametry jsou ztrátovost hlasových paketů, distribuce ztrát v měřeném intervalu, transportní zpoždění RTCP paketů a několik parametrů popisujících jitter.

Dle analýzy výsledků měření poslechového R-faktoru R_{LQ} navrženým algoritmem je celkový Pearsonův korelační koeficient ρ při srovnání s měřeními POLQA roven 0,892, tj. pro parametrickou metodu měření velmi vysoké hodnotě. Při převodu měřených hodnot na MOS-LQE_{SW} je ρ rovno dokonce 0,907. Celková střední kvadratická chyba měření MOS vyšla rovna 0,192. Přes celkové pozitivní výsledky je ale zejména v oblasti střední kvality měřených vzorků hodnota kvadratické chyby podstatně vyšší (až 0,617). Podobným způsobem byla analyzována i úspěšnost odhadu M2E zpoždění nezbytného pro výpočet konverzačního R-faktoru R_{CQ} . Korelační koeficient s měřeními POLQA byl roven 0,839, střední kvadratická chyba *rmse* byla rovna 11,15 ms.

Poslední část práce prezentuje možnosti implementace navrženého algoritmu pro neintrusivní měření hlasové kvality na zařízeních s OS Android a důvody, proč v zadání uvedený jednoduchý měřicí skript není možné realizovat. Hlavním problémem jsou totiž omezení OS Android plynoucí ze striktních bezpečnostních pravidel, parametry nutné pro výpočet R-faktoru tak nejsou ve vývojovém prostředí Android běžně vůbec k dispozici.

Výsledky práce jsou důkazem možné aplikace parametrických neintrusivní modelů pro odhad hlasové kvality VoIP technologie VoLTE a přeneseně také technologie VoWiFi, která však nebyla kvůli omezenému počtu podporovaných zařízení v době měření testována. Vhodným námětem pro pokračování práce by bylo právě ověření vlastností navrženého algoritmu u technologie VoWiFi a také při použití jiných typů uživatelských zařízení, které se mohou svými vlastnostmi lišit. Zároveň se nabízí možnost hlubšího prozkoumání možností OS Android pro zachytávání RTP komunikace během hovoru, které by mělo být přes řadu technických překážek teoreticky uskutečnitelné.

LITERATURA

- [1] HOLUB, J. Měření a hodnocení QoS v IP telefonii. In: *Teorie a praxe IP telefonie - 1. dvoudenní odborný seminář, Kongesové centrum Hotelu Olšanka, 26. a 27. května 2004* [online]. 2004 [cit. 2015-10-30]. Dostupné z: http://www.researchgate.net/publication/251886832_MEN_A_HODNOCEN_QOS_V_IP_TELEFONII
- [2] Rec. ITU-T P.800. *Methods for subjective determination of transmission quality*. (08/96). Geneva: ITU, 1996.
- [3] Rec. ITU-T P.805. *Subjective evaluation of conversational quality*. (04/2007). Geneva: ITU, 2007.
- [4] Rec. ITU-T P.830. *Subjective performance assessment of telephone-band and wideband digital codecs*. (02/96). Geneva: ITU, 1996.
- [5] Rec. ITU-T P.800.1. *Mean Opinion Score (MOS) terminology: Methods for objective and subjective assessment of quality*. (07/2006). Geneva: ITU, 2006.
- [6] SLAVATA, O. a J. HOLUB. Evaluation of objective speech transmission quality measurements in packet-based networks. *Computer Standards & Interfaces* [online]. 2014, vol. 36, issue 3, s. 626-630 [cit. 2015-11-17]. DOI: 10.1016/j.csi.2013.09.003. Dostupné z: <http://www.sciencedirect.com/science/article/pii/S0920548913001244>
- [7] Rec. ITU-T P.862. *Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs*. (02/2001). Geneva: ITU, 2001.
- [8] Rec. ITU-T P.862.1. *Mapping function for transforming P.862 raw result scores to MOS-LQO*. (11/2003). Geneva: ITU, 2004.
- [9] Rec. ITU-T P.862.2. *Wideband extension to Recommendation P.862 for the assessment of wideband telephone networks and speech codecs*. (11/2007). Geneva: ITU, 2008.
- [10] Rec. ITU-T P.863. *Perceptual objective listening quality assessment*. (09/2014). Geneva: ITU, 2015.
- [11] Rec. ITU-T P.863.1. *Application guide for Recommendation ITU-T P.863*. (09/2014). Geneva: ITU, 2015.

