

NADEŚLANE ARTYKUŁY

METODY OCENY JAKOŚCI DŹWIĘKU W TELEFONII VOICE OVER IP

Łukasz Apiecionek

*Uniwersytet Kazimierza Wielkiego
Instytut Mechaniki i Informatyki Stosowanej
doktorant
ul. Chodkiewicza 30, 85-064 Bydgoszcz
e-mail: lukaszapi@gmail.com*

Streszczenie: *Algorytmy oceny jakości dźwięku pracują z sygnałem wzorcowym bądź bez niego. Metody charakteryzują się różnymi wymaganiami na moc obliczeniową procesora oraz pamięć danych. Determinuje to możliwość zastosowania metod w trybie monitorowania stanu pracy sieci VoIP. Niniejszy artykuł przedstawia opis metod oceny jakości dźwięku, które można próbować zastosować dla telefonii VoIP.*

Słowa kluczowe: *VoIP, jakość dźwięku, PESQ, E-model*

Quality evaluation methods of a sound in Voice over IP telephony

Abstrakt: *Voice quality measurement algorithms could work with or without reference signal. They have different needs for CPU performance and data memory. This needs could disturb possibility of using algorithms in real VoIP network. This paper contains descriptions of quality measurement algorithms in case of using its in VoIP networks.*

Keywords: *VoIP, voice quality, PESQ, E-model*

1. WSTĘP

Metody oceny jakości dźwięku można podzielić na:

- metody subiektywne,
- metody obiektywne.

Metody subiektywne polegają na ocenie systemu poprzez jego użytkownika w laboratorium testowym. Metody obiektywne oceniają wartość jakości przez porównanie próbek mowy nadanej z mową u odbiorcy.

Ważną grupę stanowią metody estymowane, w których obliczana jest wartość MOS (przewidująca wartość obiektywną) na podstawie metryk różnorodnych

parametrów np. opóźnienia, jitter, utraty pakietów dźwięku, czy zastosowanego sposobu jego kodowania.

Zarówno metody subiektywne jak i obiektywne mogą być realizowane w trybie odsłuchowym jak i konwersacyjnym. Metody odsłuchowe polegają na transmisji jednokierunkowej dźwięku, podczas gdy metody konwersacyjne wykorzystują dźwięk w przeprowadzanej rozmowie.

Uwzględniając sposób pomiaru jakości sygnału możemy rozpatrywać następujące tryby pomiaru:

- na obu końcach systemu (ang. double-ended),
- na jednym końcu systemu (ang. single-ended),
- na podstawie parametrów systemu,
- hybrydowo - wykorzystując parametry systemu jak i pomiar sygnałów.

Ze względu na algorytm działania metody oceny jakości można również podzielić na:

- ingerujące w strukturę badanego systemu,
- nie ingerujące w badany system (ang. *non-intrusive*) [1] [2].

Metody oceny jakości dźwięku, których algorytm działania ingeruje w badany system, potrzebują sygnału wzorcowego – sygnału odniesienia, który w warunkach laboratoryjnych można w dość prosty sposób dostarczyć do systemu pomiarowego, podczas gdy w systemie rzeczywistym, wymaga to specjalnych rejestratorów dźwięku, które po podłączeniu do systemu pomiarowego muszą być synchronizowane w czasie, ze względu na opóźnienia w wykorzystywanej sieci. Bazują one na porównaniu różnic w sygnale, który został przesłany przez system z sygnałem, który został do badanego systemu przyłożony. Na podstawie tych różnic określają wynik pomiaru jakości transmisji sygnału w systemie. Metody oceny jakości dźwięku, które polegają tylko na „podśluchu” badanego systemu, nie ingerują w niego, pozwalając na jego normalne funkcjonowanie, przez co spełniają wymagania na metody służące do nadzorowania pracy systemu. Bazują one na parametryzacji sygnału, obliczaniu jego cech charakterystycznych, a następnie na ich podstawie określają wynik pomiaru.

