

# *Jakość usługi głosowej w sieciach telekomunikacyjnych*

*Ryszard Kobus,  
Marian Kowalewski, Bogdan Mucha*

*Dokonano przeglądu najważniejszych obiektywnych metod porównawczych, stosowanych przy ocenie jakości mowy transmitowanej we współczesnych sieciach telekomunikacyjnych. Wyszczególniono i zbadano wpływ czynników degradujących jakość mowy transmitowanej w wybranych sieciach oraz przedstawiono ogólny proces testowania jakości mowy transmitowanej. Zaproponowano optymalne i akceptowane przez użytkowników wartości wskaźnika jakości mowy.*

*usługa głosowa w sieciach telekomunikacyjnych, jakość transmisji mowy, metody badań jakości mowy*

## **Wprowadzenie**

Dokonując oceny jakości mowy transmitowanej w sieciach pakietowych czy komórkowych, należy brać pod uwagę czynniki, które nie występowały w tradycyjnych sieciach PSTN (*Public Switched Telephone Network*), takie jak:

- opóźnienie stałe,
- straty pakietów,
- fluktuacja, czyli zmienność opóźnienia (*jitter*).

Na jakość przekazu mowy postrzeganą przez użytkownika końcowego mają również wpływ parametry terminali, takie jak: charakterystyki kodeka, echo, wielkość bufora fluktuacji oraz jego parametry akustyczne. Ze względu na ich powiązania oraz wzajemny wpływ, prognozowanie staje się zadaniem trudnym, jednak widać stałe postępy w algorytmach oceny jakości mowy. Zapewniają one wysoki poziom korelacji wyników z pomiarów obiektywnych z wynikami z badań subiektywnych.

W niniejszym artykule, oprócz omówienia podstawowych metod pomiarowych jakości mowy, podano praktycznie otrzymywane wyniki dla sieci heterogenicznych oraz propozycje możliwych do uzyskania wartości wskaźników.

## **Metody badania jakości mowy transmitowanej**

Wyodrębniono dwie kategorie metod pomiaru jakości sygnału mowy:

- metody subiektywne, w których ocena jest dokonywana przez człowieka;
- metody obiektywne, w których ocena jest oparta na podstawie wielkości fizycznych.

Dalej zostaną one krótko opisane.

## Metody subiektywne

Pierwszą metodą stosowaną przy badaniach jakości mowy transmitowanej były testy subiektywne, polegające na dokonaniu oceny przez ściśle określoną i przygotowaną do tego celu grupę słuchaczy. Metoda ta została określona jako średnia wartość oceny wszystkich słuchaczy MOS (*Mean Opinion Score*) – por. tabl. 1.

Specyfikacje i opisy sposobów przeprowadzania wielu rodzajów testów subiektywnych przedstawiono w zaleceniu ITU-T P.800 [1].

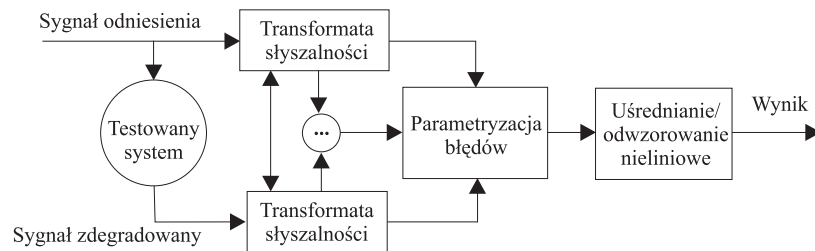
Tabl. 1. Ocena MOS w zależności od kategorii

Skala	Ocena	Kategoria		
		bezwzględna ocena	wysiłek słuchowy	głośność
A	5	Doskonała	Całkowity relaks	Znacznie głośniej niż potrzeba
B	4	Dobra	Niewielka uwaga	Głośniej niż potrzeba
C	3	Dość dobra	Umiarkowana uwaga	Głośność taka jak potrzeba
D	2	Słaba	Znaczny wysiłek słuchowy	Ciszej niż potrzeba
E	1	Zła	Brak zrozumienia, duża uwaga	Znacznie ciszej niż potrzeba

Metody subiektywne uwzględniają wszystkie czynniki odczuwane przez człowieka, wymagają jednak przeszkolonej i licznej ekipy, są kosztowne oraz czasochłonne.

## Obiektywne metody porównawcze

W przeciwieństwie do metod subiektywnych metody obiektywne są tanie, a ich największą zaletą jest prostota. Nie wymagają wcześniejszych przygotowań i angażowania dużej liczby przeszkolonych osób. Badania polegają na wyznaczeniu różnicy, według przyjętych kryteriów, między sygnałem wzorcowym (odniesienia) i sygnałem zdegradowanym na skutek przetwarzania i jej subiektywnej ocenie. Metody te mogą jednak nie uwzględniać wszystkich czynników odczuwanych przez człowieka. Ogólna koncepcja szacowania jakości mowy została zaproponowana przez Karjalainena [2] w 1985 r. (rys. 1).



Rys. 1. Ogólna koncepcja metody porównawczej szacowania jakości mowy

Do najczęściej wykorzystywanych obiektywnych metod porównawczych należą:

- metoda PSQM (*Perceptual Speech Quality Measure*),

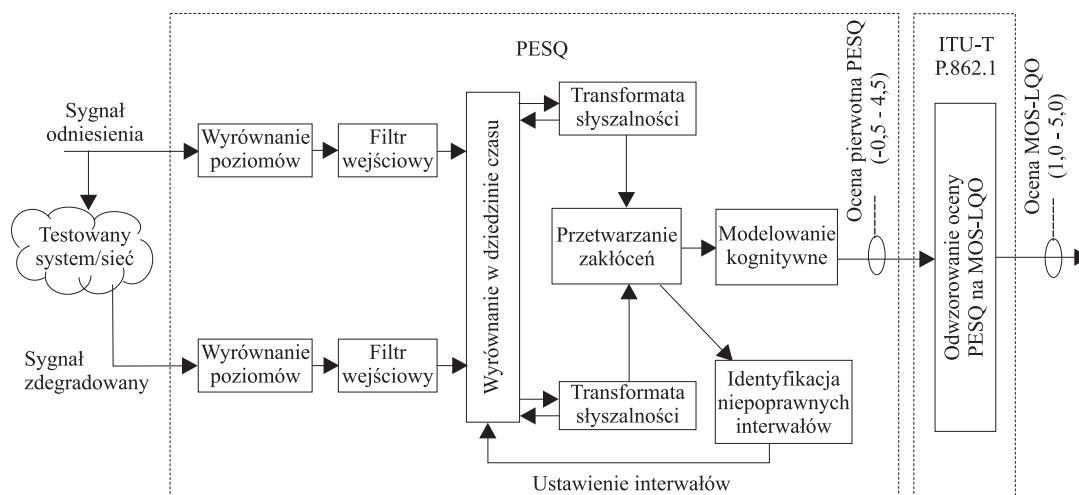
- metoda PAMS (*Perceptual Analysis Measurement System*),
- metoda PESQ (*Perceptual Evaluation of Speech Quality*).

**Metoda PSQM.** W latach 1993–1996 organizacja ITU podjęła prace nad oceną istniejących obiektywnych metod badania jakości mowy z punktu widzenia ich korelacji z subiektywną percepcją człowieka. Najwyższą ocenę uzyskała metoda PSQM, opracowana przez zespół w składzie J. G. Beerend i J. A. Stemerding z instytutu KPN Research w Holandii [3]. W konsekwencji metoda ta została zatwierdzona przez ITU i opublikowana w 1996 r. jako zalecenie P.861 [4].

