

Instytut Podstawowych Problemów Techniki
Polskiej Akademii Nauk

Metoda oceny jakości transmisji głosowej w telefonii VoIP

mgr inż. Łukasz Apiecionek

Rozprawa doktorska napisana pod kierunkiem
Prof. dr. hab. inż. Ryszarda Chorasia

Bydgoszcz, Kwiecień 2010

Podziękowania

*Pragnę złożyć serdeczne podziękowania promotorowi **prof. Ryszardowi Chorasowi** za opiekę naukową, poświęcony czas i cenne uwagi, bez których nie powstała by niniejsza praca.*

Pracę tę dedykuję moim Rodzicom i najbliższym

Łukasz Apiecionek

Spis treści

Spis oznaczeń	5
Spis skrótów	9
Tezy i cele pracy	13
Rozdział I Telefonía VoIP	18
1.1. Konwergencja sieci	18
1.2. Protokół IP	23
1.3. Sieci IP	25
1.4. Sieć telefoniczna VoIP	26
1.5. Protokół H323	29
1.6. SIP	33
1.7. MEGACO i MGCP	36
1.8. IAX	37
1.9. Podsumowanie protokołów sygnalizacji	37
Rozdział II Jakość dźwięku w telefonii	39
2.1. Ocena jakości w sieci telefonicznej VoIP	39
2.2. Jednostki ocen jakości systemów transmisji dźwięku	39
2.3. Problemy sieci telefonicznej VoIP	43
2.3.1. Opóźnienia pakietów	44
2.3.2. Jitter	49
2.3.3. Natłoki	50
2.3.4. Pozostałe parametry wpływające na jakość transmitowanego dźwięku ..	51
2.4. Ocena jakości standardów kodowania dźwięku w telefonii VoIP	53
2.5. Adaptacyjna sieć oferująca usługę VoIP	56
Rozdział III Metody oceny jakości transmisji dźwięku w sieci oferującej usługę VoIP	59
3.1. Kategorie metod	59
3.2. Pomiar jakości sygnału na podstawie wyrazistości logatomowej	61
3.3. PAMS	64
3.4. PESQ	66
3.5. PSQM	67
3.6. E-model	68

3.7.	P.563	78
3.8.	K-factor	82
3.9.	Podsumowanie przeglądu metod oceny jakości transmisji dźwięku	82
Rozdział IV Proponowana metoda <i>Oceny Jakości Transmitowanego Dźwięku</i> (OJTD) w telefonii VoIP		91
4.1.	Przygotowanie danych eksperymentalnych i przyjęcie metodologii badań	91
4.1.1.	Badanie wpływu szumów w sygnale, stosowanego języka i mówcy	92
4.1.2.	Badania wpływu sposobu kodowania dźwięku na jakość jego transmisji	94
4.1.3.	Badanie wpływu przerw w sygnale dźwięku na jakość jego transmisji ..	95
4.1.4.	Badanie wpływu rozmiaru i straty pakietów IP na jakość transmisji dźwięku	97
4.1.5.	Badanie wpływu echa na jakość transmisji dźwięku	100
4.1.6.	Badanie wpływu opóźnień w sieci IP na jakość transmisji dźwięku	101
4.1.7.	Wybór metody do porównania wyników metody <i>Oceny Jakości Transmitowanego Dźwięku</i>	102
4.2.	Proponowana metoda OJTD w telefonii VoIP	102
4.3.	Algorytm proponowanej metody OJTD	108
4.4.	Wyniki uzyskiwane za pomocą zaproponowanej metody OJTD	113
4.4.1.	Wpływ standardów kodowania dźwięku na jego jakość	114
4.4.2.	Wpływ przerw w sygnale dźwięku na jego jakość	115
4.4.3.	Wpływ echa na jakość transmitowanego dźwięku	118
4.4.4.	Wpływ opóźnień pakietów IP na jakość transmitowanego dźwięku	118
4.4.5.	Wpływ rozmiaru i strat pakietów IP na jakość transmitowanego dźwięku	119
4.5.	Wymagania sprzętowe związane z proponowanym algorytmem OJTD	123
4.6.	Możliwości zastosowań proponowanej metody OJTD	124
4.7.	Podsumowanie wyników badań	126
Wnioski		129
Dodatek A		132
Bibliografia		139

Spis oznaczeń

0 dBr – punkt odniesienia wartości parametrów w E-modelu

A – współczynnik oczekiwania jakości systemu transmisji dźwięku będący wielkością dodatnią, reprezentujący korzyści wynikające z używania różnych systemów kodowania mowy

B – bajt

$d(n)$ - sygnał akustyczny rozbrzmiewający w środowisku – wnoszący echo

D_r – współczynnik czułości mikrofonu w miejscu słuchacza

D_s – współczynnik czułości mikrofonu w miejscu mówcy

$EB1$ - pierwszy parametr proponowanej metody wykrywania przerw w sygnale

$EB2$ - drugi parametr proponowanej metody wykrywania przerw w sygnale

$e(n)$ - wartość błędu sygnału

f_g - częstotliwość graniczna sygnału [Hz]

f_i - częstotliwość próbkowania sygnału [Hz]

I_d – współczynnik pogorszenia jakości transmitowanej mowy wywołanego echem odnoszącym się do sygnału mowy

I_{dd} – pogorszenie jakości transmitowanej mowy spowodowane przez bezwzględne opóźnienie transmisji próbek dźwięku

I_{dle} – pogorszenie jakości transmitowanej mowy spowodowane przez efekt echa lokalnego u słuchacza

I_{dte} – pogorszenie jakości transmitowanej mowy spowodowane przez efekt echa lokalnego u mówcy

I_e – współczynnik pogorszenia jakości transmitowanej mowy spowodowanego zastosowaniem w łączy specjalnych urządzeń, tj. koderów o małej przepływności binarnej i innych

I_{otr} – pogorszenie jakości transmitowanej mowy spowodowane za dużym poziomem głośności dźwięku podczas połączenia

I_q – pogorszenie jakości transmitowanej mowy spowodowane przez szum kwantyzacji

I_s – współczynnik pogorszenia jakości transmitowanej mowy reprezentujący zniekształcenia mowy podobne jak przy odbiorze zbyt dużego poziomu mocy sygnału dźwięku, przy nieoptymalnym efekcie lokalnego echa lub na skutek zniekształceń kwantyzacji

I_{st} – pogorszenie jakości transmitowanej mowy spowodowane efektem echa lokalnego

J_i – wartość jitter [ms]

K - liczba nadanych testowych list logatomów

K_i – współczynnik wyrażający wpływ i -tego kodera na jakość transmitowanego dźwięku

$LSTR$ – tłumienność głośności efektu lokalnego echa u słuchacza [dB]

LT_q - absolutny próg słyszenia

MOS – wynik oceny jakości dźwięku

MOS_{EB} - wartość współczynnika MOS według proponowanej metody

$MOS_{P.563}$ - wartość współczynnika MOS według metody P.563

MOS_{PESQ} - wynik oceny jakości dźwięku próbki za pomocą metody PESQ

N - liczba słuchaczy

N_c – moc szumu wniesionego przez obwody elektryczne [dBm0p]

N_o – całkowite szумы odniesione do punktu odniesienia 0 dBr [dBm0p]

OLR – całkowita tłumienność głośności [dB]

o_i - liczba strat pakietów o długości i

O_i – czas przybycia i -tego pakietu

o_o – ilość dostarczonych pakietów

p – prawdopodobieństwo utraty pakietu gdy poprzedni pakiet był przesłany poprawnie

P_j - moc niesiona przez j -tą ramkę [dB]

$P_{n,k}$ - liczba prawidłowo odebranych logatomów k -tej listy testowej przez n -tego słuchacza

P_r – poziom hałasu otoczenia w miejscu słuchacza [dB(A)]

P_s – poziom hałasu otoczenia w miejscu mówcy $[dB(A)]$
 P_{sr} - średnia mocy w sygnale $[dB]$
 $PSQM$ – wynik oceny jakości dźwięku według metody PSQM
 q – prawdopodobieństwo utraty pakietu gdy poprzedni pakiet został również utracony
 qdu – liczba urządzeń wprowadzających zniekształcenia kwantyzacji
 R – współczynnik jakości transmitowanego dźwięku
 RLR – tłumienność głośności $[dB]$
 RLR_R – tłumienność głośności przy odbiorze $[dB]$
 R_o – podstawowy współczynnik sygnał/szum SNR transmisji mowy w punkcie odniesienia 0 dBr
 s - odchylenie średniokwadratowe wyrazistości logatomowej
 S_i - czas nadania i -tego pakietu
 SLR – tłumienność głośności przy nadawaniu odniesiona do punktu odniesienia 0 dBr najbliższego położonego miejsca R $[dB]$
 SLR_S – tłumienność głośności przy nadawaniu $[dB]$
 SLR_R – tłumienność głośności przy odbieraniu $[dB]$
 SNR – współczynnik sygnał/szum
 $STMR$ – tłumienność głośności efektu lokalnego echa w pozycji słuchacza $[dB]$
 T – średnie opóźnienie transmitowanego dźwięku w jedną stronę $[ms]$
 T_a – bezwzględne opóźnienie transmisji dźwięku od mówcy do słuchacza $[ms]$
 $TELR$ – tłumienność głośności echa mówcy $[dB]$
 T_k - liczba nadanych logatomów k -tej listy testowej
 T_p – czas trwania przerwy w sygnale $[ms]$
 T_r – całkowite opóźnienie $[ms]$
 T_s – czas trwania sygnału $[ms]$
 $WEPL$ – ważona strata efektu echa podczas transmisji w systemie $[dB]$
 $W_{k,n}$ - wyrazistość logatomowa uzyskana podczas odsłuchiwania k -tej listy testowej przez n -tego słuchacza

W_L - średnia wyrazistość logatomowa [%]

W_i - współczynnika zmian sygnał-cisza

$W(n)$ - reprezentuje filtr adaptacyjny

x_i - wartość sygnału w i -tej chwili czasowej w ramce

$x(n)$ - sygnału dźwięku na wejściu układu

Y - stała wprowadzana do eliminacji szumu pomiaru

$y(n)$ - sygnału dźwięku na wyjściu filtru adaptacyjnego

Spis skrótów

ACELP – Algebraic Code Excited Linear Prediction
ACR – Absolute Category Rating
ADPCM – Adaptive Differential Pulse Code Modulation
ATM – Asynchronous Transfer Mode
BGP – Border Gateway Protocol
CAC – Call Admission Control
CS-CELP – Conjugate Structure –Code Excited Linear Prediction
CELP – Code Excited Linear Prediction
DCME – Digital Circuit Multiplication Equipment
DCR – Degradation Category Rating
DHCP – Dynamic Host Configuration Protocol
DNS – Domain Name System
DSCP – Differentiate Service Code Point
DSP – Digital Signal Processor
ETH – Ethernet
ETSI – Europejski Instytut Standardów Telekomunikacji
FIFO – First In First Out
GOB – Good or Better
GRE – General Routing Encapsulation
GSM – Global System for Mobile Communications
HTTP – HyperText Transfer Protocol
IAX – Inter Asterix eXchange
ICMP – Internet Control Message Protocol
IETF MMUSIC – Internet Engineering Task Force Multiparty Multimedia Session Control
IP – Internet Protocols
IPSec – Internet Protocol Security
IPv4 – Internet Protocols version 4
ISDN – Integrated Services Digital Network

ISO OSI – International Organization for Standardization Open System Interconnection
ITU-T – International Telecommunication Union - Telecommunication Standardization Sector
LAN – Local Area Network
LD-CELP – Low Delay-Code Excited Linear Prediction
LPC – Linear Predictive Coding
L2TP – Layer Two Tunneling Protocol
MCU – Multipoint Control Unit
MFLOPS – Mega Floatingpoint Instruction Per Second
MG – Media Gateway
MGC – Media Gateway Controller
MGCP – Media Gateway Control Protocol
MIB – Management Information Base
MIME – Multipurpose Internet Mail Extensions
MOS – Mean Opinion Score
MPLS – MultiProtocol Label Switching
MP-MLQ – Multi Pulse-Maximum Likelihood Quantization
NAT – Network Address Translation
OID – Object Identifier
OSI – Open System Interconnection
OSPF – Open Shortest Path First
OLR – Overall Loudness Rating)
PAMS – Perceptual Analysis Measurement System
PAT – Port Address Translation
PC – Personal Computer
PCM – Pulse Code Modulation PCM
PESQ – Perceptual Evaluation of Speech Quality
POW – Poor or Worse
PSTN – Public Switched Telephony Network
PSQM – Perceptual Single ended Objective Measure
QoS – Quality of Service

RAS – Registration Admission Status
RED – Random Early Detection
RIP – Routing Information Protocol
RFC – Request For Comments
RLR – Receiving Loudness Rating
RPE-LTP – Regular Pulse Excitation with Long-Term Prediction
RSVP - Resource ReserVation Protocol
RTCP – Real Time Transfer Control Protocol
RTP – Real-time Transport Protocol
RTSP – Real Time Streaming Protocol
SCCP – Skinny Client Control Protocol
SCUI – System Control User Interface
SDP – Session Description Protocol
SG – Signalling Gateway
SIP – Session Initiation Protocol
SLR – Sending Loudness Rating
SMTP – Simple Mail Transfer Protocol
SNMP – Simple Network Management Protocol
SNR – Signal Noise Ratio
STMR – SideTone Masking Rating
TCP – Transmission Control Protocol
TELR – Talker Echo Loudness Rating
TETRA – TERrestrial TRuncked RADio
TLS – Transport Layer Security
TME – TerMinating Early
ToS – Type of Service
UA – User Agent
UDP – User Datagram Protocol
URI – Uniform Resource Identifier
URL – Uniform Resource Locator
WEPL – Weighted Echo Path Lost

VAD – Voice Active Detection

VoIP – Voice over Internet Protocols

VPE – Voice Packeting Equipment

VPN – Virtual Private Network

Tezy i cele pracy

Wprowadzenie

W chwili obecnej konwergencja sieci: komputerowej, komórkowej, telefonicznej i telewizyjnej stała się faktem. Spotyka się dostawców świadczących usługi we wszystkich wymienionych sieciach, przy wykorzystaniu jako protokołu transmisji protokołów Internetu. W takim przypadku sieci nazywane są sieciami IP (ang. Internet Protocols). Stosując protokoły IP można przesyłać dane stanowiące reprezentację cyfrowego dźwięku, obrazów, plików, stosując transmisję gwarantującą dostarczenie danych jak i transmisję bez potwierżeń ich dostarczenia. Tryb transmisji jest zależny od wymagań związanych z usługą jaką świadczy sieć. Sieć IP stosuje komutację pakietów, dla których w zależności od potrzeb zapewnia odpowiednią jakość ich przesyłania, tzw. QoS (ang. Quality of Service). Sieci IP są szeroko stosowane dzięki temu, iż potrafią wykorzystywać praktycznie dowolną technologię przesyłania danych zgodnie z modelem referencyjnym OSI (ang. Open Systems Interconnection). W sieci IP wykorzystuje się szereg usług dla użytkowników końcowych, od prostej transmisji danych, po telewizję na żądanie. Wśród świadczonych usług jest telefonia VoIP (ang. Voice over Internet Protocols). Polega ona na przesyłaniu w trybie rzeczywistym pakietów danych zawierających cyfrową reprezentację dźwięku nagrany i zakodowaną w terminalach abonentów sieci. Terminale mogą być zarówno dedykowanymi telefonami IP, jak i komputerami z zainstalowanym specjalnym oprogramowaniem do przeprowadzania rozmów pomiędzy ich użytkownikami. Dużą zaletą sieci oferującej usługę VoIP są niskie koszty jej eksploatacji.

Usługa VoIP, jak każda ze świadczonych użytkownikowi usług, wymaga stałego monitorowania i nadzoru. W tym celu muszą istnieć systemy i metody kontroli stanu pracy urządzeń i sieci. Poza tym, usługa VoIP wymaga kontroli jakości połączeń, która potrafi stwierdzić czy w przeprowadzanej rozmowie wystąpiły zakłócenia utrudniające korzystanie z usługi. O ile istnieją metody monitorowania stanu pracy urządzeń składających się na sieć IP, o tyle gorzej jest z metodami oceny jakości

połączeń głosowych. Brak jest skutecznej metody pozwalającej na ocenę jakości przeprowadzanych rozmów w pracującej sieci bez ingerencji w jej strukturę. Istniejące metody oceny jakości przeprowadzanych rozmów podają wynik w skali określonej jako MOS (ang. Mean Opinion Score). MOS jest wynikiem subiektywnej oceny odsłuchowej próbek dźwięku, dokonywanej przez reprezentatywną liczbę oceniających. Metody oceny jakości transmisji próbek dźwięku pozwalają na laboratoryjny pomiar jakości w skali MOS, bądź na oszacowanie jakości dla sieci budowanej z określonych elementów, przy zdefiniowanych warunkach jej pracy. W tradycyjnej sieci telefonicznej PSTN (ang. Public Switched Telephony Network) istnieją metody oszacowania przyczyny występowania zakłóceń w prowadzeniu rozmów na zasadzie pomiarów traktów, testów samych central telefonicznych jak i wykonywaniu rozmów testowych z określonym numerem poprzez wybrane centrale. Pozwala to na nadzór sieci telefonicznej i usuwanie usterek. Natomiast dla telefonii VoIP potrzeba metod oceny jakości transmisji próbek dźwięku w pracującej sieci, które mogłyby badać jakość połączeń bez naruszania struktury sieci. W związku z tym pojawiło się pytanie: *czy możliwa jest automatyczna ocena jakości przesyłanych próbek dźwięku w skali MOS?* Jeżeli takie rozwiązanie byłoby dostępne oraz gdyby istniał algorytm oceny jakości transmitowanych próbek dźwięku będących cyfrową reprezentacją rozmowy przeprowadzanej w sieci oferującej usługę VoIP, to można byłoby go wykorzystać do wykrywania zakłóceń w pracującej sieci z usługą VoIP. Takie rozwiązanie pozwoliło by na zbudowanie koniecznego narzędzia do zdalnego monitorowania stanu pracy sieci oferującej usługę VoIP, automatycznego informowania o zaistniałych zakłóceniach oraz podjęcie późniejszych działań w celu wykrycia przyczyny i naprawy zaistniałej anomalii.

Zadanie oceny jakości transmitowanych próbek dźwięku można by zrealizować poprzez agentów pracujących w sieci oferującej usługę VoIP, których celem byłoby nadzorowanie jakości w swoim punkcie pracy i przesłanie wyników do stacji nadzorującej. Zadaniem stacji nadzorującej byłoby prezentowanie wyników oraz informowanie o zaistniałych zakłóceniach w sieci poprzez analizę wyników zebranych od pracujących w sieci agentów.

Istotne trudności pojawiają się przy próbie oceny jakości transmitowanego dźwięku za pomocą różnych systemów kodowania, w różnych warunkach pracy sieci, a w szczególności bez sygnału odniesienia – sygnału wzorcowego. Częste straty pakietów powodują utratę fragmentów transmitowanego dźwięku. Ponadto heterogeniczna charakterystyka sieci IP, w której działa usługa VoIP, nie pozwala na zbieranie próbek dźwięku w różnych punktach sieci i przesyłanie ich do innego miejsca w sieci.

Rozwiązaniem tych wszystkich trudności jest zaproponowany w tej pracy algorytm oceny jakości transmisji dźwięku zakodowanej mowy ludzkiej nie ingerujący w pracującą sieć i pracujący bez sygnału wzorcowego. Rozwiązanie umożliwia predykcję jakości sygnału dźwiękowego za pomocą estymacji parametrów dźwięku i ich wzajemnej korelacji. Algorytm może zostać zaimplementowany w sieci oferującej usługę VoIP, gdzie istnieje dostęp do sygnalizacji rozmowy jak i przesyłanych próbek dźwięku. Agent bazując na wspomnianym powyżej algorytmie oceny jakości transmitowanego sygnału dźwięku przesyła wynik oceny jakości dźwięku za pomocą sieci IP do stacji nadzorującej, gdzie może zostać wygenerowany stosowny alarm dla operatora sieci IP z usługą VoIP.

Celem jaki postawił sobie autor niniejszej rozprawy było opracowanie metody oceny jakości transmisji dźwięku, która pozwoli monitorować sieć telefoniczną VoIP bez ingerencji w jej strukturę. Metody oceny jakości jakie wykorzystywane są obecnie nie nadają się do zastosowania w pracującej sieci oferującej usługi VoIP, gdyż albo wymagają sygnału referencyjnego, albo są bardzo podatne na zewnętrzne czynniki zakłócające, co powoduje, że wynik pomiaru jakości nie odpowiada rzeczywistej jakości rozmowy. Opracowana metoda pozwoli na detekcję miejsca występowania problemu związanego z pogorszeniem jakości transmitowanego dźwięku w sieci. Uzyskane wyniki pomiaru jakości powinny wskazywać rzeczywiste miejsca wystąpienia anomalii w sieci oraz umożliwić administratorowi sieci z usługą VoIP lokalizację występującej bądź potencjalnej awarii sieci.

Szczegółowo cele dotyczą:

- określenia wpływu parametrów sieci IP na jakość transmisji dźwięku w sieci oferującej usługę VoIP,

- opracowania algorytmu oceny jakości transmisji dźwięku w sieci oferującej usługę VoIP bez ingerencji w jej strukturę,
- wskazania praktycznych zastosowań opracowanego algorytmu.

Uwzględniając dyskusję celów postawiono następujące tezy:

1. Możliwa jest ocena jakości rozmowy realizowanej w telefonii VoIP za pomocą metody nie ingerującej w działającą sieć. Tego typu metoda umożliwi sprawdzanie stanu pracy sieci z punktów zdalnych.
2. Proponowaną w punkcie pierwszym metodę będzie można zastosować w działającej sieci oferującej usługę VoIP bez znacznego obciążania jej zasobów. Oznacza to, że implementacja algorytmu zaproponowanej metody oceny jakości transmisji próbek głosu może być zrealizowana w urządzeniach sieci oferującej usługę VoIP wykorzystując ich wolne zasoby sprzętowe.
3. Zaproponowana metoda oceny jakości transmisji próbek dźwięku pozwala na ocenę wyniku według skali MOS. Wynik metody winien być przedstawiany w skali MOS, która jest najszerzej stosowaną skalą oceny jakości sygnału, dzięki czemu systemy jest łatwiej porównywać między sobą. Wynik będzie więc można łatwo odnosić do ogólnie przyjętych kryteriów.

Aby osiągnąć postawione cele zrealizowano następujące zadania:

- zbadano zagadnienia transmisji danych w sieciach IP,
- zbadano architektury sieci telefonicznej VoIP,
- zbadano protokoły sygnalizacji połączeń w sieci telefonicznej VoIP,
- określono zagadnienie jakości rozmów głosowych przeprowadzanych w sieci telefonicznej VoIP,
- zdefiniowano główne problemy sieci telefonicznej VoIP,
- dokonano przeglądu metod pomiaru jakości przeprowadzanych rozmów głosowych,
- zaproponowano algorytm pomiaru jakości,
- dokonano implementacji i weryfikacji proponowanego algorytmu,
- przeanalizowano możliwości wykorzystania zaproponowanego algorytmu.

Rozprawa składa się z czterech rozdziałów, spośród których rozdziały trzeci i czwarty stanowią jej główną część.

W rozdziale I wprowadzono do zagadnień sieci IP. Omówiono zagadnienie konwergencji sieci i genezę powstania sieci telefonicznej VoIP. Przedstawiono sposób tworzenia sygnału cyfrowego. Scharakteryzowano najczęściej spotykane sposoby kodowania dźwięku w telefonii VoIP. Przedstawiono strukturę sieci dla dwóch podstawowych protokołów sygnalizacji przeprowadzanych rozmów oraz omówiono główne protokoły sygnalizacji spotykane w sieci telefonicznej VoIP.

W rozdziale II przedstawiono pokrótce podejścia do pojęcia jakości w sieci telefonicznej jak i podstawowe problemy jakie napotyka w swoim działaniu sieć telefoniczna. Dokonano analizy i przeglądu źródeł zakłóceń jakości przesyłanego dźwięku.

W rozdziale III dokonano przeglądu metod pomiaru jakości rozmów głosowych z nastawieniem na metody stosowane najczęściej uwzględniając ich wady i zalety.

Natomiast w rozdziale IV przedstawiono opracowaną metodę oceny jakości rozmów przeprowadzanych w sieci telefonicznej VoIP. Omówiono algorytm jej działania, przygotowania danych eksperymentalnych jak i osiągnięte wyniki badań. Szczegółowe wyniki badań zamieszczono w Dodatku A do niniejszej rozprawy. Przedstawiono wymagania sprzętowe urządzeń do pracy z zaproponowanym algorytmem oraz jego możliwości wykorzystania w praktyce.

Wnioski zawierają weryfikację osiągnięcia celów pracy oraz omówienie zasadności postawionych tez.

Rozdział I Telefonia VoIP

1.1. Konwergencja sieci

Początków telekomunikacji można szukać w starożytności, kiedy zastosowano optyczne kodowanie liter alfabetu. Na całym świecie różne kultury wykształciły swoje własne metody komunikacji. Pierwszym rozwiązaniem technicznym przenoszącym informacje był telegram. W dniu 24 maja 1844 roku pierwsza linia telegraficzna połączyła Waszyngton z odległym o 60 km Baltimore. Natomiast 10 marca 1876 roku dokonano pierwszej rozmowy telefonicznej, w której przekazano historyczne zdanie wypowiedziane przez Bella: „Watson, proszę przyjść tutaj, pilnie pana potrzebuję”. Nieco później wynaleziono technikę radio telegrafu i radio telefonii. Wszystkie sposoby komunikacji rozwijały się równolegle stale ewoluując. Pojawiły się nowe usługi, które egzystowały odrębnie:

- radio,
- telewizja,
- telefonia.

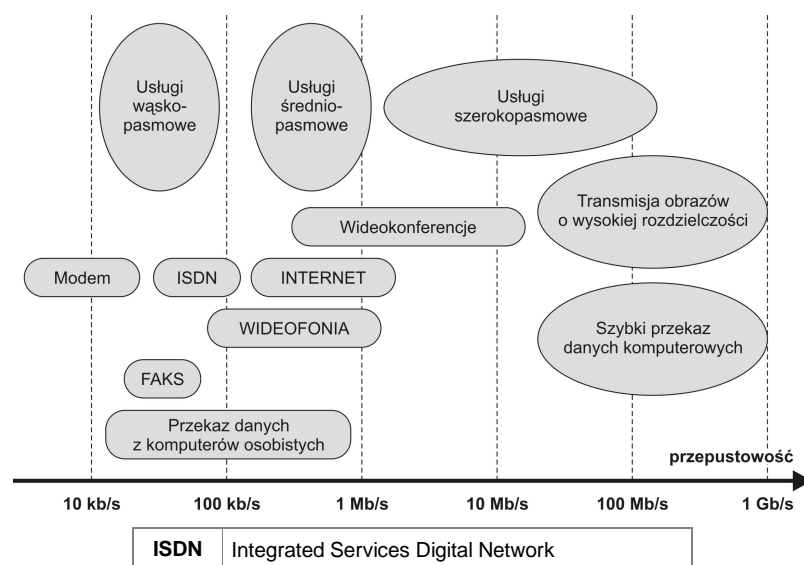
Każda z wymienionych usług wymagała swojej infrastruktury – sieci urządzeń i operatorów, przez co rozwijała się niezależnie.

Wraz z pojawieniem się komputerów, rozwojem zastosowań technik komputerowych, pojawiła się kolejna sieć – sieć komputerowa, która umożliwiała transfer danych pomiędzy komputerami. Początkowo było to tylko współdzielenie danych cyfrowych. Rozwój technologiczny pozwolił z czasem komputerom stać się nie tylko maszynami liczącymi. Wzrost ich mocy obliczeniowych zwiększył możliwości zastosowań. Pojawiły się możliwości graficzne i dźwiękowe komputerów. Jednocześnie następował rozwój sieci poprzez zwiększanie szybkości przesyłanych danych. Nowe usługi jak poczta elektroniczna na stałe zagościły jako kolejne medium komunikacji międzyludzkiej, a nawet między systemami. Większe możliwości transmisyjne sieci pozwoliły na przesyłanie nie tylko danych do obliczeń, ale na współdzielenie plików, zasobów dyskowych.

W pierwszych sieciach komputerowych rodzaj przesyłanych danych był dość ograniczony ze względu na niewielką szybkość ich transmisji. Szybki i masowy wzrost użytkowników sieci przyczynił się do wzrostu zapotrzebowania na pasmo do transferu danych generowanych przez tychże użytkowników. Sieci komputerowe początkowo pracowały jako odrębne sieci z różnymi protokołami transmisji, bez możliwości ich połączenia. Rozwój konieczności współdzielenia danych wymusił sytuację, gdzie należało dostosować się do jednej z technologii, aby użytkownicy byli ze sobą kompatybilni. Technologia, która wyparła inne technologie, bądź jak twierdzą niektórzy operatorzy wciąż dopiero wypiera, jest to technologia transmisji IP wywodząca się z sieci Internet.

Wraz ze wzrostem przepustowości łączy sieci komputerowych pracujących w technologii IP pojawiły się możliwości przesyłania danych w trybie rzeczywistym. Przy czym pod pojęciem trybu rzeczywistego należy tutaj rozumieć opóźnienie sygnału akceptowalne dla wszystkich stron korzystających z systemu transmisji. Dzięki temu, pojawiły się możliwości przesyłania danych stanowiących zakodowany dźwięk rozmowy ludzkiej, czy obraz z kamery video. Pojawiła się więc możliwość przeprowadzania rozmów telefonicznych jak i video rozmów czy wręcz video konferencji poprzez sieci IP. Do dobrze znanych sieci telefonicznych analogowych, cyfrowej sieci ISDN (ang. Integrated Services Digital Network), dołączyła sieć telefoniczna VoIP.

Pojawiła się nowa sieć telefonii mobilnej – sieć GSM (ang. Global System for Mobile Communications), której początkową funkcją było przeprowadzanie rozmów telefonicznych przez aparaty ruchome. Obecnie poprzez GSM możliwa jest transmisja danych, a aparaty GSM stały się modemami umożliwiającymi transmisję danych do komputerów.



Rysunek 1.1. Wykaz zapotrzebowania usług na pasmo transmisji danych

Pojawienie się nowych usług w sieciach komputerowych przyczyniło się do zacierania się podstawowej funkcji sieci GSM, telefonicznej, czy telewizyjnej, a proces ten określa się mianem konwergencji. Konwergencja sieci odbywa się w czterech obszarach [1]:

- rynkowym,

Mianem konwergencji rynkowej określa się proces łączenia podmiotów gospodarczych lub ich wzrost posiadania. Celem tego działania jest zwiększenie przewagi wobec konkurencji poprzez podniesienie skali podmiotu.

- technologicznym,

Pod pojęciem konwergencji technologicznej rozumie się przesyłanie wszelkich danych elektronicznych, bez względu na źródło ich pochodzenia, za pomocą jednolitego medium transmisji.

- usługowym,

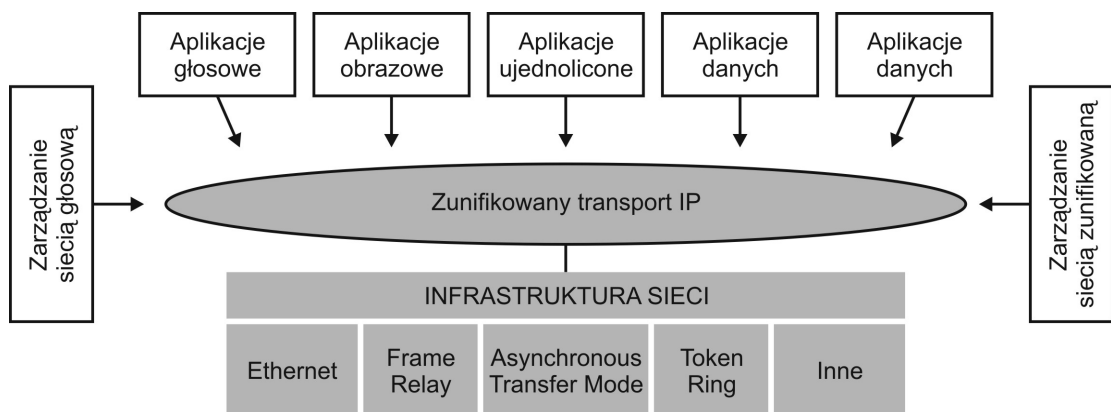
Jako konwergencję usługową rozumie się migrację usług między sektorami, które wcześniej nie były w nich wykorzystywane, nowe usługi oferowane poprzez nową konwergentną sieć, oraz usługi upodabniające działanie sieci do innej.

- regulacyjnym.

Zakres regulacyjny konwergencji określa konieczność powstania unormowań prawnych, ujednolicenia przepisów stosowanych wcześniej dla różnych sektorów działania.

Konwergencję technologiczną można rozpatrywać w trzech płaszczyznach [1]:

- jako migrację technologii między obszarami wcześniej zarezerwowanymi dla jednej usługi (np. telefonii, telewizji) – w tym zakresie znajdują się sieci oferujące usługę VoIP,
- jako powstawanie nowych technologii, które ułatwiają świadczenie usług konwergentnych,
- jako powstanie nowych technologii wspierających konwergencję.