2. PRZEGLĄD METOD OCENY JAKOŚCI DŹWIĘKU

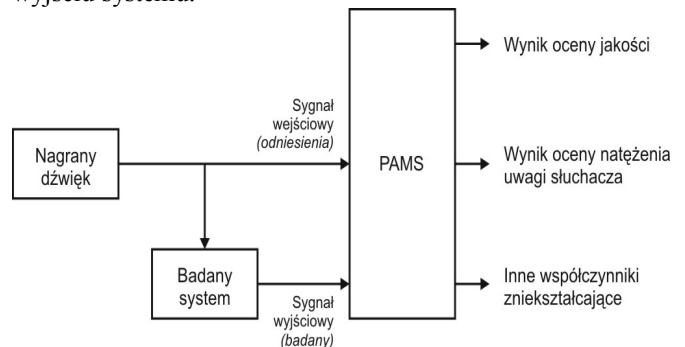
I. PAMS

Intencją opracowania metody PAMS (ang. Perceptual Analysis Measurement System) było znalezienie metody pomiaru jakości dźwięku w sieciach VoIP w przypadku występowania takich zdarzeń jak:

- czasowe przerwy w sygnale,
- straty pakietów,
- jitter,
- zniekształcenia wniesione poprzez różne metody kodowania dźwięku.

Metoda została przedstawiona w 1998 roku poprzez firmę PsyTechnics i pracuje z sygnałem odniesienia [3]. Oznacza to, że system może być badany:

- w warunkach laboratoryjnych,
- poprzez nagrywanie próbek dźwięku na wejściu i wyjściu systemu.



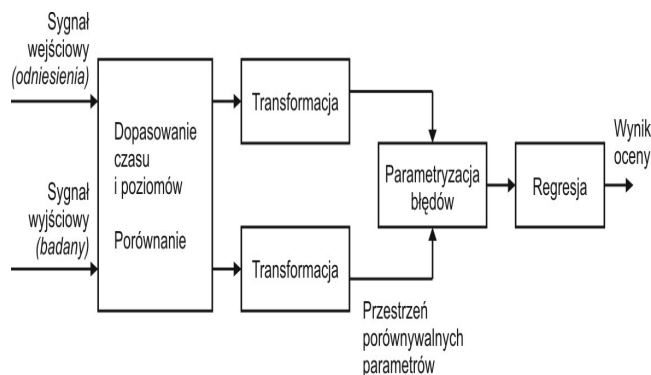
Rysunek. 1 Schemat pomiaru za pomocą metody PAMS (ang. Perceptual Analysis Measurement System)

Metoda poza wynikiem oceny jakości przesłanego dźwięku dostarcza informacji na temat prawdopodobnego wymaganego natężenia uwagi słuchacza oraz innych współczynników określających zniekształcenia dźwięku w badanym sygnale. Wynik oceny jakości przesłanego dźwięku dokonywany metodą PAMS przedstawiony jest w skali MOS. Metoda PAMS jest w stanie zmierzyć wpływ jitter'a, przesunięcia w czasie i poziomów sygnału oraz poziom szumów, natomiast nie uwzględnia opóźnienia w transmisji sygnału [3].

Warunkami brzegowymi użycia metody PAMS jest:

- brak dużych poziomów szumów,
- stały poziom mocy sygnału badanego,
- brak zbyt dużego poziomu jitter'a.

Krytycznym elementem metody jest odwzorowywanie wyników na skalę MOS [4].