**Metoda PSQM+.** W związku z dynamicznym rozwojem sieci nowej generacji (*Next Generation Network – NGN*) wzrosło zapotrzebowanie na tania metody badania jakości mowy przesyłanej przez sieć. Metoda PSQM, jako łatwa i tania w realizacji, uzyskała akceptację. Stała się popularna nie tylko w badaniach jakości mowy przy użyciu poszczególnych kodeków, ale podjęto też próby adaptacji tej metody do badania jakości mowy przesyłanej przez sieć, w tym również sieci IP. Jednak przeszkodą było nieuwzględnienie wpływu utraty pakietów na jakość przesyłanej mowy. Dlatego J. G. Beerend oraz E. J. Mejer opracowali rozszerzony model PSQM, który – po akceptacji ITU Study Group 12 – został opublikowany w 1997 r. jako COM 12-24-E [5]. Otrzymał on nazwę PSQM+ i stał się preferowaną metodą badania jakości mowy w otoczeniu sieciowym.

**Metoda PAMS.** Opracował ją zespół PsyTechnics, utworzony przez British Telecommunications [6]. W wielu aspektach jest ona podobna do metody PSQM, ma jednak inną technikę przetwarzania sygnałów oraz inny model percepcyjny. Wyniki pomiarów metodą PAMS zawierają się w przedziale od 0 do 5 i korelują ze skalą MOS, a wyniki jakości odsłuchu LQS (*Listening Quality Score*) oraz wyniki wysiłku wkładanego w zrozumiałość LES (*Listening Effort Score*) odpowiadają skali ocen, stosowanej w metodzie bezwzględnej oceny jakości mowy.

**Metoda PESQ.** Jest ona rozwinięciem metod PSQM oraz PAMS. Opisano ją w zaleceniu ITU-T P.862 [7]. Warunki i technika badania jakości mowy tą metodą są w większości zbliżone do metody PSQM. Różnica polega na tym, że w metodzie PESQ dodano operacje wyrównania opóźnienia oraz mocy między sygnałem odniesienia i sygnałem ocenianym (rys. 2).



Rys. 2. Schemat blokowy pomiaru jakości mowy metodą PESQ

W pierwszym etapie badania dokonuje się pomiaru obu sygnałów pod względem opóźnienia oraz mocy. Po wyznaczeniu różnicy tych parametrów zostają one wyrównane, co umożliwia stosowanie tej metody do badania jakości mowy podczas jej transmisji. Pozostała część badania przebiega podobnie jak w metodzie PSQM. Sygnały są zamieniane na reprezentacje wewnętrzne, po czym wyznaczone różnice służą do określenia jakości mowy. Otrzymany wynik jest prezentowany w skali oceny PESQ, od  $-0,5$  do  $4,5$  (kalibrowanej za pomocą bazy wyników uzyskanych z testów subiektywnych) lub według skorelowanej skali MOS od 1 do 5, zgodnie z zaleceniem ITU-T P.800 [1], lub MOS-LQO (*Mean Opinion Score-Listening Quality Objective*), zgodnie z zaleceniem ITU-T P.862.1 [8].

**Zastosowania oraz ograniczenia metody PESQ.** Zalecenie ITU-T P.862 [7] zawiera trzy tabele z opisem zastosowań metody PESQ. W niniejszym artykule w tabl. 2 połączono je i uzupełniono zastosowaniami, które nie są zalecane, a także dla których metoda nie została jeszcze zbadana.

**Tabl. 2. Zastosowania oraz ograniczenia metody PESQ**

Zastosowania zalecane	Zastosowania nie zalecane	Zastosowania nie zbadane
<ul style="list-style-type: none"> <li>● Ocena:                             <ul style="list-style-type: none"> <li>– sieci w warunkach rzeczywistych przez interfejsy cyfrowe lub analogowe</li> <li>– sieci pilotowych lub emulowanych</li> <li>– kodeków</li> <li>– wpływu poziomu mowy na jakość kodowania</li> </ul> </li> <li>● Uwzględnienie wpływu:                             <ul style="list-style-type: none"> <li>– błędów w kanale telekomunikacyjnym</li> <li>– utraty pakietów</li> <li>– krótkotrwałych zniekształceń sygnałów audio</li> <li>– długotrwałych zniekształceń sygnałów audio</li> <li>– niwelacji utraty pakietów dla kodeków CELP</li> <li>– szumów otoczenia po stronie nadawczej</li> <li>– zmienności opóźnienia</li> <li>– kodeków kształtu fali G.711, G.726, G.727</li> <li>– kodeków hybrydowych oraz CELP o szybkości większej niż 4 kbit/s</li> <li>– innych kodeków: FR, HR, EFR, AMR, CDMA-EVRC, TDMA-ACELP, VSELP, TETRA</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>● Echo mówcy</li> <li>● Efekt lokalny</li> <li>● Tłumienie głośności</li> <li>● Dwukierunkowa jakość rozmowy</li> <li>● Zastąpienie ciszą 25% wycinka mowy aktywnej (skrajne obciążenia czasowe)</li> <li>● Wpływ opóźnienia na jakość rozmowy</li> </ul>	<ul style="list-style-type: none"> <li>● Utrata pakietów dla kodeków PCM</li> <li>● Niwelacja utraty pakietów dla kodeków PCM</li> <li>● Wzmocnienie/stłumienie sygnałów w sieci</li> <li>● Niedopasowanie szybkości bitowej między koderem i dekoderem</li> <li>● Sygnały informacyjne sieci jako sygnał wejściowy dla kodeka</li> <li>● Sztuczna mowa jako sygnał wejściowy dla kodeka</li> <li>● Muzyka jako sygnał wejściowy dla kodeka</li> <li>● Echo słuchacza</li> <li>● Wpływ tłumików echa</li> <li>● Wpływ algorytmów redukcji szumów</li> <li>● Kodeki hybrydowe oraz CELP o szybkości mniejszej niż 4 kbit/s</li> <li>● Standard MPEG 4 HVXC</li> <li>● Badanie terminali akustycznych z zastosowaniem sztucznego torsu HATS</li> </ul>

Przy badaniach metodą PESQ są rozróżniane opóźnienia stałe i zmiana tych opóźnień (*jitter*). Jednak podczas szacowania wartości MOS opóźnienia stałe nie są uwzględniane w metodzie PESQ, jest natomiast brana pod uwagę zmienność opóźnień. Nie jest również uwzględniany wpływ echa. Ze względu na wymienione ograniczenia, metodę PESQ stosuje się do badań jakości mowy tylko w jednym kierunku, nie do oceny jakości konwersacji.

Metoda PESQ nie została zweryfikowana dla sieci, w których wykorzystuje się układy eliminacji szumów, tłumiki echa oraz algorytmy kompensacji utraty pakietów. Przy stosowaniu tych urządzeń w sieci ocena PESQ jest zwykle niższa niż ocena subiektywna.

Obecnie ITU-T Study Group 12 pracuje nad następną wersją metody szacowania jakości mowy P.OLQA (*Objective Listening Quality Assessment*), w której zostanie uwzględniony wpływ powyższych czynników na końcową ocenę jakości mowy.

### **Obiektywna metoda bezwzględna P.563**

W 2004 r. organizacja ITU-T znormalizowała, w formie zalecenia P.563 [9], metodę obiektywnej oceny jakości mowy transmitowanej, jako jednopunktowy pomiar, w prawie każdym punkcie sieci, odbieranego strumienia mowy, bez konieczności stosowania sygnału odniesienia.