Rysunek 1.2. Struktura sieci po ich konwergencji

W wyniku konwergencji sieci powstała jedna sieć pracująca w technologii IP. Stosując siedmiowarstwowy model OSI sieci komputerowej do projektowania aplikacji i usług, możliwe jest korzystanie z tych usług i aplikacji z użyciem dowolnej infrastruktury fizycznej sieci, niezależnie od technologii. Sieć IP może doskonale pracować z wykorzystaniem tak różnorodnych technologicznie sieci jak:

- Ethernet,
- Frame Relay,
- Token Ring,
- ATM (ang. Asynchronous Transfer Mode).

Konwergencja sieci umożliwiła realizację takich usług jak:

- telewizja w aparatach komórkowych GSM,
- dostęp do sieci Internet bezpośrednio w aparatach komórkowych GSM,

- możliwość wykonywania połączeń telefonicznych do i z komputera sieci stacjonarnej Internet,
- aplikacje użytkownika sieci oferujące usługę VoIP w aparatach komórkowych sieci GSM.

Trudno jest dziś definitywnie określić przynależność urządzenia do jakiegokolwiek sieci, skoro np. telefon komórkowy potrafi być komputerem przenośnym z dostępem do sieci Internet za pomocą różnych mediów takich jak bezprzewodowa komputerowa sieć radiowa czy transmisja danych w sieci GSM, a możliwość wykonywania połączeń jest aplikacją działająca w systemie telefonu np. Windows Mobile. Takie technologiczne hybrydy, które jednocześnie pozwalają na wykonywanie połączeń video, słuchanie radia, oglądanie telewizji, czy przesyłanie emaili, stanowią namacalny dowód konwergencji sieci, a telefonia VoIP znalazła szerokie zastosowanie.

Konwergencja sieci wymusiła rozpatrywanie zagadnień związanych z jej bezpieczeństwem. Nowe usługi, nowe możliwości, nowe typy połączeń, wymagają nowych sposobów podejścia do zabezpieczenia sieci przed nieupoważnionym dostępem oraz rozpatrzenie zagadnień związanych z jej niezawodnością i dostępnością usług. Narodziło się więc nowe zagadnienie – bezpieczeństwo sieci konwergentnej [2].

Jeszcze w 2004 roku w szeregu publikacjach wskazywano, że główną przeszkodą wykorzystania wszystkich możliwości jakie dają sieci IP są wysokie koszty łączy internetowych koniecznych do zapewniania odpowiednich przepływności [3]. Obecnie przepływności łączy stosowanych w gospodarstwach domowych przekraczają już znacznie 1Mbit/s, a czasami wynoszą już po kilkanaście Mbit/s, co znaczy, że stosowanie usługi VoIP czy videokonferencji nawet w domowych zastosowaniach jest już dostępne.

W dzisiejszych czasach nie mówi się już o telekomunikacji, ale o aplikacjach telekomunikacyjnych. W ich skład wlicza się takie jak:

- telefonia,
- telegrafia (teleks i telegram), faks,
- wideotelefonia, zestawianie telekonferencji,
- poczta elektroniczna, wymiana dokumentów,
- transmisja plików,

- łączenie sieci lokalnych LAN (ang. Local Area Network),
- dołączanie stacji roboczych lub sieci LAN do publicznych sieci komputerowych,
- dołączanie kas fiskalnych i bankomatów do sieci,
- bazy danych i elektroniczne wiadomości,
- transmisja nieruchomych obrazów,
- multimedia,
- telewizja,
- biuro w domu (telepraca),
- telezakupy,
- teleusługi socjalne i medyczne,
- teleakcje (telenadzór, telealert, telemetria itp.).

1.2. Protokół IP

Protokół IP jest opisem standardu przesyłania danych w sieciach komputerowych. Został szeroko przyjęty i jest obecnie najczęściej stosowany. Dane przesyłane są w postaci zbiorów odpowiednio oznaczonych bajtów, zwanych pakietami IP. Przesyłane dane mogą stanowić dowolną reprezentacją obiektów cyfrowych obrazów, plików, rysunków czy dźwięków. Każdy pakiet zawiera dane stanowiące opis stanu protokołu IP – tzw. nagłówek IP.

Definicja protokołu, znana pod nazwą IPv4 (IP wersja 4), powstała w 1981 na Uniwersytecie Południowej Kalifornii [4]. Protokół IP, zaprojektowany do łączenia systemów w sieciach pracujących z komutacją pakietów. Do transmisji danych używane są pakiety, stanowiące bloki danych, przesyłane od źródła do celu. Cel i źródło identyfikowane są adresem, zwanym adresem IP, który musi być unikalny w skali sieci.

Protokół IP nie zapewnia pewności dostarczenia danych do celu, co oznacza, że dane mogą być utracone, a protokół IP nie dostarczy żadnej informacji o tym fakcie. Ponadto protokół IP nie zapewnia kontroli przepływu pakietów w sieci. Przesyłane pakiety są od siebie niezależne. Protokół IP poprzez zawarte w nagłówku dane dostarcza czterech mechanizmów kontroli i sterowania.

Mechanizmy są dostępne poprzez następujące pola danych w nagłówku:

- Type of Service – służy do oznaczenia pakietu odpowiednim priorytetem ważności danych, co pozwala na zdefiniowanie polityki jakości transmisji danych,
- Time to Live – ustawiane przez nadawcę określające czas życia pakietu, za każdym razem gdy pakiet zostaje przesłany przez router łączący ze sobą oddzielne sieci, wartość pola jest zmniejszana, gdy osiągnie wartość zero zanim pakiet dotrze do celu, to pakiet zostaje usunięty z sieci,
- Options – pole opcji, które może być wykorzystywane do ustawień bezpieczeństwa i znaczników czasowych,
- Header Checksum – pole sumy kontrolnej, które pozwala stwierdzić, czy dane zostały przesłane prawidłowo.

Sam protokół IP nie zapewnia transmisji danych, ani pewności ich dostarczenia. Dlatego wymagał zdefiniowania protokołów warstwy wyższej modelu OSI. Najczęściej stosowanymi protokołami są protokoły TCP (ang. Transmission Control Protocol) i UDP (ang. User Datagram Protocol).

Protokół TCP jest protokołem połączeniowym, tj. zapewnia, że między źródłem, a odbiorcą zostaje zestawiony kanał transmisyjny. Protokół utrzymuje połączenie za pomocą transmisji pakietów do zestawienia i potwierdzenia sesji. W momencie gdy nie są transmitowane dane, między odbiorcą i nadawcą wymieniane są pakiety, które mają za zadanie utrzymać działanie sesji wymiany danych. Zwiększa to ilość danych koniecznych do przesłania. Ponadto do każdego pakietu dodawany jest nagłówek protokołu TCP, który zawiera odpowiednie dane m.in. pozwalające na ustalenie kolejności dotarcia pakietów. Odbiorca za pomocą odpowiednich pakietów sterujących potwierdza poprawne odebranie danych. Protokół TCP jest używany wszędzie tam, gdzie wymaga się pewności dostarczenia danych.

Protokół UDP jest protokołem bezpołączeniowym. Przesyła on dane do celu bez poświadczenia o ich dostarczeniu. Nie są stosowane żadne mechanizmy kontroli dostarczenia danych oraz kolejności otrzymania pakietów tak jak jest to w protokole TCP. Może więc wystąpić w sieci sytuacja, gdy pakiet, który został wysłany później dotrze do celu przed swoim poprzednikiem. Protokół ten cechuje się mniejszym

narzutem danych dodatkowych w stosunku do protokołu TCP, dzięki brakowi mechanizmów kontroli. Stosuje się go wszędzie tam, gdzie dane muszą być wysyłane z dużą częstotliwością, a ich utrata nie jest kosztowna, gdyż z upływem czasu tracą na wartości.

Wymaganie posiadania adresu IP przez każdą ze stacji roboczych przy wzroście ich ilości spowodowało konieczność rozszerzenia protokołu zdefiniowanego w 1981 roku. W 1995 roku powstał dokument opisujący standard IP zwany wersją szóstą, która znacznie rozszerza możliwości adresowania urządzeń w sieci [5].

1.3. Sieci IP

Sieci działające z protokołem IP składają się z czterech głównych elementów:

- stacji roboczych – na których pracują aplikacje użytkowników, a każda stacja ma swój adres IP,
- routerów – zadaniem tych urządzeń jest łączenie sieci o różnych adresacjach, posiadają tablice routingu, które pozwalają określić w którym kierunku należy wysłać pakiet, aby został dostarczony do urządzenia o danym adresie IP,
- przełączników – do ich zadań należy łączenie stacji roboczych użytkowników,
- bram – zadaniem bram jest łączenie sieci IP z sieciami, gdzie protokół IP nie jest protokołem transmisji, ale są w stanie zapewnić tunel do przesyłania pakietów IP poprzez swoją infrastrukturę do innej bramy.

W celu łączenia sieci IP w większą całość za pomocą routerów i wymiany ich tablic routingu opracowano specjalne protokoły. Każdy z protokołów działa nieco inaczej, niosąc ze sobą pewne wady i zalety. Najczęściej stosowane to:

- OSPF (ang. Open Shortest Path First),
- BGP (ang. Border Gateway Protocol),
- RIP (ang. Routing Information Protocol),

Do kontroli urządzeń stosuje się protokoły:

- ICMP (ang. Internet Control Message Protocol),
- SNMP (ang. Simple Network Management Protocol).

W celu zestawiania bezpiecznych połączeń opracowano protokoły takie jak:

- L2TP (ang. Layer Two Tunneling Protocol),
- IPSec (ang. Internet Protocols Security).

Ponadto opracowano wiele protokołów tunelowania pakietów, celem ukrycia adresacji IP działającej sieci, takie jak:

- tunele IP w IP,
- tunele MPLS (ang. MultiProtocol Label Switching),
- tunele GRE (ang. General Routing Encapsulation).

1.4. Sieć telefoniczna VoIP

W celu przesłania dźwięku poprzez sieć IP musi on zostać przekształcony z postaci analogowej do postaci cyfrowej. Realizuje się to poprzez operacje próbkowania, kwantowania a następnie kodowania pobranych próbek dźwięku.

Zgodnie z twierdzeniem Kotielnikowa-Shannona [6] częstotliwość próbkowania musi być co najmniej dwa razy większa bądź równa górnej częstotliwości sygnału:

$$f_i \geq 2f_g \quad (1.1)$$

gdzie:

f_i - częstotliwość próbkowania sygnału [Hz],

f_g - częstotliwość maksymalnej składowej sygnału analogowego [Hz].

Dźwięk, jaki towarzyszy mowie ludzkiej składa się z wielu częstotliwości. Zależą one od wielu czynników np. od wieku czy też sposobu intonacji. Ucho ludzkie posiada pewien zakres percepcji, w którym mowa ludzka jest najlepiej zrozumiała [7]. Ponieważ w wyniku badań uznano, że zakres częstotliwości mowy ludzkiej od 300 do 3400 Hz jest wystarczający dla zrozumienia rozmówcy, jako częstotliwość próbkowania w sieci telefonicznej przyjęto 8 kHz.

Aby przekazywany dźwięk nie posiadał częstotliwości większych niż 4 kHz jest on poddawany filtrowaniu polegającemu na przekazywaniu sygnałów znajdujących się tylko w paśmie częstotliwości od 300 do 3400 Hz. Dlatego w telefonii tradycyjnej przyjmuje się budowę traktów przenoszących dane tak, aby nie było zakłóceń w zakresie wymienionego pasma. Takie podejście pozwala na pomiary parametrów łączy

z poziomu centrali dostawcy usługi, a przez to umożliwia lokalizację miejsca ewentualnego problemu z jakością sygnału w sieci.

W sieci oferującej usługę VoIP dźwięk poddawany jest kodowaniu według różnych standardów. Wybór standardu wiąże się z określoną metodą kodowania, co gwarantuje różną ilość danych koniecznych do przesłania. Standardy kodowania przedstawiono w tabeli 1.1.

Tabela 1.1. Najczęściej stosowane standardy kodowania dźwięku w telefonii VoIP [8]

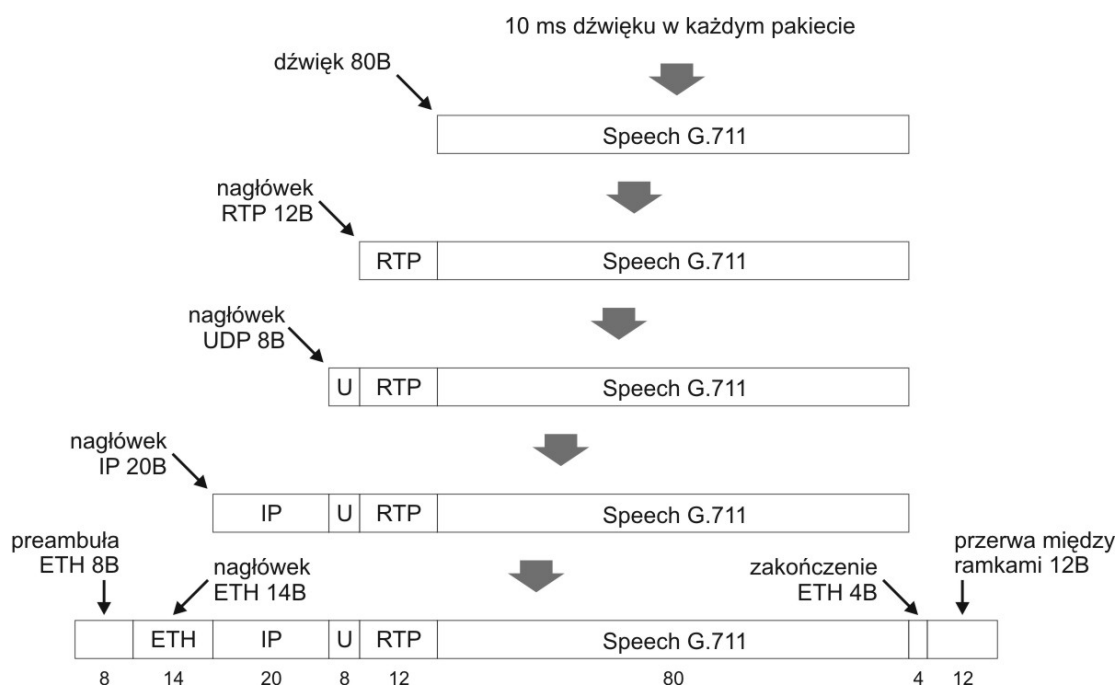
Standard	Algorytm kompresji dźwięku	Wymagane pasmo transmisji [kbps]
G.711	Modulacja impulsowo-kodowa (ang. PCM)	64
G.723.1A	Algebraiczna pobudzana kodem predykcja liniowa (ang. ACELP)	5,3; 6,3
G.723.1A	Modulacja wielo impulsowa maksymalnego prawdopodobieństwa kwantyzacji (ang. MP-MLQ)	5,3; 6,3
G.726	Adaptacyjna różnicowa modulacja impulsowo-kodowa (ang. ADPCM)	16; 24; 32; 40
G.728	Pobudzana kodem predykcja liniowa z małym opóźnieniem (ang. LD-CELP)	16
G.729a	Sprzężona algebraiczna pobudzana kodem predykcja liniowa (ang. CS-CELP)	8
GSM	Regularne pobudzanie impulsowe – predykcja długoterminowa (ang. RPE-LTP)	13,2

Dane reprezentujące zakodowany dźwięk są następnie przesyłane zgodnie z protokołami Internetu. Proces budowy pakietu IP został przedstawiony na rysunku 1.3 i jest nazywany procesem enkapsulacji danych. Zakodowany dźwięk jest przesyłany za pomocą protokołu RTP (ang. Real-time Transport Protocol) definiującego sposób przesyłania danych w czasie rzeczywistym. Protokół RTP zapewnia oznaczanie pakietów czasem nadania, co pozwala na ich identyfikację u celu i odpowiednie uszeregowanie w zadanej poprzez nadawcę kolejności. Protokół RTP, tak jak każdy kolejny, dodaje do pakietu swoje dane stanowiące m.in. sygnały sterujące, czy stan pracy. Zwiększa to ilość bajtów koniecznych do przesłania. Protokół RTP korzysta z protokołu UDP. Protokół UDP pracuje bez potwierżeń. Nie stosuje ponadto powtórzeń. Dane nie dostarczone są po prostu tracone. Z punktu widzenia danych czasu rzeczywistego nie ma to znaczenia, gdyż dane dostarczone ze zbyt dużym opóźnieniem

i tak tracą swoją wartość. Protokół UDP działa używając protokołu IP, który dodaje kolejne bajty do przesłania stanowiące dane adresowe odbiorcy oraz sumy kontrolne. Jeśli dane przesyłane są w sieci Ethernet, która jest obecnie najczęściej stosowana, to w warstwie fizycznej są opakowane w nagłówek i zakończenie ramki wymienionego typu sieci. Ponadto ramka musi zostać poprzedzona preambułą, a po jej transmisji musi nastąpić wymagana przerwa o długości wynoszącej czas trwania transmisji dwunastu bajtów danych. Taka enkapsulacja niesie za sobą konieczność uwzględnienia opakowania danych przy projektowaniu sieci telefonicznej. Wymagane przepływności muszą być odpowiednio większe. Wielkość dodatkowej przepływności sieci jest zależna więc od:

- stosowanego sposobu kodowania,
- definiowalnego odstępu między ramkami,
- sposobu transmisji pakietów IP.

Aby jednak nastąpiło przeprowadzenie rozmowy, musi zaistnieć, analogicznie jak w telefonii tradycyjnej, sygnalizacja zestawianych i przeprowadzanych rozmów. Dla telefonii VoIP zdefiniowano już wiele standardów zestawiania połączeń, ale najszerszej stosowanymi są protokoły H323 i SIP (ang. Session Initiation Protocol) [9].



Speech G.711	dane będące wynikiem zakodowania próbki dźwięku według standardu G.711
RTP	Real-Time Transport Protocol – protokół transmisji danych w czasie rzeczywistym
UDP	ang. User Datagram Protocol – protokół transmisji danych bez potwierżeń odebrania
IP	ang. Internet Protocol – protokół Internetu
ETH	ang. Ethernet – standard przesyłania danych IP poprzez przewód symetryczny cztero-parowy
B	bajt

Rysunek 1.3. Proces enkapsulacji zakodowanego dźwięku:

1.5. Protokół H323

Protokół H323 jest pierwszym szeroko stosowanym standardem, który został zdefiniowany w 1996 roku. Opisuje system telefoniczny i urządzenia końcowe, które są wykorzystywane do transmisji danych multimedialnych za pomocą lokalnej sieci komputerowej, jednak nie gwarantuje stałej jakości usług [9]. Standard H323 został opracowany przez ITU-T (ang. International Telecommunication Union – Telecommunication Standardization Section). W następnych latach prace nad standardem przyniosły nowe usługi. W roku 1998 pojawiła się druga wersja protokołu, w 1999 trzecia, a w 2001 roku czwarta, rok 2003 przyniósł wersję piątą [10][11][12][13][14]. Standard H323, zwany protokołem sygnalizacji, w swojej pracy wykorzystuje kilka innych protokołów określonych stosownymi dokumentami. Protokół

H323 jest szeroko stosowany i implementowanym przez producentów sprzętu, co gwarantuje dość małą ilość błędów w jego działaniu.

Sieć telefoniczna zgodnie z protokołem H323 składa się z następujących elementów:

- H323 Terminal

Definicja określa terminal H323 jako punkt końcowy umożliwiający komunikację w trybie rzeczywistym z innym terminalem, bramą lub mostkiem konferencyjnym.

Terminal może realizować:

- rozmowy telefoniczne poprzez transmisję próbek dźwięku za pomocą pakietów IP,
- rozmowy telefoniczne poprzez transmisję próbek dźwięku za pomocą pakietów IP wraz z transmisją danych,
- rozmowy telefoniczne poprzez transmisję próbek dźwięku za pomocą pakietów IP wraz z transmisją obrazu,
- wszystkie trzy możliwe połączenia jednocześnie: głos, dane i obraz.

H323 Terminal jest więc telefonem VoIP, Video-Telefonem, lub programem komputerowym, który działa według opisanego standardu sygnalizacji realizując funkcję połączenia głosowego, video, czy transmisji danych. Wyposażony jest w interfejsy pokazane na rysunku 1.4.

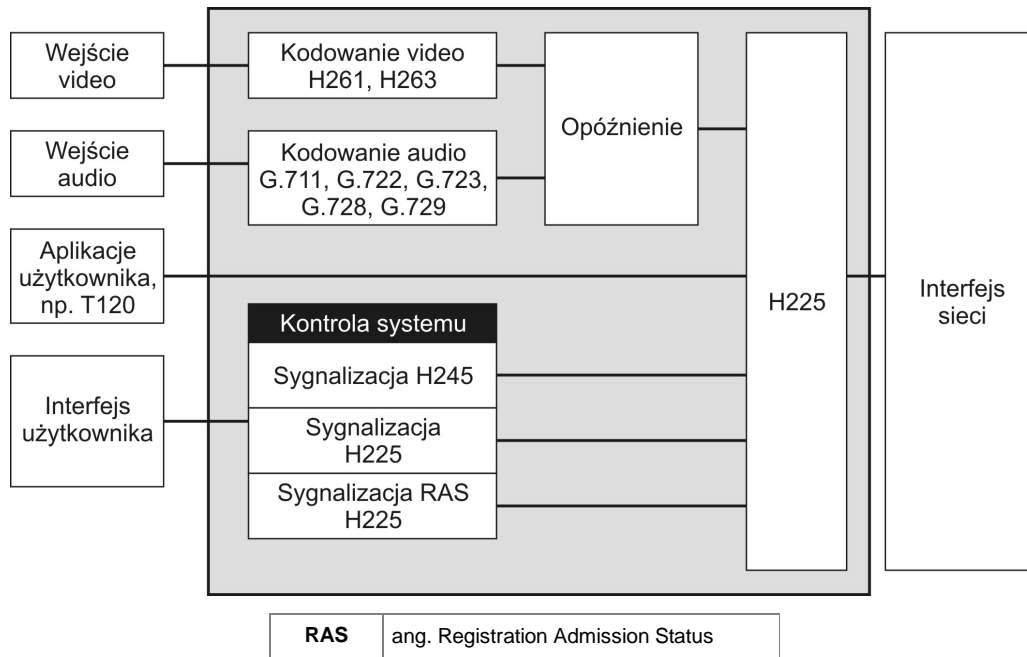
- Brama H323

Brama H323 jest elementem sieci, który pozwala łączyć sieć telefoniczną VoIP pracująca w standardzie H323 z siecią telefoniczną pracującą w innym standardzie. Zadaniem bramy jest zapewnienie odpowiedniej translacji protokołów sygnalizacyjnych i kontrolnych. Często konieczna jest konwersja zakodowanego sygnału dźwięku do formatu kompresji stosowanej w sieci podłączanej i odwrotnie. Urządzenie to pozwala więc na łączenie sieci oferującej usługi VoIP z siecią telefoniczną stacjonarną PSTN, ISDN, czy siecią oferującą usługę VoIP działającą z innym protokołem sygnalizacji.

- Mostek konferencyjny (ang. MCU - Multipoint Control Unit)

Mostek konferencyjny jest urządzeniem końcowym, którego zadaniem jest umożliwienie dowolnym minimum trzem różnym urządzeniom końcowym zestawienie jednoczesnego połączenia – połączenia tzw. konferencyjnego. Taką funkcjonalnością

może cechować się zarówno Terminal H323 lub może być to osobne urządzenie. Zestawiona konferencja może dotyczyć trzech rzeczy: głosu, obrazu i wymiany danych. Urządzenie MCU nie jest urządzeniem koniecznym do pracy sieci z sygnalizacją H323, dostarcza tylko usługi dodatkowej.



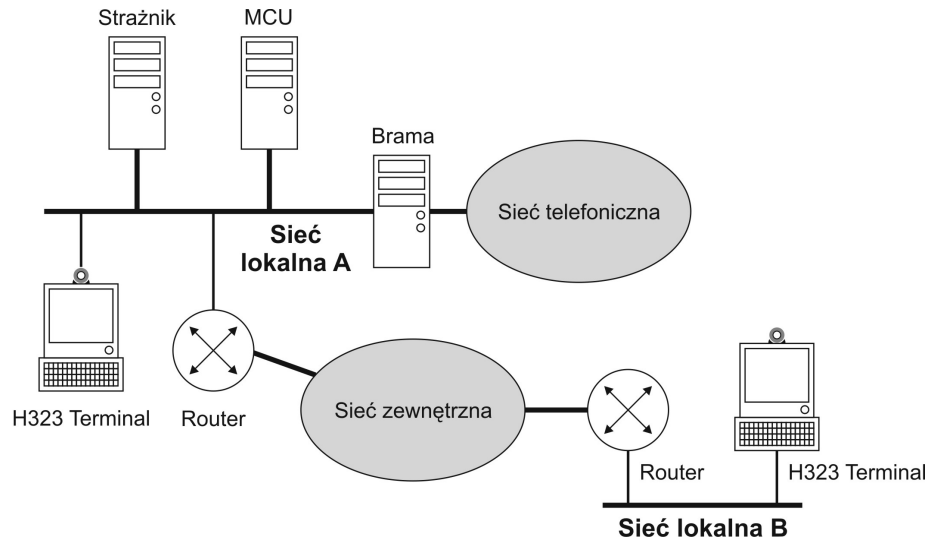
Rysunek 1.4 Struktura terminala H323

- Strażnik (ang. Gatekeeper)

Strażnik jest kluczowym elementem sieci działającej według protokołu H323. Jego zadaniem jest kontrola i zarządzanie siecią. Pozwala on na nadzór nad urządzeniami końcowymi. Dostarcza takich funkcji jak: bilingowanie połączeń, co jest niezbędnego do prowadzenia rozliczeń kosztów rozmów, nadzór nad możliwościami nawiązywania połączeń CAC (ang. Call Admission Control), translacje adresów, zarządzanie pasmem. Jedną z najważniejszych jego funkcji jest zarządzanie numeracją telefoniczną w sieci. Urządzenia końcowe – terminale, bramy, MCU, rejestrują się do strażnika za pomocą protokołu H225. Podczas rejestracji mogą być dokonywane metody autoryzacji użytkowników, a następnie przypisywanie im odpowiednich numerów telefonów. W sieci pracującej w standardzie H323 bez tego urządzenia możliwe jest wykonywanie połączeń tylko poprzez wybieranie adresów IP terminali czy bram. Dopiero wprowadzenie strażnika do sieci pozwala na zastosowanie numeracji zgodnej ze

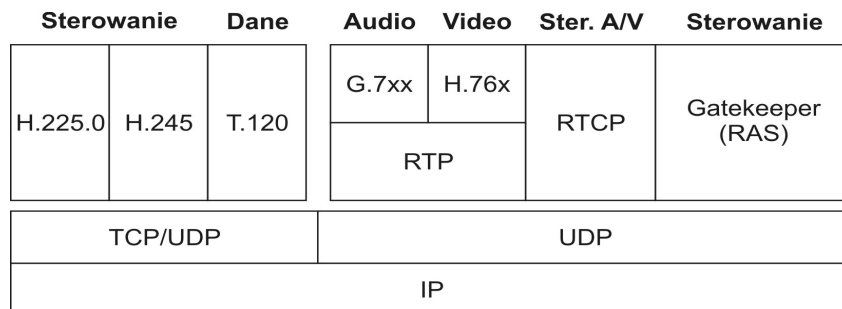
standardem E164 lub nazewnictwa według wyobraźni projektanta sieci – np. imion czy nazw użytkowników.

Przykładową strukturę sieci H323 pokazano na rysunku 1.5.



Rysunek 1.5. Przykładowa struktura sieci H323

Protokół H323 do swojej pracy wykorzystuje inne protokoły opisane również poprzez ITU-T oraz protokoły Internetu, które można określić mianem stosu protokołów H323.



TCP	ang. Transmission Control Protocol
UDP	ang. User Datagram Protocol
IP	ang. Internet Protocols
RTP	ang. Real-time Transfer Protocol
RTCP	ang. Real-time Transfer Control Protocol
RAS	ang. Registration Admission Status

Rysunek 1.6. Stos protokołów H323

W trakcie nawiązywania połączenia terminale H323 wymieniają informacje w jakim standardzie kodowania dźwięku i obrazu preferują pracować oraz listę wszystkich obsługiwanych metod kompresji – proces wymiany tych danych i ustalenia kodowania nazywa się procesem negocjacji parametrów. Od wybranego sposobu kodowania zależy zapotrzebowanie na pasmo transmisyjne. Oczywiście jest, że im większa kompresja danych jest stosowana, tym wymagane pasmo transmisji jest mniejsze, jednakże wzrasta zapotrzebowanie na zasoby mocy obliczeniowej terminala konieczne do dokonania stosownej kompresji. Standard T.120 określa metodę wymiany danych poprzez użytkownika. SCUI (ang. System Control User Interface) służy do sterowania terminalem – użytkownik uruchamia odpowiednie procesy protokołów H245 i H225 np. w momencie gdy chce nawiązać połączenie, odebrać rozmowę przychodzącą czy przełączyć rozmowę na inny numer w sieci. Zadaniem protokołu H245 jest głównie wymiana możliwości technicznych terminali oraz informacji o otwieranych kanałach transmisji danych przy nawiązywaniu połączeń. Natomiast zadaniem protokołu H225 jest sygnalizacja i kontrola połączenia. Komunikaty RAS (ang. Registration Admission Status) protokołu H225 służą do komunikacji terminala ze strażnikiem. Pozwala on strażnikowi na rejestrację terminali, zarządzania siecią i przechowywanie stanu pracy sieci. Wersja druga protokołu H323 wprowadziła możliwość tunelowania komunikatów protokołu H245 w komunikatach protokołu H225 celem zmniejszenia ilości pakietów IP koniecznych do wymiany przy nawiązywaniu połączenia.

Aby zapewnić niezawodność sygnalizacji i kontroli stosowane są w zależności od potrzeb protokoły TCP lub UDP. Dźwięk i obraz są przesyłane w trybie rzeczywistym wykorzystując protokół RTP i są kontrolowane za pomocą protokołu RTCP (ang. Real-time Transport Control Protocol). Protokoły RTP i RTCP działają wykorzystując protokół UDP, ponieważ chwilowy brak danych dotyczących dźwięku lub obrazu nie wymaga ich ponownej transmisji, po prostu w danej chwili wystąpi cisza bądź brak obrazu.

1.6. SIP

Protokół sygnalizacji SIP został opracowany w 1999 roku przez IETF MMUSIC (ang. The Internet Engineering Task Force Multiparty Multimedia Session Control)

i opisany w dokumencie RFC2543 (ang. Request For Comments) [15]. W 2002 roku po dokonaniu analizy został opisany w dokumencie RFC3261. Dużą zaletą protokołu jest fakt, iż używa adresacji i składni URL (ang. Uniform Resource Locator), oraz to, że jest protokołem tekstowym, co ułatwia jego implementację jak i kontrolę poprawności działania. Podobnie do protokołu H323 pozwala użytkownikom nawiązywać rozmowy telefoniczne, zestawiać konferencje multimedialne, dystrybuować multimedia. Standard SIP został opracowany dla środowiska IP pracującego zarówno w wersji czwartej (IPv4) jak i szóstej (IPv6). Przewidziano pracę w trybie rozgłoszeniowym multicast jak i docelowym unicast. Silnym punktem standardu SIP jest praca z wykorzystaniem takich protokołów jak:

- HTTP 1.1 (ang. HyperText Transfer Protocol) RFC2616 do formatowania komunikatów,
- SDP (ang. Session Description Protocol) RFC2327 do negocjacji sposobów kodowania,
- RTP RFC 1889 i RTSP (ang. Real Time Streaming Protocol) RFC2326 do transmisji multimediiów,
- URL RFC1738 i URI (ang. Uniform Resource Identifier) RFC2396 do adresacji terminali użytkowników,
- DHCP (ang. Dynamic Host Configuration Protocol) RFC2131 i DNS (ang. Domain Name System) RFCs1034/1035 do rozwiązywania nazw i zapewnienia mobilności użytkowników,
- MIME (ang. Multipurpose Internet Mail Extensions) RFC2046 do kodowanie,
- TLS (ang. Transport Layer Security) RFC2246 i IPSec RFC2401 do zapewniania bezpieczeństwa.

Do działania systemu z protokołem SIP definiuje się pięć funkcji:

- odnajdowanie systemu, z którym ma być dokonywana komunikacja,
- rozpoznawanie możliwości dokonywania komunikacji przez system,
- rozpoznawanie typu mediów, które mają być wykorzystywane oraz ich parametrów,
- dzwonięcie, ustalenie parametrów po obu stronach,
- nadzorowanie sesji.

Elementami sieci pracującej z protokołem SIP są:

- UA (ang. User Agent)

UA jest odpowiednikiem terminala H323, który inicjuje i zgłasza żądania oraz odpowiada na połączenia.