Rysunek. 2 Schemat strukturalny metody PAMS

II. PESQ

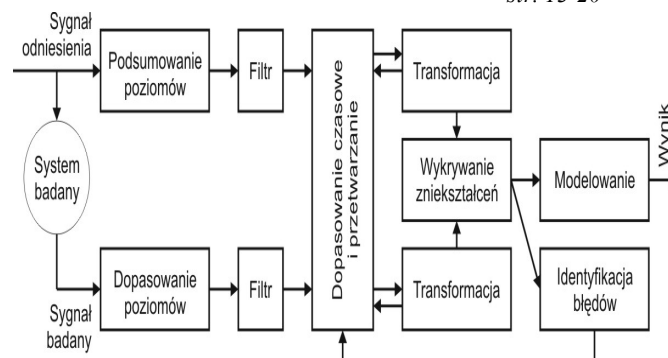
Metoda PESQ (ang. Perceptual Evaluation of Speech Quality) została opracowana i zaakceptowana przez ITU w 2001 roku jako rekomendacja P.862 i szybko stała się szeroko używaną metodą do oceny jakości systemów [5]. Ocenia jakość dźwięku z uwzględnieniem:

- czasowych przerw w sygnale,
- straty pakietów,
- efektów filtracji sygnału,
- jitter'a,
- zniekształceń wniesionych poprzez kanał transmisyjny,
- transkodowania,
- szumów dodanych przez system transmisyjny.

Metoda ta nie uwzględnia i nie nadaje się do pomiarów [3]:

- opóźnień, utraty poziomu sygnału jak i całkowitego poziomu szumów, co spowodowane jest wprowadzeniem bloków dopasowujących sygnał w czasie i w jego poziomie,
- wpływu jednoczesnej rozmowy w dwóch kierunkach,
- niedopasowania kodowania sygnału,
- szumów w sygnale wejściowym,
- muzyki w sygnale wejściowym,
- dźwięku po kompresji o przepływnościach mniejszych od 4 kbs,
- echa,
- sprzężenia zwrotnego sygnałów.

Braki w implementacji wymienionych efektów w metodzie PESQ powodują, iż czasami wynik oceny jakości sygnału może być dobry (wysoki wskaźnik jakości MOS) podczas gdy sygnał faktycznie jest bardzo słabej jakości.



Rysunek. 3 Schemat algorytmu metody PESQ (ang. Perceptual Evaluation of Speech Quality)

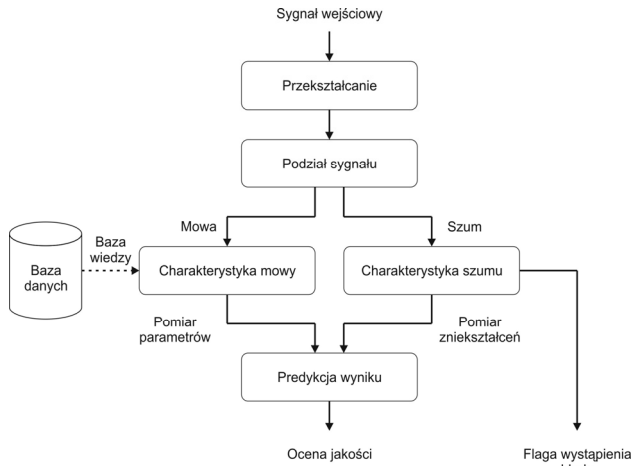
III. PSQM

Metoda PSQM (ang. Perceptual Single ended Objective Measure) została zarekomendowana przez ITU-T w 1996 roku. Zniekształcenia sygnału są mierzone na podstawie porównań i odniesień do statystycznych wyników zawartych w bazie danych. Dysponowanie odpowiednią bazą danych jest ograniczeniem i główną wadą metody PSQM. Metoda PSQM bazuje na analizie parametrów w dziedzinie czasowej i częstotliwościowej. Sygnał mowy jest oddzielany od szumu, a następnie oceniany z wykorzystaniem statystycznej bazy danych. Pomiar poziomu zniekształceń jest dokonywany na podstawie pomiaru szumów. Korelacja poziomu zniekształceń wraz z parametrami mowy pozwala na predykcję wyniku jakości sygnału mowy w systemie. Wynik metody PSQM jest wykorzystywany do obliczenia wyniku skali MOS według wzoru [6]:

$$MOS = \frac{4}{1 + e^{0,66PSQM-2,2}} + 1$$

gdzie:

- $PSQM$ - wynik oceny jakości dźwięku przez metodę PSQM.