Metoda ta jest zalecana do monitorowania sieci podczas jej normalnej eksploatacji. Uwzględnia ona większość występujących w niej zniekształceń. Wynikiem końcowym jest ocena jakości mowy podawana w skali MOS-LQO, zgodnie z zaleceniem ITU-T P.800.1<sup>①</sup>. Metoda ta ma jednak dość dużo ograniczeń i nie umożliwia analizowania wpływu takich czynników, jak:

- efekt lokalny;
- wpływ opóźnienia w testach konwersacyjnych;
- echo mówcy;
- muzyka lub sygnały tonowe sieci jako sygnały wejściowe;
- techniki kodowania o przepływności poniżej 4 kbit/s, np. IMBE, AMBE, LPC10e;
- niedopasowanie szybkości bitowej między koderem i dekerem,
- obecność szumu tła w sygnale wejściowym;
- muzyka jako sygnał wejściowy;
- niedopasowanie bitowe między koderem i dekerem.

Dlatego też, aby otrzymać wiarygodne wyniki, jest niezbędne analizowanie dużej liczby pomiarów.

## **Porównanie metod**

Metoda PESQ jest najczęściej stosowaną metodą badania jakości mowy transmitowanej w różnych sieciach telekomunikacyjnych. Uwzględnia większość czynników degradujących jakość mowy, w tym szczególnie w sieciach VoIP (*Voice over Internet Protocol*). Przeznaczona jest głównie do badań i projektowania kodeków, doboru sprzętu dla sieci VoIP oraz do testowania sieci. Umożliwia kontrolę jakości usługi telefonii w testach połączeń realizowanych „od końca do końca”.

Metoda ta, spośród stosowanych obiektywnych metod badania jakości mowy, ma najwyższy współczynnik korelacji, wynoszący 0,9 – 0,95, w stosunku do metod subiektywnych. Dlatego, dalej opisane, badania jakości przekazu głosu zostały przeprowadzone tą metodą w warunkach laboratoryjnych i rzeczywistych.

<sup>①</sup> ITU-T Rec. P.800.1 (07/2006): „Mean Opinion Score (MOS) terminology”.

Analiza wyników badań metodą PESQ umożliwiła określenie limitów wskaźnika jakości głosu (MOS, PESQ) dla różnych konfiguracji połączeń. Minimalne dopuszczalne wartości wskaźnika jakości głosu MOS zostały ustalone pod kątem możliwości ich uwzględnienia w opracowaniach zarówno projektu nowelizacji rozporządzenia dotyczącego usługi powszechnej [10; art. 81], jak i rozporządzenia na temat innych publicznie dostępnych usług telekomunikacyjnych [10; art. 63 ust. 3].

## Sygnały testowe

W badaniach, prowadzonych zgodnie z zaleceniem ITU-T P.862 [7], sygnałami testowymi powinny być specjalnie przygotowane próbki mowy. Próbki mowy rzeczywistej powinny być bardzo dobrej jakości, z małą zawartością szumów (dla mowy żeńskiej i męskiej). Ponadto muszą one być zorganizowane w segmenty o czasie trwania od 8 do 15 s dla mowy aktywnej (tzn. z wyłączeniem okresów ciszy). Próbki mowy powinny być zarejestrowane zgodnie z zaleceniem ITU-T P.50 [11].

Zaleca się korzystanie z próbek sztucznej mowy ASTS (*Artificial Speech Test Stimulus*), przygotowanych profesjonalnie, uwzględniających właściwości lingwistyczne i charakterystyczne cechy sygnału mowy. Zawierają one dużo mniej nadmiarowości niż mowa rzeczywista (15 s ASTS odpowiada ok. 2 h mowy rzeczywistej), co umożliwia radykalne skrócenie czasu pomiarów bez pogorszenia dokładności badań.

Podczas badań powinno się wykorzystać co najmniej po dwie próbki głosów męskich oraz dwie próbki głosów żeńskich, zawierających różne sentencje. Uzyskane wyniki pomiarów powinny zostać uśrednione.

Sygnały testowe sztucznej mowy ASTS powinny spełniać niżej wymienione warunki.

**Czas trwania sygnału.** W metodzie PESQ powinny być stosowane sygnały o czasie trwania od 8 do 12 s. Jednak dopuszcza się też sygnały o czasie trwania do 30 s.

**Stosunek mowy do ciszy.** Aktywność mowy w sygnale powinna być od 40% do 80% oraz zawierać co najmniej 3,2 s mowy aktywnej.

**Struktura czasowa sygnału.** Sygnał testowy, oprócz normalnej mowy, powinien zawierać wydzielone okresy wzmożonej aktywności głosowej (wyraźnie głośniejszych dźwięków), jak również okresy ciszy, aby jak najdokładniej odzwierciedlać naturalną mowę. Okresy głośniejsze powinny jednorazowo trwać od 1 do 3 s, natomiast okresy ciszy co najmniej 300 ms.

**Poziom mowy aktywnej.** Powinien zostać ustawiony możliwie wysoko, jednak nie może przesterować układów. Sieć PSTN została zoptymalizowana dla poziomu mowy aktywnej ASL (*Adaptive Speed Leveling*) równej  $-18$  dBm<sub>0</sub>.

**Wybór języka.** Wybór języka próbek mowy ma duży wpływ na otrzymane wyniki pomiarów. Dostępne są próbki głosu dla wielu języków. Naturalny wydaje się wybór języka rodzimego. Rozwiązanie takie zapewnia w zasadzie lepsze poznanie parametrów jakościowych transmisji krajowych, ale uniemożliwia porównanie wyników z otrzymanymi w innych krajach. Dlatego też wydaje się rozsądniejszy wybór powszechnie stosowanych próbek języka angielskiego, a próbek języka polskiego jedynie do testów uzupełniających.

**Wielkość próby testowej.** Wielkość próby testowej, niezbędnej do zapewnienia wymaganej wiarygodności i dokładności pomiaru, zależy od powtarzalności otrzymywanych wyników. Zalecane jest przyjęcie metody doboru wielkości próby testowej podanej w normie ETSI EG 202 057-2 [12; aneks F].

**Interpretacja wyników.** Bardzo duże wahania wyników pomiarów (MOS rzędu  $\pm 1$ ) mogą wskazywać na występowanie przerw w zestawionym połączeniu, nieodpowiednio zastosowaną technikę pomiaru lub uszkodzenia w sieci. W sieci komórkowej może to być, np. przeciążenie komórki (sektora), przełączenie

lub problemy z odbiorem sygnału (zasięg), a w sieci VoIP – niewłaściwe ustawienia poziomu QoS (*Quality of Service*), w wyniku czego może wystąpić znacząca utrata pakietów lub wzrost fluktuacji.