- Brama SIP

Brama SIP jest elementem nieobowiązkowym, który podobnie jak brama H323, służy do translacji formatu rozmowy przeprowadzanej w standardzie SIP do formatu stosowanym w innej sieci, z którą sieć SIP chcemy połączyć. Zadaniem bramy jest zapewnienie odpowiedniej translacji protokołów sygnalizacyjnych i kontrolnych. Podobnie jak w standardzie H323, gdy konieczna jest konwersja głosu do formatu kompresji stosowanej w sieci podłączanej i odwrotnie trzeba użyć bramy. Urządzenie to pozwala więc na łączenie sieci oferującej usługę VoIP z siecią PSTN, ISDN, czy siecią oferującą usługę VoIP działającą z innym protokołem sygnalizacji

- SIP Registrar Server

Zadaniem SIP Registrar Servera jest ułatwienie lokalizacji użytkowników, gdyż nie są oni przywiązani do konkretnego miejsca (adresu IP). SIP Registrar Server przechowuje w swojej bazie powiązania połączenia użytkowników z ich numeracją,

- SIP Proxy Server

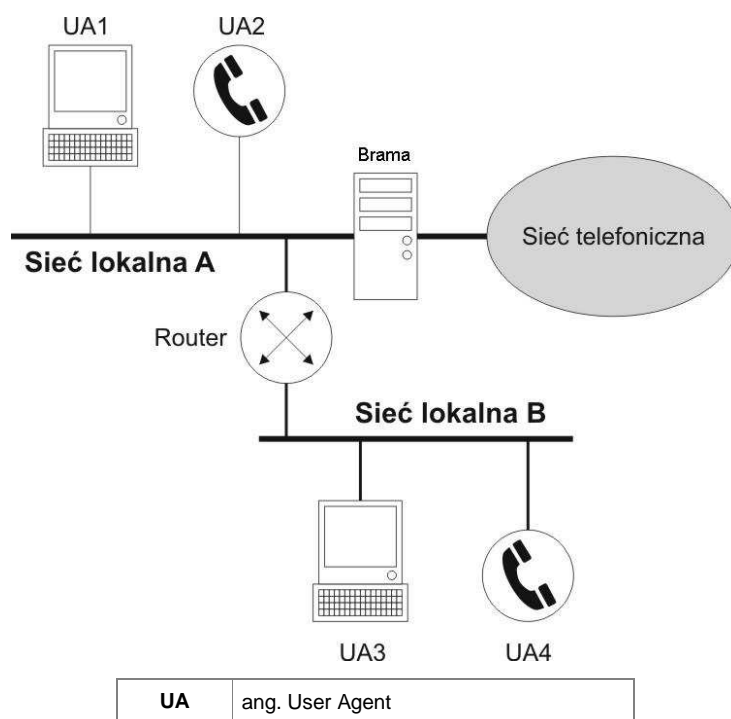
SIP Proxy Server kieruje i przesyła komunikaty SIP, jest przydatny przy prowadzeniu bilingu i nadzorze sieci. Sygnalizacja rozmów przy użyciu SIP Proxy Servera jest przesyłana właśnie poprzez ten element sieci.

- SIP Redirect Server

Element ten ma za zadanie wykrywanie zmian lokalizacji użytkowników w sieci.

W standardzie SIP adresacja użytkowników może być:

- adresacją w pełni domenową np.: sip:lukasz.wp.pl,
- adresacją domenową wzorowaną na adresacji SMTP (ang. Simple Mail Transfer Protocol) np.: sip:lukasz@wp.pl,
- adresacją według standardu E164 np.: 14056775432@brama.com,
- połączeniem wszystkich powyższych typów adresacji.



Rysunek 1.7. Przykładowa struktura sieci SIP

1.7. MEGACO i MGCP

Protokoły MEGACO i MGCP (ang. Media Gateway Control Protocol) zostały opracowane w celu określenia metody połączenia dwóch tradycyjnych sieci telefonicznych poprzez sieć IP [15]. Definiują one zasady działania bram do sieci zewnętrznych oraz ich kontrolę i sposoby zestawiania połączeń. Protokół MGCP jest wynikiem rewizji protokołów SIP i H323, a powstał w grupie opracowującej MEGACO IETF [16][17] oraz ITU-T [18]. Protokoły te skupiają swoją uwagę na budowie bramy, którą dzielą na trzy warstwy:

- kontroler mediów MGC (ang. Media Gateway Controller) odpowiedzialny za sygnalizacją w sieci IP,
- brama właściwa MG (ang. Media Gateway) odpowiedzialna za kodowanie i przesyłanie dźwięku,
- kontroler sygnalizacji SG (ang. Signalling Gateway) odpowiedzialny za sygnalizację w sieci dołączanej do sieci IP.

1.8. IAX

Protokół IAX (ang. Inter-Asterix eXchange) został opracowany przez Marka Spencera dla oprogramowania Asterix służącego do budowy i łączenia sieci oferujących usługę VoIP pracujących z różnymi typami sygnalizacji. Oprogramowanie Asterix często nazywane jest centralą. W 2009 roku został opublikowany dokument opisujący protokół IAX w wersji drugiej [19]. Zaletą protokołu jest praca z użyciem protokołu UDP i tylko jednego portu do przesyłania zarówno sygnalizacji jak i danych stanowiących zakodowany dźwięk. Pozwala to na budowanie strumieni danych, zmniejszenie ilości nagłówków pakietów koniecznych do przesłania.

1.9. Podsumowanie protokołów sygnalizacji

Istnieje jeszcze co najmniej kilka szerzej znanych protokołów sygnalizacji, jak choćby protokół SCCP (ang. Skinny Client Control Protocol). Wymienione w rozdziale protokoły są jednak najszerzej stosowanymi obecnie. Największa przyszłość wydaje się mieć protokół SIP dlatego iż jest on łatwo poszerzalny oraz zastosowano w nim nowoczesne podejście do numeracji. Sieci zbudowane w oparciu o standard SIP są łatwo skalowalne. Natomiast budowa sieci w oparciu o standard H323 jest stosunkowo skomplikowana. Wymaga konfiguracji strażników, bram, urządzeń MCU. Jednak standard H323 będąc starszym protokołem gwarantuje większą dostępność różnorodnego sprzętu do budowy sieci. Należy zaznaczyć, że kompatybilność sprzętu ze standardem należy raczej sprawdzić osobiście, gdyż często można znaleźć błędy w implementacji protokołu.

Przy projektowaniu sieci należy zwrócić uwagę na jej bezpieczeństwo. Należy tutaj wspomnieć o zabezpieczeniu poufności rozmów, zapewnieniu dostępności sieci na odpowiednim poziomie. Przy dostępie do sieci przez zewnętrznych użytkowników zagadnienie bezpieczeństwa staje się szczególnie ważne. Jeśli zastosuje się ukrywanie adresacji sieciowej poprzez usługę NAT (ang. Network Address Translation) lub PAT (ang. Port Address Translation) trzeba zapewnić urządzenie, które obsłuży poprawnie sygnalizację stosowanego protokołu usługi VoIP, co często jest problemem. Przy

zastosowaniu protokołu SIP, z tej racji, iż korzysta ze znanych wcześniej protokołów, zazwyczaj nie ma z tym problemów.

Należy wspomnieć również o komunikatorach internetowych różnych firm, które nie opisują jawnie standardu sygnalizacji jaki stosują. Komunikatory Skype, Gadu-gadu, czy Tlen, pozwalają na przeprowadzanie rozmów głosowych z ustawieniami różnych standardów kodowania dźwięku, ale nie jest do końca wiadomy sposób sygnalizacji przeprowadzenia rozmowy. Jest to zabieg marketingowy, gdyż zapewnia, iż użytkownik nie jest w stanie skorzystać z innego komunikatora niż danego dostawcy.

Rozdział II Jakość dźwięku w telefonii

2.1. Ocena jakości w sieci telefonicznej VoIP

Z punktu widzenia użytkownika końcowego zazwyczaj nie ważne jest jak pakiety dotarły do celu oraz w jaki sposób dźwięk był kodowany. Ważne są przede wszystkim trzy aspekty gwarantujące poprawną pracę:

- opóźnienie w przesyłaniu dźwięku,
- jakość przesyłanego dźwięku,
- echo.

Protokół RTCP pozwala na określenie opóźnienia pakietów. Często pod pojęciem jakości telefonii VoIP błędnie rozumie się politykę zapewnienia odpowiedniej transmisji pakietów IP [21] [22] [23]. Natomiast ocena jakości przesyłanego dźwięku jest zadaniem o wiele bardziej złożonym. O jakości dźwięku transmitowanego przez łącze telefoniczne świadczą te parametry, które umożliwiają identyfikowanie mówcy, decydują o naturalności brzmienia głosu, określają stopień swobody wymiany myśli i zrozumiałość przekazywanego tekstu oraz wpływają na natężenie uwagi rozmówców oraz częstość pytań zwrotnych [24].

2.2. Jednostki ocen jakości systemów transmisji dźwięku

Aby oceniać jakość dźwięku w systemach transmisji, potrzebna jest skala i jednostka, kryterium, które pozwala badane systemy porównać. Metody pomiarowe umożliwiają ocenę jakości przesyłanego dźwięku w sposób bardzo różnorodny.

Metody oceniające jakość transmisji dźwięku w systemie często bazują na ocenie jakości transmitowanych dźwięków o określonym czasie trwania prezentowanych słuchaczom bezpośrednio po sobie. Zadaniem słuchaczy jest ocena zniekształceń sygnału w stosunku do prezentowanego poprzednio bądź bezpośrednio jego jakości. Wartości ocen zniekształceń sygnału oraz jego jakości wraz z opisem ich znaczenia zostały zebrane w tabelach 2.1 i 2.2. W obu przypadkach im większa jest wartość oceny, tym jakość sygnału jest lepsza.

Tabela 2.1. Ocena zniekształceń sygnału

Wartość współczynnika jakości	Zniekształcenia sygnału w stosunku do próbki dźwięku prezentowanej poprzednio
5	Niezauważalna
4	Dobra
3	Dostateczna
2	Mierna
1	Nieodpowiednia

Tabela 2.2. Tabela ocen związana z oceną jakości sygnału

Wartość współczynnika jakości	Jakości sygnału w stosunku do próbki dźwięku prezentowanej poprzednio
3	Dużo lepsza
2	Lepsza
1	Nieco lepsza
0	Tożsama
-1	Nieco gorsza
-2	Gorsza
-3	Dużo gorsza

Jedną z najczęściej stosowanych skal oceny jakości transmitowanego dźwięku w telefonii VoIP jest skala zdefiniowana poprzez ITU-T określana jako MOS [25] [26]. Skala ta przyjmuje wartości od 1 do 5, przy czym dźwięk oceniony jako 5 ma jakość idealną, natomiast im nota jest niższa, tym jakość dźwięku jest gorsza. Oczywiście badanym dźwiękiem jest mowa ludzka transmitowana poprzez system telefoniczny. Wartość współczynnika MOS i jego znaczenie zostały przedstawione w tabeli 2.3. Wartości współczynnika MOS zostały określone w drodze eksperymentów przeprowadzanych przez szereg lat poprzez ITU-T. Reprezentatywnej grupie słuchaczy prezentowano nagrania dźwięków poddanych i nie poddanych zniekształceniom, a następnie notowano oceny jakości dźwięku jakie przedstawili słuchacze. W drodze uśrednienia wyników uzyskano skalę MOS, która uzyskała szeroką akceptację oraz opis, który pozwala łatwo interpretować wynik oceny. Należy zaznaczyć, że słuchacze, którzy brali udział w eksperymentach nie byli specjalistami w dziedzinie badania jakości dźwięku.

Tabela 2.3. Wartości współczynnika *MOS* [25]

Wartość współczynnika MOS	Jakość mowy (opis wartości)
5	Doskonała
4	Dobra
3	Dostateczna
2	Mierna
1	Nieodpowiednia

Innym ze stosowanych wskaźników jakości transmitowanego dźwięku jest współczynnik *R*. Wskaźnik ten jest wykorzystywany przy planowaniu mającej powstać sieci telefonicznej. Wartości współczynnika *R* decyduje o poziomie jakości systemu uwzględniając przewidywaną jakość mowy. Dla tego współczynnika również przyjęto zasadę, iż im wyższa jego wartość, tym jakość sygnału jest lepsza.

Współczynnik *R* może być odzwierciedlany w postaci wartości współczynnika MOS. Zależność pomiędzy współczynnikiem *R* i współczynnikiem MOS można przedstawić następująco:

$$MOS = \begin{cases} 1 & \text{dla } R < 0 \\ 1 + 0,035R + R(R - 60)(100 - R)^7 * 10^{-6} & \text{dla } 0 < R < 100 \\ 4,5 & \text{dla } R > 100 \end{cases} \quad (2.1)$$

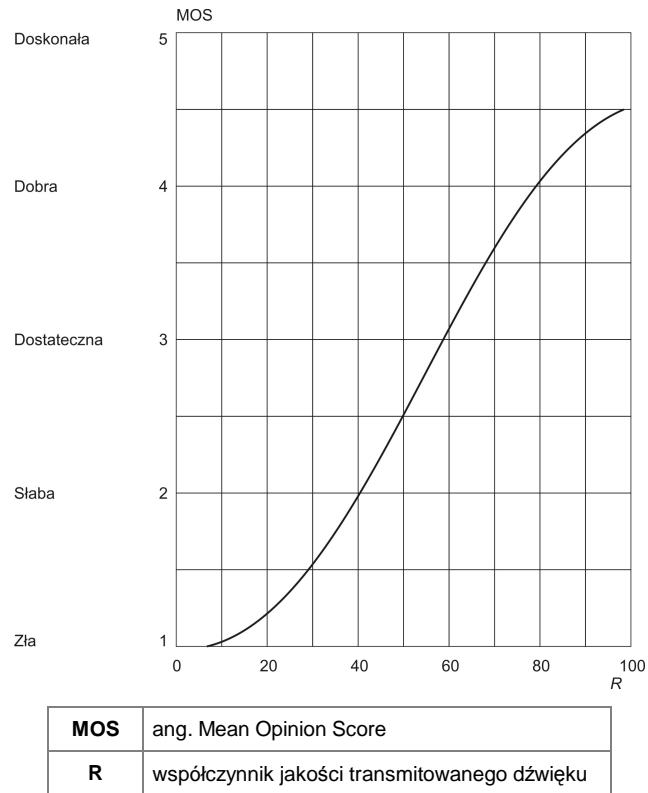
Zależność współczynnika *MOS* od współczynnika *R* została przedstawiona graficznie na rysunku 2.1.

Tabela 2.4. Wartości współczynnika jakości transmitowanego dźwięku *R* [24]

Wartość współczynnika R	Jakość transmitowanej mowy w systemie telefonicznym
90	Bardzo dobra
80	Dobra
70	Odpowiednia
60	Dostateczna
45	Nieodpowiednia
35	Reklamacje klientów

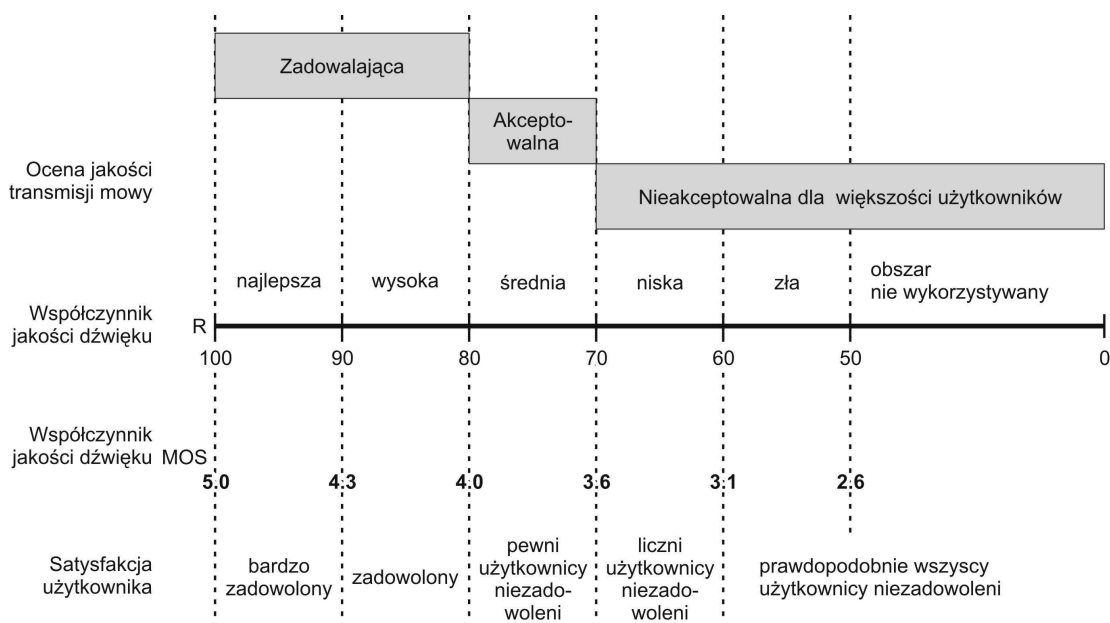
Ponieważ najczęściej stosowaną skalą jest skala MOS, można przyjąć, że najlepszą metodą pomiaru systemów cyfrowej transmisji dźwięku jest metoda,

która przedstawia swój wynik właśnie skali MOS. Takie wyniki można później łatwo porównywać. Pożądane jest więc aby wynik oceny jakości dźwięku w telefonii VoIP był właśnie wynikiem w skali MOS.



Rysunek 2.1. Zależność współczynnika *MOS* od współczynnika *R* [24]

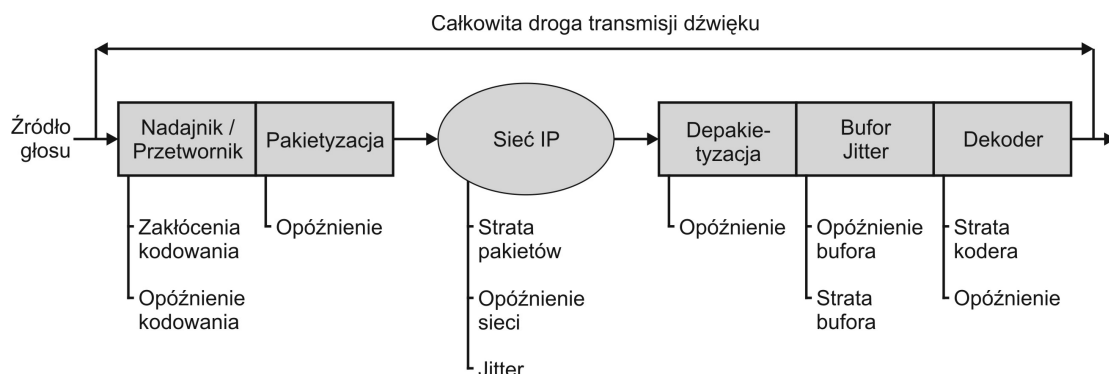
Wzajemne powiązanie ocen jakości transmitowanego dźwięku zostało przedstawione na rysunku 2.2. Niezależnie od stosowanej skali ocen ważne jest uzyskanie odpowiednio wysokiego poziomu zadowolenia użytkowników korzystających z systemu telefonicznego.



Rysunek 2.2. Wzajemne powiązanie ocen jakości transmitowanego dźwięku

2.3. Problemy sieci telefonicznej VoIP

Bez względu na wybrany sposób sygnalizacji połączeń w telefonii VoIP, każda sieć napotyka na problemy przesyłania danych reprezentujących dźwięk nagrany i transmitowany w czasie rzeczywistym za pomocą protokołów IP. Sieć IP nie była projektowana z myślą o telefonii, ale mechanizmy, które dostarcza, przy odpowiednim wykorzystaniu pozwalają w zupełności na budowę sieci telefonicznej. Przy rozważaniu problemów jakości w sieci ważnym aspektem w analizie jest kwestia właściciela sieci, mianowicie, czy sieć jest siecią prywatną czy publiczną. W sieci prywatnej, całe zarządzanie mechanizmami jej działania zależą od decyzji jej administratorów. W sieci publicznej mechanizmy są określane przez wielu administratorów sieci IP i nie zawsze są one zbieżne, co znacznie utrudnia osiągnięcie odpowiedniej jakości transmisji w sieci. Obecnie często sieci oferujące usługę VoIP łączą jednostki jednej organizacji: firm czy instytucji, będąc sieciami prywatnymi. Dlatego w niniejszej pracy analizie poddawana jest właśnie prywatna sieć IP oferująca usługę VoIP.



Rysunek 2.3. Parametry i wpływ sieci oferującej usługę VoIP na jakość transmitowanego sygnału

Główne problemy jakie występują w sieci z usługą VoIP zostały zobrazowane na rysunku 2.3. Należą do nich:

- opóźnienia – kodowania jak i dekodowania, transmisji pakietów, przez co generowane jest opóźnienie transmisji dźwięku,
- jitter,
- natłoki,
- echo,
- szумы
- straty sygnału dźwięku.

Na samą jakość sygnału ponadto wpływ ma również:

- płéć mówcy,
- stosowany język,
- wiek mówcy.

2.3.1. Opóźnienia pakietów

Aby przesłać dźwięk poprzez sieć IP, musi on zostać odpowiednio przetworzony do postaci cyfrowej, przesłany, a następnie musi zostać dokonane przetwarzanie odwrotne – z postaci cyfrowej do analogowej. W każdym z elementów tego przetwarzania powstają opóźnienia [27].

Na całkowite opóźnienie wpływa mają cząstkowe składniki odpowiedzialne za:

- próbkowanie i kodowanie dźwięku,
- przygotowanie pakietu IP,
- transmisję pakietu IP,
- dekodowanie pakietu IP,
- przetwarzanie do postaci analogowej.

Od wyboru kodowania dźwięku zależy współczynnik kompresji i czas potrzebny na jego zakodowanie. Do uzyskania większej kompresji potrzebna jest większa moc obliczeniowa. Przy braku tej mocy obliczeniowej, wkrada się większe opóźnienie w przygotowaniu danych dla pakietu IP. Kolejnym krokiem jest przygotowanie pakietu IP, jego enkapsulacja w odpowiednie nagłówki protokołów. Jeśli stosowane jest kodowanie samego pakietu IP, należy tutaj wliczyć również czas potrzebny na wykonanie tej operacji i uwzględnić fakt, że w odbiorniku będzie musiał zajść proces odwrotny. Następnie wykonywana jest sama transmisja pakietu (czasami zwana serializacją pakietu), która wnosi swoje opóźnienie zależne od rodzaju sieci, jej przepływności binarnej. Dla sieci IP pracującej w technologii GigabitEthernet czasy transmisji pakietów będą pomijalne, natomiast dla sieci pracującej w technologii Frame Relay czas transmisji pakietu jest już znaczny. Przykładowe czasy transmisji pakietu w zależności od przepływności łącza zostały przedstawione w tabeli 2.5. W odbiorniku dochodzi czas konieczny na dekodowanie i dekompresję pakietu jeśli wystąpiła kompresja, a następnie przetwarzanie dźwięku z postaci cyfrowej do postaci analogowej. Ogromny wpływ na wartość całkowitego opóźnienia ma czas transmisji pakietów. O ile czasy kodowania czy dekodowania dźwięku i pakietów są przewidywalne w zależności od stosowanego urządzenia i metody kompresji, to na czasy transmisji pakietów wpływ ma wiele czynników. Czasy opóźnień i ich wpływ na rozmowę zostały przedstawione w tabeli 2.6.

Pierwszą sprawą, która wpływa na opóźnienia samej transmisji pakietu są przepływności łącza [29]. Od przepływności łącza zależy czas serializacji pakietu, czyli czas przesłania wszystkich bitów pakietu IP poprzez łącze. Im szybsze łącze, tym czas serializacji pakietu jest krótszy. Oprócz czasu serializacji w sieci występuje również czas propagacji samego sygnału przenoszącego poszczególne bity pakietu. Czas

ten jest zależny od zastosowanego medium transmisyjnego. Obecnie najszerzej stosowane są łącza światłowodowe oraz łącza elektryczne za pomocą przewodów symetrycznych czteroparowych.

Tabela 2.5 Czas transmisji pakietu IP dla różnych przepływności sieci

Przepływności binarna sieci [bit/s]	Czas transmisji pakietu [ms]		
	Wielkość pakietu [B]		
	100	200	300
64k	12	24	36
128k	6	12	18
256k	3	6	9
1M	0,8	1,6	2,4
10M	0,08	0,16	0,24
100M	0,008	0,016	0,024
1000M	0,0008	0,0016	0,0024

Tabela 2.6. Czasy opóźnień w transmisji dźwięku w sieci oferującej usługę VoIP [28]

Opóźnienie t [ms]	Akceptowalność
$t < 150$	Akceptowalne
$150 \leq t \leq 400$	Akceptowalne, w miejscach gdzie nie ma wpływa jakość usług
$400 < t$	Nieakceptowalne

Tabela 2.7. Czasy propagacji sygnału w przewodach

Typ przewodu	Czas propagacji sygnału na odcinku 1 m w [μ s]
Światłowod	0,005
Przewód elektryczny symetryczny czteroparowy	0,0056
Przewód koncentryczny	0,006 – 0,007

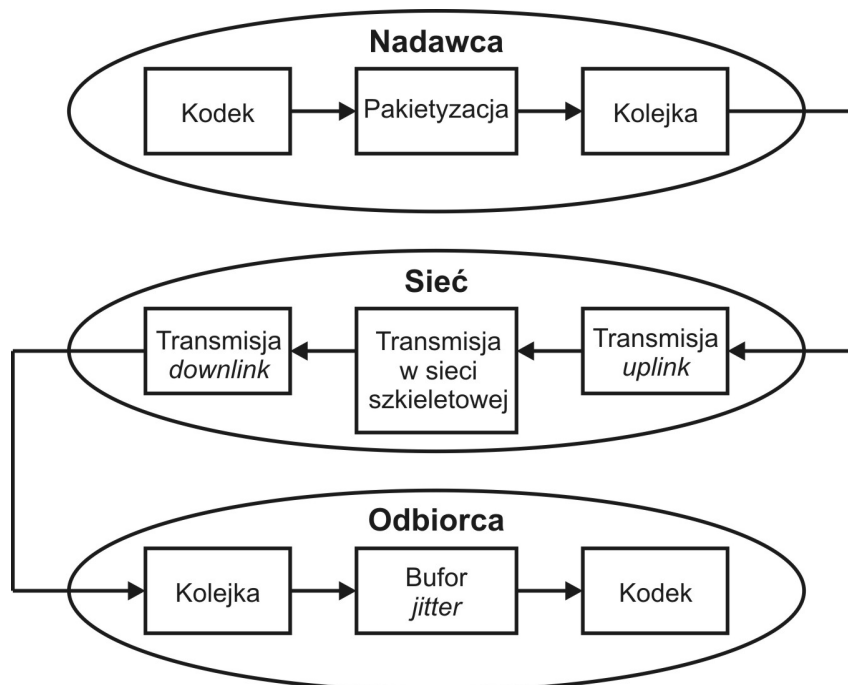
Kolejnym składnikiem opóźnienia jest czas przetwarzania pakietu poprzez routery. Router, którego zadaniem jest połączenie sieci, ma swoją ograniczoną wydajność. Od obciążenia jego procesora, architektury, mocy obliczeniowej, zależy czas, jaki musi upłynąć od otrzymania przez niego pakietu, poprzez znalezienie odpowiedniej drogi, modyfikację pakietu, do przesłania go interfejsem wyjściowym. Router może być obciążony wieloma zadaniami takimi jak zapewnienie szyfrowania, ukrywania adresacji czy wdrażania samej polityki zapewnienia jakości. Od struktury

sieci, ilości routerów i zasad jej funkcjonowania zależy więc czas opóźniania, jaki wniosą routery, które występują na drodze pakietu do odbiorcy. Wielkość opóźnienia transmisji pakietu może być również zależna od pory dnia. W określonych porach dnia, ruch w sieci może być znacznie większy niż średnie obciążenie sieci w czasie działania, co może znacząco wpłynąć na jej pracę.

Aby zapewnić jak najszybsze przekazywanie pakietów usług czasu rzeczywistego, jaką jest telefonia VoIP, korzysta się z mechanizmów gwarancji jakości przesyłanych pakietów IP - QoS [8] [31]. Wyróżnia się trzy rodzaje mechanizmów QoS [31]:

- Best Effort: W mechanizmie tym nie implementuje się żadnej polityki implementacji usług. W urządzeniach stosowane są kolejki FIFO (ang. First In First Out) polegające na tym, że pakiet, który został przysłany do urządzenia, zostanie przetworzony i odesłany jako pierwszy. Jest to najczęściej domyślny tryb pracy sieci stosowany w sieci Internet.
- InntServ: Mechanizm ten polega na gwarancji odpowiedniej przepływności sieci dla dostarczania pakietów – co za tym idzie czasu opóźnienia – poprzez zestawienie kanału transmisji z rezerwacją pasma. Warstwa aplikacji modelu ISO OSI (ang. International Organization for Standardization Open System Interconnection) musi przesłać odpowiednie pakiety mówiące o tym, iż routery mają zarezerwować pasmo dla transmisji danych. Protokołem, który jest wykorzystywany do dokonywania rezerwacji jest protokół RSVP (ang. Resource ReserVation Protocol). Wadą tego rozwiązania jest mała skalowalność.
- DiffServ: Mechanizm ten polega na wykorzystaniu pola ToS (ang. Type of Service) nagłówka pakietu IP do oznaczenia ważności pakietu. Pole to posiada osiem bitów. W pierwszej wersji implementacji usług wykorzystywano tylko trzy bity, co dawało osiem możliwości ustalania priorytetu pakietów. Dwa z nich zostały zarezerwowane do oznaczania pakietów koniecznych do pracy sieci. Natomiast pozostałe sześć priorytetów można było dowolnie wykorzystywać. Wykorzystywane trzy bity określa się mianem IP Precedence. Wersja późniejsza wprowadziła wykorzystanie sześciu bitów i określa je mianem DSCP (ang. Differentiate Service Code Point). Pozwala to w sumie na określenie

64-ech klas pakietów. Dla zapewnienia kompatybilności z urządzeniami pracującymi z wykorzystaniem tylko trzech bitów, klasy te dzieli się na podgrupy, które poprzez stosowanie odpowiedniej maski bitowej mogą być odwzorowywane na grupy pakietów określone poprzez bity IP Precedence. Po oznaczeniu pakietu stosuje się różne metody kolejkowania pakietów, ustalania ich kolejności do transmisji, przydzielania grupom pakietów określonego pasma fizycznego łącza czy decyduje się o kolejności usuwania w momencie zbyt dużego obciążenia sieci. Zaletą tego rozwiązania jest dość duża skalowalność. Mechanizm ten napotyka jednak swoje problemy. Przy budowie tuneli wirtualnych, stosowanych przy zestawianiu bezpiecznych łączy VPN (ang. Virtual Private Network) pola ToS są zastępowane przez pola nagłówka tunelu. Podobna sytuacja występuje przy zastosowaniu tuneli GRE w celu wymiany tablic routingu poprzez sieci nie pracujące z usługą multicast. Występuje tutaj konieczność odpowiedniego przekazywania parametrów nagłówka (pola ToS) do pakietu routera. Nie wszystkie routery mają takie możliwości.



Rysunek 2.4. Czynniki mające wpływ na wielkość opóźnienia w sieci IP [33]

Największym problemem budowy sieci oferującej usługę VoIP jest przesyłanie pakietów poprzez zewnętrznego usługodawcę – providera – sieci. Problem ten występuje w momencie, gdy firmowa sieć telefoniczna VoIP musi przesyłać swoje pakiety poprzez łącza dostarczone przez inną firmę. Konieczne jest wówczas zdefiniowanie z dostawcą łącza odpowiednich zasad polityki transmisji pakietów tak, aby zewnętrzna sieć reagowała na ustawione znaczniki priorytetu pakietów transmitując je w odpowiedniej kolejności. Problemem jaki spotyka się w sieci Internet, jest brak polityk zapewnienia jakości na routerach, z których jest zbudowana.

2.3.2. Jitter

Jitter jest zjawiskiem związanym z opóźnieniem w transmisji pakietów. Otóż w sieci IP występujące opóźnienia mają zmienną wartość. Trasa pakietów IP może być zależna od ruchu występującego w sieci. Nie każdy pakiet nadany od nadawcy A do odbiorcy B zostanie przesłany tą samą drogą. Wystąpią różne opóźnienia w transmisji pakietów od tego samego nadawcy do tego samego odbiorcy, a w najgorszym przypadku, pakiety wysłane wcześniej, dotrą do odbiorcy później od kolejnego nadanego pakietu. Zmienny czas stanowiący opóźnienie dostarczania pakietów określa się nazwą jitter. W odbiorniku musi nastąpić proces sortowania pakietów według kolejności ich nadania. Do określenia kolejności służą znaczniki nadane przez protokół RTP w procesie enkapsulacji danych do wysłania. Dodane znaczniki pozwalają również obliczyć wartość jitter [32]. Wartość jitter J_i dla i -tego pakietu oblicza się ze wzoru:

$$J_i = J_{i-1} + \frac{D_{(i-1,i)} + J_{i-1}}{Y} \quad (2.2)$$

$$D_{(i-1,i)} = (O_i - S_i) - (O_{i-1} - S_{i-1}) \quad (2.3)$$

gdzie:

- O_i - czas przybycia i -tego pakietu,
- S_i - czas nadania i -tego pakietu.
- Y - stała wprowadzana do eliminacji szumu pomiaru, zwykle przyjmuje wartość 16.

Odbiornik posiada bufor, który zanim nastąpi dekodowanie i odtwarzanie przesłanego dźwięku, najpierw zbiera określoną ilość pakietów [33]. Bufor ten określa się mianem dejitter-buffer. Im większa jest wartość jitter, tym większy musi być bufor, co spowoduje zwiększenie opóźnienia w odtwarzaniu dźwięku. W skrajnym przypadku pakiet może dotrzeć do odbiornika z opóźnieniem większym niż wynosi wielkość bufora przeliczona na długość posiadanego w nim dźwięku, co za tym idzie, w odpowiedniej dla niego chwili czasowej wystąpiła przerwa w odtwarzanym dźwięku jak dla utraconego pakietu. Zakres wartości jitter i jego wpływ na jakość przesyłanego dźwięku został przedstawiony w tabeli 2.8. Z zakresów wartości wynika, że wartości jitter do 75 ms są do zaakceptowania w sieci oferującej usługę VoIP. Natomiast wartości większe są niepożądane.