Rysunek. 4 Algorytm działania metody PSQM (ang. Perceptual Single ended Objective Measure)

IV. E-model

E-model został zaproponowany przez ETSI (Europejski Instytut Standardów Telekomunikacji) [9]. ITU wydało stosowne rekomendacje w 1998 roku [7][8]. Zadaniem modelu było dostarczenie dla operatorów i projektantów narzędzia do projektowania sieci IP uwzględniającego czynniki obiektywne, wynikające z konfiguracji sieci, metody transmisji, czy stosowanych kodeków. E-model jest modelem matematycznym, który zakłada, że pomiędzy ustami osoby mówiącej, a uchem słuchacza mamy do czynienia z kanałem telekomunikacyjnym składającym się z wielu torów i elementów, które jako całość tworzą łącze telefoniczne. Jakość transmitowanej mowy jest zależna od następujących parametrów i czynników:

- rodzaju elementów końcowych, mianowicie telefony analogowe i cyfrowe:
 - stacjonarne,
 - bezprzewodowe,
 - mobilne,
- rodzaju elementów połączeń, analogowe i cyfrowe centrale:
 - abonenckie,
 - międzymiastowe,
- rodzaju elementów transmisji, łącza:
 - przewodowe,
 - światłowodowe,
 - cyfrowe,
- konfiguracji połączenia:
 - długości połączenia,
 - typu połączenia,
 - rodzaju składników sieci,
- parametrów mowy i słuchu człowieka,

- tłumienia transmisji mowy między mówcą a słuchaczem,
- tłumienia linii abonenckiej,
- tłumienia obwodów 4-przewodowych,
- efektu lokalny echa mówcy i słuchacza,
- echa,
- stabilności systemu,
- czasu transmisji – opóźnienia,
- szumu i zniekształceń kwantyzacji,
- przesłuchów.

Ważnym ogniwem jest też występowanie urządzeń specjalnych, takich jak kodery mowy o małej przepływności binarnej, systemy mobilne, systemy wykrywania mowy VAD (ang. Voice Active Detection) czy systemy eliminacji echa za pomocą procesorów DSP. W E-modelu zakłada się, że pomiar jakości mowy przesyłanej przez łącze telekomunikacyjne odbywa się w słuchawce telefonicznej słuchacza dla pasm częstotliwości od 300 Hz do 3400 Hz. Wynik oceny jakości mowy w E-modelu wyznacza się za pomocą współczynnika jakości transmitowanego dźwięku R , który określa zadowolenie bądź niezadowolenie użytkowników z jakości mowy w badanym systemie. Należy podkreślić, że stopień zadowolenia bądź jego braku jest zależny zarówno od sposobu mówienia mówcy jak i koncentracji osoby słuchającej.

Współczynnik R wyrażany jest wzorem:

$$R = R_O - I_s - I_d - I_e + A$$

gdzie:

- R_O - podstawowy współczynnik sygnał/szum SNR (ang. Signal Noise Ratio) transmisji mowy w punkcie odniesienia 0 dBr;
- I_s - współczynnik pogorszenia jakości transmitowanej mowy, reprezentujący zniekształcenia mowy podobne jak przy odbiorze zbyt głośnego sygnału mowy, przy nieoptymalnym efekcie lokalnym echa lub na skutek zniekształceń kwantyzacji;
- I_d - współczynnik pogorszenia jakości transmitowanej mowy, wywołanego echem odnoszącym się do sygnału mowy, tj. tłumienności głośności echa mówcy TELR (ang. Talker Echo Loudness Rating) i ważonej straty efektu echa podczas transmisji w systemie WEPL (ang. Weighted Echo Path Lost) oraz trudności w porozumiewaniu się, spowodowane zbyt dużym bezwzględnym opóźnieniem transmitowanych próbek dźwięku T_a ;
- I_e - współczynnik pogorszenia jakości transmitowanej mowy, spowodowanego zastosowaniem w łączu specjalnych urządzeń, tj. koderów o małej przepływności binarnej, DCME (ang. Digital Circuit Multiplication