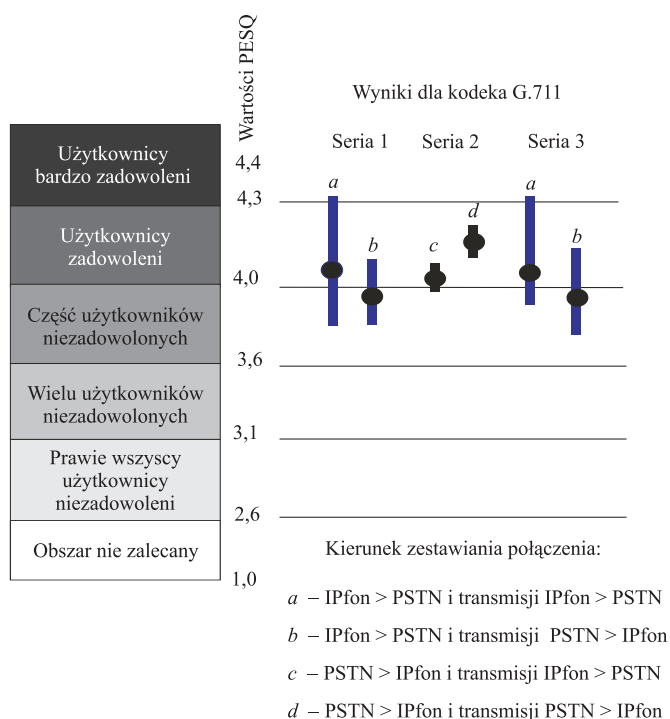
## Badania jakości transmisji mowy w rzeczywistych sieciach

Badania przeprowadzono na wybranych i dostępnych rzeczywistych sieciach przewodowych oraz bezprzewodowych, uwzględniając głównie sieci i technologie, w których występuje degradacja jakości mowy, wynikająca z kompresji danych lub ze strat informacji. W badaniach wzięto pod uwagę technologie stosowane i przewidziane dla usługi telefonicznej. Badaniami nie zostały objęte tzw. komunikatory głosowe, które zapewniają możliwość komunikacji głosowej, lecz są wyłączone z działań urzędów regulacyjnych (usługa ta nie jest zaliczana do usług telefonicznych). We wszystkich badaniach stosowano aparaturę pomiarową MultiDSL A firmy Malden Electronic.

### Jakość transmisji mowy w sieciach IP i hybrydowych

Badania prowadzono między numerem IP (dostawca usługi telefonicznej IPfon, jako sieć dostępową wykorzystano sieć telewizji kablowej o przepustowości 512/128 kbit/s) a analogowym numerem stacjonarnym TP SA ze strefy numeracyjnej SN22. Badania wykonywano w godzinach dużego ruchu i bez stosowania mechanizmów QoS, tzn. transmisję traktowano jak zwykłe dane. W czasie badań łącze dostępowe nie było obciążane dodatkowym ruchem danych. W sieci stosowano protokół transmisji, zgodny z zaleceniem ITU-T G.711 [13].

Otrzymane wyniki przeanalizowano, biorąc pod uwagę kierunek transmisji i kierunek zestawiania połączenia.



Rys. 3. Jakość transmisji mowy dla trzech serii testowych (IP, kodek G.711)

Na rys. 3 pokazano wartości średnie i odchylenia standardowe dla poszczególnych serii testowych z uwzględnieniem kierunku transmisji. Dla połączeń zestawianych z sieci PSTN do bramki VoIP IPfon osiągnięto bardzo dobre wyniki. Blisko 80% wyników przekracza 4, 2,8% mieści się w zakresie 3,6 – 3,1, a 17,2% – w zakresie 3,6 – 4,0. Nie stwierdzono wyników poniżej 3,1. Dla kodeka G.711 jest możliwe więc uzyskanie średnich wartości PESQ powyżej 4. Wymaga to jednak zapewnienia transmisji pakietów na odpowiednim poziomie QoS. Przeciążenie łącza dostępowego, przy braku gwarantowanego pasma na transmisję głosu, powoduje drastyczne spadki jakości mowy, nawet do wartości ok. 1,3.

Zastosowanie stratnego kodeka mowy, np. G.729, zapewnia znaczne oszczędności potrzebnego pasma, ale wiąże się z następującymi mankamentami:

- spadkiem jakości transmisji mowy,
- brakiem możliwości transmisji cyfrowych (faksów),
- koniecznością zastosowania mechanizmów wspomagania sygnałów DTMF (*Dialing Tone Multi-frequency Signaling*).

Pomiary przeprowadzone na instalacji testowej, zestawionej przez firmę Cisco, wykazały średni spadek jakości transmisji mowy (PESQ) do poziomu 3,5, przy czym zarejestrowane wartości nie przekroczyły 3,89.

### **Jakość transmisji mowy w sieciach GSM**

Połączenia wychodzące z sieci PSTN do sieci GSM (*Global System for Mobile Communications*) są zaliczane do usługi powszechnej. Ponadto operatorzy sieci GSM skłaniają się do zastosowania jej jako sieci dostępowej dla sieci PSTN (usługa ta działa pod nazwą G4F). Dlatego w testach wykorzystano terminale GSM produkowane dla usługi G4F. Są one wyposażone w port analogowy, umożliwiający proste dołączenie zestawu pomiarowego. Można też w nich stosować karty SIM (*Subscriber Identity Module*) skonfigurowane dla systemu G4F oraz standardowe karty SIM dowolnych operatorów – zarówno abonamentowe, jak i *prepaid*.

Jakość transmisji mowy w sieciach GSM zasadniczo zależy od typu wykorzystywanego kodeka. W sieciach GSM są stosowane cztery standardowe kodeki, określane jako:

- GSM-EFR (GSM – *Enhanced Full Rate*): kodujący do przepływności 12,2 kbit/s i zapewniający najwyższą w systemie jakość transmisji mowy;
- GSM-FR (GSM – *Full Rate*): pierwszy kodek zastosowany w GSM, kodujący sygnał o przepływności 13 kbit/s i zapewniający średnią jakość transmisji mowy;
- GSM-HR (GSM – *Half Rate*): wprowadzony do systemu GSM na początku lat 90., generujący sygnał o przepływności 5,6 kbit/s, a więc wymagający połowy pasma kodeka FR; zapewniający – kosztem jakości transmisji mowy – podwojenie pojemności sieci, co zmniejsza straty w przypadku niespodziewanego zwiększenia ruchu; wymagający o 30% mniej energii zasilania, co oszczędza baterię telefonu;
- AMR (*Adaptive Multi-Rate*): kodek adaptacyjny, optymalizujący jakość dźwięku; przyjęty przez 3GPP w 1998 r. (stosowany w sieciach GSM i UMTS); definiujący 8 trybów pracy, kompresując mowę do sygnału o przepływnościach od 12,2 do 4,75 kbit/s.

W tabelicy 3 zestawiono wyniki zbiorcze testów prowadzonych dla sieci GSM. Wyniki te potwierdzają wartości podawane w literaturze. Na uwagę zasługują wartości otrzymane dla kodeka EFR (tylko 12,5% wyników było poniżej 4), jak również niewielka różnica między wynikami jakości mowy dla transmisji między siecią PSTN (PCM64) a siecią GSM z kodekiem HR i między siecią GSM z kodekiem EFR



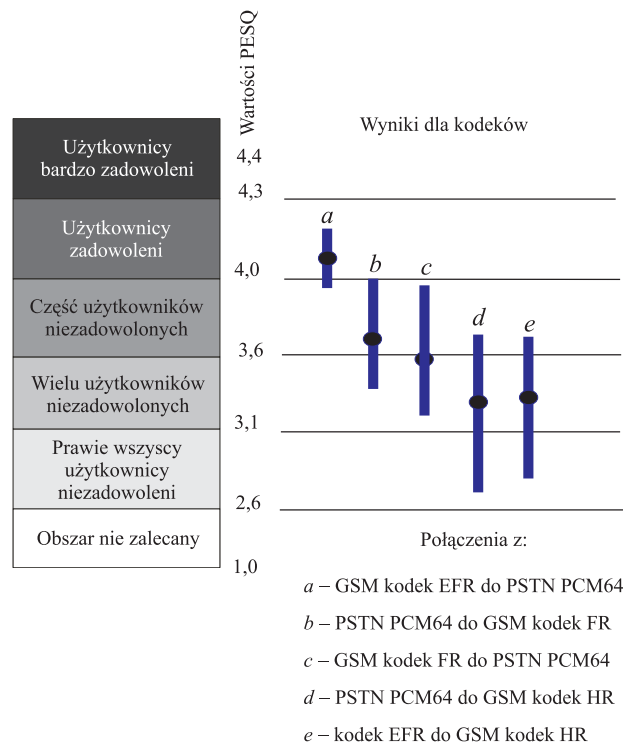
a siecią GSM z kodekiem HR. Kodek EFR jest jedynym kodekiem stosowanym w sieciach GSM, zapewniającym jakość transmisji mowy na poziomie zadowalającym 90% użytkowników. Wyniki potwierdzają też znaczącą różnicę między wynikami otrzymanymi dla kodeka EFR i FR.