Tabela 2.8. Wartości jitter i jego wpływ na jakość transmisji dźwięku w sieci oferującej usługę VoIP [30]

Wartość jitter J_i [ms]	Wpływ na jakość transmisji dźwięku
$J_i < 40$	Jitter nie wykrywalny – dobra jakość
$40 \leq J_i \leq 75$	Dobra jakość
$75 < J_i$	Nieakceptowalne

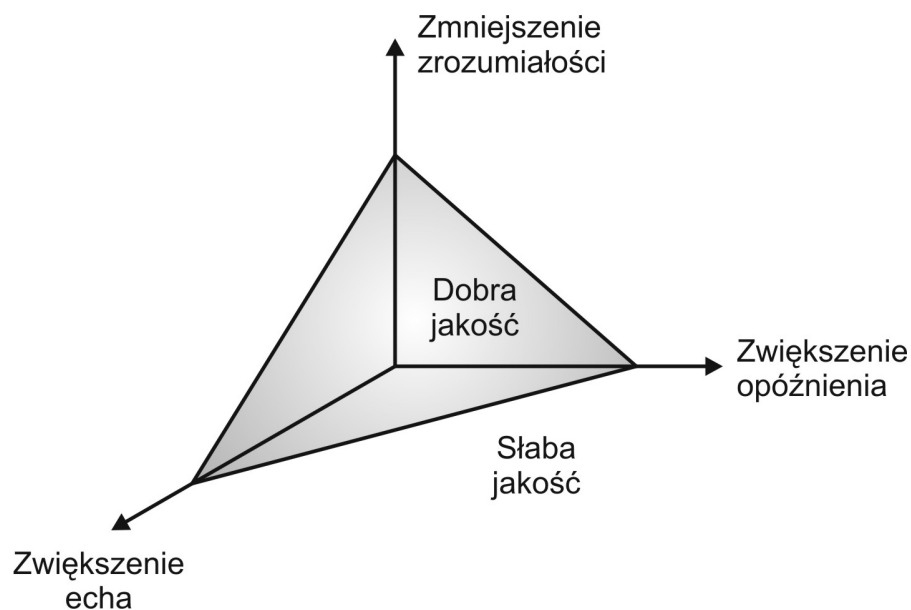
2.3.3. Natłoki

Jednym z problemów sieci IP są występujące w niej natłoki. Zjawisko to jest dość często poruszane w literaturze [31][34][35]. Polega ono na zbyt dużej ilości pakietów, która napływa do routera w stosunku do możliwości ich wysłania przez posiadane łącza. Opracowywane są mechanizmy redukcji natłoków, usuwania ich niepożądanych efektów – jak m.in., ponowna synchronizacja transmisji TCP, co zmniejsza poziom wykorzystania łącza. Głównym efektem tych prac są metody usuwania pakietów zanim natłok w sieci powstanie oraz w momencie, gdy natłok już powstał. Dla sieci oferującej usługę VoIP usuwanie pakietów ma ogromne znaczenia na jakość przeprowadzanych rozmów. Każdy utracony pakiet, to strata pewnego odcinka

czasowego przesyłanego dźwięku. Dlatego mechanizmy usuwania i zapobiegania natłokom są bardzo ważne, ale dla pakietów sieci oferującej usługę VoIP natłoki nie mogą w ogóle występować.

2.3.4. Pozostałe parametry wpływające na jakość transmitowanego dźwięku

W sieci oferującej usługę VoIP, tak jak w każdej sieci telefonicznej, pojawiają się problemy związane z echem, kompensacją utraconych pakietów danych, szumami czy różnicą poziomów przesyłanych sygnałów. Wpływ różnorodnych zjawisk na jakość sygnału można przedstawić na wspólnym wykresie (rysunek 2.5). Wraz ze wzrostem opóźnienia, echa, czy zmniejszeniem zrozumiałości mowy zmniejsza się jakość transmitowanego dźwięku w przeprowadzanej rozmowie, a ich wzajemne oddziaływanie dodatkowo potęguje wrażenie utraty tejże jakości [36].



Rysunek 2.5. Przestrzeń jakości transmitowanego dźwięku

Do zjawisk wpływających na zmniejszenie zrozumiałości należą:

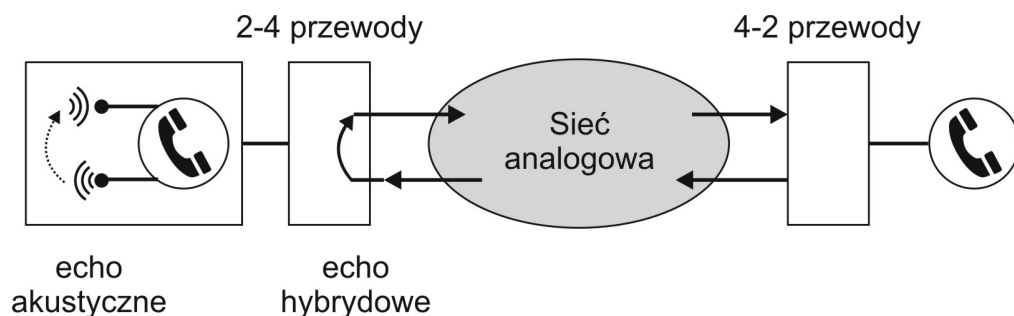
- przerwy w sygnale,
- różnice w poziomach głośności,
- zniekształcenia sygnału,

- jitter,
- przerywanie rozmów,
- dostępności sieci,
- jakość aparatów telefonicznych,
- szумы,
- możliwości prowadzenia rozmowy dwukierunkowo.

Istnieje wiele rozwiązań redukcji szumów w sygnale [20][37][38][39]. W sieci oferującej usługę VoIP szумы mogą zostać naniesione na sygnał transmitowanego dźwięku w momencie przetwarzania sygnału analogowego do postaci cyfrowej: w aparatach telefonicznych i bramach do sieci zewnętrznych.

Echo, które jest sygnałem zwrotnym, wracającym do nadawcy ale opóźnionym w czasie, powstaje w wyniku m.in.:

- niedopasowania impedancyjnego w sieci,
- słabych jakościowo aparatów telefonicznych,
- stosowania urządzeń do syntezy mowy.

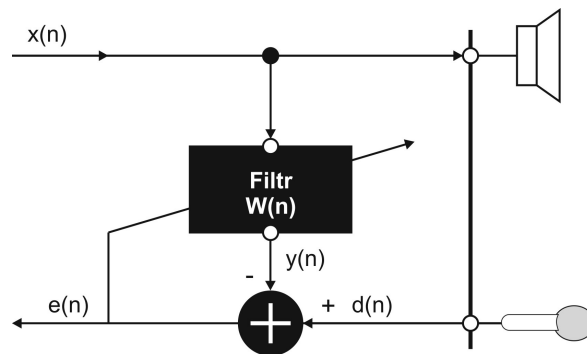


Rysunek 2.6. Źródło echa w telefonii analogowej

W sieci oferującej usługę VoIP echo wynikające z niedopasowania w sieci może powstać na styku z siecią zewnętrzną. Pakiety IP przesyłane w sieci, transmitujące zakodowany sygnał akustyczny, nie mają na siebie wpływu, czyli nie mogą wytworzyć echa. Dlatego w sieci oferującej usługę VoIP zjawisko echa może pochodzić głównie z aparatów telefonicznych – terminali użytkowników jak i sieci zewnętrznych podłączonych za pomocą bram do sieci oferującej usługę VoIP. Analogowe terminale telefoniczne mogą wnosić do sygnału echo akustyczne, wynikające z tego, iż część sygnału odtwarzana przez głośnik jest ponownie poddawana przetwarzaniu poprzez

obwód wejściowy – mikrofon. Ponadto w centralach telefonicznych może pojawić się echo elektryczne, które jest spowodowane przejściem z sygnału transmitowanego za pomocą jednej pary przewodów do systemu cztero przewodowego.

Aby zmniejszyć echo, a nawet je usunąć, stosuje się algorytmy kasowania echa, najczęściej zapisane sprzętowo w procesorach DSP (ang. Digital Signal Processor) wykorzystujące filtry do przetwarzania sygnału. Układ eliminacji echa (rysunek 2.7) wykorzystuje metodę adaptacyjną. Filtr adaptacyjny zmienia iteracyjnie swoje parametry w celu zmniejszenia różnicy pomiędzy wartością na wyjściu, a wartością uzyskaną w wyniku predykcji. Podczas każdej iteracji błąd sygnału jest podawany do filtru i wpływa na zmianę jego parametrów.



Rysunek 2.7. Przykład schematu eliminacji echa w sygnale dźwięku [40]

Na rysunku 2.7 występują następujące oznaczenia:

- $x(n)$ - sygnału dźwięku na wejściu,
- $y(n)$ - sygnału dźwięku na wyjściu filtru adaptacyjnego,
- $e(n)$ - wartość błędu sygnału,
- $d(n)$ - sygnał akustyczny rozbrzmiewający w środowisku – wnoszący echo,
- $W(n)$ - reprezentuje filtr adaptacyjny.

2.4. Ocena jakości standardów kodowania dźwięku w telefonii VoIP

Na jakość dźwięku przesyłanego poprzez system wpływ ma metoda kodowania dźwięku. Jak już wspomniano w telefonii VoIP stosuje się różną kompresję dźwięku – a co za tym idzie – wprowadzana jest różna strata w jakości sygnału. Jakość

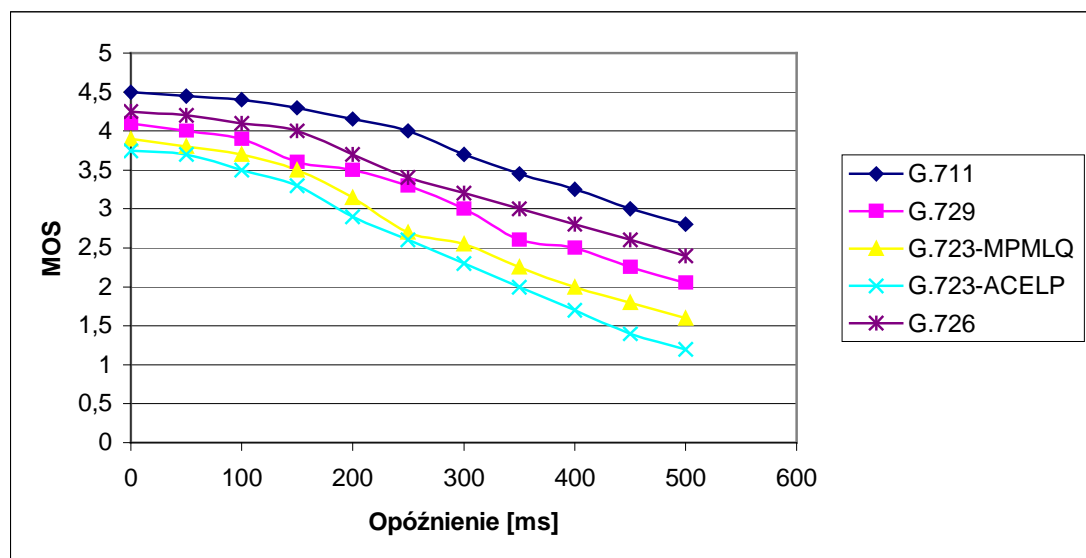
systemu dla warunków idealnych w transmisji można więc prognozować na samym etapie budowy sieci wybierając odpowiedni sposób kodowania dźwięku.

Tabela 2.9. Najczęściej stosowane metody kodowania dźwięku i ich ocena w skali MOS

Oznaczenie metody kompresji	Metoda kompresji	Pasma transmisyjne [kbps]	MOS
G.711	Modulacja impulsowo-kodowa (ang. PCM)	64	4,1
G.723.1A	Algebraiczna pobudzana kodem predykcja liniowa (ang. ACELP)	5,3	3,65
G.723.1A	Modulacja wielo impulsowa maksymalnego prawdopodobieństwa kwantyzacji (ang. MP-MLQ)	6,4	3,9
G.726	Adaptacyjna różnicowa modulacja impulsowo-kodowa (ang. ADPCM)	32	3,85
G.728	Pobudzana kodem predykcja liniowa z małym opóźnieniem (ang. LD-CELP)	16	3,61
G.729a	Sprzężona algebraiczna pobudzana kodem predykcja liniowa (ang. CS-CELP)	8	3,7
GSM	Regularne pobudzanie impulsowe – predykcja długoterminowa (ang. RPE-LTP)	13,2	4,0

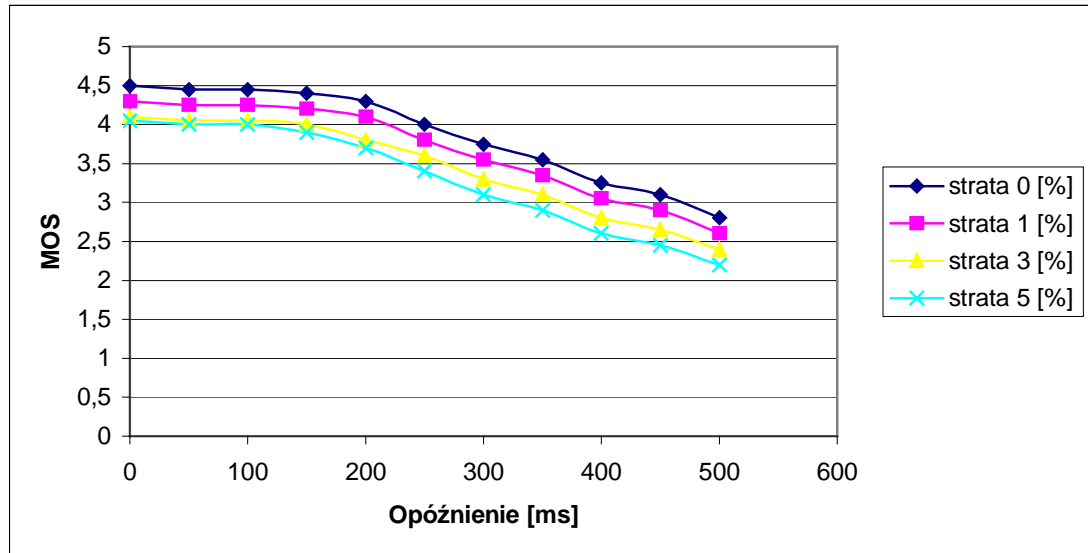
Ocenę wpływu opóźnienia jak i utraty pakietów można znaleźć w literaturze [41] [42].

Im większe opóźnienie pojawia się w sieci tym mniejsza jest jakość sygnału.

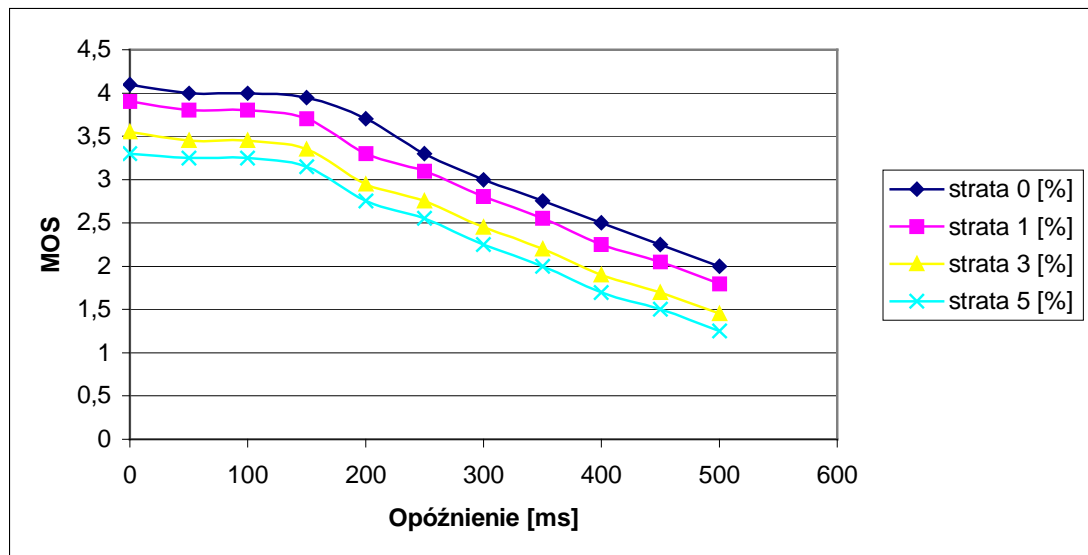


Rysunek 2.8. Ocena MOS metod kompresji dźwięku w funkcji opóźnienia transmisji pakietów [41]

Podobna sytuacja występuje przy stracie pakietów. Niezależnie od tego jaką zastosujemy kompresję dźwięku, strata pakietów spowoduje pogorszenie jakości sygnału.



Rysunek 2.9. Ocena metody kompresji dźwięku G.711 w funkcji opóźnienia dla różnych wartości procentowej utraty pakietów [41]



Rysunek 2.10. Ocena metody kompresji dźwięku G.729 w funkcji opóźnienia pakietów dla różnych wartości procentowej utraty pakietów [41]

Zwiększanie jakiegokolwiek z wymienionych czynników mających wpływ na jakość sygnału powoduje jej pogorszenie. Badania [41] potwierdzają zobrazowany na wykresach wpływ zniekształceń sygnału na jego jakość.

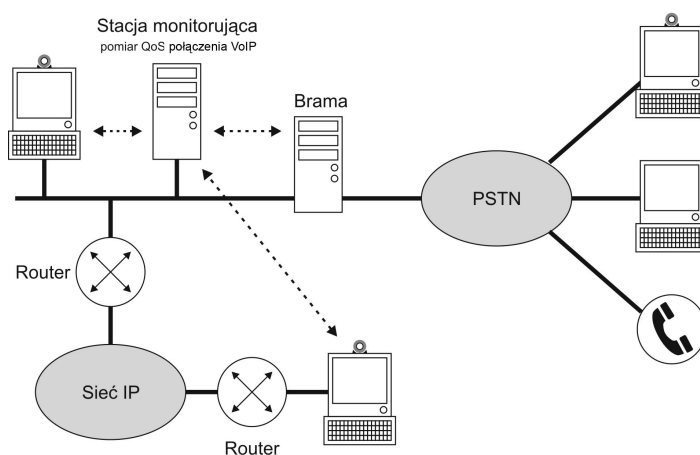
2.5. Adaptacyjna sieć oferująca usługę VoIP

Tradycyjna sieć oferująca usługę VoIP działa według ściśle określonych reguł polityk wykonywania połączeń. Połączenia wykonywane są przy ustalonych parametrach:

- typu sposobu kodowania dźwięku,
- określonej przepływności,
- określonej maksymalnej ilości połączeń.

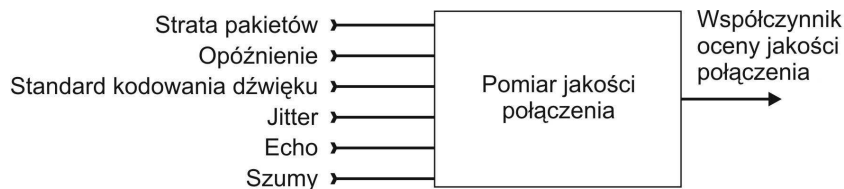
W momencie gdy w sieci wystąpią anomalie, przestaje ona działać prawidłowo, a administrator może tylko otrzymać informację od jej użytkowników, że w sieci wystąpił jakiś problem.

Sieć oferująca usługę VoIP można by spróbować konfigurować do działania w sposób adaptacyjny do anomalii jakie w niej wystąpiły. Jeśli sieć potrafiła by monitorować politykę jakości połączeń, to mogła by z dodatkiem pewnej inteligencji podejmować decyzje o reakcji na występujący problem. Taki punkt decyzyjny mógłby działać w centralnej stacji monitorującej prywatną sieć oferującą usługę VoIP (rysunek 2.11).



Rysunek 2.11. Sieć oferująca usługę VoIP ze stacją monitorującą jakość połączeń głosowych

Stacja monitorująca pracę sieci oferującej usługę VoIP za pomocą pomiaru jakości połączeń oraz parametrów wpływających na jakość (rysunek 2.12) dokonywała by decyzji, czy w sieci muszą nastąpić jakieś zmiany w jej działaniu.



Rysunek 2.12. Pomiar jakości połączenia głosowego

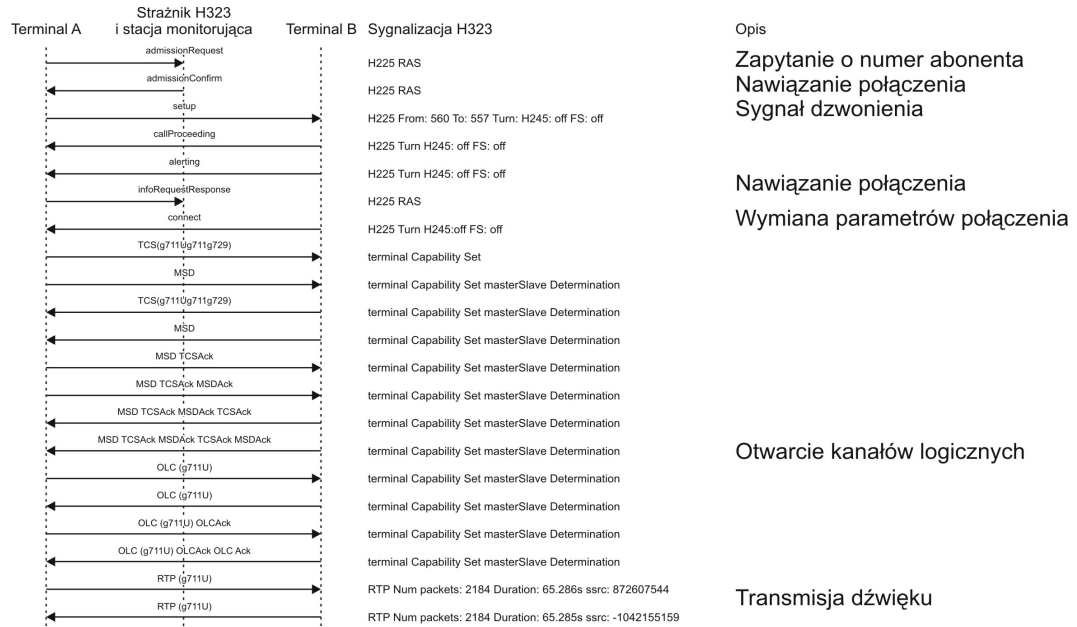
Na podstawie wyniku pomiaru jakości transmitowanego dźwięku w przeprowadzanej rozmowie można dokonać renegocjacji ustalonych parametrów rozmowy, np.:

- standardu kodowania dźwięku, wybierając standard o mniejszych wymaganiach na przepływności binarną w przypadku gdy problemem są natłoki w sieci,
- wielkości transmitowanego pakietu.

Proces negocjacji parametrów połączenia odbywa się na początku każdej rozmowy przeprowadzanej w technologii VoIP. Terminale VoIP wymieniają się informacjami jakich parametrów wymagają do poprawnego zestawienia połączenia. W momencie wykrycia anomalii w sieci, można rozbudować protokołu sygnalizacji, o renegocjacje parametrów połączenia, zamiast zwyczajowo przerwać rozmowę, której przeprowadzanie stałoby się utrudnione. Przykład sygnalizacji połączeń oraz zachowania terminali w momencie wystąpienia anomalii w adaptacyjnej sieci oferującej usługę VoIP został przedstawiony na rysunku 2.13. Po wykryciu anomalii w sieci terminale VoIP dokonują renegocjacji warunków połączenia adaptacyjnie do możliwości sieci.

Możliwa jest w ten sposób budowa adaptacyjnej sieci oferującej usługę VoIP, która w zależności od warunków pracy sieci może regulować jakość połączeń, zapewniając ją na odpowiednim poziomie. Taka sieć może w początkowej fazie pracy używać jak najlepszych sposobów kompresji dźwięku zapewniając wysoką jakość połączeń. W momencie wykrycia problemów z jakością – np. zwiększonego

zapotrzebowania na pasmo transmisyjne w sieci, adaptacyjna sieć oferująca usługę VoIP może wymusić zmianę sposobu kodowania dźwięku, tak aby zmniejszyć wartość zapotrzebowania przeprowadzanych połączeń na pasmo transmisyjne. Jednakże aby zbudować taką sieć, niezbędna jest metoda oceny jakości połączeń.



Wystąpiła anomalia w sieci, terminale otrzymują informacje ze stacji monitorującej



Rysunek 2.13. Przykład sygnalizacji w adaptacyjnej sieci oferującej usługę VoIP

Rozdział III Metody oceny jakości transmisji dźwięku w sieci oferującej usługę VoIP

3.1. Kategorie metod

Metody oceny jakości dźwięku można podzielić na:

- metody subiektywne,
- metody obiektywne.

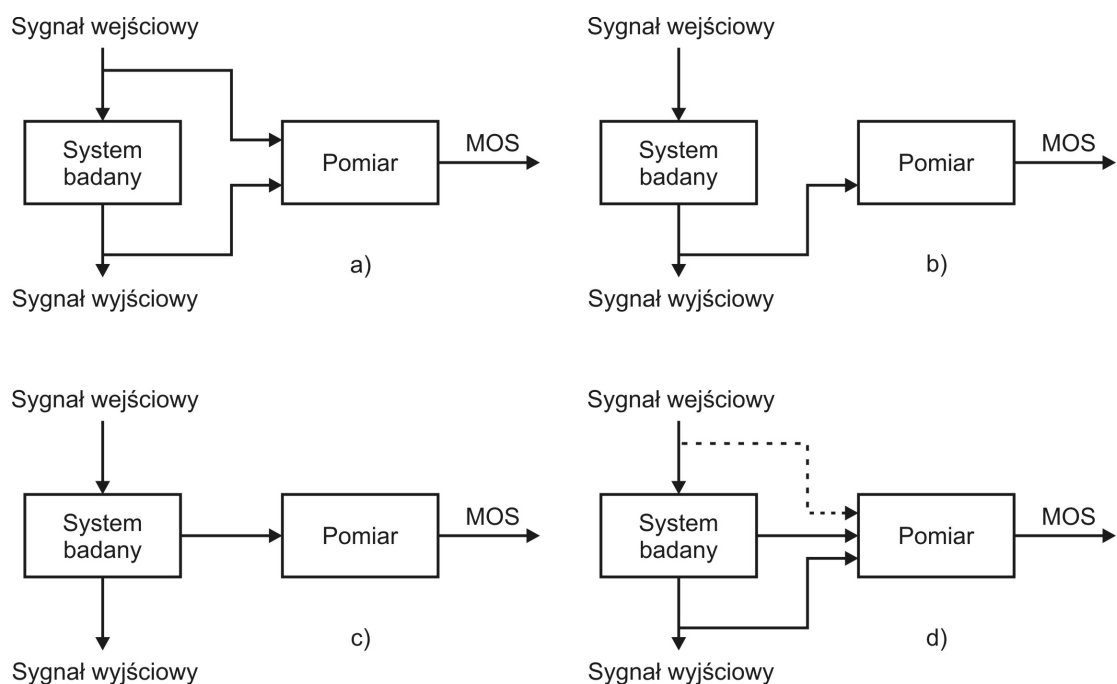
Metody subiektywne polegają na ocenie systemu poprzez jego użytkownika w laboratorium testowym. Metody obiektywne oceniają wartość jakości przez porównanie próbek mowy nadanej z mową u odbiorcy.

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.



Rysunek 3.1. Metody oceny jakości przesyłanych próbek dźwięku w trybie pomiaru:

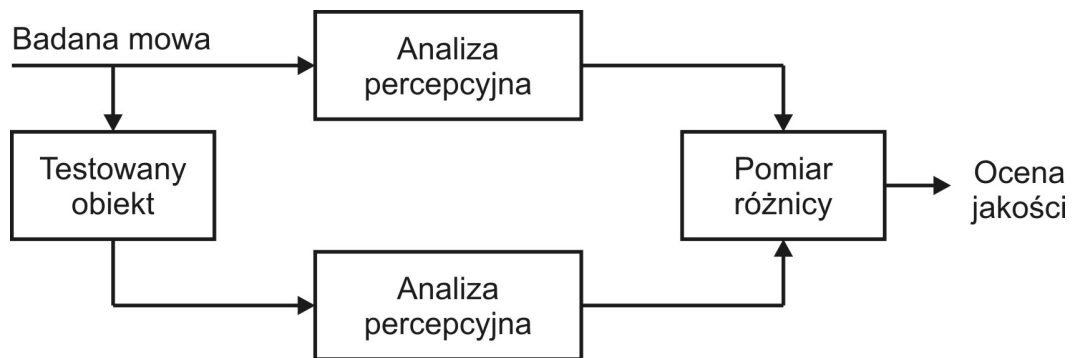
- a) na obu końcach systemu,
- b) na jednym końcu systemu,
- c) na podstawie parametrów systemu,
- d) hybrydowo - wykorzystując parametry systemu jak i pomiar sygnałów.

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

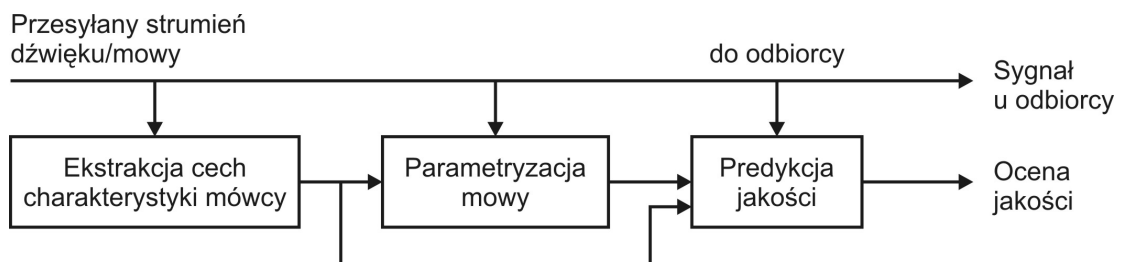
- ingerujące w strukturę badanego systemu,
- nie ingerujące w badany system [42] [43].

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 „podsluchu” 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.



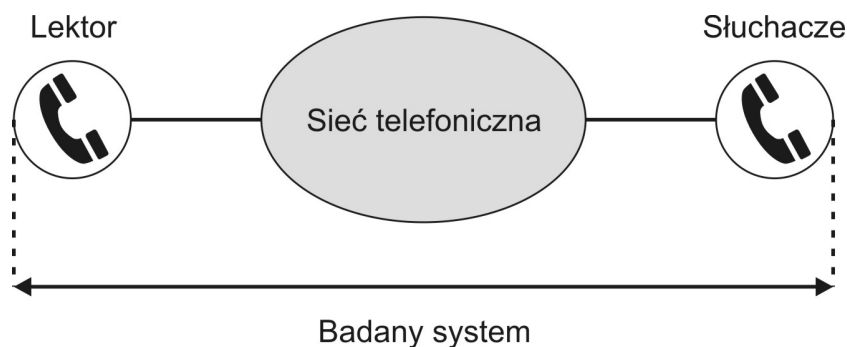
Rysunek 3.2. Schemat zasady działania metody oceny jakości dźwięku ingerującej w badany system



Rysunek 3.3. Schemat zasady działania metody oceny jakości dźwięku nie ingerującej w badany system

3.2. Pomiar jakości sygnału na podstawie wyrazistości logatomowej

Logatom to charakterystyczne dla danego języka złożenie fonemów nie posiadające znaczenia. Metoda oceny wyrazistości logatomowej polega na odczytywaniu przez lektora logatomów. Słuchacze zapisują logatomy, które usłyszeli po przetransmitowaniu poprzez badany system.



Rysunek 3.4. Struktura środowiska pomiarowego metody wyrazistości logatomowej

Każdy pomiar winien być poprzedzony sekwencją treningową. Następnie obliczana jest średnia wyrazistość logatomowa dla dla k-tej listy i n-tego słuchacza według wzoru:

$$W_L = \frac{1}{NK} \sum_{n=1}^N \sum_{k=1}^K W_{n,k} [\%] \quad (3.1)$$

gdzie:

- N - liczba słuchaczy,
- K - liczba nadanych list testowych,
- $W_{n,k}$ - wyrazistość logatomowa uzyskana podczas odsłuchiwania k-tej listy testowej przez n-tego słuchacza,

przy czym:

$$W_{n,k} = \frac{P_{n,k}}{T_k} \cdot 100 [\%] \quad (3.2)$$

gdzie:

- $P_{n,k}$ - liczba prawidłowo odebranych logatomów k-tej listy testowej przez n-tego słuchacza,
- T_k - liczba nadanych logatomów k-tej listy testowej.

Ponadto oblicza się odchylenie średniokwadratowe wyrazistości:

$$s = \left[\frac{1}{NK - 1} \sum_{n=1}^N \sum_{k=1}^K (W_{n,k} - W_L) \right]^{1/2} \quad (3.3)$$

Na podstawie tych parametrów oszacowana jest wartość MOS. Należy tutaj zaznaczyć, że inne są wartości oceny jakości dla różnych języków. Dla każdego języka potrzebny jest inny zestaw list logatomów i inny zestaw słuchaczy, którzy dobierani są

statystycznie, tak aby wynik pomiaru jakości przesyłanego dźwięku był reprezentatywny.