Equipment), VPE (ang. Voice Packeting Equipment) i innych; wpływ tych urządzeń na jakość mowy jest bardzo złożony i trudno jest go analizować, biorąc pod uwagę poszczególne parametry;

- A – współczynnik oczekiwania jakości dźwięku w systemie telekomunikacyjnym, będący wielkością dodatnią, reprezentującą korzyści wynikające z używania różnych systemów transmisji mowy. Np. w niektórych okolicznościach systemy radiowe mają znaczną przewagę nad systemami kablowymi, pomimo gorszej jakości transmisji mowy. Dotyczy to rozwoju telefonii mobilnej w trudno dostępnych regionach, gdzie względy ekonomiczne okazały się ważniejsze niż jakość połączenia [9]. Skrócony opis modelu E pokazuje jego przydatność i uniwersalność do oceny różnorodnych projektów sieci i występujących w nich możliwych połączeniach [9]. Wykorzystując E-model można przewidywać jakość mowy w budowanym systemie, jednak konieczne jest odpowiednie dobranie współczynników do wzorów obliczeniowych. Model ten jest dobrym modelem do badań w środowiskach testowych, gdzie w idealnych warunkach można dokonywać pomiarów wymaganych parametrów. Natomiast nie nadaje się do użycia w działającej sieci oferującej usługę VoIP, ponieważ nie ma wówczas możliwości pomiaru wszelkich niezbędnych do prawidłowego działania E-modelu parametrów.

V. P.563

Metoda P.563 ta została zatwierdzona przez ITU-T w maju 2005 roku [10]. Wydana rekomendacja opisuje metodę, nie ingerującą w badane środowisko, do predykcji subiektywnej oceny jakości przesyłanej mowy w wąskopasmowej telefonii (pasmo 3,1 kHz). Należy tutaj podkreślić, iż algorytm metody jest aplikacją przewidującą jakość systemu bez użycia sygnału odniesienia. ITU zaleca tą metodę do monitorowania sieci przy użyciu nieznanego źródła dźwięku pochodzącego z dalekiego miejsca w sieci. Jest to pierwsze podejście ITU do metod pomiaru jakości bez sygnału odniesienia.

Algorytm był testowany z różnorodnymi źródłami dźwięku w laboratoriach różnych firm, co skutkowało publikacjami dokumentów i wyników tych badań [11]. Metoda P.563 została opracowana w taki sposób, żeby istniała możliwość pomiaru takich cech

w transmitowanym dźwięku jak:

- charakterystyka akustyczna środowiska,
- szum środowiska po stronie nadawczej,
- charakterystyka akustyczna interfejsu nadawczego,
- zniekształcenia spowodowane elektroniką po stronie nadawczej,

- charakterystyka kodowania,
 - poziom sygnału nadawanego,
 - błędy w kanale transmisyjnym,
 - straty pakietów i ukryte straty w wyniku stosowania kodowania CELP (ang. Code Excited Linear Prediction),
 - przepływność standardu kodowania próbek dźwięku jeśli posiada więcej niż jedną,
 - zmiana sposobów kodowania próbek dźwięku,
 - efekt jitter w odniesieniu do testów ACR,
 - krótkotrwałe wypaczenia w dźwięku,
 - długotrwałe wypaczenia w dźwięku,
 - systemy transmisyjne eliminacji echa i szumów w warunkach rozmowy jednokierunkowej w odniesieniu do testów ACR,
 - różne technologie kodowania, np. G.711; G.726; G.727, G.729, G.723.1, G.728, Metodę P.563 rekomenduje się do użycia w następujących środowiskach:
 - monitorowanie stanu pracy sieci w czasie rzeczywistym stosując interfejs cyfrowy bądź analogowy,
 - testowanie całościowe (od nadawcy do odbiorcy) systemu transmisji w czasie rzeczywistym,
 - testowanie całościowe (od nadawcy do odbiorcy) systemu transmisji w czasie rzeczywistym przy braku znajomości sygnału nadawanego.
- W dokumentacji metody P.563 [10] podane są efekty, które wpływają na niewłaściwe oszacowanie wyniku:
- poziom słyszanego dźwięku,
 - straty poziomu sygnału,
 - efekt opóźnienia w teście rozmowy dwukierunkowej,
 - echo,
 - muzyka w sygnale wejściowym,
 - kodowanie według standardu LPC (kodowanie z predykcją liniową, ang. Linear Predictive Coding) o przepływności mniejszej niż 4 kbit/s,
 - niedopasowanie przepływności standardu kodowania dźwięku na wejściu i wyjściu systemu,
 - stosowanie syntetyzowanej mowy.