- [12] SWISSQUAL. *Voice Quality with ITU-T P.863 'POLQA': Application Note* [online]. 2012 [cit. 2015-11-26]. Dostupné z: <http://www.swissqual.com/en/expertise/technical-publications/>
- [13] OPTICOM. *Perceptual Objective Listening Quality Analysis* [online]. 2011 [cit. 2015-11-29]. Dostupné z : http://www.opticom.de/download/POLQA_White_Paper_1011.pdf
- [14] Rec. ITU-T G.107. *The E-model: a computational model for use in transmission planning*. (06/2015). Geneva: ITU, 2015.
- [15] Rec. ITU-T P.561. *In-service non-intrusive measurement device - Voice service measurements*. (07/2002). Geneva: ITU, 2003.
- [16] Rec. ITU-T P.562. *Analysis and interpretation of INMD voice-service measurements*. (05/2004). Geneva: ITU, 2005.
- [17] Rec. ITU-T P.563. *Single-ended method for objective speech quality assessment in narrow-band telephony applications*. (05/2004). Geneva: ITU, 2005.
- [18] MALFAIT, L., J. BERGER a M. KASTNER. P.563—The ITU-T Standard for Single-Ended Speech Quality Assessment. *IEEE Transactions on Audio, Speech and Language Processing* [online]. 2006, vol. 14, issue 6, s. 1924-1934 [cit. 2015-11-24]. DOI: 10.1109/TASL.2006.883177. Dostupné z: <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=1709882>
- [19] ETR 250. *Transmission and Multiplexing (TM): Speech communication quality from mouth to ear for 3,1 kHz handset telephony across networks*. ETSI, 1996.
- [20] CÔTÉ, N. *Integral and diagnostic intrusive prediction of speech quality*. Berlin: Springer, 2011, xvii, 248 s. T-labs series in telecommunication services. ISBN 978-3-642-18462-8.
- [21] Rec. ITU-T G.107.1. *Wideband E-model*. (06/2015). Geneva: ITU, 2015.
- [22] MYAKOTNYKH, E. *Adaptive speech quality in voice-over-IP communications* [online]. Ann Arbor, 2008. ISBN 978-0-549-73252-5. Ph.D. Dissertation. University of Pittsburgh, Faculty of the Telecommunications, Graduate School of Information Sciences. Dostupné z: <http://d-scholarship.pitt.edu/7789/>.
- [23] WALTERMANN, M., I. TUCKER, A. RAAKE a S. MOLLER. Extension of the E-model towards super-wideband speech transmission. In: *2010 IEEE International Conference on Acoustics, Speech and Signal Processing* [online].

- IEEE, 2010, s. 4654-4657 [cit. 2015-12-01]. ISBN 978-1-4244-4295-9, ISSN 1520-6149. DOI: 10.1109/ICASSP.2010.5495199. Dostupné z: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5495199>
- [24] Rec. ITU-T G.108. *Application of the E-model: A planning guide*. (09/99). Geneva: ITU, 2000.
- [25] Rec. ITU-T G.109. *Definition of categories of speech transmission quality*. (09/1999). Geneva: ITU, 2000.
- [26] Rec. ITU-T G.113. *Transmission impairments due to speech processing*. (11/2007). Geneva: ITU, 2008.
- [27] POIKSELKÄ, M., H. HOLMA, J. HONGISTO, J. KALLIO a A. TOSKALA. NOKIA SIEMENS NETWORKS. *Voice over LTE: VoLTE*. First Edition. Finland: John Wiley & Sons Ltd, 2012. ISBN 978-1-119-95168-1.
- [28] GSMA IR.92. *IMS Profile for Voice and SMS*. Version 9.0. GSM Association, 2015.
- [29] ANRITSU. *Understanding IMS* [online]. Issue 2, 09/2013. Anritsu, 2013 [cit. 2015-12-06]. Dostupné z: <http://dataedge.ie/wp-content/uploads/2013/07/Network-Testing-Understanding-IMS.pdf>
- [30] SPIRENT. *IMS Architecture: The LTE User Equipment Perspective* [online]. Rev B. 04/14. Spirent, 2014 [cit. 2016-03-07]. Dostupné také z: http://www.spirent.com/White-Papers/Mobile/LTE_User_Equipment_Perspective
- [31] 3GPP TS 23.203. *Technical Specification Group Services and System Aspects; Policy and charging control architecture*. V13.7.0 (2016-03). 3GPP, 2016.
- [32] GSMA IR.51. *IMS Profile for Voice, Video and SMS over Wi-Fi*. Version 2.0. GSM Association, 2015.
- [33] SPIRENT. *VoLTE Deployment and the radio access network: The LTE User Equipment Perspective* [online]. Rev. A 08/12. Spirent, 2012 [cit. 2016-03-07]. Dostupné také z: http://www.spirent.com/~media/White%20Papers/Mobile/VoLTE_Deployment_and_the_Radio_Access_Network.pdf
- [34] ROSENBERG, J. et al. *IETF RFC 3261: SIP: Session Initiation Protocol*. June 2002. The Internet Society, 2002.
- [35] HANDLEY, M., V. JACOBSON a C. PERKINS. *IETF RFC 4566: SDP: Session Description Protocol*. July 2006. The Internet Society, 2006.