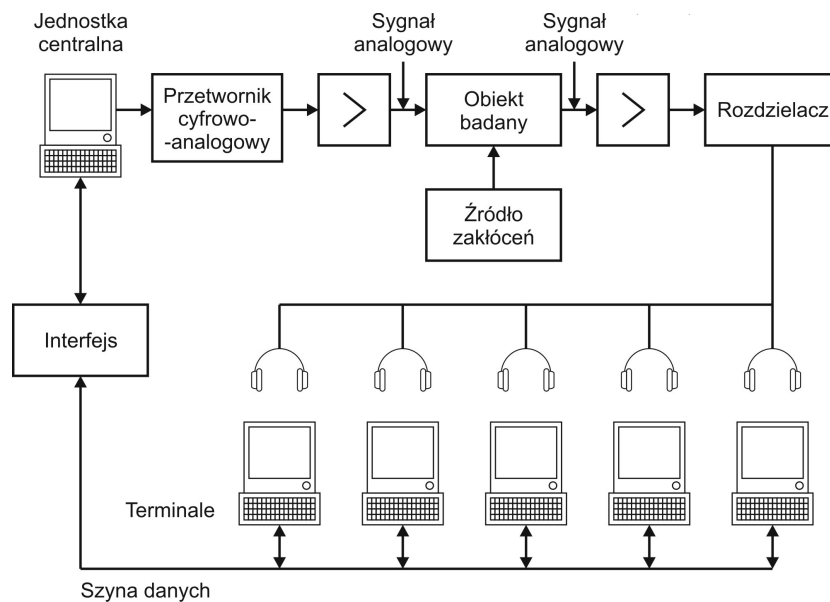
Metoda ACR (ang. Absolute Category Rating) jest metodą zalecaną przez ITU do oceny jakości sygnałów mowy systemów analogowych i cyfrowych [44]. W metodzie tej wykorzystuje się proste zdania, nie powiązane ze sobą semantycznie. Zdania są wcześniej przygotowane i zarejestrowane na nośniku poprzez nagranie mówcy, który winien cechować się poprawną dykcją. Wyróżnia się tutaj badania dla zdań nagranych zarówno przez mężczyznę jak i kobietę. Wyniki takie winno się analizować oddzielnie. Podobnie jak w metodzie wyrazistości logatomowej, badania przeprowadza się na grupie słuchaczy, którzy nie mogą być ekspertami w dziedzinie badania dźwięku. Ocenę jakości dokonuje się według trzech wybieranych skal ocen przedstawionych w tabeli 3.1.

Tabela 3.1. Pięciostopniowa skala ocen metody ACR [44]

Skala jakości	Skala wysiłku słuchowego	Skala preferowanej głośności
5 - doskonała 4 - dobra 3 - dostateczna 2 - mierna 1 - niedostateczna	5 - Rozumienie mowy bez najmniejszego natężenia uwagi. 4 - Rozumienie mowy bez trudności z lekkim natężeniem uwagi. 3 - Rozumienie mowy z umiarkowanym natężeniem uwagi. 2 - Rozumienie mowy z dużym natężeniem uwagi. 1 - Brak możliwości pełnego rozumienia mowy.	5 - Dużo głośniejsza niż optymalna 4 - Głośniejsza niż optymalna 3 - Optymalna 2 - Cichsza niż optymalna 1 - Zbyt cicha

W metodzie DCR (ang. Degradation Category Rating) stosowany jest taki sam materiał testowy jak w metodzie ACR, natomiast inny jest sposób oceny poprzez słuchaczy. Badane próbki mowy łączone są w pary i odtwarzane po sobie. Celowe jest stosowanie tzw. par zerowych, tj. dwóch identycznych próbek występujących po sobie. Pozwala to na określenie wrażliwości słuchaczy. Słuchacze oceniają degradacje sygnałów odtwarzanych w parach w stosunku do pierwszego z nich.

Metody wyrazistości logatomowej, ACR jak i DCR, bazują na ocenie dokonywanej przez słuchaczy. Ich zastosowanie wymaga poświęcenia czasu i środków finansowych, przez co nie nadają się do monitorowania stanu pracy sieci telefonicznej VoIP. Można spotkać pewne rozwiązania automatyzacji tych procesów pomiarowych [44]. W różnych podejściach można zauważyć wykorzystanie systemów komputerowych do odtwarzania próbek dźwięku i zbierania wyników. Pozwala to skrócić czas otrzymania wyników dla metod subiektywnych, jednak można je stosować praktycznie tylko w środowisku laboratoryjnym.



Rysunek 3.5. Przykład systemu automatyzacji metod subiektywnych [44]

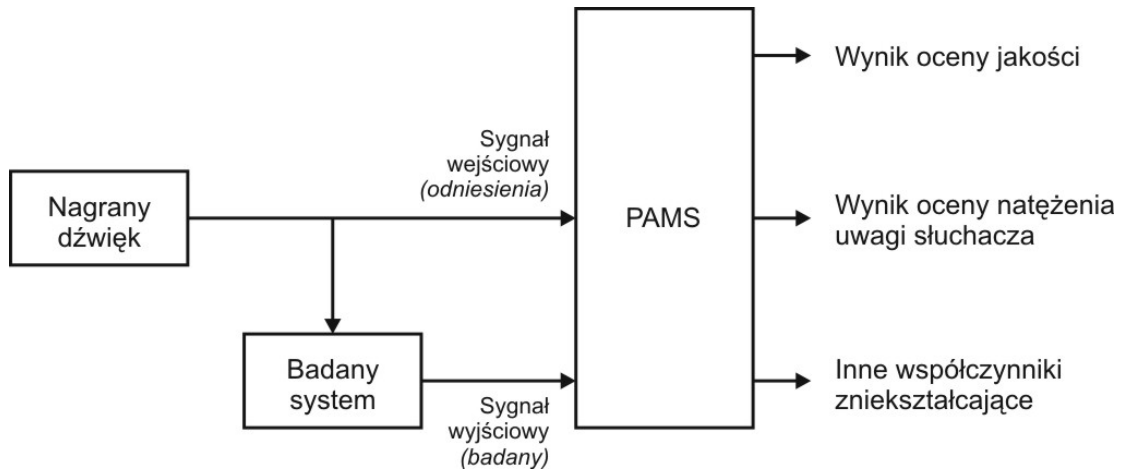
3.3. PAMS

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 [45]. Oznacza to, że system może być badany:

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



Rysunek 3.6. Schemat pomiaru za pomocą metody PAMS (ang. Perceptual Analysis Measurement System)

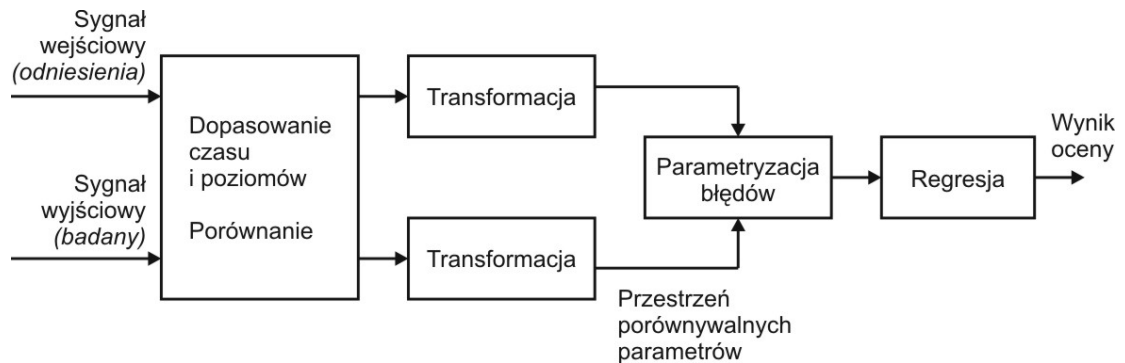
Praca z nagrami próbkami dźwięku na wejściu i wyjściu systemu wymaga specjalnego podejścia do pomiaru. Nagrane dźwięki muszą zostać poddane synchronizacji w czasie. Ponadto trzeba dopasować ich poziom mocy. Następnie można przeprowadzić porównanie nagranych dźwięków poprzez ich transformację do przestrzeni porównywanych parametrów. Ekstrakcja badanych parametrów uwzględnia aspekt ludzkiej percepcji. Obliczone parametry są porównywane w celu znalezienia różnic, które są odwzorowywane na wynik oceny jakości przesłanego dźwięku. 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 [45].

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 [46].



Rysunek 3.7. Schemat strukturalny metody PAMS

3.4. PESQ

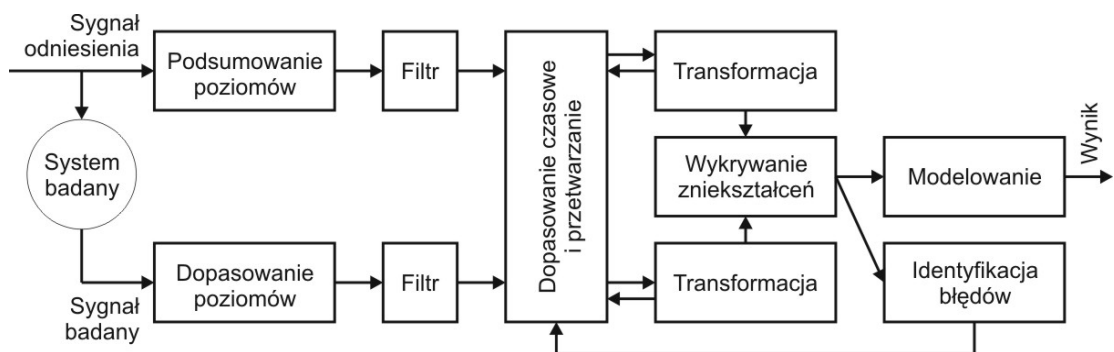
Metoda PESQ (ang. Perceptual Evaluation of Speech Quality) została opracowana i zaakceptowana przez ITU w 2001 roku jako rekomendacja P.862 i szybko stała się szeroko używaną metodą do oceny jakości systemów [47]. Ocenia jakość dźwięku z uwzględnieniem:

- 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 [45]:

- 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.



Rysunek 3.8. Schemat algorytmu metody PESQ (ang. Perceptual Evaluation of Speech Quality)

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.

3.5. 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.

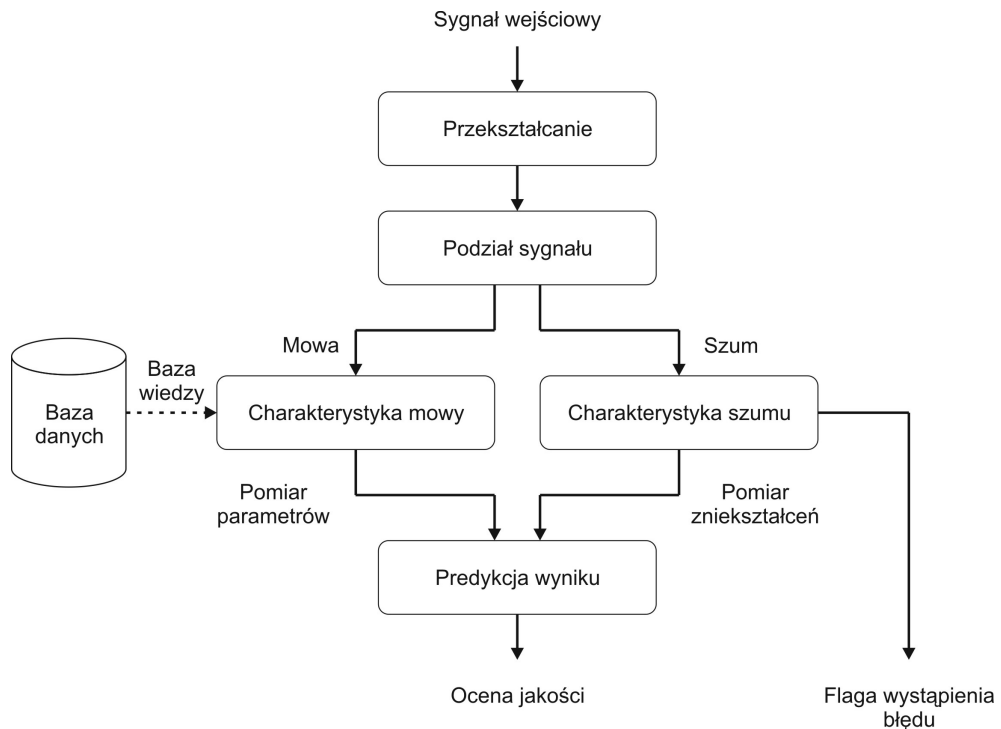
Metoda PSQM bazuje na analizie parametrów w dziedzinie czasowej i częstotliwościowej. Sygnał mowy jest oddzielany od szumu, a następnie oceniany z wykorzystaniem statystycznej bazy danych. Pomiar poziomu zniekształceń jest dokonywany na podstawie pomiaru szumów. Korelacja poziomu zniekształceń wraz

z parametrami mowy pozwala na predykcję wyniku jakości sygnału mowy w systemie. Wynik metody PSQM jest wykorzystywany do obliczenia wyniku skali MOS według wzoru (3.4) [48]:

$$MOS = \frac{4}{1 + e^{0,66PSQM-2,2}} + 1 \quad (3.4)$$

gdzie:

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



Rysunek 3.9. Algorytm działania metody PSQM (ang. Perceptual Single ended Objective Measure)

3.6. E-model

E-model został zaproponowany przez ETSI (Europejski Instytut Standardów Telekomunikacji) [24]. ITU wydało stosowne rekomendacje w 1998 roku [49][50]. 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 poprzez łącze telekomunikacyjne odbywa się w słuchawce telefonicznej słuchacza dla pasm częstotliwości od 300 Hz do 3400 Hz. Schemat struktury połączenia oraz mierzone parametry przedstawiono na rysunku 3.10. 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.

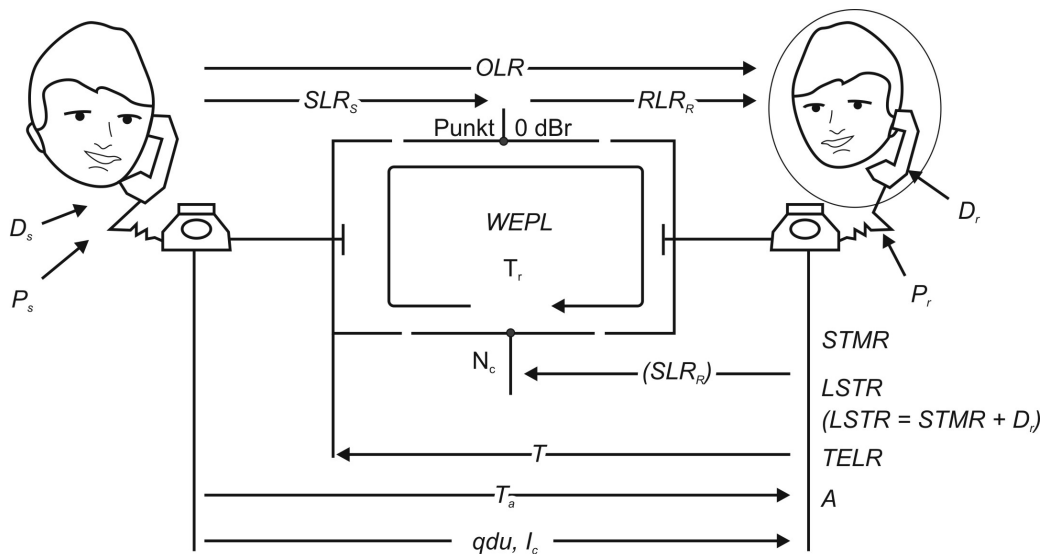
Współczynnik R wyrażany jest wzorem:

$$R = R_o - I_s - I_d - I_e + A \quad (3.5)$$

gdzie:

- R_o - podstawowy współczynnik sygnał/szum SNR (ang. Signal Noise Ratio) transmisji mowy w punkcie odniesienia 0 dBr (oznaczonym na rysunku 3.10, stanowiącym punkt odniesienia do wyznaczania pozostałych parametrów);
- 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_d ;
- I_e - współczynnik pogorszenia jakości transmitowanej mowy, spowodowanego zastosowaniem w łączy 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 [24].



Rysunek 3.10. Struktura pomiaru w E-model [24]

Na rysunku 3.10 występują następujące oznaczenia:

- OLR – całkowita tłumienność głośności (ang. Overall Loudness Rating),
- SLR_s – tłumienność głośności u nadawcy (ang. Sending Loudness Rating Sending),
- RLR_r – tłumienność głośności u odbiorcy (ang. Receiving Loudness Rating Receiving),
- $WEPL$ – ważonej straty efektu echa podczas transmisji w systemie (ang. Weighted Echo Path Lost)
- D_s - współczynnik czułości mikrotelefonu w miejscu mówcy (ang. D-factor Sending Side),

- D_r - współczynnik czułości mikrotelefonu w miejscu słuchacza (ang. D-factor Receiving Side),
- P_s - poziom hałasu otoczenia w miejscu mówcy (ang. Room Noise Sending Side),
- P_r - poziom hałasu otoczenia w miejscu słuchacza (ang. Room Noise Receiving Side),
- T_r - całkowite opóźnienie (ang. Round Trip Delay),
- T - średnie opóźnienie transmitowanego dźwięku w jedną stronę (ang. Mean One-Way Delay),
- T_a - bezwzględne opóźnienie od mówcy do słuchacza (ang. Absolute Delay),
- qdu – liczba urządzeń wprowadzających zniekształcenia kwantyzacji,
- I_e – współczynnik pogorszenia jakości transmitowanej mowy, spowodowanego zastosowaniem w łączu specjalnych urządzeń,
- STMR – tłumienność głośności efektu lokalnego echa w pozycji słuchacza (ang. Sidetone Masking Rating),
- LSTR – tłumienność głośności lokalnego echa słuchacza (ang. Listener Sidetone Rating),
- TELR – tłumienność głośności echa mówcy (ang. Talker Echo Loudness Rating).

Współczynnik R_o można wyrazić jako:

$$R_o = 15 - 1,5(SLR + N_o) \quad (3.6)$$

gdzie:

- SLR – tłumienność głośności (ang. Sending Loudness Rating) przy nadawaniu odniesiona do punktu odniesienia 0 dBr najbliższej położonego miejsca R,
- N_o – całkowite szумы odniesione do punktu odniesienia 0 dBr .

Całkowite szумы N_o otrzymuje się poprzez sumowanie mocy:

- szumu obwodu elektrycznego. Szum ten otrzymuje się poprzez dodanie mocy szumów generowanych przez różne źródła w łączu, przy czym wszystkie te szумы odniesione są do punktu 0 dBr,

- równoważnego szumu obwodu spowodowanego hałasem otoczenia w miejscu S,
- równoważnego szumu obwodu spowodowanego hałasem otoczenia w miejscu R,
- hałasu tła występującego w miejscu R.

Współczynnik I_S jest wyrażony przez:

$$I_S = I_{olr} + I_{st} + I_q \quad (3.7)$$

i jest sumą składników pogorszenia jakości transmitowanej mowy spowodowanej:

- I_{olr} – zbyt dużym poziomem mocy sygnału, tj. wtedy, gdy tłumienność OLR jest zbyt mała;
- I_{st} – niewłaściwą wartością efektu lokalnego echa;
- I_q – szumem kwantyzacji.

Współczynnik I_{olr} obliczany jest następująco:

$$I_{olr} = 20 \left[\left[1 + (X/8)^8 \right]^{1/8} - X/8 \right] \quad (3.8)$$

gdzie X określone przez równanie:

$$X = OLR + 0,2(64 + N_t) \quad (3.9)$$

gdzie:

- OLR – całkowita tłumienność głośności (ang. Overall Loudness Rating),

a N_t określamy tak:

$$N_t = N_o - RLR \quad (3.10)$$

gdzie:

- RLR – tłumienność głośności u odbiorcy (ang. Receiving Loudness Rating).

Współczynnik I_{st} można wyrazić następująco:

$$I_{st} = 10 \left[1 + \left\{ (STMR_o - 12) / 5 \right\}^{1/6} \right] + 46 \left[1 + \left\{ STMR_o / 23 \right\}^{1/10} \right] + 36 \quad (3.11)$$

przy czym $STMR_o$ oblicza się ze wzoru:

$$STMR_o = -10 \lg \left[10^{-STMR/10} + e^{-T/4} 10^{-TELR/10} \right] \quad (3.12)$$

gdzie:

- $STMR$ – tłumienność głośności efektu lokalnego echa w pozycji słuchacza (ang. SideTone Masking Rating),

- T – średnie opóźnienie transmitowanego dźwięku w jedną stronę (ang. Mean One-Way Delay),
- $TELR$ – tłumienność głośności echa mówcy

Należy zauważyć, że zmienna T reprezentuje tutaj opóźnienie transmisji w jednym kierunku podawane w [ms], natomiast $TELR$ reprezentuje tłumienność głośności echa mówcy podawane w [dB].

Współczynnik I_q definiujemy następująco:

$$I_q = 151\lg[1 + 10^Y] \quad (3.13)$$

gdzie:

$$Y = (R_o - 100)/15 + (46 - G)/10 \quad (3.14)$$

$$G = 1,07 + 0,258Q + 0,0602Q^2 \quad (3.15)$$

$$Q = 37 - 15\lg(qdu) \quad (3.16)$$

a qdu – liczba urządzeń wprowadzających zniekształcenia kwantyzacji.

Współczynnik I_d pogorszenia jakości transmitowanej mowy, spowodowany opóźnieniami w łączu, wynosi:

$$I_d = I_{dte} + I_{dle} + I_{dd} \quad (3.17)$$

jest więc sumą pogorszeń jakości transmitowanej mowy spowodowanych przez:

- I_{dte} – echo mówcy (wyznaczone poprzez tłumienność głośności echa mówcy $TELR$ i średnie opóźnienie echa – w jedną stronę),
- I_{dle} – echo słuchacza (wyznaczone poprzez tłumienie ścieżki echa $WEPL$ i opóźnienie dla echa – w obie strony),
- I_{dd} – bezwzględne opóźnienie występujące nawet po bardzo dobrym wyeliminowaniu echa.

Współczynnik I_{dle} obliczamy następująco:

$$I_{dle} = \left[(R_{oe} - R_e)/2 + \sqrt{(R_{oe} - R_e)^2/4 + 100} - 1 \right] (1 - e^{-T}) \quad (3.18)$$

gdzie:

$$R_{oe} = -1,5(N_o - RLR) \quad (3.19)$$

$$R_e = 80 + 2,5(TERV - 14) \quad (3.20)$$

$$TERV = TELR - 40 \lg \frac{1 + T/10}{1 + T/150} + 6e^{-0,3T^2} \quad (3.21)$$

Dla $T < 1$ ms echo mówcy powinno być rozważane jako efekt lokalnego echa i wówczas $I_{dte} = 0$. Powyższe zależności mają zastosowanie wtedy, kiedy efekt lokalny zawiera się w przedziale:

$$7dB < STMR < 15dB \quad (3.22)$$

Dla wartości $STMR$ mniejszych od 7 dB echo mówcy jest częściowo maskowane przez efekt lokalny, wówczas $TERV$ w równaniu (3.20) jest zastępowane przez $TERV_s$:

$$TERV_s = TERV + I_{st} / 2 \quad (3.23)$$

Przy większych wartościach efektu lokalnego echa (dla $STMR > 15dB$), echo mówcy jest bardziej zauważalne niż w normalnym zakresie wartości efektu lokalnego echa, wówczas w równaniu (3.17) I_{dle} należy zastąpić przez I_{dtes} :

$$I_{dtes} = \sqrt{I_{dte}^2 + I_{st}^2} \quad (3.24)$$

Współczynnik I_{dle} oblicza się następująco:

$$I_{dle} = (R_o - R_{le}) / 2 + \sqrt{(R_o - R_{le})^2 / 4 + 169} \quad (3.25)$$

gdzie:

$$R_{le} = 10,5(WEPL + 7)(T_r + 1)^{-0,25} \quad (3.26)$$

a T_r - całkowite opóźnienie.

Współczynnik I_{dd} dla bezwzględnego opóźnienia sygnału od mówcy do słuchacza, oznaczanego T_a , mniejszego od 100 ms wynosi 0, natomiast dla opóźnienia większego od 100 ms wyraża się wzorem:

$$I_{dd} = 25 \left\{ (1 + X^6)^{1/6} - 3(1 + [X/3]^6)^{1/6} + 2 \right\} \quad (3.27)$$

gdzie:

$$X = \frac{\lg(T_a / 100)}{\lg 2} \quad (3.28)$$

Współczynnik I_e jest stosowany w przypadku urządzeń przetwarzających mowę w sposób bardziej skomplikowany m.in. tam, gdzie zastosowano kodery o małej przepływności binarnej. Wpływ koderów w modelu jest uwzględniany przez współczynnik K . Jeśli w złożonym systemie występuje więcej połączonych koderów, wówczas ich wpływ na pogorszenie jakości mowy wyraża współczynnik I_e , który jest sumą poszczególnych wartości K dla stosowanych koderów:

$$I_e = \sum_{i=1}^n K_i \quad (3.29)$$

gdzie:

- n – ilość zastosowanych koderów dźwięku w systemie,
- K_i – współczynnik wyrażający wpływ i -tego kodera na jakość transmitowanego dźwięku.

W tabeli 3.2 przedstawiono wartości współczynników K dla różnych typów koderów dźwięku. Za pomocą współczynnika I_e można również modelować wpływ straty pakietów w zależności od zastosowanej metody kodowania dźwięku [51]. Współczynniki oczekiwania A zestawiono w tabeli 3.3.

Jak wspomniano jakość systemu w E-modelu wyrażana jest poprzez wartość współczynnika R , który zawiera się w przedziale od 0 do 100. Wartość 0 oznacza system o bardzo złej jakości, natomiast wartość 100 wskazuje na system o wysokiej jakości połączenia.

Współczynnik R może zostać przekształcony w inną miarę jakości transmisji, np.:

- procentowy wskaźnik użytkowników oceniających jakość połączenia jako dobrą lub bardzo dobrą GOB (ang. Good or Better) i słabą lub złą POW (ang. Poor or Worse);
- procentowy wskaźnik użytkowników przedwcześnie przerywających połączenia z powodu złej jakości TME (ang. TerMinating Early);
- wskaźnik oceny jakości MOS.

Tabela 3.2. Wartość współczynnika K [49]

Metoda kompresji dźwięku	Przepływność [kbit/s]	Wartość K	Oznaczenie
Adaptacyjna różnicowa modulacja impulsowo-kodowa (ang. ADPCM)	40	2	G.726, G.727
	32	7	G.721 (1988), G.726, G.727
	24	25	G.726, G.727
	16	50	G.726, G.727
Pobudzana kodem predykcja liniowa z małym opóźnieniem (ang. LD-CELP)	16	7	G.728
	12,8	20	
Sprzężona algebraiczna pobudzana kodem predykcja liniowa (ang. CS-CELP)	8	12	G.729
	8	13	G.729-A+Voice Activity Detection
Regularne pobudzanie impulsowe – predykcja długoterminowa (ang. RPE- LTP)	13	20	GSM 06.10, Full-rate
Algebraiczna pobudzana kodem predykcja liniowa (ang. ACELP)	12,2	3	GSM 06.60, Enhanced full-rate
Algebraiczna pobudzana kodem predykcja liniowa (ang. ACELP)	5,3	19	G.723.1
Modulacja wielo impulsowa maksymalnego prawdopodobieństwa kwantyzacji (ang. MP-MLQ)	6,3	15	G.723.1

Tabela 3.3. Wartość współczynnika A [49]

System komunikacji	Współczynnik A
System konwencjonalny – przewodowy	0
Sieci komórkowe w budynkach	5
Sieci mobilne w wolnej przestrzeni i w pojazdach	10
System połączeń do trudno dostępnych obszarów, np. wielostopniowe połączenia satelitarne	20

Skrócony opis modelu E pokazuje jego przydatność i uniwersalność do oceny różnorodnych projektów sieci i występujących w nich możliwych połączeniach [24]. Wykorzystując E-model można przewidywać jakość mowy w budowanym systemie, jednak konieczne jest odpowiednie dobranie współczynników do wzorów obliczeniowych m.in. K , A , $WELP$, $TELR$. Model ten jest dobrym modelem do badań w środowiskach testowych, gdzie w idealnych warunkach można dokonywać pomiarów

parametrów przedstawionych na rysunku numer 3.10. 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.

3.7. P.563

Metoda P.563 ta została zatwierdzona przez ITU-T w maju 2005 roku [52]. 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 [53]. 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 [52] 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 przedstawiono na rysunku 3.11. Obliczane są parametry 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 szumy. Ponadto odpowiada za wykrywanie powtarzających się próbek w dźwięku. Jednym z zadań elementów tej klasy jest porównanie sygnału z wygenerowanym sztucznie sygnałem odniesienia.

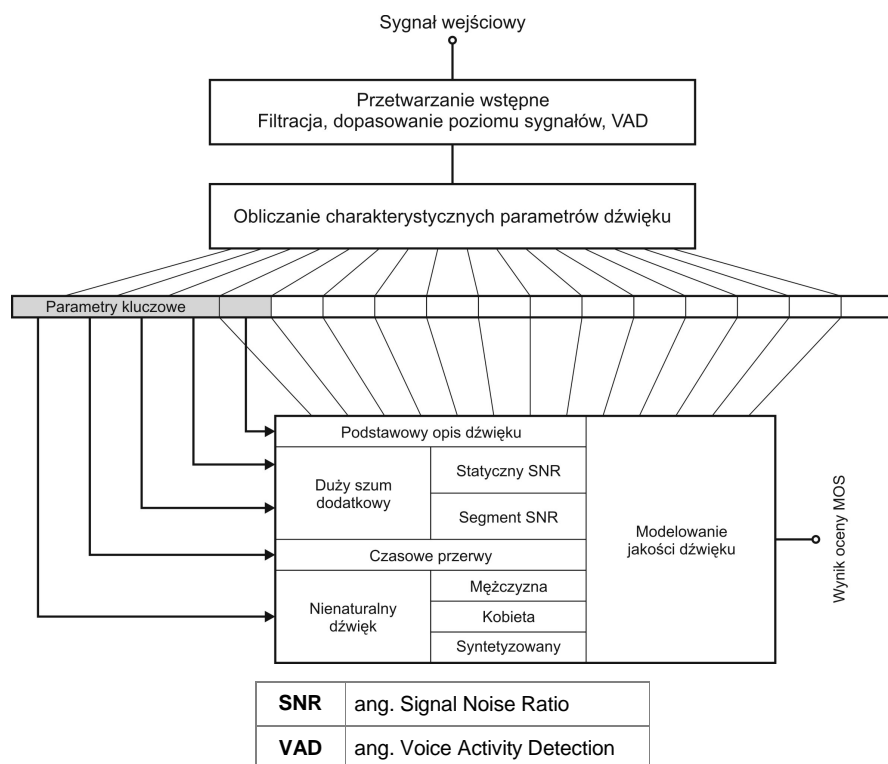
- analiza dodatkowego szumu:
 - statyczny pomiar SNR,
 - segment pomiaru SNR,

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.

Ponadto metoda P.563 stosuje deskryptory opisujące poziom dźwięku, aktywność mowy i jej zmiany. Wykorzystywane są funkcje wykrywania głosu VAD i analiza danych. Dla prawidłowego pomiaru jakości dźwięku sygnał wejściowy musi spełniać wymagania określone w tabeli 3.4.



Rysunek 3.11. Algorytm metody P.563

Tabela 3.4. Wymagane parametry sygnału dla prawidłowego pomiaru jakości dźwięku metodą P.563 [49]

Parametr dźwięku	Wartość
Częstotliwość próbkowania	8000 Hz
Rozdzielczość kwantowania	16 bit
Kodowanie	Modulacja impulsowo kodowa (PCM)
Minimalna długość mowy w badanej próbce	3 s
Maksymalna długość badanej próbki	20s
Minimalny procent aktywnej mowy	25
Maksymalny procent aktywnej mowy	75
Zakres poziomu aktywnej mowy (poziom odniesienia dla metody wynosi -26 dBov)	-36,0 do -16,0 dBov

Zalecenie dotyczące metody P.563 zostały opublikowane wraz z kodem źródłowym. Zakodowany algorytm cechuje się dość małymi wymaganiami sprzętowymi (tabela 3.5). Algorytm metody P.563 może być jeszcze optymalizowany, tak aby wymagania sprzętowe były jeszcze mniejsze [54].

Tabela 3.5. Wymagane sprzętowe metody P.563 [53]

Parametr	Wartość
Ilość możliwych operacji zmiennoprzecinkowych do wykonania przez procesor obliczeniowy na sekundę (ang. MFLOPS - Mega Floatingpoint Instruction Per Second) – dla pracy w czasie rzeczywistym	Minimum 175 max 262
Pamięć programu	Od 200 do 490 kB
Pamięć danych	Około 4MB dla 8 sekund dźwięku – zależy od długości badanej próbki

3.8. 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.

3.9. Podsumowanie przeglądu metod oceny jakości transmisji dźwięku

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

Metody ACR i DCR

Metody ACR i DCR wymagają pracy z reprezentatywną grupą słuchaczy. Nie nadają się przez to w ogóle do oceny jakości transmitowanego dźwięku pracujących systemów. Praktycznie można je stosować tylko w środowisku laboratoryjnym.

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 w 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 najszerszej stosowaną metodą w laboratoriach. Posiada jednak swoje wady.

Przeprowadzono testy metody PESQ, których wyniki były bardzo zaskakujące [55]. 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. Wyniki pomiaru jakości metodą PESQ zostały zestawione w tabeli 3.6.