Algorytm metody P.563 wymaga obliczenia parametrów do oszacowania różnorodnych zjawisk w transmisji dźwięku (ludzkiej mowy) i na ich podstawie dokonywane jest modelowanie jakości dźwięku. W metodzie P.563 można wydzielić trzy główne klasy pomiarowe:

- analiza wokalna traktu oraz nienaturalności dźwięku (mowy):
 - głos mężczyzny,
 - głos kobiety,
 - głos syntetyzowany,

Klasa ta jest głównym elementem testów występowania nienaturalności w transmitowanej mowie. Zawiera komponenty wykrywania w sygnale danych, które mogą

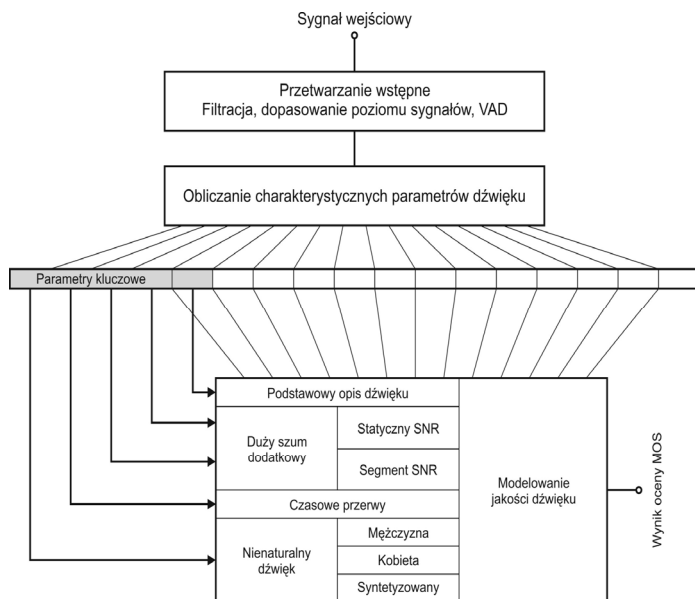
być interpretowane jako głos oraz osobno wykrywają dane reprezentujące szумы. Ponadto odpowiada za wykrywanie powtarzających się próbek w dźwięku. Jednym z zadań elementów tej klasy jest porównanie sygnału z wygenerowanym sztucznie sygnałem odniesienia.

- analiza dodatkowego szumu:
 - statyczny pomiar SNR,
 - segment pomiaru SNR,

Zadaniem elementów w tej klasie jest ocena charakterystyk różnych szumów, które mogą wystąpić w transmitowanym dźwięku. Oszacowują również czy występujący szum jest zdarzeniem lokalnym, czy globalnie ma wpływ na ogólną jakość sygnału.

- przerwy, cisza i czasowe zaniki dźwięku

Zadaniem elementów tej klasy jest wykrywanie przerw w sygnale, które mogą być spowodowane stosowaniem mechanizmów VAD, DCME albo zakłóceń w kanale transmisyjnym. Algorytm stosowany w metodzie zaprojektowany jest do wykrywania zarówno normalnego zakończenia słów jak i ich urwania.



Rysunek. 5 Algorytm metody P.563

Ponadto metoda P.563 stosuje deskryptory opisujące poziom dźwięku, aktywność mowy i jej zmiany. Wykorzystywane są funkcje wykrywania głosu VAD i analiza danych.