**Tabl. 3. Wyniki zbiorcze testów prowadzonych dla sieci GSM**

Test	Połączenie		Wartość PESQ (P.862.1)				Część wyników poniżej wartości [%]			
	od	do	średnia	odchylenie standardowe	min	max	4,0	3,6	3,1	2,6
619	GSM-EFR	PCM64	4,07	0,14	3,51	4,21	77,5	27,5	7,5	0
387	PCM64	GSM-FR	3,66	0,35	2,74	4,14	92,9	39,3	14,3	1,79
725	GSM-FR	PCM64	3,55	0,4	2,49	4,05	12,5	2,5	0	0
594	PCM64	GSM-HR	3,22	0,51	2,13	3,94	100	70,0	40,0	12,5
586	GSM-EFR	GSM-HR	3,25	0,45	2,18	3,92	100	33,0	10,0	5,0

Kodek AMR zapewnia generowanie sygnału, w jednym z 8 trybów o różnych stopniach kompresji, ustawianego w zależności od dostępnego pasma i zakłóceń transmisji. Tryb o najwyższej przepływności jest równoważny kodekowi EFR.

Na rys. 4 przedstawiono wyniki jakości transmisji mowy dla połączeń do GSM oraz z GSM.



Rys. 4. Jakość transmisji mowy dla połączeń do GSM oraz z GSM

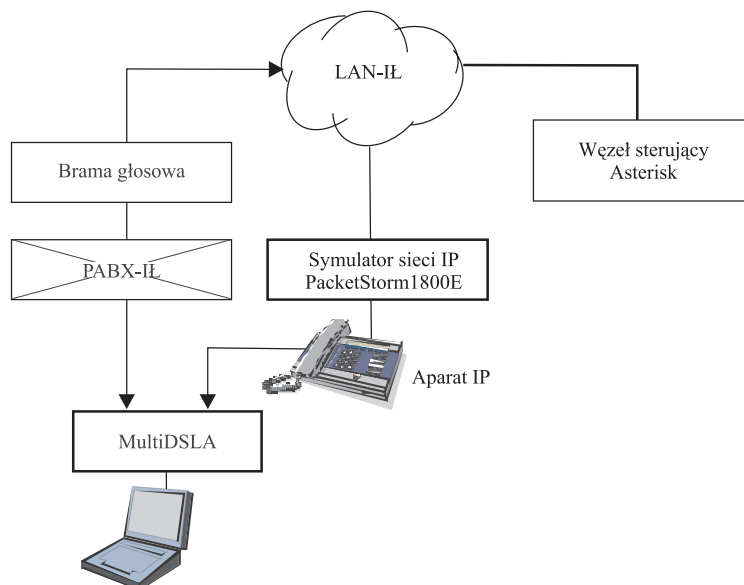
## Badania jakości transmisji mowy w warunkach symulowanych

Jakość transmitowanej mowy w sieciach IP zależy od zastosowanego kodeka i poziomu degradacji pakietów IP. W literaturze są podawane możliwe do otrzymania wartości PESQ dla poszczególnych kodeków i rodzajów degradacji pakietów, mimo tego przeprowadzono testy dla wybranych przypadków. Badania ograniczono do kodeka G.711, który jako jedyny zapewnia jakość transmisji mowy na poziomie osiąganym w sieciach PSTN/ISDN (PSTN/*Integrated Services Digital Network*).

Badania wykonano w laboratorium Zakładu Zastosowań Technik Łączności Elektronicznej w Instytucie Łączności w Warszawie, w warunkach gwarantujących pełną kontrolę nad parametrami transmitowanych pakietów, dla połączeń zestawianych między portem centrali PABX-IŁ (*Private Automated Branch Exchange-IŁ*) a aparatem Cisco IP Phone 7905, dołączonym do sieci LAN-IŁ (*Local Area Network-IŁ*). Degradacje pakietów mowy zapewniał symulator sieci (PacketStorm1800E firmy Packet Storm Inc.), wyposażony w 2 porty pomiarowe 1000/100/10 MB/s. Umożliwił on symulację takich zniekształceń sygnałów, jak:

- opóźnienie stałe;
- *jitter*, opóźnienie zmienne z możliwością wyboru rozkładu;
- ograniczenie przepływności;
- straty pakietów;
- zmiany kolejności;
- błędne bity w transmisji.

Schemat instalacji testowej zaprezentowano na rys. 5.



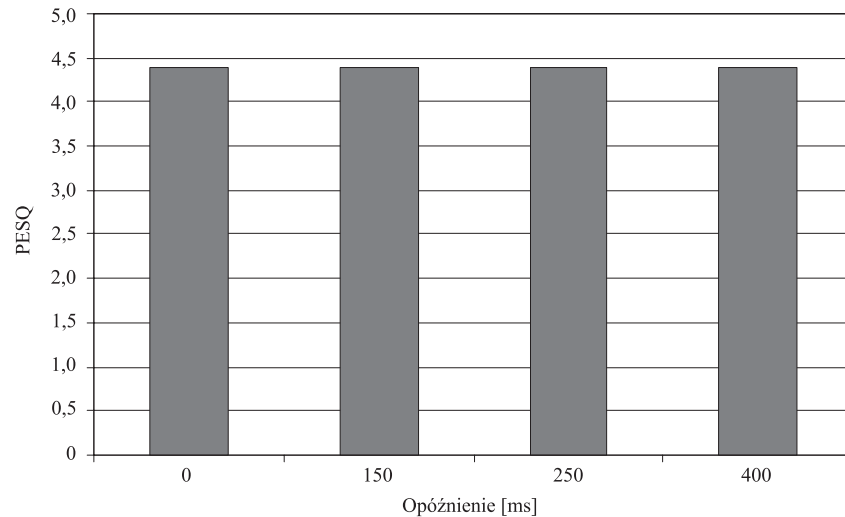
Rys. 5. Układ do badania jakości transmisji mowy w symulowanych warunkach

System do badania jakości transmisji mowy dołączono do portu analogowego centrali PABX-IŁ oraz gniazda mikrofonu (*handset*) aparatu telefonicznego Cisco IP Phone 7905.

Pomiary przeprowadzono dla 4 podstawowych parametrów jakościowych sygnałów pakietowych.

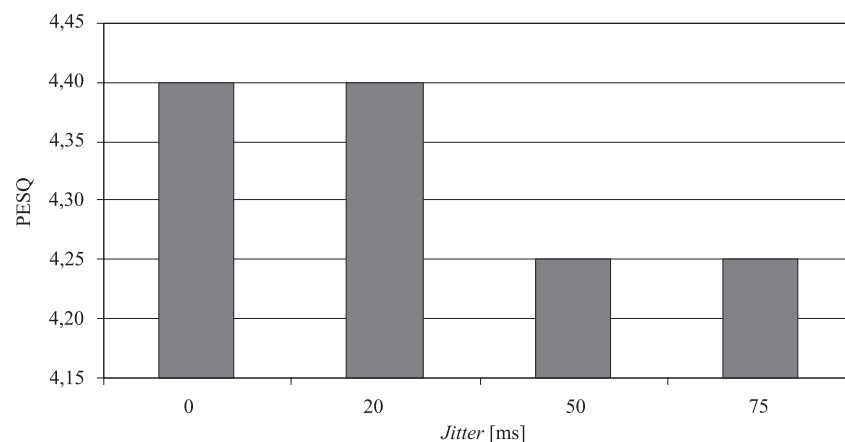
**Analiza wyników.** Otrzymane wyniki przedstawiono w postaci wykresów (rys. 6–9). Pokazują one wpływ poszczególnych parametrów na jakość transmisji mowy.