- [36] ZURAWSKI, R. *The Industrial Information Technology Handbook*. Boca Raton: CRC Press, 2005, 1 v. ISBN 08-493-1985-4.
- [37] SCHULZRINNE, H., S. CASNER, R. FREDERICK a V. JACOBSON. *IETF RFC 3550: RTP: A Transport Protocol for Real-Time Applications*. July 2003. The IETF Trust, 2003.
- [38] SJOBERG, J., M. WESTERLUND, A. LAKANIEMI a Q. XIE. *IETF RFC 4867: RTP Payload Format and File Storage Format for the Adaptive Multi-Rate (AMR) and Adaptive Multi-Rate Wideband (AMR-WB) Audio Codecs*. April 2007. The IETF Trust, 2007.
- [39] ERICSSON AB. *Evolved HD Voice for LTE* [online]. Uen 284 23-3160 Rev B. 2014 [cit. 2016-03-13]. Dostupné z: <http://www.ericsson.com/res/docs/whitepapers/wp-evolved-hd-voice-for-lte.pdf>
- [40] Rec. ITU-T G.722.2. *Wideband coding of speech at around 16 kbit/s using Adaptive Multi-Rate Wideband (AMR-WB)*. (07/2003). Geneva: ITU, 2004.
- [41] Rec. ITU-T G.722.2. *Wideband coding of speech at around 16 kbit/s using Adaptive Multi-Rate Wideband (AMR-WB): Annex E: Frame structure*. (01/2002). Geneva: ITU, 2002.
- [42] ETSI TR 126 976. *Performance characterization of the Adaptive Multi-Rate Wideband (AMR-WB) speech codec: 3GPP TR 26.976. V12.0.0 (2014-09)*. ETSI, 2014.
- [43] RAY, N. Patent-Pending: T-Mobile's Next Network Upgrade With Enhanced Voice Services. In: *T-Mobile US, News & Blogs* [online]. 2016 [cit. 2016-04-14]. Dostupné z: <https://newsroom.t-mobile.com/news-and-blogs/volte-enhanced-voice-services.htm>
- [44] NOKIA SOLUTIONS AND NETWORKS. *The 3GPP Enhanced Voice Services (EVS) codec* [online]. 2015 [cit. 2016-04-14]. Dostupné z: http://networks.nokia.com/sites/default/files/document/nokia_3gpp_evs_codec_white_paper_0.pdf
- [45] RAAKE, A. *Speech Quality of VoIP: Assessment and Prediction*. Chichester, England: Wiley, 2006. ISBN 978-0-470-03060-8.
- [46] RACHWALSKI, J. a Z. PAPIR. Burst Ratio in Concatenated Markov-based Channels. *Journal of Telecommunications and Information Technology* [online]. 2014, **2014**(1), 3-9 [cit. 2016-05-04]. ISSN