VI. K-factor

Metoda ta jest wykorzystywana przez firmę Cisco w rozwiązaniach do budowy sieci oferujących usługę

VoIP. Telefony firmy Cisco potrafią na bieżąco informować użytkownika o wartości wskaźnika MOS przeprowadzanej rozmowy. Ponadto są w stanie podawać bieżącą wartość opóźnienia oraz jitter. Uwzględniając informacje o sposobie kodowania dźwięku, jaki jest używany w przeprowadzanej rozmowie, telefon umożliwia uzyskanie dość dużej ilości danych na temat stanu pracy sieci. Informacje o wartości wskaźnika MOS są przechowywane w bazach danych Cisco Call Managera, który pełni rolę „centrali” dla urządzeń firmy Cisco. Według informacji od producenta metoda K-factor estymuje wartość wskaźnika MOS na podstawie:

- stosowanej metody kodowania dźwięku,
- liczby utraconych pakietów.

Metoda nie uwzględnia w swoich obliczeniach żadnych parametrów „kanału” związanego choćby z opóźnieniem pakietów transmitowanych w sieci. Algorytm metody nie jest znany, przez co trudno mieć zaufanie do wyników jakie przedstawia

3. PODSUMOWANIE I WNIOSKI

Analizując metody oceny jakości transmisji dźwięku postanowiono dokonać ich oceny pod kątem przydatności do zastosowań w monitorowaniu stanu pracy sieci oferujących usługę VoIP.

E-Model

Głównym celem opracowania E-modelu była konieczność posiadania narzędzia do projektowania sieci telefonicznych. Stosując wzory obliczeniowe przy ekstrakcji parametrów sieci rzeczywistej można pokusić się pod pewnymi warunkami o predykcję jakości. Otóż algorytm musi posiadać parametry bazowe – zmierzone w środowisku pracy terminala VoIP czy parametry samego terminala VoIP. Ponieważ w sieci nie można określić i przewidzieć wszystkich warunków pracy, nie jest możliwe przygotowanie parametrów do obliczeń. Ponadto trudno jest przewidzieć jakie parametry będzie miał sam terminal i nie można określić czy jego parametry nie uległy zmianie w stosunku do parametrów zakładanych. Gdyby zbudować sieć firmową tylko z urządzeń jednego producenta, to i tak zastosowanie modelu E byłoby trudne.

K-factor

Metoda bazująca na wskaźniku K jest metodą opracowaną przez producenta urządzeń i może być wykorzystana tylko

przy pracy z urządzeniami pochodzącymi od niego. Gdy sieć oferująca usługę VoIP nie jest siecią zamkniętą – tj. jednego producenta, gdy łączy się z sieciami publicznymi, określanie jakości dźwięku przy wykorzystaniu metody K-factor jest już ograniczone. Brak pełnej dokumentacji metody sprowadza ją do roli wskaźnika w zamkniętym systemie telefonów wymienionej firmy. Ponadto, jak już wspomniano, metoda nie uwzględnia w swoich obliczeniach żadnych parametrów związanych choćby z opóźnieniem pakietów transmitowanych w sieci, mimo, że terminale posiadają takie informacje.

PSQM

Wadą zastosowania metody PSQM jest konieczność posiadania statystycznej bazy danych służącej do określania wpływu parametrów na jakość sygnału. Metodę można by zastosować w sieci z usługą VoIP, gdyby parametry dźwięku były zbierane w punktach sieci, a następnie przesyłane do stacji zarządzającej. Dopiero stacja zarządzająca posiadając odpowiednie zasoby na bazę danych dokonywała by oceny jakości przeprowadzanej rozmowy. Takie rozwiązanie stawia jednak duże wymagania na moc obliczeniową stacji zarządzania. Zamiast tylko zbierać informacje i przekazywać je operatorowi sieci, zajmowała by się nieustannie dokonywaniem obliczeń i predykcją wyników. Konieczna moc obliczeniowa rosła by wraz ze wzrostem ilości użytkowników w sieci co wiąże się z ilością przeprowadzanych rozmów. Podejście to powoduje konieczność budowy skomplikowanej stacji zarządzającej. Nie będzie wówczas możliwości korzystania z oprogramowania do monitorowania stanu pracy sieci z dowolnego terminala w sieci IP.