Tabela 3.6. Wynik oceny jakości dźwięku metodą PESQ

Testowany fragment dźwięku	Wynik oceny PESQ
Wzorzec degradowana szumem białym	4,431
Wzorzec poddany filtracji niskich i wysokich częstotliwości	4,431

Metoda P.563

Wadą metody P.563 jest stosunkowo duży błąd w przypadku krótkotrwałej utraty dźwięku, co nie powinno być zauważone gdy długości trwania utraty dźwięku są mniejsze od 20 ms [7]. Przeprowadzone badania wykazały, że metoda w takich przypadkach przedstawia błędne wyniki oceny jakości, w związku z czym nie spełnia postawionych przed nią wymagań. Rysunki od 3.15 do 3.22 obrazują wyniki pomiarów metodami P.563 i PESQ próbek dźwięku, które zawierały nagrane fragmenty ludzkiej mowy z częstotliwością próbkowania 8kHz. Z nagrań usunięto fragmenty dźwięku z odstępem 200 i 1000 próbek przez określoną długość czasu (ilość próbek), czyli co 25 ms i co 125 ms. Dla odstępu 125 ms symulowano utratę dźwięku przez czas trwania wynoszący:

- 2,5 ms,
- 12,5 ms,
- 25 ms,
- 50 ms.

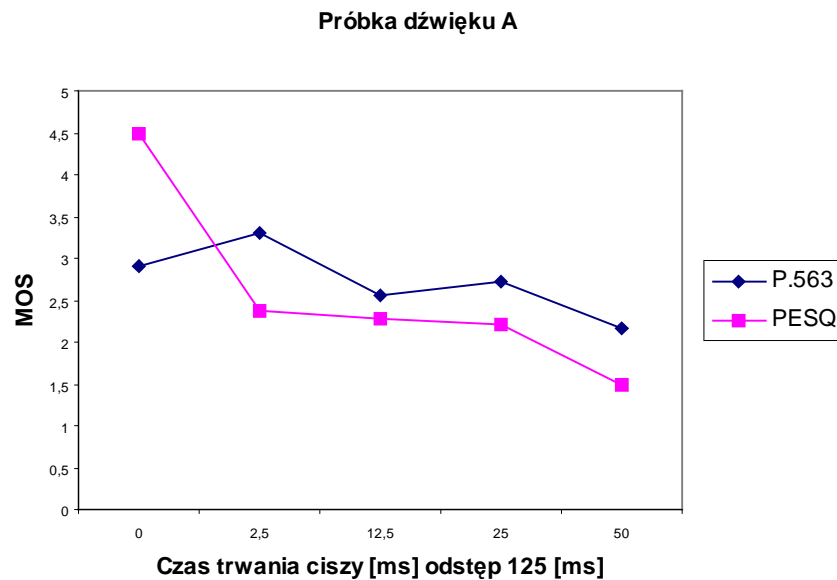
Natomiast dla odstępu 25 ms symulowano utratę dźwięku przez czas trwania wynoszący:

- 2,5 ms,
- 8,75 ms,
- 18,75 ms.

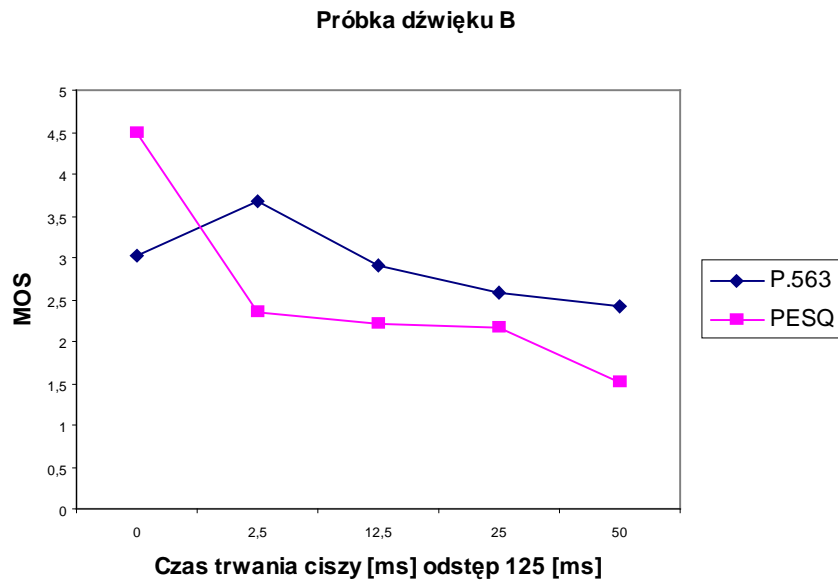
Można zauważyć, że ocena jakości transmitowanego dźwięku obliczona metodą P.563 jest czasami:

- większa od wyniku metody PESQ, co może mylić użytkowników systemu, że urządzenia działają poprawnie,
- mniejsza od wyniku metody PESQ co zaniża wynik pomiaru myląc użytkownika, że problem leży w badanym urządzeniu,
- zwiększana mimo wydłużenia czasu trwania straty sygnału.

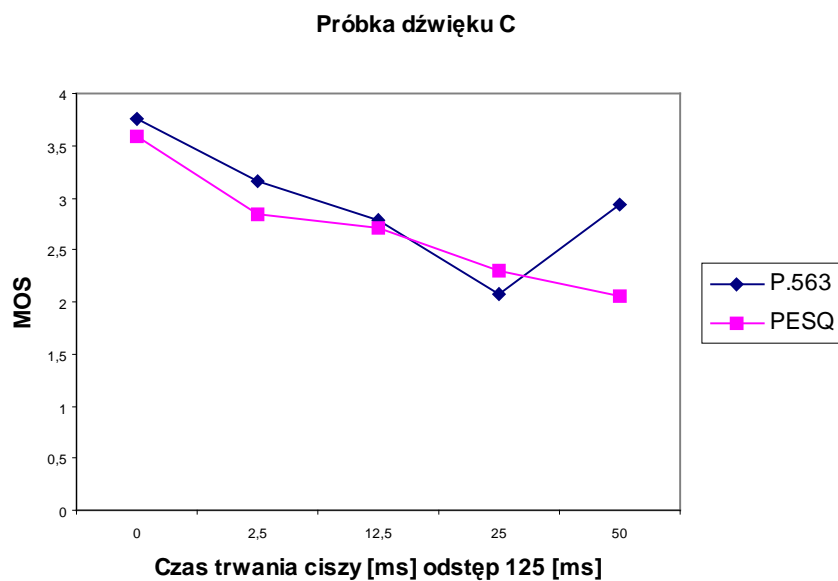
Na wykresach zostały zaprezentowane wyniki dla czterech losowo wybranych nagrań (próbek) dźwięku oznaczonych jako A, B, C i D.



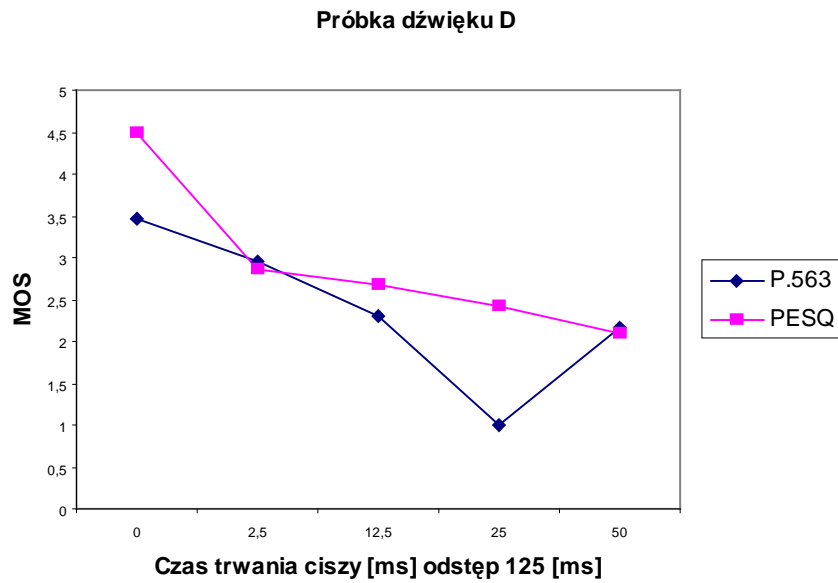
Rysunek 3.12. Wyniki oceny jakości dźwięku próbki A w funkcji utraty sygnału przy odstępach 125 [ms]



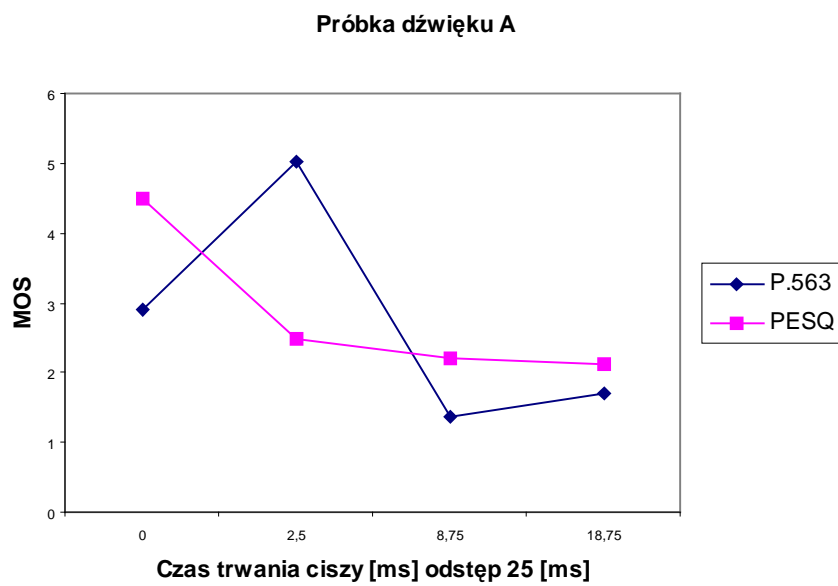
Rysunek 3.13. Wyniki oceny jakości dźwięku próbki B w funkcji utraty sygnału przy odstępie 125 [ms]



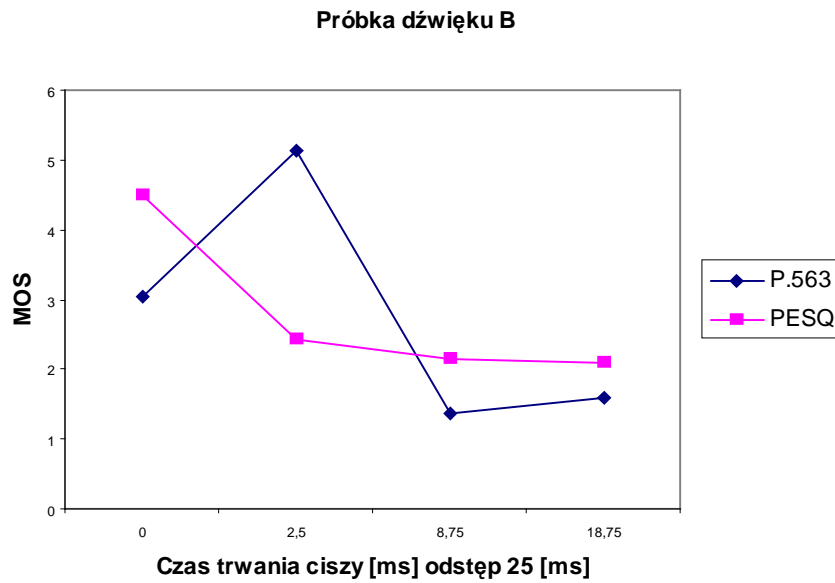
Rysunek 3.14. Wyniki oceny jakości dźwięku próbki C w funkcji utraty sygnału przy odstępie 125 [ms]



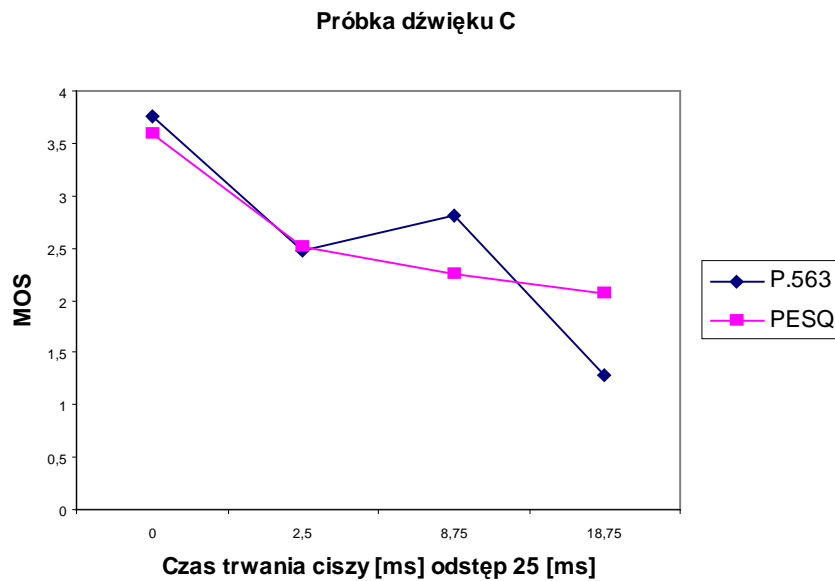
Rysunek 3.15. Wyniki oceny jakości dźwięku próbki D w funkcji utraty sygnału przy odstępnie 125 [ms]



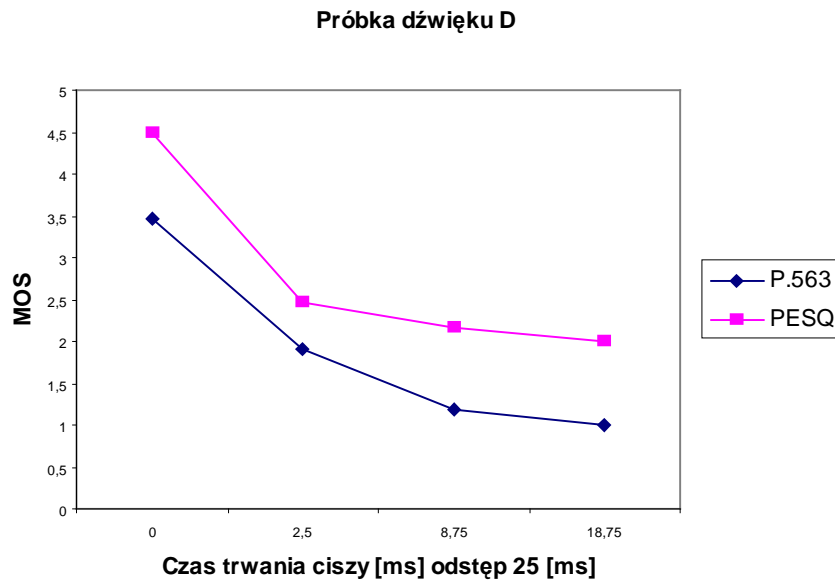
Rysunek 3.16. Wyniki oceny jakości dźwięku próbki A w funkcji utraty sygnału przy odstępnie 25 [ms]



Rysunek 3.17 Wyniki oceny jakości dźwięku próbki B w funkcji utraty sygnału przy odstępnie 25 [ms]



Rysunek 3.18 Wyniki oceny jakości dźwięku próbki C w funkcji utraty sygnału przy odstępnie 25 [ms]



Rysunek 3.19 Wyniki oceny jakości dźwięku próbki D w funkcji utraty sygnału przy odstępnie 25 [ms]

Jak widać na załączonych rysunkach 3.12 - 3.19, metoda P.563 również nie nadaje się do rzetelnego nadzoru sieci telefonicznej VoIP.

Podsumowując przegląd przedstawionych w rozdziale 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 opracowania algorytmu metody oceny jakości transmisji dźwięku dla sieci telefonicznej VoIP. Wyniki oceny jakości transmitowanego dźwięku uzyskane za pomocą opracowanej metody winny być zbliżone do wyników metody PESQ, która jest najszerzej stosowaną metodą w środowiskach laboratoryjnych.

Rozdział IV Proponowana metoda Oceny Jakości Transmitowanego Dźwięku (OJTD) w telefonii VoIP

4.1. Przygotowanie danych eksperymentalnych i przyjęcie metodologii badań

Do badań nad metodami oceny jakości dźwięku potrzebne są odpowiednio przygotowane próbki nagranej mowy, które będzie można wykorzystać wielokrotnie w celu porównania wyników jakie uzyskuje się w wybranych metodach. Uwzględniając cele pracy, do badań nad algorytmem oceny jakości transmisji dźwięku w sieci oferującej usługę VoIP należało przygotować próbki dźwięku z uwzględnieniem wpływu parametrów sieci IP na jakość transmitowanego dźwięku. Ponadto, aby algorytm spełniał swoją rolę w sieci heterogenicznej, należało uwzględnić parametry dodatkowe, które mogą zostać wygenerowane na punktach styku sieci IP z sieciami zewnętrznymi. Zdecydowano więc, że próbki dźwięku muszą uwzględniać wpływ takich parametrów jak:

- szumy w sygnale,
- echo,
- stosowany język,
- mówca,
- opóźnienia w sieci,
- rozmiar pakietu IP,
- strata pakietów IP,
- typ sposobu kodowania dźwięku,
- przerwy w sygnale dźwięku.

Proponowany algorytm powinien poprawnie rozpoznawać pogorszenie jakości dźwięku w wyniku wprowadzenia szumów w sygnale. Podobna sytuacja powinna wystąpić w momencie pojawienia się echa w przesyłanym sygnale. Algorytm metody musi uwzględniać różne języki jakie będą stosowane w przeprowadzanych rozmowach oraz różną płęć i wiek mówców. Ponadto powinien pracować poprawnie z różną wielkością

pakietów IP transmitowanych poprzez sieć. Zwiększający się poziom straty pakietów powinien być zauważany i skutkować pogorszeniem oceny jakości rozmowy. Zarówno opóźnienia występujące w sieci jak i sposób kodowania dźwięku powinny wpływać na wynik zgodnie z teorią przedstawioną w poprzednich rozdziałach. Przerwy w sygnale winny wpływać na pogorszenie jakości sygnału w momencie gdy są one zauważalne [7]. Dlatego celem przygotowania próbek wykonano kilka eksperymentów.

4.1.1. Badanie wpływu szumów w sygnale, stosowanego języka i mowy

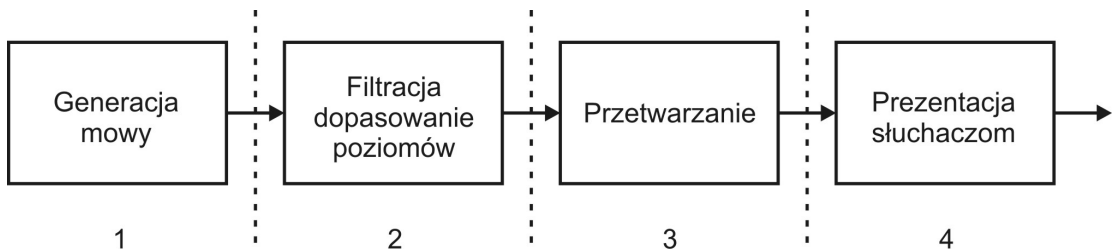
Jako próbki wzorcowe do badań z uwzględnieniem szumów, wpływu języka i mowy w niniejszej pracy wybrano bazę danych próbek dostarczanych przez ITU [56]. Dostarczone przez ITU próbki zostały przygotowane przez niezależne laboratoria. Laboratoria te miały za zadanie przygotowanie danych służących do badania metod pod kątem zjawisk zachodzących w kanale transmisyjnym. Każde laboratorium przygotowywało próbki do badania wpływów innych parametrów. Próbki te zostały poddane przez ITU standaryzacji.

Próbki zostały opracowane i nagrane do postaci cyfrowej, dostarczane w postaci plików zawierających wartości chwilową dźwięku próbkowaną z częstotliwością 16kHz i 16 bitowej rozdzielczości zapisu wartości. Ponadto ustalono, że maksymalna długość sygnału wzorcowego będzie wynosiła 12 sekund. Nagrane przez ITU dźwięki zostały przygotowane w trzech zbiorach:

- Pierwszy zbiór danych służy do badania wpływu różnego rodzaju kodowania mowy oraz współpracy systemów z różnymi standardami kodowania dźwięku,
- Drugi zbiór danych służy do testowania wpływu efektów środowiska takich jak: szumy, w tym szum pojazdów, szum występujący w biurze, na ulicy, w pokoju, czy muzyka w tle rozmowy,
- natomiast trzeci zbiór danych zawiera próbki dźwięku, które powstały w wyniku degradacji kanału transmisyjnego poprzez błędy, takie jak np. przypadkowa utrata warunków transmisyjnych.

Schemat przygotowania próbek został zobrazowany na rysunku 4.1. W pierwszym kroku nagrywano osobę mówiącą. Następnie dokonywano filtracji i dopasowania

poziomu sygnału. W trzecim kroku daną próbkę przetwarzano, tak aby uzyskać zamierzony efekty degradacji. Takie próbki rejestrowane były w postaci cyfrowej, a następnie prezentowane słuchaczom.



Rysunek 4.1. Przygotowanie próbek przez ITU [56]

W sumie wygenerowano 424 próbki wejściowe, które poddano przetworzeniu. Jako mówców w eksperymentach wykorzystano cztery osoby – dwie kobiety i dwóch mężczyzn. Przygotowane dane zawierają próbki w sześciu językach.

Tabela 4.1. Zestawienie próbek ITU P.Sup 23

Parametr	Wartość
Ilość próbek mowy:	
- pierwsza partia próbek	188
- druga partia próbek	28
- trzecia partia próbek	208
Język:	
- pierwsza partia próbek	francuski, japoński, amerykański angielski, niemiecki, norweski
- druga partia próbek	francuski, japoński, amerykański angielski, niemiecki
- trzecia partia próbek	francuski, japoński, amerykański angielski, niemiecki, włoski

Do badań wpływu:

- szumów w sygnale,
- stosowanego języka,
- mówcy,

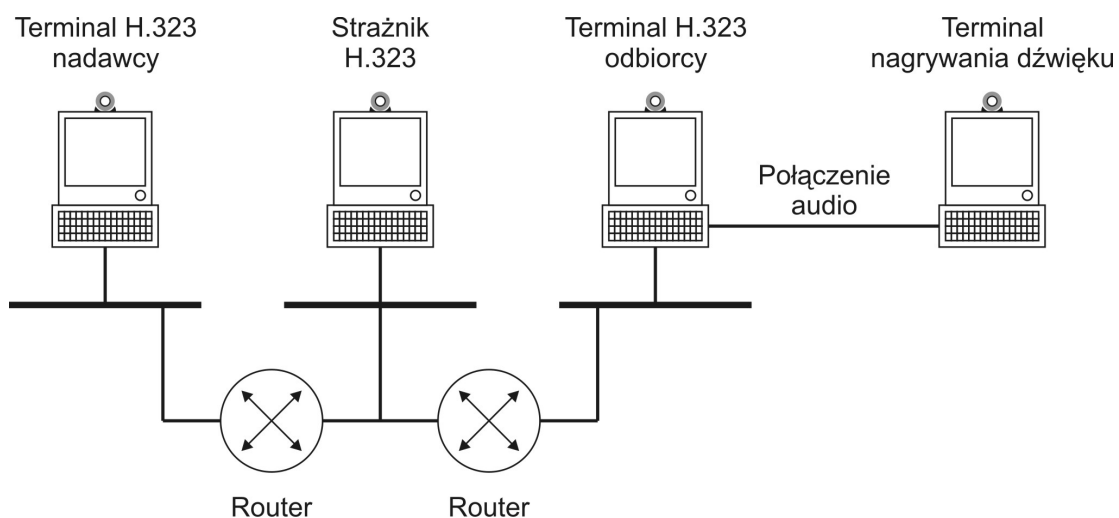
wybrano 1024 próbki dźwięku w sześciu językach przygotowane w każdym wariantcie przez dwie kobiety i dwóch mężczyzn.

4.1.2. Badania wpływu sposobu kodowania dźwięku na jakość jego transmisji

W celu przygotowania próbek do badania wpływu systemu kodowania dźwięku na jakość transmitowanego dźwięku przeprowadzono następujący eksperyment. Zbudowano sieć oferującą usługę VoIP składającą się z dwóch terminali H323. Były to komputery klasy PC (ang. Personal Computer) z systemem operacyjnym Windows XP i oprogramowaniem terminala H323 – programem OpenPhone pochodzącym z pakietu OpenH323 pobranym ze strony www.openh323.org. Oprogramowanie to jest zgodne z oprogramowaniem Netmeeting, które poprzez twórców urządzeń pracujących w standardzie sygnalizacji H323 uważane jest jako wzorzec implementacji protokołu H323. Z terminala H323 nadawcy dokonywano połączeń głosowych w technologii VoIP do terminala odbiorcy. Jako strażnika H323 użyto oprogramowania GNU Gatekeeper [57]. Terminal odbiorcy poprzez swoje wyjście audio karty dźwiękowej był połączony z terminalem, którego zadaniem było nagranie odtwarzanego dźwięku. Podczas tego eksperymentu dokonywano połączeń, w których jako sygnał wejściowy w terminalu nadawcy wykorzystywano próbki otrzymane z ITU. Rozmowy były wykonywane w czterech standardach kodowania dźwięku:

- G.711 A-law,
- G.711 μ -law,
- G.728,
- GSM 06-10.

W sumie do eksperymentu wybrano 1044 próbki przygotowane przez ITU, co przy czterech rozpatrywanych wariantach kodowania wygenerowało 4176 próbek do badania wpływu sposobu kodowania dźwięku na jakość jego transmisji.



Rysunek 4.2. Sieć oferująca usługę VoIP do przygotowania próbek eksperymentalnych

4.1.3. Badanie wpływu przerw w sygnale dźwięku na jakość jego transmisji

Jednym z decydujących czynników wpływających na jakość transmisji dźwięku są zniekształcenia wynikające z przerw w sygnale, dlatego też zbadano wpływ utraty sygnału w transmisji. W tym celu próbki zostały podane degradacji według dwóch algorytmów.

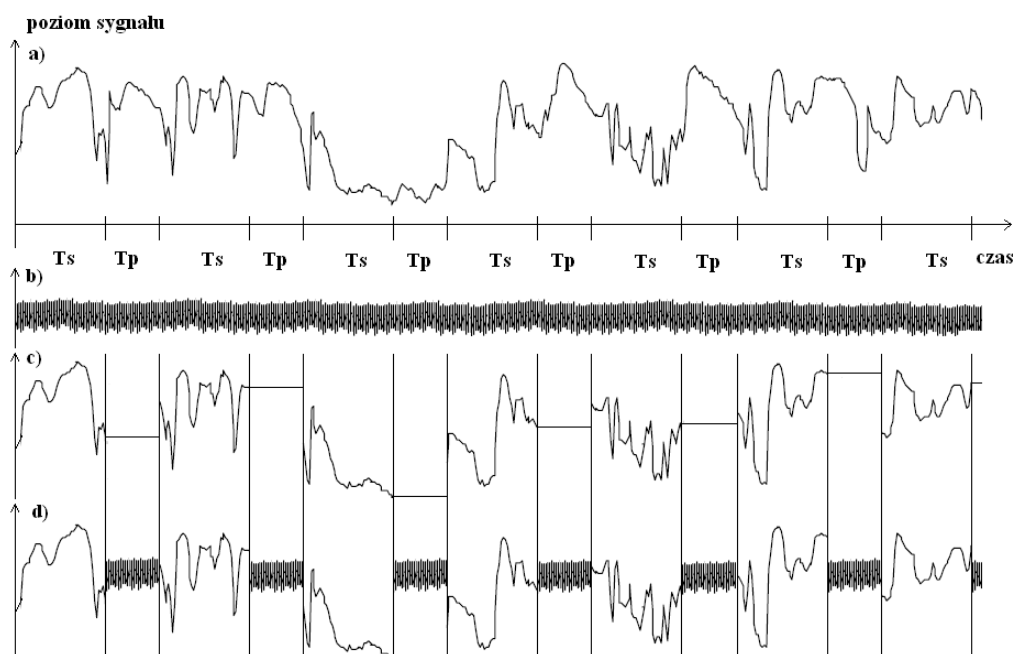
W pierwszym podejściu w plikach źródłowych usuwano wartości próbek dźwięku (x) poprzez zastąpienie ich wartością poprzednią:

$$x[t] = x[t-1] \quad (4.1)$$

przez założony czas trwania utraty dźwięku.

W drugim podejściu w plikach źródłowych usuwano wartości postaci cyfrowej dźwięku (x) zastępując je wartościami dźwięku nagranych w pomieszczeniu testowym (N), co ma symulować szum występujący w badanym pomieszczeniu:

$$x[t] = N[t-1] \quad (4.2)$$



Rysunek 4.3. Wykres sposobów obcinania dźwięku w dziedzinie czasu –
 a) przetwarzana próbka, b) sygnał nagranych szumu, c) sygnał obcięty przez zastąpienie poprzednią wartością, d) sygnał obcięty przez zastąpienie sygnałem szumu (T_s - czas trwania sygnału, T_p - czasu trwania przerwy)

Rysunek 4.3 ilustruje sposób usuwania fragmentów dźwięku według omówionych wyżej sposobów. Przez ustalony czas trwania sygnału T_s próbki dźwięku nie były poddawane degradacji. Natomiast w okresie czasu trwania przerwy T_p próbki dźwięku były poddawane degradacji bądź poprzez zastąpienie wartością poprzednią bądź poprzez zastąpienie próbką sygnału szumu. Zastosowane wartości czasów trwania sygnału T_s i przerwy T_p zostały zestawione w tabeli 4.2.

Przyjęto cztery różne wartości czasu trwania sygnału T_s . Dla poszczególnych czasów trwania sygnału wybrano od czterech do sześciu wariantów czasu trwania przerwy w sygnale T_p . W sumie uzyskano 20 wariantów kombinacji czasu trwania sygnału i przerwy. Dla każdej kombinacji wprowadzono dwa wspomniane warianty usuwania dźwięku, co w sumie podwaja ilość próbek dźwięku. Dodatkowo należy przebadать próbkę dźwięku bez jej degradacji, co daje jeden kolejny wariant badanego dźwięku. Podsumowując ilość kombinacji otrzymano 41 wariantów dźwięku do przebadania.

Tabela 4.2. Zestawienie czasów obcinania sygnału

Tp [ms]	Ts [ms]			
	25	62,5	125	250
2,5	X	X	X	X
6,25	X	X	X	X
12,5	X	X	X	X
18,75	X	X	X	X
31,25		X	X	X
62,5				X

Poddając wybrane próbki dźwięku wariantom degradacji otrzymano 43848 próbek do badań wpływu utraty dźwięku.

W literaturze można znaleźć różne sposoby symulacji utraty pakietów [58][59]. Podejście zastosowane w niniejszej pracy reprezentuje praktyczną możliwość wystąpienia awarii w skomplikowanej sieci oferującej usługę VoIP. Wartości czasu zostały wybrane z uwzględnieniem faktu, iż ucho ludzkie nie rozpoznaje w dźwięku przerw krótszych od 20 ms [7]. Pojawia się pytanie skąd w sieci oferującej usługę VoIP, gdzie pakiety RTP przenoszą próbki danych o czasach od 20 ms, może pojawiać się przerwa o czasie krótszym. Otóż takie zjawisko może wystąpić w centralach analogowych lub na bramach dostępu do sieci – czyli w miejscu styku sieci oferującej usługę VoIP z innym systemem. Przyczyną takiego błędu może być problem z przetwornikiem analogowo-cyfrowym bądź przeciążenie samego urządzenia. Ponieważ sieci telefoniczne VoIP nie są sieciami zamkniętymi, należy to zagadnienie również uwzględnić przy planowaniu metody monitorowania stanu pracy sieci.

4.1.4. Badanie wpływu rozmiaru i strat pakietów IP na jakość transmisji dźwięku

W celu przygotowania próbek do badania wpływu rozmiaru i strat pakietów IP na jakość transmitowanego dźwięku przeprowadzono eksperyment. Do eksperymentu

wykorzystano sieć oferującą usługę VoIP, której schemat został przedstawiony na rysunku 4.2. W zbudowanej sieci przeprowadzano rozmowy telefoniczne w technologii VoIP z użyciem dwóch systemów kodowania dźwięku:

- G.711 - μ -law,
- G.711 - a-law.

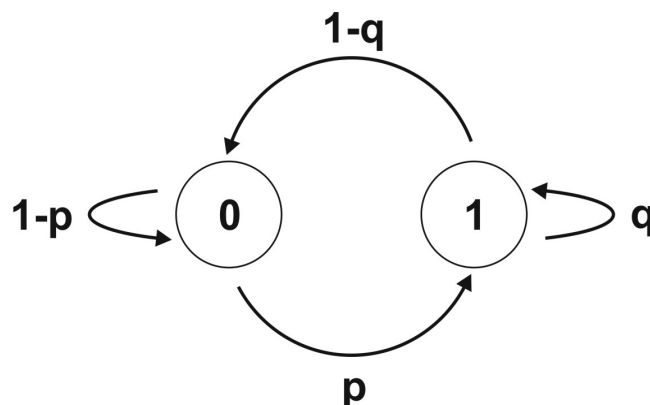
Terminale VoIP wysyłały pakiety:

- co 10 ms,
- co 20 ms,
- co 30 ms.

W efekcie przy kodowaniu G.711 uzyskiwano trzy różne wielkości pakietów IP transmitujących dźwięk:

- 120 bajtów,
- 200 bajtów,
- 300 bajtów.

Dwustanowy proces Markowa - model Gilberta – jest jednym z najczęściej stosowanych mechanizmów modelowania strat w torze transmisyjnym [60][61][62][63][64]. Dlatego na jednym z routerów zaimplementowano model Gilberta do generacji losowej straty pakietów. Algorytm działania modelu został przedstawiony na rysunku 4.4. Stan 0 oznacza, że pakiet będzie transmitowany poprawnie. Natomiast stan 1 oznacza, że pakiet zostanie usunięty z transmisji. Zmienna p oznacza prawdopodobieństwo utraty pakietu gdy poprzedni pakiet był przesłany poprawnie, a zmienna q oznacza prawdopodobieństwo utraty pakietu gdy poprzedni pakiet był przesłany poprawnie.