- 1509-4553. Dostupné z: <https://www.infona.pl/resource/bwmeta1.element.baztech-d856592c-f1cf-4090-9e3e-1505e659a47b/content/partContents/91bf172e-713c-35f5-a3dc-b9e8095e5041>
- [47] PANG, L. a L. BÖSZÖRMENYI. Adaptive Jitter Buffer based on Quality Optimization under Bursty Packet Loss. *The International Journal on Advances in Telecommunications* [online]. IARIA, 2012, **2012**(vol. 5, no. 1 & 2), 1-10 [cit. 2016-05-04]. ISSN 1942-2601. Dostupné z: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.638.7957&rep=rep1&type=pdf#page=10>
- [48] ABDULLAH, M. B. On a Robust Correlation Coefficient. *The Statistician* [online]. 1990, **39**(4), 455-460 [cit. 2016-05-05]. DOI: 10.2307/2349088. ISSN 00390526. Dostupné z: <http://www.jstor.org/stable/10.2307/2349088?origin=crossref>
- [49] OPTICOM. *POLQA 2015 (V2.4) Investigated: Technical White Paper* [online]. October 2014. The POLQA Coalition, 2014 [cit. 2016-05-05]. Dostupné také z: <http://www.abctrd.com/ksboard/skin/board/basic/download.php?bid=5&id=460&file=2>
- [50] *Android Developers: Package Index* [online]. Android 6.0 r1. [cit. 2016-05-07]. Dostupné z: <http://developer.android.com/reference/packages.html>
- [51] TrafficStats: public class. *Android Developers* [online]. [cit. 2016-05-07]. Dostupné z: <http://developer.android.com/reference/android/net/TrafficStats.html>
- [52] Android Debug Bridge. *Android Developers* [online]. [cit. 2016-05-07]. Dostupné z: <http://developer.android.com/tools/help/adb.html>

SEZNAM SYMBOLŮ, VELIČIN A ZKRATEK

A	Faktor zvýhodnění, E-model
B_{pl}	Odolnost proti ztrátovosti paketů, E-model
$BundlR$	Míra spojování paketů
$BurstR$	Shlukovitost ztrát paketů
$DLSR$	Doba od poslední SR RTCP zprávy
D_r	Hodnota D telefonu, přijímací strana, E-model
D_s	Hodnota D telefonu, vysílací strana, E-model
E	Gaussova chybová funkce
f_s	Vzorkovací kmitočet
GoB	Podíl osob s hodnocením kvality dobrá nebo lepší
I_d	Degradace vlivem zpoždění, E-model
$I_{d,WB}$	Degradace vlivem zpoždění, Wideband E-model
I_{dd}	Vliv úplného (absolutního) zpoždění T_a , E-model
$I_{dle,WB}$	Vliv echa na straně posluchače, Wideband E-model
$I_{dte,WB}$	Vliv echa na straně mluvčího, Wideband E-model
I_e	Faktor degradace použitou technikou, E-model
$I_{e,eff}$	Efektivní faktor degradace použitou technikou, E-model
$I_{e,eff,WB}$	Efektivní faktor degradace použitou technikou, Wideband E-model
$I_{e,WB}$	Faktor degradace použitou technikou, Wideband E-model
I_j	Faktor degradace jitterem
I_s	Zkreslení synchronní se signálem, E-model
$I_{s,WB}$	Zkreslení synchronní se signálem, Wideband E-model
I_{tot}	Souhrn všech zdrojů degradace kvality, E-model
J	Jitter dle definice RFC 3550
J_X	Jitter nezávislý na hodnotě $BundlR$
J_{M2E}	Jitter před začátkem měřeného úseku
J_{p94}	94. percentil jitteru
$LSTR$	Místní vazba posluchače, E-model
mbl	Průměrná délka shluku ztrát paketů
me	Průměrná chyba