Zgodnie z metodą badania PESQ, przyrząd podaje wartości stałego opóźnienia pakietów (rys. 6), ale są one pomijane w obliczeniach PESQ. W E-modelu [14] przyjmuje się, że opóźnienie stałe do 150 ms nie degraduje jakości transmisji mowy.



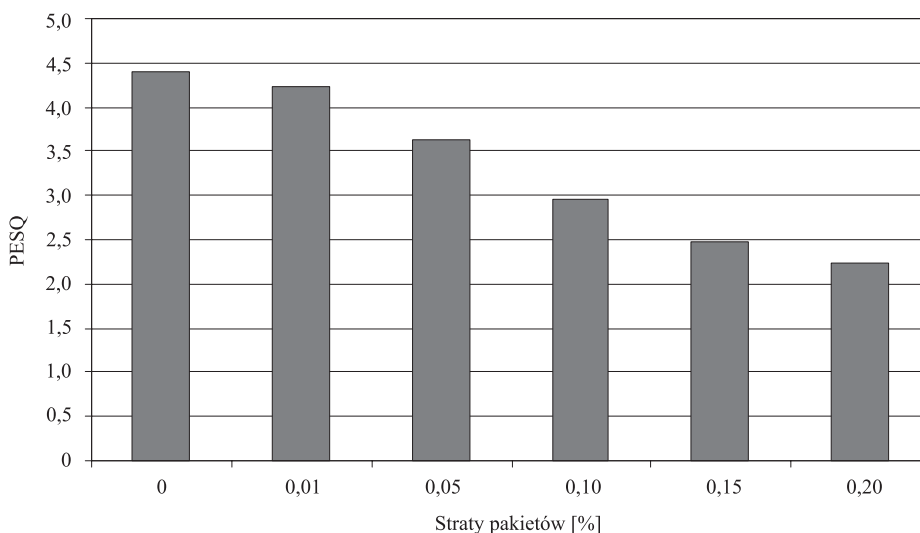
Rys. 6. Wartości PESQ w funkcji opóźnienia pakietów

*Jitter* o niewielkiej wartości nie powoduje zauważalnego spadku jakości transmisji mowy, gdyż jest on buforowany, o większych wartościach średnich (50 – 75 ms) zaś powoduje zauważalny, chociaż niezbyt duży spadek jakości mowy (rys. 7). Zwiększenie pojemności bufora jest możliwe i powoduje zwiększenie odporności systemu na *jitter*, ale równocześnie zwiększa opóźnienie stałe.

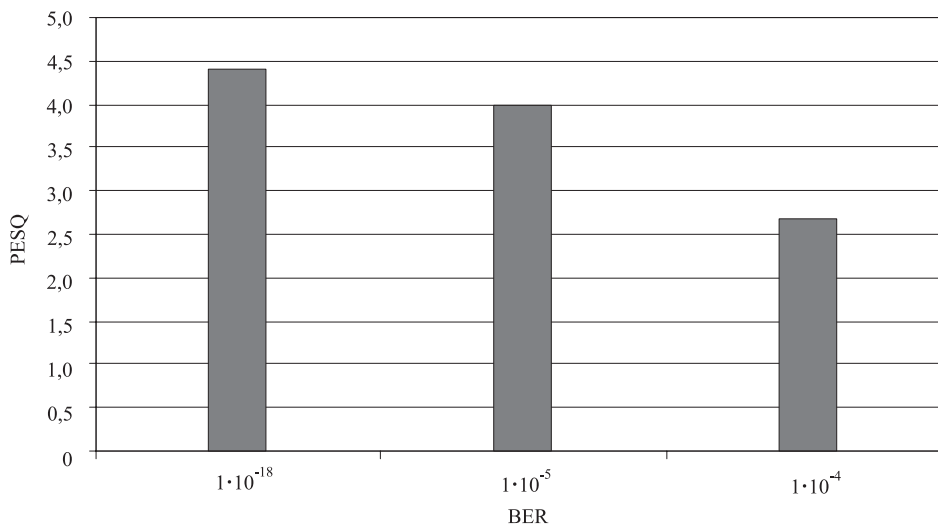


Rys. 7. Wartości PESQ w funkcji wartości średniej jittera

Niewielkie straty pakietów (do 1%) powodują minimalny spadek jakości mowy. Jednak ich wzrost wyraźnie pogarsza jakość transmisji mowy, a przy stratach przekraczających 10% wartość PESQ spada poniżej 3,0, co może być już nieakceptowane przez większość użytkowników (rys. 8).



Rys. 8. Wartości PESQ w funkcji strat pakietów



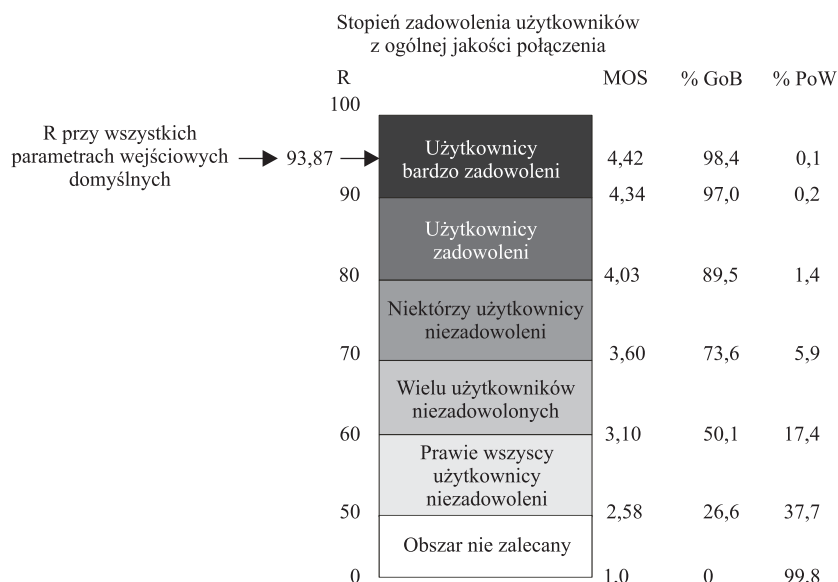
Rys. 9. Wartości PESQ w funkcji bitowej stopy błędów (BER)

Straty pakietów i bitowa stopa błędów (*Bit Error Rate* – BER) są parametrami najsilniej wpływającymi na jakość transmisji mowy. Stwierdzono, że niewielka stopa błędów nie powoduje zauważalnego spadku jakości mowy, występuje on dla stopy błędów równej  $1 \cdot 10^{-5}$ , a przy dalszym zwiększaniu stopy błędów spada gwałtownie (rys. 9).

## Propozycje wartości wskaźnika jakości mowy

Parametry jakościowe sieci nowej generacji powinny być ustalane w fazie tworzenia założeń i wymagań na sieć. W odniesieniu do jakości transmitowanej mowy należy kierować się zaleceniem ITU-T G.107 [15], definiującym E-model.