Metoda PAMS i PESQ

Metody te nie nadają się do zastosowania w monitorowaniu stanu pracy sieci oferujących usługę VoIP ze względu na konieczność posiadania sygnału wzorcowego. Niemożliwe wręcz jest opracowanie sposobów na dostarczenie do każdego badanego punktu w sieci sygnału odniesienia, przeprowadzenie rozmowy a następnie dokonanie oceny jej jakości. Obecnie metoda PESQ jest najszerzej stosowaną metodą w laboratoriach. Posiada jednak swoje wady. Przeprowadzono testy metody PESQ, których wyniki były bardzo zaskakujące [12]. Otóż przygotowano dwa fragmenty dźwięku poprzez modyfikację wzorca. W pierwszym podejściu wzorzec zmiksowano z szumem białym o niskim poziomie mocy. Po odsłuchu uzyskanego nagrania nie stwierdzono

problemów z jakością. Natomiast drugie nagranie do badań uzyskano poprzez poddanie wzorca obróbce polegającej na zmniejszeniu ilości składowych harmonicznym o małej i wysokiej częstotliwości. Odsłuchiwane nagranie znacznie odbiega od oryginału i bez problemów można stwierdzić utratę jakości. Wyniki pomiaru jakości nie wykazały różnic.

Metoda P.563

Wadą metody P.563 jest stosunkowo duży błąd w przypadku krótkotrwałej utraty dźwięku. Przeprowadzone przez autora badania wykazały, że metoda w takich przypadkach przedstawia błędne wyniki oceny jakości (wynik jest zaniżony bądź zawyżony), w związku z czym nie spełnia postawionych przed nią wymagań

Podsumowując przegląd przedstawionych w artykule metod oceny jakości transmitowanego dźwięku, można stwierdzić, że nie ma dostatecznie dobrej metody, która nadawała by się do zastosowania w sieci telefonicznej VoIP. W związku z czym, istnieje uzasadniona konieczność prac nad algorytmami oceny jakości transmisji dźwięku dla sieci telefonicznej VoIP.

Literatura

1. J. Davidson, J. Peters, VOICE OVER IP PODSTAWY, Mikom, 2005
2. D. Picovici, A. E. Mahdi, Towards Non-intrusive Speech Quality Assessment for Modern Telecommunications, First Joint IEI/IEE Symposium on Telecommunications Systems Research, Dublin, 2001
3. E. Bernex, A. Gatineau, Quality of service in VoIP environments, White Paper, www.neotip.com
4. M.P. Hollier, PAMS: Measuring speech quality over networks, as the customers hear it, Psytechnics white paper, 2001
5. J. Anderson, Methods for measuring perceptual speech quality, Agilent Technologies white paper, 2000
6. R. Dai, A Technical White Paper on Sage's PSQM Test, White paper, 2000
7. ITU-T Recommendation G.107, The E-model, a computational model for use in transmission planning, grudzień 1998
8. ITU-T Recommendation G.109, Definition of categories of speech transmission quality, wrzesień 1999
9. M.J. Trzaskowska, Matematyczny model do oceny jakości mowy przesyłanej w sieciach

telekomunikacyjnych, Telekomunikacja i techniki informacyjne, nr 1, 2001

10. ITU T Recommendation P.563, Single-ended method for objective speech quality assessment in narrow-band telephony applications, maj 2004

11. Opticom GmbH, P.563 – Single Sided Speech Quality Measure, www.opticom.de, 2006

12. P. Ordas, B. Fox, Perceptual Evaluation of Speech Quality (PESQ), Microtronix Systems Ltd, 2004, <http://www.microtronix.ca/pesq-disc.html>