mlr	Ztrátovost paketů, $P_{pl}/100$
$N_{0,WB}$	Suma příspěvků zdrojů šumu vtažených k bodu 0 dBr, WB E-model
N_c	Šum v okruhu vztažený k bodu 0 dBr, E-model
$N_{os,WB}$	Ekvivalent šumu místnosti na vysílací straně, Wideband E-model
N_{or}	Ekvivalent šumu místnosti na přijímací straně, E-model
$N_{fo,WB}$	Šumové pozadí na přijímací straně, Wideband E-model
N_{for}	Vychozí šumové pozadí na přijímací straně, E-model
$N_{for,WB}$	Vychozí šumové pozadí na přijímací straně, Wideband E-model
OLR	Celková hlasitost, E-model
PoW	Podíl osob s hodnocením kvality špatná nebo horší
P_{pl}	Ztrátovost paketů v procentech, E-model
P_r	Šum místnosti na přijímací straně, E-model
P_s	Šum místnosti na vysílací straně, E-model
P_{re}	Efektivní šum místnosti na přijímací straně, E-model
qdu	Počet jednotek kvantizačního zkreslení, E-model
ρ	Pearsonův korelační koeficient
R	R-faktor, E-model
R_0	Poměr signálu k šumu, E-model
$R_{0,max}$	Maximální poměr signálu k šumu, E-model
$R_{0,max,NB}$	Maximální poměr signálu k šumu, E-model
$R_{0,max,WB}$	Maximální poměr signálu k šumu, Wideband E-model
$R_{0,WB}$	Poměr signálu k šumu, Wideband E-model
R_{CQ}	Konverzační R-faktor
R_{LQ}	Poslechový R-faktor
$rmse$	Střední kvadratická chyba
R_x	R-faktor transformovaný z WB kontextu na NB, Wideband E-model
RLR	Hlasitost přijímací strany, E-model
RTT	Obousměrné zpoždění
SLR	Hlasitost vysílací strany, E-model
SN	Pořadové číslo RTP paketu
SNR	Poměr signálu k šumu
$STMR$	Maskování místní vazby, E-model

T	Průměrné jednosměrné zpoždění echa, E-model
T_a	Úplné (absolutní, M2E) zpoždění, E-model
$T_{\text{AMR-WB}}$	Aditivní zpoždění kodéru AMR-WB
T_{buffer}	Zpoždění vkládané jitter bufferem na vstupu hlasového dekodéru
T_{codec}	Zdrojové zpoždění, zpoždění kodeku
$TELR$	Hodnocení hlasitosti echa mluvího, E-model
T_{IA}	Paketové zpoždění
T_{LA}	Algoritmické zpoždění
T_{network}	Sítové zpoždění
$T_{\text{one-way}}$	Jednosměrné (transportní) zpoždění
T_r	Obousměrné zpoždění ve smyčce, E-model
TS	Časová značka
TS_{RTP}	RTP časová značka
TS_{UE}	Časová značka UE, tzv. chipset time
T_{frame}	Délka rámce
$WEPL$	Vážený útlum echa, E-model

2G	Mobilní síť druhé generace (např. GSM)
3G	Mobilní síť třetí generace (např. UMTS)
3GPP	3rd Generation Partnership Project
3SQM	Single Sided Speech Quality Measure, ITU-T P.563
AAA	Authentication, Authorization and Accounting
ACELP	Algebraic Code Excited Linear Prediction
ACR	Absolute Category Rating
ADB	Android Debug Bridge
AMR	Adaptive Multi-Rate kodek
AMR-NB	Adaptive Multi-Rate Narrowband kodek, stejný jako AMR
AMR-WB	Adaptive Multi-Rate Wideband kodek
AP	Access Point
API	Application Programming Interface
BGCF	Breakout Gateway Control Function
CCI	Call Clarity Index
CP	Cellular processor
CS	Circuit-Switched
CSFB	Circuit-Switched Fallback
CQE	Conversational Quality Estimated
CSCF	Call Session Control Function
CSRC	Contributing Source Identifier
CRC	Cyclic Redundancy Check
DCR	Degradation Category Rating
DM	Diagnostic Mode
DMOS	Degradation Mean Opinion Score
DTW	Dynamic Time Warping
DTX	Discontinuous Transmission
E2E	End-to-End
EPC	Evolved Packet Core
EPS	Evolved Packet System
ESP	Encapsulating Security Payload
ePDG	evolved Packet Data Gateway