E-model umożliwia obserwowanie różnicy w jakości, wynikającej z parametrów transmisji sygnałów i kompresji sygnału. Ten komputerowy model może być narzędziem dla planistów oraz operatorów nowoczesnych sieci telefonicznych, pomagającym w ocenie jakości połączenia. Dostarcza on bowiem informacji, jaki procent użytkowników sieci będzie zadowolony z usługi transmisji mowy od początku do końca. Jest to jednak ocena jakości tylko do celów planowania oraz projektowania sieci telefonicznych i nigdy nie będzie ona odwzorowaniem faktycznej opinii użytkowników.



Rys. 10. Skala jakości E-modelu zdefiniowana w zaleceniu ITU-T G.109 [16]

Ze względu na złożoność sieci telekomunikacyjnych, przy jej planowaniu wiele parametrów transmisji powinno być rozpatrywanych nie tylko osobno, ale również w powiązaniu z innymi parametrami. Wartość wyjściowa modelu jest opisywana przez skalarny współczynnik jakości transmisji R, na podstawie którego można też wywnioskować teoretyczną ocenę reakcji użytkownika, np. określić procent użytkowników, oceniających modelowaną łączność jako *dobrą* lub *bardzo dobrą* (GoB) albo jako *slabą* lub *złą* (PoW) – por. rys. 10. Poza tym współczynnik R można uzależnić niemalże od każdego pojedynczego parametru transmisji i tym samym oszacować wielkość zmian w jakości połączenia, spowodowanych wartością danego parametru.

Szeroki zakres uwzględnianych czynników sprawia, że E-model [14], [17] stanowi bardzo dobre narzędzie do prognozowania jakości mowy. Głównie używają go planiści i projektanci sieci, a nie bezpośrednio operatorzy. Niestety, wymaga bardzo dużej liczby danych wejściowych, a tym samym kosztownych pomiarów. Mapę poziomu zadowolenia klientów przedstawiono w zaleceniu ITU-T G.107 [15].

**Jako w pełni zadowolający poziom jakości transmisji mowy przyjmuje się wartości MOS (PESQ) powyżej 4,0 i taka średnia wartość pomiarów powinna być w sieci PSTN/ISDN oraz VoIP, pracujących z protokołem ITU-T G.711. Dla sieci ISDN można oczekiwać osiągnięcia wartości MOS ok. 4,4.**

Podane wartości powinny dotyczyć jedynie usługi telefonii oferowanej w ramach usługi powszechnej i usługi publicznie dostępnej. Nie odnoszą się do systemów, do których nie można zestawiać połączeń z sieci PSTN, np. tzw. komunikatorów, takich jak Skype, oraz innych systemów wydzielonych. Systemy takie nie są objęte działaniami regulacyjnymi, zresztą w wielu przypadkach jest to niemożliwe ze względu na siedzibę dostawcy usługi poza granicami UE.

Dostawca usługi/operator VoIP często nie jest dostawcą sieci, która może być dostarczona przez innego dostawcę/operatora na podstawie umowy zawartej z użytkownikiem. W takim przypadku dostawca/operator usługi VoIP może odpowiadać tylko za parametry usługi dostarczanej przez własne oraz dzierżawione przez niego urządzenia i sieci.

Proponując wymagania dotyczące jakości transmisji mowy, należy uwzględnić wiele czynników ekonomicznych, fizjologicznych i kulturowych. Należy pamiętać, że:

- zapewnienie jakości transmisji mowy większej niż to jest niezbędne, powoduje zbędne koszty;
- koszt podniesienia jakości transmisji mowy zależy w dużym stopniu od stosowanych technologii;
- trzeba uwzględniać również dostępność innych usług realizowanych w pasmie akustycznym kanału.

Wydaje się, że przede wszystkim należy rozgraniczyć jakość transmisji zapewnianej przez systemy stacjonarne i mobilne. Użytkownik telefonii mobilnej godzi się na gorszą jakość transmisji mowy w zamian za dodatkowe usługi i funkcjonalność. Natomiast trudniej jest skłonić użytkownika sieci stacjonarnej do zgody na gorszą jakość (często godzi się dopiero w przypadku znacznej różnicy w cenie). Należy pamiętać też o tym, że użytkownik podświadomie będzie oczekiwał takiej jakości transmisji mowy, jaka kojarzy mu się z używanym obecnie terminalem.

Z analizy wyników badań i dokumentów źródłowych, np. [18], wynika, że **dla sieci stacjonarnej podstawowa średnia wartość MOS (PESQ) powinna być powyżej 4,0.**

Wyniki prowadzonych pomiarów potwierdzają, że jest to możliwe dla sieci VoIP z zastosowaniem kodeka G.711. Wyniki testów symulacyjnych oraz podane w [18] świadczą o tym, że rozwiązanie to zapewnia:

- wysoką jakość transmitowanej mowy;
- małe opóźnienia wnoszone przez łańcuch telekomunikacyjny;
- dużą odporność na straty pakietów i błędy transmisji;
- transparentny kanał 300 – 3400 Hz, zapewniający realizację wszystkich usług, podobnie jak w sieci PSTN, przez przesyłanie dowolnych sygnałów akustycznych.

Ostatni warunek jest szczególnie istotny przy migracji użytkowników klasycznych systemów telefonii, stosujących kanał PCM64, do innych, np. pracujących w systemach VoIP. Wiadomo, że w sieciach są używane różne urządzenia, takie jak centraliki alarmowe, w których zastosowano niestandardowe sygnały akustyczne. Nie będą one prawidłowo pracować w sieciach z kompresją głosu.

Z prowadzonych badań wynika, że przy zastosowaniu kodeka G.711 zakłócenia, takie jak *jitter* o wartości nie przekraczającej pojemności bufora, straty pakietów poniżej 1% oraz elementowa stopa błędów mniejsza niż  $10^{-5}$ , nie powodują dokuczliwego spadku jakości transmisji mowy.

**Można dopuścić, dla pewnych przypadków, niższą wartość MOS. Dotyczyć to może, np. połączeń międzynarodowych, przy których użytkownik może opcjonalnie zrezygnować z wysokiej jakości w zamian za niższą cenę. Zaleca się jednak, aby nawet wtedy średnia wartość wskaźnika wynosiła co najmniej 3,6.**

Dla sieci mobilnej sytuacja jest bardziej złożona. Ograniczenia dostępnego pasma częstotliwości i inne powodują, że trzeba wykorzystywać kodeki z silną kompresją pasma. W systemie GSM są dodatkowe mechanizmy, poprawiające warunki odbioru sygnału radiowego i tym samym zmniejszające zniekształcenia sygnału mowy, takie jak:

- przepłot sygnału wyjściowego kodeka;
- *hooping* (skakanie), czyli zmiany częstotliwości.

Pomimo stosowania tych mechanizmów występują duże zakłócenia odbieranego sygnału mowy. Potwierdzają to nawet pomiary wykonywane w zasadzie w optymalnych warunkach, tzn. w niedużej odległości od stacji bazowej oraz stałej lokalizacji terminalu.

Proponowane wartości wskaźnika jakości głosu MOS (PESQ) dla sieci VoIP i GSM przedstawiono w tablicy 4.