Rysunek 4.4. Dwustanowy proces Markowa

Na rysunku 4.4 występują następujące oznaczenia:

- 0 –pakiet transmitowany poprawnie,
- 1 –pakiet do usunięcia z transmisji,
- p – prawdopodobieństwo utraty pakietu gdy poprzedni pakiet był przesłany poprawnie,
- q – prawdopodobieństwo utraty pakietu gdy poprzedni pakiet został również utracony.

Niech π_0 i π_1 oznaczają prawdopodobieństwa stanów odpowiednio 0 i 1 badanego modelu:

$$\pi_0 = P(X = 0) \quad (4.3)$$

$$\pi_1 = P(X = 1) \quad (4.4)$$

Można wyznaczyć je z zależności:

$$\begin{bmatrix} 1-p & q \\ p & 1-q \end{bmatrix} \begin{bmatrix} \pi_0 \\ \pi_1 \end{bmatrix} = \begin{bmatrix} \pi_0 \\ \pi_1 \end{bmatrix} \quad (4.5)$$

$$\pi_0 + \pi_1 = 1 \quad (4.6)$$

Wyznaczone prawdopodobieństwa dane są wzorami:

$$\pi_0 = \frac{1-q}{p+1-q} \quad (4.7)$$

$$\pi_1 = \frac{p}{p+1-q} \quad (4.8)$$

Prawdopodobieństwo p_k wystąpienia postępującej po sobie straty pakietów o długości indeksu k wyrażone jest wzorem:

$$p_k = (1-q)^{k-1} q \quad (4.9)$$

Prawdopodobieństwo straty pakietów w modelu Gilberta ma więc rozkład geometryczny. Jeśli oznaczymy jako Z zmienną losową opisującą rozkład strat pakietów w transmisji, to wartość oczekiwana (średnia) strat przedstawia się następująco:

$$E[Z] = \sum_{k=1}^{\infty} kp_k = \sum_{k=1}^{\infty} k(1-q)q^{k-1} = \frac{1}{1-q} \quad (4.10)$$

Prawdopodobieństwa p i q można obliczyć również ze statystyki ruchu. Jeśli oznaczymy jako o_0 ilość dostarczonych pakietów, natomiast jako o_i oznacza liczbę strat pakietów o długości i , przy czym $i=1,2,\dots,n-1$, gdzie $n-1$ jest wartością najdłuższego wystąpienia ciągłej straty pakietów, to [60]:

$$p = \frac{\sum_{i=1}^{n-1} o_i}{o_0} \quad (4.11)$$

$$q = 1 - \frac{\sum_{i=2}^{n-1} o_i(i-1)}{\sum_{i=1}^{n-1} o_i i} \quad (4.12)$$

Gdy $p = q$, to model Gilberta redukuje się do modelu Bernoulliego.

Jako wartości straty pakietów do przeprowadzania badań przyjęto straty na poziomie 0, 10, 20, 30 i 40%. Do badań wybrano 1024 próbki dźwięku pochodzące z ITU. Dla trzech wariantów wielkości pakietów IP przeprowadzono eksperyment z utratą pakietów w pięciu wartościach procentowej straty. Uwzględniając dwa warianty kodowania dźwięku uzyskano 30720 próbek do przeprowadzenia badań wpływu wielkości i strat pakietów IP.

4.1.5. Badanie wpływu echa na jakość transmisji dźwięku

Do badania wpływu echa na wyniki oceny uzyskiwane poprzez metody oceny jakości transmitowanego dźwięku próbki otrzymane z ITU poddano degradacji w programie GoldWave v.5.25 uzyskanym ze strony www.goldwave.com. Jako źródła wzorcowe dźwięku wybrano 156 próbek po 52 sentencje nagrane przez dwie kobiety i jednego mężczyznę. Za pomocą programu do wzorcowej próbki dźwięku wprowadzono echo z opóźnieniem 0,5 sekundy i poziomem -6dB w stosunku do sygnału oryginalnego. Przygotowane w ten sposób próbki dźwięku miały za zadanie ukazanie prawidłowości działania algorytmów oceny jakości dźwięku.

4.1.6. Badanie wpływu opóźnień w sieci IP na jakość transmisji dźwięku

W celu przygotowania próbek do badania wpływu opóźnień pakietów IP na jakość transmitowanego dźwięku przeprowadzono eksperyment, w którym wykorzystano sieć oferującą usługę VoIP przedstawioną na rysunku 4.2. W takiej sieci zestawiano połączenia między terminalami H323 i wysyłało próbki dźwięku dostarczone przez ITU. W sieci za pomocą zastosowanych routerów wprowadzono dwa rodzaje opóźnień:

- opóźnienie o stałej wartości,
- opóźnienie o zmiennej wartości generowanej co piąty pakiet IP.

Dla opóźnienia o stałej wartości wygenerowano próbki dźwięku dla trzech wariantów opóźnień:

- 20 ms,
- 100 ms,
- 250 ms.

Natomiast przy generacji zmiennej wartości opóźnień, co piąty pakiet IP został opóźniony w stosunku do poprzedniego o:

- 20 ms,
- 50 ms,
- 150 ms.

Celem generacji zmiennej wartości opóźnienia było zasymulowanie efektu jitter, polegającego na tym, iż pakiety IP docierają do odbiorcy z różnym opóźnieniem.

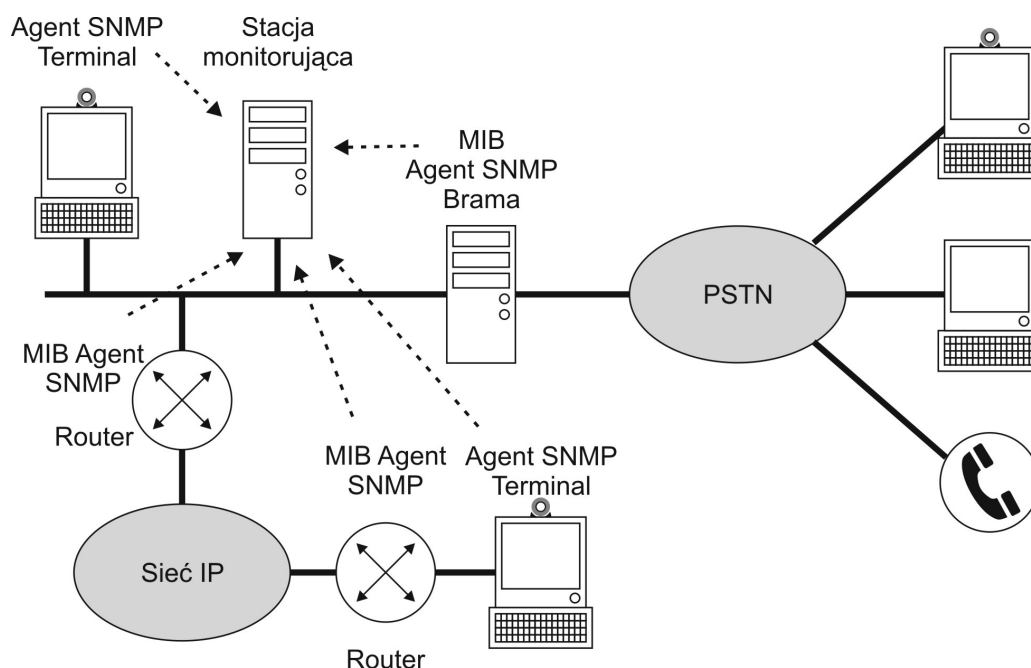
Jako źródła wzorcowe wybrano próbki nagrane przez dwóch mężczyzn i dwie kobiety. Każda z osób dostarczyła 52 próbki dźwięku. Uwzględniając warianty opóźnień oraz jeden wariant, gdzie opóźnienia nie występowały, uzyskano 832 próbki dźwięku dla opóźnienia o stałej wartości oraz próbki dźwięku dla opóźnienia o zmiennej wartości.

4.1.7. Wybór metody do porównania wyników metody *Oceny Jakości Transmitowanego Dźwięku*

Aby określić poprawność działania metody Oceny Jakości Transmitowanego Dźwięku (OJTD – Ocena Jakości Transmitowanego Dźwięku), należy dokonać jej sprawdzenia. W tym celu konieczna jest znajomość wartości poprawnej oceny jakości badanego sygnału dźwięku. Najczęściej używaną metodą oceny jakości w środowiskach laboratoryjnych jest metoda PESQ. Ponieważ sposób przeprowadzania badań pozwala na zastosowanie metody PESQ, właśnie ją wybrano do określania wartości poprawnych ocen jakości badanych próbek uzyskanych w wyniku przeprowadzonych eksperymentów.

4.2. Proponowana metoda OJTD w telefonii VoIP

Do celów monitorowania stanu pracy sieci oferującej usługę VoIP potrzebna jest metoda pomiaru jakości transmitowanego dźwięku pozwalająca na ocenę jakości rozmowy w punkcie sieci oferującej usługę VoIP gdzie dostępne są informacje dotyczące połączenia: sygnalizacja połączenia oraz transmitowany dźwięk. Oznacza to, że zachodzi konieczność opracowania algorytmu pomiaru jakości dźwięku, który będzie można zastosować w urządzeniach aktywnych sieci – routerach, bramach do innych systemów, czy terminalach abonenckich. Wyniki oceny jakości powinny być dostępne w stacji nadzorowania sieci. Efekt ten można uzyskać poprzez zastosowanie protokołu SNMP [65]. SNMP pracuje z bazami danych MIB (ang. Management Information Base) i pozwala na zgłaszanie zdarzeń zaistniałych w jego bazach – np. zmiany parametrów oznaczających wystąpienie anomalii w sieci IP. Zasada działania polega na pracy agentów i menadżerów. Zadaniem agentów SNMP jest monitorowanie stanu pracy sieci IP i zgłaszanie zdarzeń polegających na wystąpieniu w urządzeniach sieciowych pewnych ściśle określonych warunków. Natomiast zadaniem menadżerów jest zbieranie informacji od agentów i jej późniejsze przetworzenie.



Rysunek 4.5. Przykładowy schemat monitorowania sieci z użyciem agentów SNMP

Rozwiązania monitorowania sieci IP pozwalają na zbieranie danych odnośnie ilości i charakteru ruchu sieciowego [66][67][68][69]. Dostarczają one administratorowi sieci informacji na temat m.in.:

- obciążenia routerów,
- zasad działania polityk QoS,
- opóźnień pakietów w sieci,
- strat pakietów w sieci,
- natłoków w sieci i działania mechanizmów zapobiegania natłokom RED (ang. Random Early Detection),
- jitter'a,
- godzin występowania zwiększenia natężenia ruchu,
- anomalii w pracy sieci pod kątem włamań do sieci IP [70] [71].

Bazy MIB protokołu SNMP są bazami rozszerzalnymi. Pola w bazach opisywane są identyfikatorami OID (ang. Object Identifier) [65]. Metoda badająca jakość transmisji sygnału może więc swój wynik oceny MOS umieszczać w bazie MIB. Agent SNMP wykrywając pogorszenie jakości – spadek oceny – może wygenerować

informacje (zdarzenie) do systemu nadzorowania siecią oferującą usługę VoIP o wystąpieniu anomalii w jej pracy.

Zgodnie z postawionym celem w niniejszej pracy dla OJTD w punkcie sieci oferującej usługę VoIP gdzie dostępne są wszystkie konieczne informacje o połączeniu, w dowolnym czasie, potrzebna jest taka metoda oceny jakości transmisji dźwięku, która nie ingeruje w badany system. Aby spełnić to kryterium należy opracować metodę należącą do jednej z klas:

- metoda pracująca na jednym końcu systemu (ang. single-ended),
- metoda bazująca na parametrach systemu,
- metoda hybrydowa.

W metodach bazujących na parametrach systemu, takich jak E-model, należy przygotować szereg parametrów, które stanowią dane wejściowe dla obliczeń związanych z oceną jakości transmisji dźwięku takich jak:

- właściwości używanych terminali,
- rodzaj elementów połączeń,
- rodzaj elementów transmisji,
- konfiguracja połączeń,
- parametry mowy i słuchu człowieka,
- tłumienia transmisji mowy między mówcą i słuchaczem,
- tłumienia linii abonenckiej,
- tłumienia obwodów 4-przewodowych,
- efekt lokalnego echa mówcy i słuchacza,
- echo,
- czas transmisji – opóźnienia,
- szумы i zniekształcenia kwantyzacji,
- przesłuchy,
- ilość i rodzaje zmian systemów kodowania dźwięku.

Ze względu na szeroki wachlarz parametrów koniecznych do wyznaczania oraz heterogeniczny charakter sieci IP, a co za tym idzie, sieci telefonicznej VoIP, klasę metod OJTD bazujących na parametrach systemu należało odrzucić. Obecnie nie ma możliwości określenia, jaki terminal będzie posiadał użytkownik oraz jaki będzie

jego typ połączenia do sieci. Nie jest więc możliwe zebranie, ani oszacowanie, wartości parametrów wejściowych, niezbędnych w tego typu klasie metod.

Metoda hybrydowa poza parametrami systemu oraz sygnałem na wyjściu układu mierzonego, wymaga sygnału referencyjnego. Ze względu na wspomniany charakter sieci oferującej usługę VoIP również ta klasa metod oceny jakości transmisji dźwięku nie jest uwzględniana w dalszych rozważaniach ponieważ nie ma możliwości dostarczenia sygnału wzorcowego w rzeczywistych warunkach pracy sieci z usługą VoIP.

Jedyną klasą metod OJTD, która może być brana pod uwagę, jest klasa metod oceny transmisji dźwięku pracujących na jednym końcu systemu. Metody te nie ingerują w badaną sieć oraz nie wymagają sygnału odniesienia. Spełniają więc wymagania współczesnej sieci oferującej usługę VoIP, ponieważ mogą być wykorzystane w punkcie sieci z usługą VoIP, gdzie mają dostęp do wszystkich danych odnośnie połączenia, monitorując jakość połączeń.

Jednym z określonych w pracy warunków algorytmu oceny jakości połączeń w sieci oferującej usługę VoIP, była możliwość jego użycia w zdefiniowanym wcześniej miejscu sieci. Dlatego algorytm metody OJTD musi być prosty obliczeniowo, co oznacza, że:

- musi działać w czasie rzeczywistym,
- nie może posiadać dużych wymagań na pamięć operacyjną (liczonych w MB),
- nie może posiadać dużych wymagań na moc obliczeniową procesora (liczonych w milionach operacji na sekundę).

Jako duże wymagania rozumie się tutaj wymagania przekraczające możliwości najłagodniejszych urządzeń w sieci oferującej usługę VoIP – terminali abonenckich. Proponowany algorytm powinien być możliwy do implementacji w urządzeniach telefonii VoIP. przy wykorzystaniu ich wolnych zasobów sprzętowych.

Ponadto jednym z celów pracy było zaproponowanie takiego algorytmu oceny jakości rozmów przeprowadzanych w sieci oferującej usługę VoIP, który umożliwi przedstawianie wyniku oceny jakości transmitowanego dźwięku w skali MOS. Analizując warunki pracy algorytmu oraz znane i opracowane już metody oceny jakości, jako metodę wyjściową do badań wybrano metodę P.563 proponowaną przez ITU.

Metoda P.563 została zaprojektowana do monitorowania stanu pracy sieci w czasie rzeczywistym stosując interfejs cyfrowy bądź analogowy [52]. Ponadto zaletą tej metody są stosunkowo małe wymagania sprzętowe [53] oraz algorytm opierający się na ekstrakcji parametrów i predykcji wyniku na podstawie tych parametrów. Takie podejście umożliwia dodawanie szeregu parametrów, które z parametrami zdefiniowanymi przez ITU pozwolą oceniać jakość transmisji dźwięku z dokładnością zbliżoną do metody PESQ.

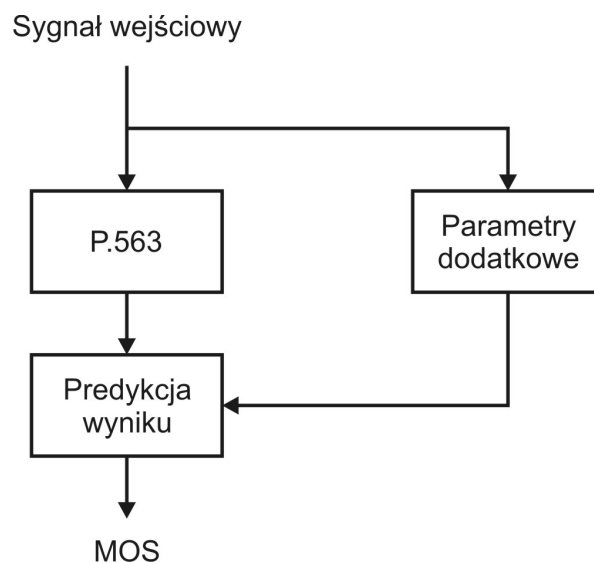
Zjawiskami, które muszą być rozważone w pomiarze jakości transmitowanego dźwięku jest echo i opóźnienia w transmitowanym sygnale, ponieważ te czynniki mają znaczący wpływ na odczucia użytkowników systemu dotyczące jakości transmisji. Obecnie opracowano wiele algorytmów eliminacji echa, które są stosowane w procesorach sygnałowych DSP pracujących w układach wejściowych terminali abonenckich. W literaturze można znaleźć dokładne algorytmy ich implementacji jak i same idee [40] [72] [73] [74] [75] [76] [77].

Kolejnym aspektem wymagającym uwagi to kwestia opóźnień w sieci oferującej usługę VoIP. Otóż jak wcześniej wspomniano, opóźnienia są jednym z kluczowych czynników pogorszenia jakości usługi związanej z transmisją dźwięku w sieci. W sieci IP pakiety mogą być transmitowane różnymi drogami, co za tym idzie, mogą być dostarczane z różnymi wartościami czasu opóźnień. Sprawą administratorów sieci IP jest zapewnienie łączności, balansu wykorzystania i obciążenia łączy i routerów. Za pomocą protokołów routingu takich jak OSPF można określać, które łącze jest preferowane [78] [79] [80]. Mechanizmy QoS pozwalają administratorom zapewnić opóźnienia dla wybranych pakietów na poziomie ustalonym na etapie projektowania sieci i monitorowanym w trakcie jej pracy [81] [82] [83]. Jednakże w rzeczywistej sieci IP przy jej obecnej złożoności zapewnienie wszystkich parametrów gwarantujących odpowiedni poziom transmisji pakietów IP jest zadaniem niezmiernie trudnym.

Podsumowując proponowana metoda OJTD ma być metodą pracującą w miejscu sieci gdzie dostępne są informacje o samym połączeniu jak i próbki transmitowanego dźwięku. Przyjmując za wzorzec metodę P.563 pozostaje do rozwiązania problem opisany w poprzednim rozdziale, a polegający na błędnej ocenie jakości dźwięku przy występowaniu zaniku sygnału. Należy więc do parametrów, które metoda OJTD

wyznacza, dodać parametr, bądź ich grupę, odpowiadający za zaniki sygnału oraz włączyć go do wzoru predykcji wyniku pomiaru jakości opracowanego przez ITU.

Do wykrywania przerw w sygnale można pośrednio wykorzystać algorytmy detekcji mowy VAD. Zadaniem tych algorytmów jest wykrywanie dźwięku będącego ludzką mową. Algorytmy VAD w ogólności wykorzystują szereg skomplikowanych obliczeniowo kryteriów wykrywania mowy np. w kanale z dużymi wartościami szumów [84] [85] [86], i mogą być implementowane zarówno w procesorach DSP [87] jak i w postaci sieci neuronowych [88]. Algorytmy VAD mają jednak pewną wadę. Ich sposób działania wymaga bądź dużej mocy obliczeniowej, bądź dedykowanych procesorów DSP. Implementacja w postaci sieci neuronowych w urządzeniach sieci oferujących usługę VoIP jest trudna. Ponadto wykrywanie bardzo krótkich przerw w sygnale wymaga innego podejścia do rozpatrywanego sygnału niż dotychczas spotykane w algorytmach VAD. Dlatego też do wykrywania przerw w sygnale postanowiono wykorzystać zmodyfikowane kryterium mocy przenoszonej przez dane próbki sygnału.



Rysunek 4.6. Schemat działania proponowanej metody OJTD

4.3. Algorytm proponowanej metody OJTD

W celu określenia wyniku ilościowego i jakościowego zaniku sygnału zaproponowano dwa parametry dodatkowe – $EB1$ i $EB2$, będących pewnym ilorazem związanych z sygnałem współczynników sygnał-cisza.

Zaproponowany algorytm metody OJTD z uwzględnieniem tych parametrów jest następujący:

1. Obliczenie średniej mocy sygnału w próbce badanego dźwięku i wskaźnika do klasyfikacji sygnał-cisza.
2. Budowa tablicy przerw sygnału – klasyfikacja sygnał-cisza
3. Określenie wartości współczynników zmian sygnał-cisza W_i
4. Wyznaczenie parametrów $EB1$ i $EB2$
5. Wyznaczanie wyniku oceny jakości transmitowanego dźwięku z wykorzystaniem macierzy odwzorowania

Badana próbka dźwięku o czasie trwania minimalnie ośmiu sekund zapisana w postaci cyfrowej z próbkowaniem o częstotliwości 8kHz i rozdzielczości 16 bitów dzielona jest na przedziały o długości 20 wartości, nazywanych dalej ramkami, czyli o czasie trwania 2,5 ms. Dla tak dobranych ramek obliczany jest pierwiastek mocy sygnału według wzoru:

$$P_j = \frac{\sqrt{\sum_{i=1}^{20} x_i^2}}{20} \quad (4.13)$$

gdzie:

- P_j - pierwiastek mocy niesiony przez j -tą ramkę,
- x_i - wartość sygnału w i -tej chwili czasowej w ramce.

Jako kryterium rozpoznania, czy w danej ramce znajduje się zakodowany sygnał dźwięku, czy jest to cisza, wybrano pierwiastek wartości średniej P_j pierwiastków mocy niesionych przez poszczególne ramki sygnału:

$$P_{sr} = \frac{\sqrt{\sum_{j=1}^{LiczbaRamek} P_j}}{LiczbaRamek} \quad (4.14)$$

gdzie:

- *LiczbaRamek* - liczba ramek w badanej próbie sygnału,
- P_{sr} - wskaźnik średniego pierwiastka mocy w sygnale, wykorzystywany do późniejszego porównania.

Na podstawie pierwiastka mocy sygnału w ramce oraz porównania ze wskaźnikiem średniej wartości dla ramek P_{sr} budowana jest tablica *BreakMap* określająca czy dana ramka reprezentuje ciszę czy dźwięk. Zastosowano następującą zasadę, iż:

- a) jeżeli pierwiastek mocy sygnału P_j zawartego w j-tej ramce jest mniejszy niż:

$$P_j < PercentPower P_{sr} \quad (4.15)$$

gdzie:

- *PercentPower* - określa ile procent średniego pierwiastka mocy należy użyć w podejmowaniu decyzji o zakwalifikowaniu sygnału jako cisza, w badaniach został ustalony na 70%,

to sygnał reprezentuje ciszę, lub też

- b) jeżeli wartość występujących po sobie wartości próbek sygnału niesionego przez ramkę zmienia się mniej niż określony procent w całej ramce sygnału:

$$\hat{i \in \langle 1,19 \rangle} \frac{x_{i+1} - x_i}{x_i} < PercentChange \quad (4.16)$$

gdzie:

- *PercentChange* - określa wskaźnik procentowy minimalnych zmian występujących po sobie wartości próbek sygnału, w badaniach został ustalony na 15%,
- x_i - wartość sygnału w i-tej chwili czasowej w ramce.

to sygnał również reprezentuje ciszę.

Określenie wartości współczynnika zmian sygnał-cisza W_i

Za pomocą opracowanej tablicy *BreakMap* określa się wartości współczynników W_i .

Współczynniki te określają liczbę zmian:

- sygnał-cisza,
- cisza-sygnał,

przy czym wartość i definiuje przez jaką ilość ramek sygnału badanego określa się występowanie wartości – sygnał bądź cisza.

Dla wyznaczenia W_i można zbudować tablicę pomocniczą *BreakMapPomoc*, której ilość wartości będzie wynosiła ilość wartości tablicy *BreakMap* podzieloną przez wartość i . Wartości tablicy *BreakMapPomoc* obliczane są na podstawie i wartości tablicy *BreakMap* według wzoru:

$$BreakMapPomoc[j] = \begin{cases} 1 & \text{dla } I_{signal} > I_{break} \\ 0 & \text{dla } I_{signal} < I_{break} \end{cases} \quad (4.17)$$

gdzie:

- j - indeks wartości, przyjmuje wartości od 1 do ilości wartości tablicy *BreakMap* podzielonej przez wartość i ,
- I_{signal} - ilość ramek zakwalifikowanych w tablicy *BreakMap* jako sygnał dla j -tego zbioru wartości,
- I_{break} - ilość ramek zakwalifikowanych w tablicy *BreakMap* jako cisza dla j -tego zbioru wartości.

Tablica *BreakMapPomoc* jest więc tablicą reprezentującą przewagę ramek zakwalifikowanych jako ramki sygnału w stosunku do ramek zakwalifikowanych jako ramki ciszy w zbiorach zawierających i ramek.

Wartości współczynników W_i oblicza się zliczając w tablicy *BreakMapPomoc* ilość zmian typu $1 \rightarrow 0$ i $0 \rightarrow 1$.

Algorytm wyznaczania współczynników W_i może być łatwo modyfikowany i optymalizowany pod kątem szybkości działania algorytmu.

Wyznaczenie parametrów $EB1$ i $EB2$.

Wartości parametrów oblicza się według wzorów:

$$EB1 = \frac{1000W_{10}}{W_{20}W_{30}W_{40}} \quad (4.18)$$

$$EB2 = \frac{1000W_{50}}{W_{100}W_{150}W_{200}} \quad (4.19)$$

Ze wzorów (4.8) i (4.9) wynika, iż do wyznaczenia parametrów $EB1$ i $EB2$ potrzebne są współczynniki W_i dla $i \in \{10,20,30,40,50,100,150,200\}$. Liczba 1000 służy do przeskalowania wartości parametru.

Wyznaczanie wyniku OJTD

Po wyznaczeniu parametrów $EB1$ i $EB2$ można dokonać predykcji wyniku jakości transmisji dźwięku. Wynik algorytmu proponowanej metody OJTD, oznaczony jako MOS_{EB} , obliczany jest ze wzoru:

$$MOS_{EB} = MOS_{P.563} [MacierzOdwzorowania]_{i,j,k} \quad (4.20)$$

gdzie:

- $MOS_{P.563}$ jest wynikiem pomiaru metody P.563,
- $[MacierzOdwzorowania]_{i,j,k}$ jest trójwymiarową macierzą służącą do odwzorowywania wyników P.563 tak, aby uwzględniały zaniki sygnału w badanej próbce dźwięku za pomocą parametrów $EB1$ i $EB2$, przy minimalizacji ilości błędnych pomiarów w odniesieniu do wyników ocen jakości uzyskanych za pomocą metody PESQ.

Indeksy macierzy $[MacierzOdwzorowania]_{i,j,k}$:

- i – określa komórkę macierzy na podstawie wyniku $MOS_{P.563}$,
- j – określa komórkę macierzy na podstawie parametru $EB1$,
- k – określa komórkę macierzy na podstawie parametru $EB2$.

Zakresy wartości indeksów i , j , k macierzy odwzorowania wyników $[MacierzOdwzorowania]_{i,j,k}$ zostały wyznaczone w drodze przeprowadzonych badań i zostały przedstawione w tabelach 4.3 - 4.5.

Tabela 4.3. Wartość indeksu i w zależności od wyniku $MOS_{p.563}$

Indeks i	Przedział wartości $MOS_{p.563}$
0	$< 1;1,25)$
1	$< 1,25;1,5)$
2	$< 1,5;1,75)$
3	$< 1,75;2)$
4	$< 2;2,25)$
5	$< 2,25;2,5)$
6	$< 2,5;2,75)$
7	$< 2,75;3)$
8	$< 3;3,25)$
9	$< 3,25;3,5)$
10	$< 3,5;3,75)$
11	$< 3,75;4)$
12	$< 4;4,25)$
13	$< 4,25;4,5)$
14	$< 4,5;5)$

Tabela 4.4. Wartość indeksu j w zależności od wyniku $EB1$

Indeks j	Przedział wartości $EB1$
0	$(0;1)$
1	$< 1;2)$
2	$< 2;3)$
3	$< 3;4)$
4	$< 4;5)$
5	$< 5;10)$
6	$< 10;100)$
7	$< 100;1000)$
8	$< 1000;10000)$

Tabela 4.5. Wartość indeksu k w zależności od wyniku $EB2$

Indeks j	Przedział wartości $EB2$
0	(0;100)
1	< 100;1000)
2	< 1000;10000)
3	< 10000;100000)

Wartości poszczególnych komórek macierzy $[MacierzOdwzorowania]_{i,j,k}$ zostały wyznaczone na podstawie wyników badań przy wykorzystaniu przygotowanych próbek dźwięków dla różnych wariantów degradacji sygnału przesyłanego. Wartości dla poszczególnych komórek macierzy oblicza się ze wzoru:

$$MacierzOdwzorowania_{i,j,k} = \frac{\sum_{l=1}^N \frac{MOS_{PESQ}}{MOS_{P.563}}}{N} \quad (4.21)$$

gdzie:

- $MacierzOdwzorowania_{i,j,k}$ - oznacza komórkę macierzy $[MacierzOdwzorowania]_{i,j,k}$ o indeksie i, j, k ,
- MOS_{PESQ} - jest wynikiem oceny jakości dźwięku próbki za pomocą metody PESQ, która została zakwalifikowana za pomocą parametrów $MOS_{P.563}, EB1, EB2$ do wyznaczania wartości w macierzy $[MacierzOdwzorowania]_{i,j,k}$ o indeksie i, j, k ,
- $MOS_{P.563}$ - jest wynikiem oceny jakości dźwięku próbki za pomocą metody P.563, która została zakwalifikowana za pomocą parametrów $MOS_{P.563}, EB1, EB2$ do wyznaczania wartości w macierzy $[MacierzOdwzorowania]_{i,j,k}$ o indeksie i, j, k .

4.4. Wyniki uzyskiwane za pomocą zaproponowanej metody OJTD

Macierz $[MacierzOdwzorowania]_{i,j,k}$ wykorzystywana podczas ekspertów powstała w wyniku prac nad próbkami przygotowanymi do opracowywania algorytmu

metody. Próbki dźwięku zostały poddane degradacji w różnorodny sposób uwzględniający wszelkie rodzaje degradacji sygnału, które mogą wystąpić w sieci oferującej usługę VoIP. Następnie tak przygotowane próbki dźwięku poddano ocenie jakości metodą PESQ, P.563 oraz wyznaczono dla nich parametry $EB1$, $EB2$. Następnie wyznaczono macierz $[MacierzOdwzorowania]_{i,j,k}$, która posłużyła do oceny proponowanego algorytmu OJTD

Zaproponowany algorytm OJTD przy uwzględnieniu wpływu:

- szumów w sygnale,
- języka,
- mówcy,
- sposobu kodowania dźwięku,
- przerw w sygnale,
- rozmiaru i strat pakietów IP,
- echa,
- opóźnień stałego i jitter'a,

na ocenę jakości transmitowanego dźwięku pozwala na uzyskanie wyników bardziej zbliżonych do wyników oceny jakości metodą PESQ, niż metoda P.563.

4.4.1. Wpływ standardów kodowania dźwięku na jego jakość

Jednym z przeprowadzonych testów oceny metody OJTD był pomiar jakości próbek dźwięku, które nie podlegały przetworzeniu polegającemu na jego degradacji, a jedynie na transmisji z użyciem różnych sposobów kodowania dźwięku. Przebadano 4176 próbek dźwięku przygotowanych poprzez nagranie przeprowadzonych rozmów z użyciem czterech sposobów kodowania dźwięku. Średni błąd bezwzględny MOS metody OJTD i P.563 był taki sam, jednakże zmniejszyło się odchylenie standardowe tegoż błędu.

Analizując wyniki przedstawione przez metodę P.563 oraz przez zaproponowany algorytm OJTD okazało się, że:

- 66% próbek nie zostało poddanych korekcie przez proponowany algorytm, wyniki uzyskane były jednakowe jak za pomocą metody P.563,
- 34% próbek zostało poddanych analizie przez proponowaną metodę OJTD i wyniki oceny w większości przypadków zostały poprawione.