ETSI	European Telecommunications Standards Institute
EVS	Enhanced Voice Services
E-UTRAN	Evolved UMTS Terrestrial Radio Access Network
FB	Full-band, 20 – 20 000 Hz
FFT	Fast Fourier Transform
FR	Full Reference
GBR	Guaranteed Bit Rate
GPS	Global Positioning System
GSM	Global System for Mobile Communications; Groupe Spécial Mobile
GSMA	GSM Association; Groupe Spéciale Mobile Association
GUI	Graphical User Interface
HSS	Home Subscriber Server
IEEE	Institute of Electrical and Electronics Engineers
IETF	Internet Engineering Task Force
IMS	IP Multimedia Subsystem
INMD	In-service non-intrusive measurement device, ITU-T P.561
IP	Internet Protocol
IPv4	Internet Protocol, verze 4
IPv6	Internet Protocol, verze 6
IPsec	Internet Protocol Security
ISIM	IP Multimedia Services Identity Module
ITU	International Telecommunication Union
ITU-T	ITU Telecommunication Standardization Sector
LPC	Linear Predictive Coding
LTE	3GPP Long Term Evolution
LQE	Listening Quality Estimated
LQO	Listening Quality Objective
M2E	Mouth-to-Ear
MCPTT	Mission Critical user plane Push To Talk
MGCF	Media Gateway Control Function
MME	Mobility Management Entity
MMTel	Multimedia Telephony

MOC	Mobile Originated Call
MOS	Mean Opinion Score
MPS	Multimedia Priority Services
MSC	Mobile Switching Center
MTC	Mobile Terminated Call
NB	Narrow-band, 300 – 3400 Hz
NR	No Reference
OS	Operating System
OSI	Open Systems Interconnection model
OTT	Over-the-top
PAMS	Perceptual Analysis/Measurement System
PC	Personal Computer
PCRF	Policy and Charging Rules Function
PDCP	Packet Data Convergence Protocol
PESQ	Perceptual Evaluation of Speech Quality, ITU-T P.862
POLQA	Perceptual Objective Listening Quality Assessment, ITU-T P.863
PS	Packet-Switched
PDN-GW	Public Data Network Gateway
PSQM	Perceptual Speech Quality Measure, ITU-T P.861
QCI	QoS Class Identifier
QoS	Quality of Service
RoHC	Robust Header Compression
SDES	Source Description Items
SDP	Session Description Protocol
SID	Silence Insertion Descriptor
SIP	Session Initiation Protocol
SLF	Subscriber Location Function
SR	Sender Report (RTCP)
SRVCC	Single Radio-Voice Call Continuity
CSRC	Synchronization Source Identifier
SWB	Super-wideband, 50 – 14 000 Hz
RR	Receiver Report (RTCP)

RTP	Real-time Transport Protocol
RTCP	RTP Control Protocol
TCP	Transmission Control Protocol
TCP/IP	Transmission Control Protocol/Internet Protocol Suite
TSL	Transport Layer Security
TWAG	Trusted WLAN Access Gateway
UDP	User Datagram Protocol
UE	User Equipment
UICC	Universal Integrated Circuit Card
UMTS	Universal Mobile Telecommunications System
URI	Uniform Resource Identifier
USB	Universal Serial Bus
VAD	Voice Activity Detector
VoIP	Voice over Internet Protocol
ViLTE	Video over LTE
VoLTE	Voice over LTE
VoWiFi	Voice over Wi-Fi
VQE	Voice Quality Enhancement
WB	Wideband, 50 – 7000 Hz
WLAN	Wireless Local Area Network

OBSAH PŘILOŽENÉHO DVD

Na následující straně je uvedena adresářová struktura obsahu DVD přiloženého k diplomové práci. Adresář **Digitální verze práce** obsahuje 3 kopie ve formátu PDF určené pro tisk (všechny obrázky vektorové), digitální prohlížení (náročné obrázky rastrové) a komprimovanou verzi pro informační systém VUT. Dále obsahuje podadresář se zdrojovými soubory všech obrázků v práci ve formátech EPS (Encapsulated PostScript), FIG (MATLAB Figure), PDF (Portable Document Format), PNG (Portable Network Graphics) a SVG (Inkscape Scalable Vector Graphics).

Adresář **Měření hlasové kvality** obsahuje data z měření popsanych v kapitole 4. Protože veškeré datové soubory z měření systémem Accuver XCAL zabírají 180 GB, podadresář **Accuver měření - ukázka** obsahuje jen soubory z jednoho zvoleného měření. Vnořené adresáře obsahují WAV soubory všech degradovaných nahrávek v průběhu daného měření.