**Tabl. 4. Proponowane wartości wskaźnika jakości głosu MOS (PESQ)**

Rodzaj sieci	Propozycje wartości wskaźnika MOS (PESQ)	Uwagi
Sieć VoIP z podstawowym kodekiem G.711	Wartość średnia > 4,0	Warunki standardowe
Sieć VoIP w przypadkach opcjonalnych	> 3,6 dla 90% połączeń <sup>1)</sup>	Opcjonalnie, jeżeli użytkownik wybierze usługę o niższej cenie, np. tańsze połączenia międzynarodowe.
Połączenia do sieci mobilnej GSM z podstawowym kodekiem EFR	> 3,6 dla 80% połączeń <sup>2)</sup>	Wartość zalecana. Jakość głosu powinna być akceptowana przez większość użytkowników. W dużym stopniu zależy od części połączeń realizowanych z kodekiem HR.
Połączenia do sieci mobilnej GSM z podstawowym kodekiem FR	> 3,1 dla 80% połączeń <sup>3)</sup>	Wartość nie zalecana, natomiast tryb stosowany przez operatorów w warunkach przeciążenia sieci. Z uwagi na parametry kodeka FR jakość mowy nie będzie akceptowana przez wielu użytkowników.
Uwagi: Zgodnie z zaleceniem ITU-T G.107 [15] warunek oznacza, że co najmniej: <sup>1)</sup> 66,3% (czyli 73,6% z 90%) użytkowników uzna jakość połączeń za dobrą (GoB), a nie więcej niż 15,3% (czyli 5,9% z 90% + 10%) za złą (PoW); <sup>2)</sup> 58,9% (czyli 73,6% z 80%) uzna jakość połączeń za dobrą (GoB), a nie więcej niż 24,7% (czyli 5,9% z 80% + 20%) za złą (PoW); <sup>3)</sup> 40,8% (czyli 50,1% z 80%) uzna jakość połączeń za dobrą (GoB), a nie więcej niż 33,9% (czyli 17,4% z 80% + 20%) za złą (PoW).		

## Podsumowanie

Tradycyjne sieci PSTN gwarantowały nie jakość dla poszczególnych usług a parametry kanału transmisyjnego, które powinny tę jakość zapewnić. Wprowadzenie innych technik transmisji sygnałów i zaawansowane metody kompresji sygnału mowy wymusiły rozwój pomiarów jakości usług. Zachodzi również potrzeba określenia wymagań dotyczących jakości transmisji mowy w sieciach stacjonarnych i mobilnych, zadowolających użytkowników.

Podstawowym celem, jaki musi realizować sieć telekomunikacyjna, jest zapewnienie usług na odpowiednim, akceptowanym przez abonenta poziomie. Szczególnie usługa głosowa, ze względu na wymóg przesyłania mowy w czasie rzeczywistym, jest narażona na degradację przez czynniki zakłócające, charakterystyczne dla sieci pakietowych. Analiza jakości transmitowanej mowy, za pomocą metod obiektywnych, umożliwia oszacowanie jakości pracy elementów sieci w zestawionym łańcuchu telekomunikacyjnym, zwłaszcza źródła zniekształceń.

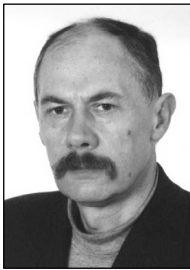
## Bibliografia

- [1] ITU-T Rec. P.800 (08/1996): *Methods for subjective determination of transmission quality*
- [2] Karjalainen M.: *A new auditory model for the evaluation of sound quality of audio systems*. W: Materiały z konferencji *IEEE ICASSP*, Tampa, USA, 1985, s. 608–611
- [3] Beerend J. G., Stemerding J. A.: *A Perceptual speech quality measurement based on psycho-acoustic sound representations*, J. Audio Eng. Soc., vol. 42, pp. 115–123, March 1994
- [4] ITU-T Rec. P.861 (08/1996): *Objective quality measurement of telephone band (300–3400 Hz) speech codecs*
- [5] Atkinson, D. J.: *Proposed Annex A to Recommendation P.861*. ITU-T Study Group 12 Contribution 24 (COM 12-24-E), Dec. 1997
- [6] *PAMS Usage Guidelines*. PsyTechnics Group, British Telecommunications, Febr. 2000
- [7] ITU-T Rec. P.862 (02/2001): *Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs plus Amendment 2 (11/2005)*, Revised Annex A – *Reference implementations and conformance testing for Recommendations P.862, P.862.1 and P.862.2*
- [8] ITU-T Rec. P.862.1 (11/2003): *Mapping function for transforming P.862 raw result scores to MOS-LQO*
- [9] ITU-T Rec. P.563 (05/2004): *Single-ended method for objective speech quality assessment in narrow-band telephony applications*
- [10] *Ustawa z dnia 16 lipca 2004 r. „Prawo telekomunikacyjne”*. Dz.U., 2004, nr 171, poz. 1800 z późniejszymi zmianami
- [11] ITU-T Rec. P.50 (09/1999): *Artificial voices*
- [12] ETSI EG 202 057-2 V1.3.1 (2009–02): *Speech Processing, Transmission and Quality Aspects (STQ); User related QoS parameter definitions and measurements; Part 2: Voice telephony, Group 3 fax, modem data services and SMS*
- [13] ITU-T Rec. G.711 (11/1988): *Pulse code modulation (PCM) of voice frequencies*
- [14] Trzaskowska M. J.: *Ocena jakości sygnału mowy w oparciu o model E w zastosowaniu do krajowej sieci telefonicznej*. Praca doktorska, Warszawa, Instytut Łączności, 2000



- [15] ITU-T Rec. G.107 (07/2002): *The E-model: A computational model for use in transmission planning*
- [16] ITU-T Rec. G.109 (09/1999): *Definition of categories of speech transmission quality*
- [17] ITU-T Rec. G.108 (09/1999): *Application of the E-model: A planning guide*
- [18] TIA Telecommunications Systems Bulletin Telecommunications – IP Telephony Equipment – Voice Quality Recommendations for IP Telephony TSB-116-A (Revision of TSB-116), March 2006

### **Ryszard Kobus**



Mgr inż. Ryszard Kobus (1951) – absolwent Wydziału Elektroniki Politechniki Warszawskiej (1975); długoletni pracownik Instytutu Łączności w Warszawie (od 1975); ekspert Komitetu Technicznego CEN/TC 331 w zakresie usług pocztowych i zastępca przewodniczącego PKN/KT 259 ds. Poczty; współautor wielu opatentowanych rozwiązań; zainteresowania naukowe: telekomunikacja, badania jakości usług telekomunikacyjnych, badania jakości usług pocztowych, normalizacja.

e-mail: R.Kobus@itl.waw.pl

### **Marian Kowalewski**



Doc. dr hab. inż. Marian Kowalewski (1951) – absolwent WSOWŁ (1975); nauczyciel akademicki, pracownik naukowy i prorektor ds. dydaktyczno-naukowych w Wyższej Szkole Oficerskiej Wojsk Łączności (1975–1997); pracownik naukowy Instytutu Łączności w Warszawie (od 1997), zastępca dyrektora ds. naukowych i ogólnych IŁ (1997–2004), kierownik projektu TETRA w IŁ (od 2002); organizator oraz współorganizator wielu seminariów i konferencji naukowych; autor wielu podręczników i skryptów akademickich, artykułów, prac naukowo-badawczych dotyczących problematyki telekomunikacyjnej; zainteresowania naukowe: planowanie i projektowanie oraz efektywność systemów telekomunikacyjnych.

e-mail: M.Kowalewski@itl.waw.pl

### **Bogdan Mucha**



Mgr inż. Bogdan Mucha (1958) – absolwent Wydziału Elektroniki Politechniki Warszawskiej (1982); długoletni pracownik Instytutu Łączności w Warszawie (od 1982); zainteresowania naukowe: jakość usługi głosowej w sieciach łączności elektronicznej, badania jakości usług telekomunikacyjnych.

e-mail: B.Mucha@itl.waw.pl