Adresář **Exportovaná měření** obsahuje vybraná data exportovaná z datových souborů měření prostřednictvím programu Accuver XCAP-U. Jedná se o všechna užitečná data, která byla v souvislosti s prací analyzována. Adresář **Exportovaná měření s přepočty** obsahuje tatáž data, jen jsou odfiltrována měření získaná ve verzi Accuver XCAL-M 3.3.4 (více v kap. 4.1.4 na str. 59) a jsou již doplněna o nezbytné výpočty vstupních parametrů navrženého algoritmu výpočtu hlasové kvality.

Adresář **Měření POLQA** obsahuje nezpracované výstupy algoritmu POLQA ze všech proběhlých měření. Obsah je totožný se souborem **POLQA - souhrn všech měření.xlsx**, ve kterém je však všech 27 835 měřených vzorků sdruženo do jednoho souboru a chronologicky seřazeno. Adresář **WinPcap** obsahuje textové soubory vygenerované z měření programem Accuver XCAL-M. Soubor **PPP Frame Message Flow.txt** obsahuje výpis příchozích a odchozích paketů během jednoho testovacího VoLTE hovoru. Soubor **RTP,RTCP pakety.txt** obsahuje výpis dvou odeslaných a dvou přijatých RTP paketů a také jednu odeslanou a jednu přijatou RTCP zprávu. Soubor **SIP zprávy.txt** obsahuje zápis z průběhu inicializace měřeného VoLTE hovoru prostřednictvím SIP zpráv včetně SDP komunikace.

Soubor **Celkový souhrn měření.xlsx** obsahuje finální výpis všech úspěšných měření, celkem 17 006 záznamů, včetně výpočtů jednotlivých parametrů navrženého algoritmu.

/(4,59 GB)

- ├─ Digitální verze práce
 - ├─ Obrázky
 - ├─ EPS
 - └─ (48 položek)
 - ├─ FIG
 - └─ (13 položek)
 - ├─ PDF
 - └─ (48 položek)
 - ├─ PNG
 - └─ (48 položek)
 - ├─ SVG
 - └─ (48 položek)
 - ├─ Měření hlasové kvality technologie VoLTE-VoWiFi_Print.pdf
 - ├─ Měření hlasové kvality technologie VoLTE-VoWiFi_Digital.pdf
 - ├─ Měření hlasové kvality technologie VoLTE-VoWiFi_IS.pdf
 - └─ Měření hlasové kvality
 - ├─ Accuver měření - ukázka
 - ├─ 20160318_2003_Napa_VoLTE_Duplex_M2E-M1
 - └─ (18 položek)
 - ├─ 20160318_2003_Napa_VoLTE_Duplex_M2E-M2
 - └─ (19 položek)
 - ├─ 20160318_2003_Napa_VoLTE_Duplex_M2E.jpg
 - ├─ 20160318_2003_Napa_VoLTE_Duplex_M2E.xls
 - ├─ 20160318_2003_Napa_VoLTE_Duplex_M2E_POLQAInfo.xls
 - ├─ 20160318_2003_Napa_VoLTE_Duplex_M2E_RTPInfo.xlsx
 - ├─ 20160318_2003_Napa_VoLTE_Duplex_M2E-M1.drm
 - ├─ 20160318_2003_Napa_VoLTE_Duplex_M2E-M1.txt
 - ├─ 20160318_2003_Napa_VoLTE_Duplex_M2E-M2.drm
 - ├─ 20160318_2003_Napa_VoLTE_Duplex_M2E-M2.txt
 - ├─ Exportovaná měření
 - └─ (84 položek)
 - ├─ Exportovaná měření s přepočty
 - └─ (56 položek)
 - ├─ Měření POLQA
 - └─ (98 položek)
 - ├─ WinPcap
 - ├─ PPP Frame Message Flow.txt
 - ├─ RTP,RTCP pakety.txt
 - ├─ SIP zprávy.txt
 - ├─ Celkový souhrn měření.xlsx
 - └─ POLQA - souhrn všech měření.xlsx