Tabela 4.6. Wyniki badania różnych standardów kodowania dźwięku

Metoda	Średni błąd bezwzględny MOS	Odchylenie standardowe błędu bezwzględnego MOS
P.563	1,23	0,81
OJTD	1,23	0,67

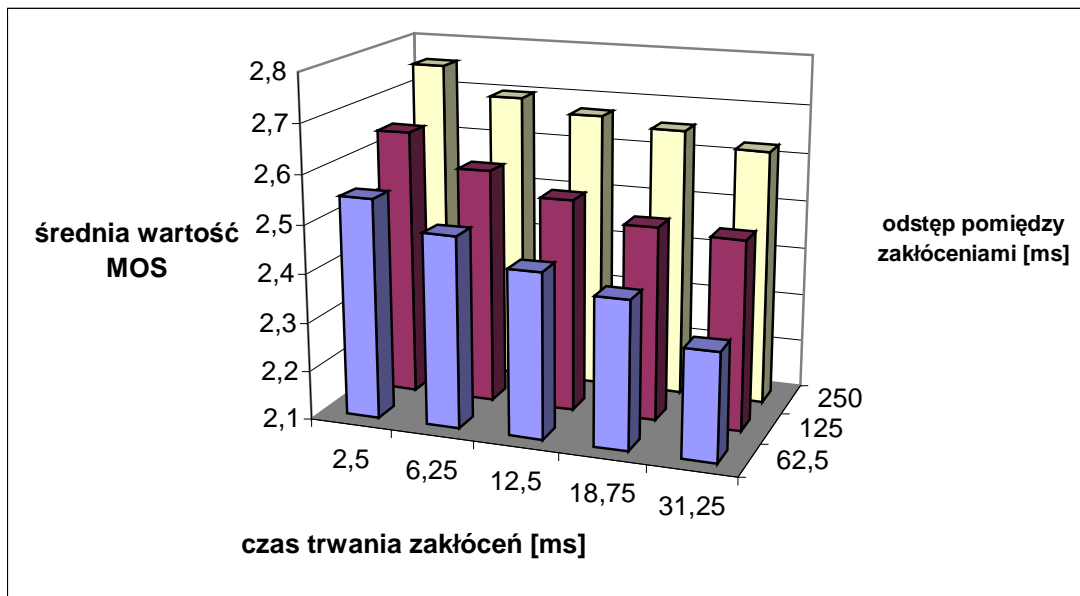
4.4.2. Wpływ przerw w sygnale dźwięku na jego jakość

Kolejnym istotnym testem był test oceny próbek dźwięku z wprowadzonymi celowo przerwami w sygnale. Z testu oceny tak zdegradowanych próbek wynika, że proponowany algorytm OJTD dokonuje predykcji wyników znacznie lepiej niż algorytm P.563 (wady tej metody oceny jakości dźwięku zostały ukazane w rozdziale III). Wynik testu potwierdził poprawność działania zaproponowanego algorytmu OJTD. Oznacza to, że metodę OJTD można zaimplementować w warunkach rzeczywistych, gdzie degradacje polegające na utracie pakietów występują względnie często, a metoda OJTD przedstawi wynik o wiele bardziej wiarygodny niż metody stosowane dotychczas. Średni błąd bezwzględny MOS metody OJTD był mniejszy niż metody P.563, jednocześnie zmniejszyło się odchylenie standardowe tegoż błędu

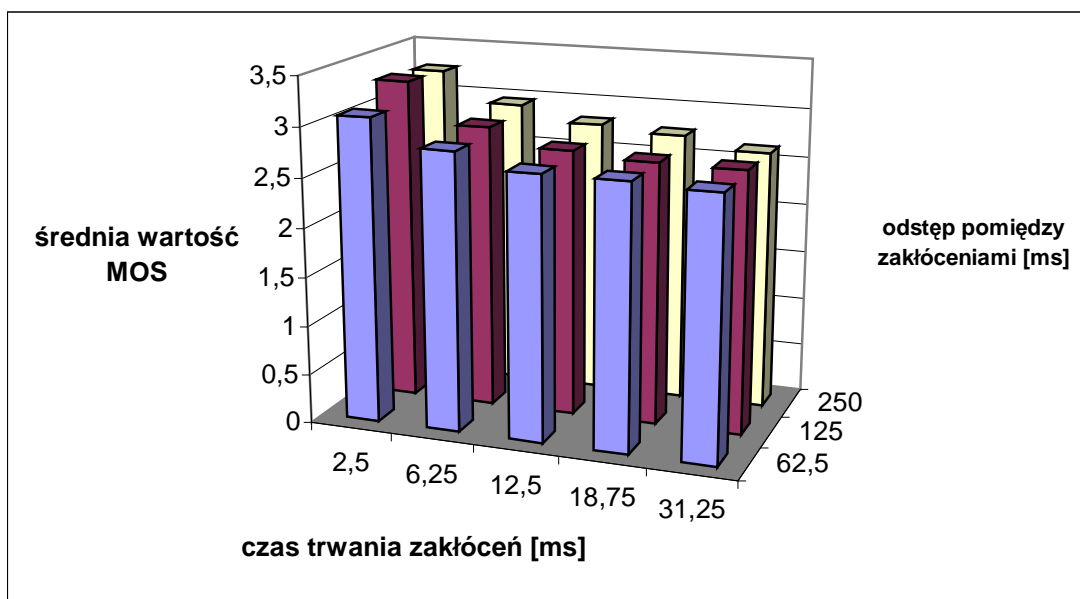
Na rysunkach 4.7-9 przedstawiono wykresy średniej wartości MOS funkcji zniekształceń polegających na utracie sygnału przez określony czas dla metod P.563, PESQ i proponowanej OJTD. Można zauważyć, że metoda OJTD bardziej przybliżyła metodę PESQ. Wartości średnie współczynnika MOS dla metody OJTD są bliższe wartościom uzyskiwanym za pomocą metody PESQ niż metody P.563 w 100% badanych przypadków.

Tabela 4.7. Wyniki badania próbek z przerwami w sygnale dźwięku

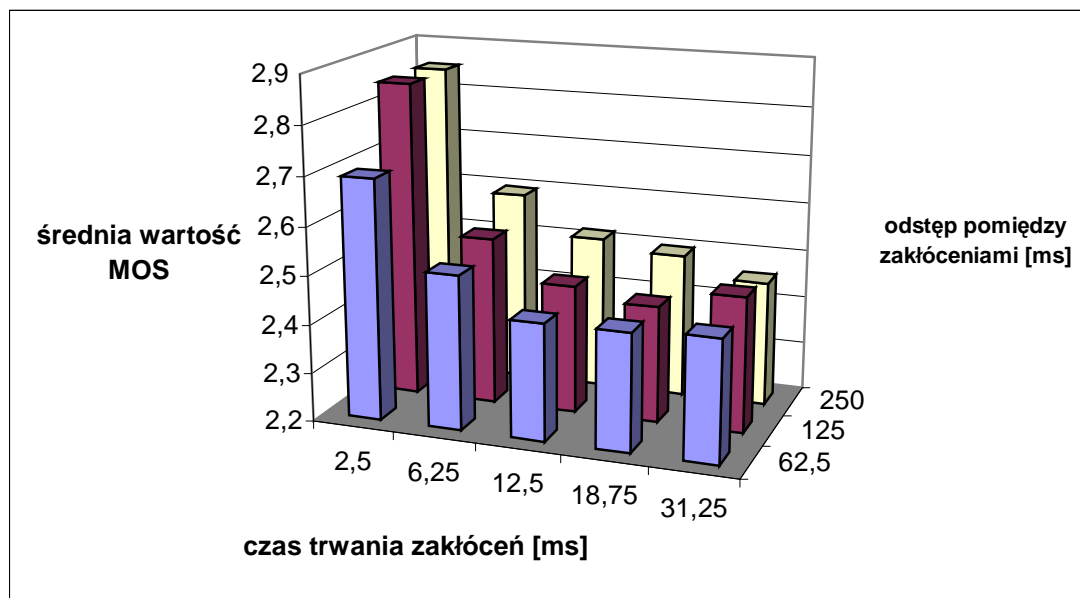
Odstęp między zakłóceniami [ms]	Czas trwania zakłócenia [ms]	Średni błąd bezwzględny MOS		Odchylenie standardowe błędy bezwzględnego MOS	
		P.563	OJTD	P.563	OJTD
25	6,25	0,52	0,41	0,37	0,29
25	12,5	0,53	0,43	0,36	0,33
25	18,75	0,60	0,57	0,44	0,39
62,5	2,5	0,70	0,49	0,48	0,46
62,5	25	0,54	0,38	0,36	0,30
62,5	12,5	0,48	0,37	0,37	0,29
62,5	18,75	0,51	0,41	0,39	0,32
62,5	31,25	0,54	0,49	0,41	0,36
125	2,5	0,80	0,56	0,60	0,57
125	6,25	0,50	0,33	0,34	0,28
125	12,5	0,44	0,31	0,33	0,26
125	18,75	0,44	0,33	0,33	0,27
125	31,25	0,44	0,35	0,33	0,27
250	2,5	0,66	0,48	0,46	0,45
250	6,25	0,47	0,31	0,32	0,26
250	12,5	0,43	0,29	0,31	0,24
250	18,75	0,43	0,29	0,30	0,24
250	31,25	0,45	0,29	0,30	0,24
250	62,5	0,45	0,31	0,30	0,24
250	125	0,46	0,32	0,31	0,25
Wartości średnie		0,52	0,39	0,37	0,32



Rysunek 4.7. Średnia wartość MOS dla metody PESQ



Rysunek 4.8. Średnia wartość MOS dla metody P.563



Rysunek 4.9. Średnia wartość MOS dla proponowanej metody OJTD

4.4.3. Wpływ echa na jakość transmitowanego dźwięku

Test wpływu echa na wynik pomiaru wykazał, że zarówno w metodzie P.563 jak i proponowanym algorytmie OJTD należało by dopracować wpływ echa na wyniki pomiarów. Jednakże średni błąd bezwzględny MOS metody OJTD był znacznie (przeszło 30%) mniejszy niż metody P.563, jednocześnie zmniejszyło się odchylenie standardowe tegoż błędu

Tabela 4.8. Wyniki badania wpływu echa na jakość transmitowanego dźwięku

Metoda	Średni błąd bezwzględny MOS	Odchylenie standardowe błędu bezwzględnego MOS
P.563	1,36	0,80
OJTD	0,95	0,62

4.4.4. Wpływ opóźnień pakietów IP na jakość transmitowanego dźwięku

Badania wpływu opóźnień na wyniki pomiarów wykazały, że zarówno metoda P.563 jak i proponowany algorytm OJTD nie w pełni uwzględniają wpływ opóźnień na jakości transmitowanego dźwięku. Należy zaznaczyć, że opóźnienia w sieci oferującej usługę VoIP na poziomie akceptowalnym powinny być utrzymane w granicach do 150

ms. Zjawisko opóźnienia, o ile ma wartość stałą dla wszystkich pakietów, jest jednak trudne do zauważenia przy metodzie odsłuchowej. Jest ono zauważalne dopiero w momencie przeprowadzania konwersacji, kiedy użytkownicy systemu są zdezorientowani czy osoba, z którą przeprowadzana jest rozmowa jest w trakcie słuchania czy mówienia. Metody odsłuchowe, aby oceniać wpływ opóźnienia, musiały korzystać z danych przekazywanych z protokołu transmisji pakietów RTP. Dodając kolejny parametr do proponowanej metody, można uwzględnić wpływ opóźnienia na jakość sygnału. Będzie to jednak parametr zewnętrzny, pochodzący spoza samej próbki dźwięku. Natomiast badanie wpływu jitter'a pokazało, że obie metody oceny jakości, czyli metoda P.563 jak i metoda OJTD, wymagają dalszych prac nad uwzględnieniem tego zjawiska. Należy jednak wspomnieć, że terminale VoIP posiadają tzw. dejitter buffer, czyli bufor gromadzący dane w postaci zakodowanych próbek dźwięku przed przekazaniem ich do odtwarzania. Eliminuje to zjawisko zmiennego opóźnienia pakietów, z kolei jednak wprowadzając opóźnienie w rozmowie.

Średni błąd bezwzględny MOS metody OJTD był nieznacznie (około 6%) większy niż metody P.563, jednakże zmniejszyło się odchylenie standardowe tegoż błędu.

Tabela 4.9. Wyniki badania wpływu jitter na jakość transmitowanego dźwięku

Metoda	Średni błąd bezwzględny MOS	Odchylenie standardowe błędu bezwzględnego MOS
P.563	1,14	1,04
OJTD	1,21	0,89

4.4.5. Wpływ rozmiaru i strat pakietów IP na jakość transmitowanego dźwięku

Przeprowadzone badania wpływu rozmiaru pakietów i ich strat na ocenę jakości transmitowanego dźwięku wykazały, że proponowany algorytm OJTD prezentuje wyniki zbliżone do wyników uzyskiwanych metodą PESQ, a algorytm metody P.563 nie jest w takim samym stopniu zgodny z metodą PESQ. Potwierdza to prawidłowość działania proponowanego algorytmu, który spełnia swoją rolę niezależnie od wielkości transmitowanych pakietów w sieci. Strata pakietów dla różnych wartości pakietów IP

była modelowana za pomocą algorytmu zaimplementowanego według dwustanowego modelu Gilberta.

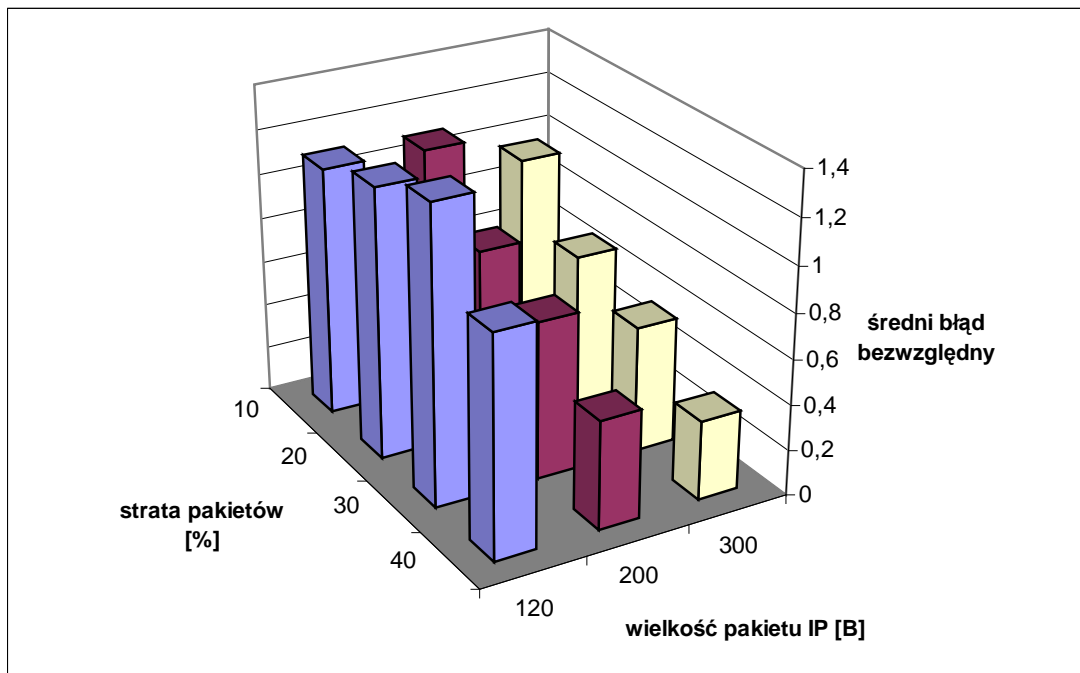
Średnia wartość błędu bezwzględny MOS metody OJTD była mniejsza niż metody P.563, jednocześnie zmniejszyło się odchylenie standardowe tegoż błędu.

Na rysunkach 4.10-11 została zobrazowana średnia wartość błędu bezwzględnego metod P.563 i proponowanej metody OJTD. Widać, że wartość błędu bezwzględnego MOS dla metody OJTD maleje wraz ze zwiększaniem się degradacji sygnału – zwiększaniem się straty pakietów.

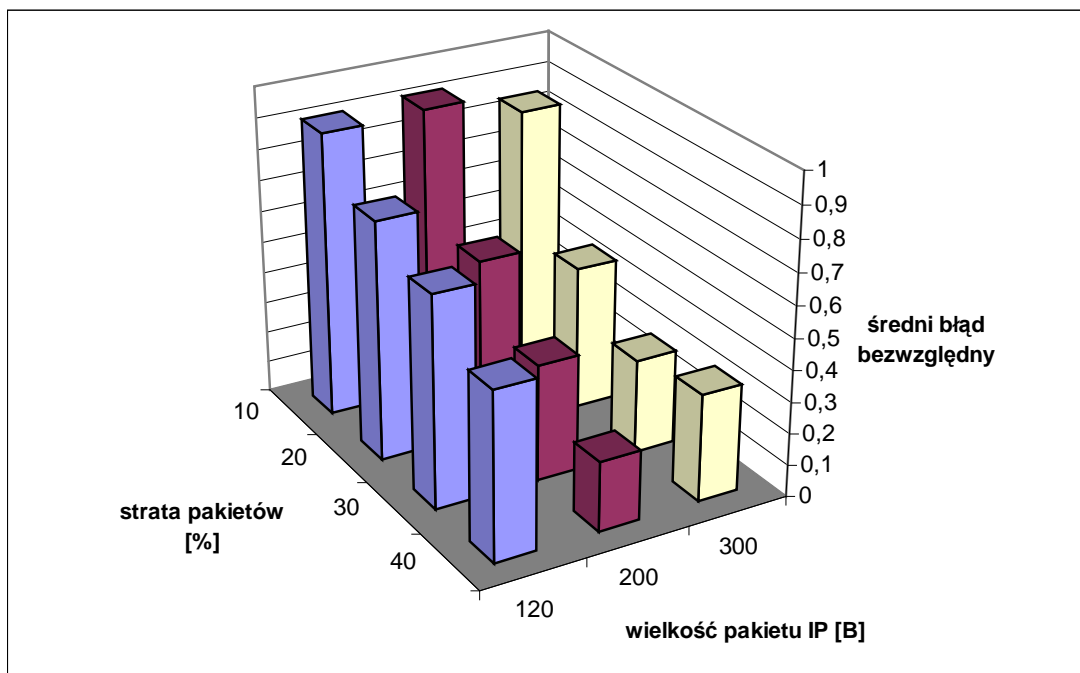
Rysunki 4.12-14 przedstawiają wykresy średniego wskaźnika MOS funkcji straty pakietów przy różnej wielkości pakietów dla metod P.563, PESQ i proponowanej OJTD. Można zauważyć, że wartości metody OJTD są bardziej zbliżone do metody PESQ niż metody P.563.

Tabela 4.10. Wyniki badania wpływu straty pakietów na jakość transmitowanego dźwięku

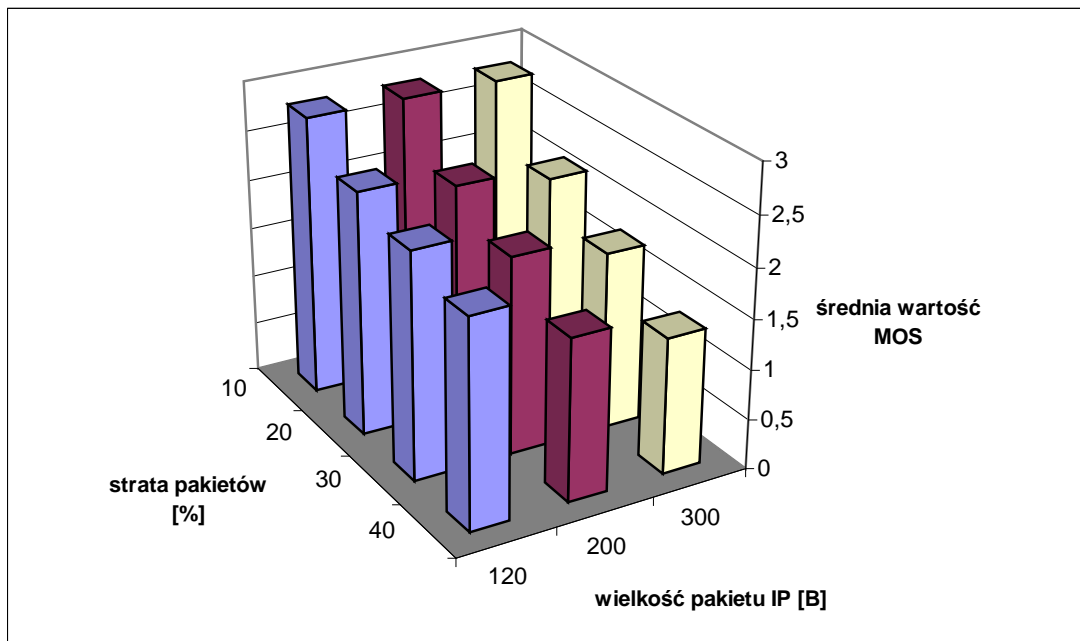
Rozmiar pakietu [B]	Strata pakietów [%]	Średni błąd bezwzględny MOS		Odchylenie standardowe błędu bezwzględnego MOS	
		P.563	OJTD	P.563	OJTD
120	10	1,11	0,91	0,78	0,60
120	20	1,20	0,76	0,56	0,53
120	30	1,31	0,67	0,49	0,53
120	40	0,97	0,53	0,47	0,51
200	10	1,10	0,92	0,68	0,49
200	20	0,82	0,56	0,49	0,32
200	30	0,70	0,37	0,38	0,26
200	40	0,48	0,22	0,23	0,19
300	10	0,96	0,85	0,67	0,48
300	20	0,69	0,46	0,43	0,31
300	30	0,56	0,30	0,31	0,25
300	40	0,35	0,34	0,26	0,24
Wartości średnie		0,85	0,57	0,47	0,39



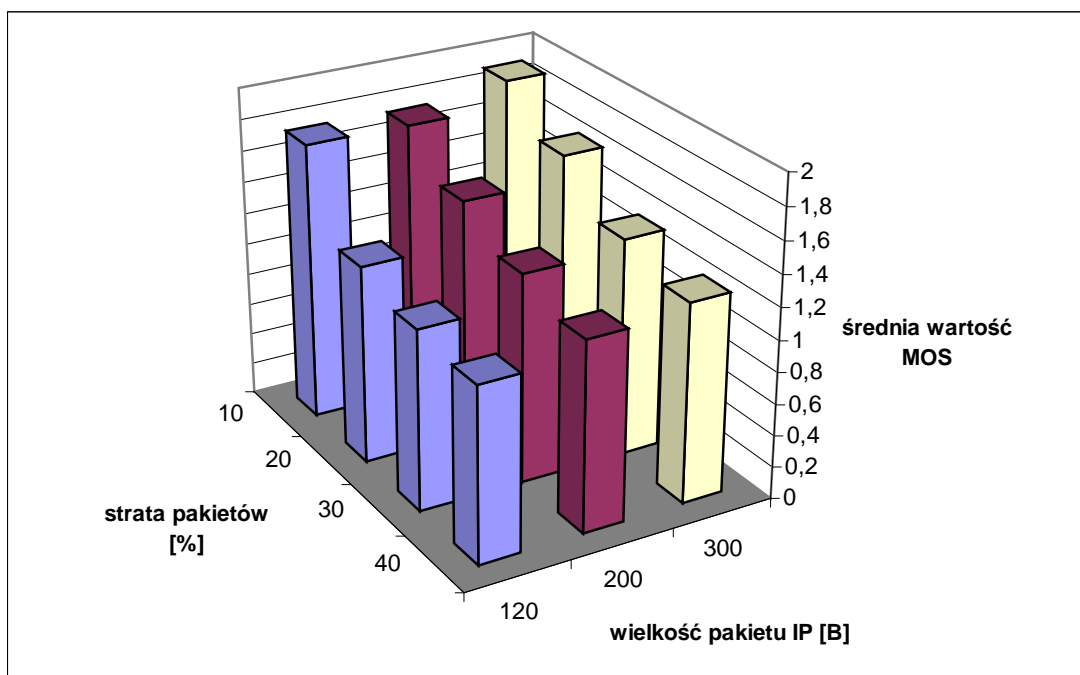
Rysunek 4.10. Średni błąd bezwzględny MOS metody P.563



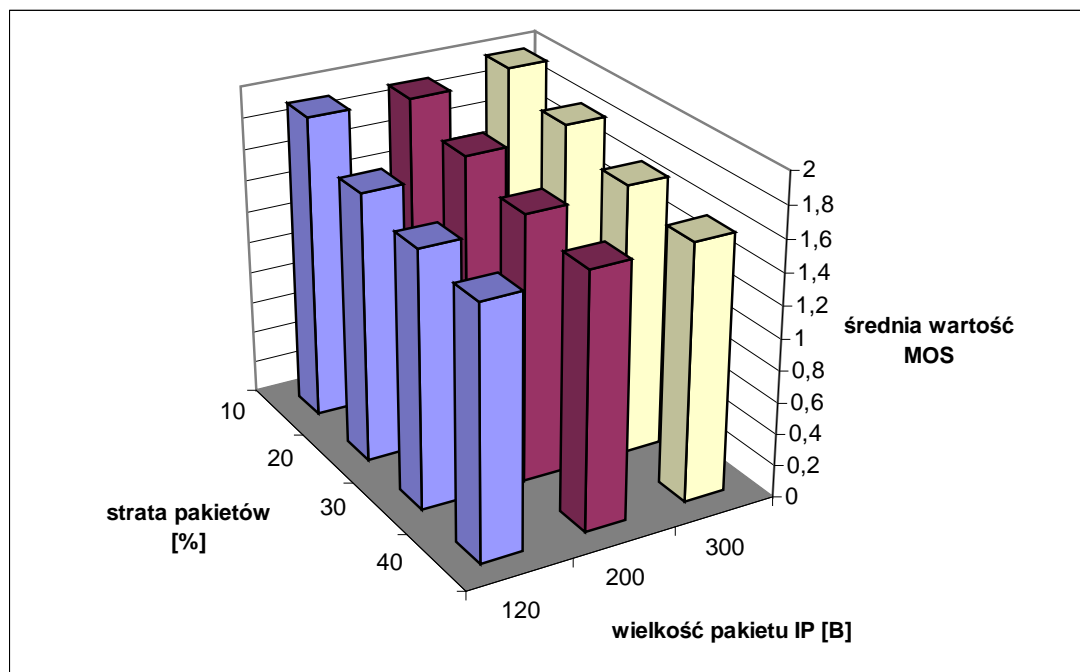
Rysunek 4.11. Średni błąd bezwzględny MOS metody OJTD



4.12. Wartość średnia MOS dla metody PESQ



4.13. Wartość średnia MOS dla metody P.563



4.14. Wartość średnia MOS dla metody OJTD

4.5. Wymagania sprzętowe związane z proponowanym algorytmem OJTD

Proponowany algorytm oceny jakości transmitowanego dźwięku w telefonii VoIP jest zmodyfikowaną metodą P.563. Metoda P.563 cechuje się stosunkowo małymi wymaganiami na zasoby sprzętowe urządzenia, na którym ma pracować. Szczegółowe wymagania metody zostały przedstawione w rozdziale III niniejszej pracy. Proponowany algorytm wymaga obliczeń parametrów dodatkowych $EB1$, $EB2$. Podczas prac badawczych zaproponowany algorytm został zaimplementowany w środowisku Linux RedHat 6.2 z użyciem języka C. Dokonywane pomiary jakości dźwięku przygotowanych próbek badawczych nie wykazały wydłużenia czasu działania algorytmu OJTD w stosunku do algorytmu P.563, z czego należy wysunąć wniosek, iż algorytm jest efektywny i nadaje się do użycia z procesorami o stosunkowo małej mocy obliczeniowej. Ponadto sposób wyznaczania parametrów $EB1$, $EB2$ można zaimplementować efektywnie np. poprzez wyznaczanie wskaźników W_i w jednej pętli obliczeń. Powinno to przyspieszyć pracę działania algorytmu.

Podsumowując, dodanie do metody proponowanej przez ITU - P.563 – nowych parametrów oraz prostego algorytmu ich obliczeń nie powoduje znacznego zwiększenia

zasobów systemowych jakie musi posiadać urządzenie. Niezbędna do obliczeń tablica odwzorowywania wyników posiada trzy indeksy o wymiarach 4, 9 i 15 elementów, co daje 540 wartości koniecznych do przechowania w pamięci. Przy zapisie wartości elementów na 8 bajtach potrzeba 4320 bajtów do zapisania całej tablicy. Dla obecnych procesorów, czy nawet mikroprocesorów, nie jest to duży zakres pamięci. Możliwe jest zastosowanie algorytmu w systemach wbudowanych w urządzenia. Implementacja algorytmu w systemie Linux z wykorzystaniem języka C++ pozwala na jego efektywne wykorzystanie w procesorach DSP czy dowolnych urządzeniach pracujących z procesorami, dla których istnieją kompilatory tego języka. Możliwe jest więc zastosowanie proponowanego algorytmu obliczeń jakości transmisji dźwięku w dowolnym miejscu sieci oferującej usługę VoIP gdzie jest dostęp do pakietów transmitujących dźwięk oraz do protokołu sygnalizacji połączenia.

4.6. Możliwości zastosowań proponowanej metody OJTD

Analizę zastosowania proponowanej metody OJTD można dokonać uwzględniając założenia jakich dokonano przed przystąpieniem do przeprowadzonych badań. Głównym założeniem metody było umożliwienie jej zastosowania w dowolnym miejscu sieci oferującej usługę VoIP, gdzie jest dostęp do protokołu sygnalizacji oraz pakietów IP transmitujących dźwięk. Należy tutaj rozważyć sposób pracy algorytmu. Czy metoda ma działać w trybie rzeczywistym pracy sieci analizując każdą przeprowadzaną rozmowę, czy ma ona być wyzwalana dla wybranych losowo rozmów przeprowadzanych w różnym czasie. Takie założenia definiują moc obliczeniową urządzeń, gdyż musi być ona zapewniona dla każdej instancji działającego algorytmu. Obecnie systemy monitorowania stanu pracy urządzeń działają w trybie rzeczywistym. Przed metodami nadzorowania ich działania stawia się warunki zgłaszania informacji o występujących anomaliach jak najszybciej. Dlatego metody oceny jakości transmitowanego dźwięku w telefonii VoIP powinny działać w czasie rzeczywistym. Wymagania sprzętowe proponowanego algorytmu OJTD pozwalają zaimplementować go w miejscu sieci oferującej usługę VoIP, gdzie można uzyskać dostęp zarówno do protokołu sygnalizacji jak i transmitowanych próbek dźwięku, pozwalają na pracę

w trybie rzeczywistym i uzyskanie oceny jakości dźwięku w poszczególnych połączeniach.

Opracowana metoda OJTD dzięki małym wymaganiom sprzętowym może pracować w dowolnym punkcie sieci oferującej usługę VoIP gdzie ma dostęp do pakietów transmitujących dźwięk oraz sygnalizacji rozmów:

- bramach VoIP,
- terminalach VoIP.
- strażnikach standardu sygnalizacji H323 w szczególnym trybie pracy sieci gdy transmisja pakietów przebiega przez to urządzenie.

Obecnie w sieci konwergentnej często stosowane są routery brzegowe bądź bramy VoIP, które dokonują zmiany systemów kodowania dźwięku. Urządzenia te łączą sieci telefoniczne VoIP z sieciami telefonicznymi pracującymi z innymi algorytmami kodowania dźwięku. Na takich punktach styku sieci może nastąpić degradacja sygnału. Ponadto urządzenia brzegowe mogą zostać przeciążone z powodu zbyt dużej ilości zadań do realizacji. Często ich zadaniem jest zapewnienie odpowiedniej polityki jakości przesyłania pakietów QoS, translacji adresów NAT, czy udostępnianie usługi zapory sieciowej, co znacznie obciąża ich systemy. Wówczas mogą występować przerwy w transmitowanym dźwięku. Proponowany algorytm oceny jakości idealnie nadaje się do wykrywania tego typu awarii w sieci.

Ważne jest, że proponowany algorytm oceny jakości transmitowanego dźwięku jest niezależny od protokołu sygnalizacji połączeń stosowanego w sieci telefonicznej VoIP. Metoda korzystająca z tego algorytmu, może więc być zastosowana w sieci telefonicznej VoIP pracującej z dowolnym systemem sygnalizacji połączeń. Może to być często używany H323 jak i wypierający go SIP, czy każdy inny.

Wyniki oceny jakości transmitowanego dźwięku mogą być przekazywane za pomocą protokołu SNMP do stacji zarządzającej informującej operatora sieci o występujących problemach i potencjalnych miejscach ich przyczyny. Przekazywanie wyników może odbywać się w czasie rzeczywistym dla każdej przeprowadzanej rozmowy, lub dla wybranych połączeń. Należy podkreślić fakt, iż wyniki prezentowane przez zaproponowany algorytm są przedstawiane według skali MOS, która jest obecnie najszerzej stosowaną skalą oceny jakości dźwięku w sieciach telefonicznych.

4.7. Podsumowanie wyników badań

Osiągnięte wyniki badań, świadczą, że zaproponowany algorytm OJTD nadaje się do wykrywania pogorszenia jakości transmitowanego dźwięku w sieci oferującej usługę VoIP przy degradacji transmitowanych sygnałów. W przypadku, gdzie rekomendowana przez ITU metoda P.563 daje zawyżony wynik pomiaru (tzn. mylnie informujący administratora o braku jakiegokolwiek problemu w danym elemencie realizującym usługę VoIP), proponowany algorytm OJTD pozwoli na stwierdzenie prawdopodobnego wystąpienia anomalii. Zaproponowany algorytm OJTD spełnił swoje zadanie sprowadzając wyniki OJTD do wartości rzeczywistych. Należy wspomnieć, że sytuacje, w których oba algorytmy zawodzą, nie powinny mieć miejsca w praktycznie działającej sieci IP. Stan wystąpienie ewentualnych anomalii w sieci IP doprowadzających do niepożądanych degradacji sygnału, w których metody prezentują błędne wyniki, może być monitorowany przez dostępne obecnie systemy nadzoru sieci IP. Nadzór stanu pracy sieci IP winien być dokonywany równoległe do nadzorowana stanu pracy usługi VoIP. W przypadku wystąpienia jakichkolwiek problemów należy rozpocząć od sprawdzenia sieci IP, a dopiero w kolejnym kroku należy szukać przyczyn w samej usłudze VoIP.

Porównując proponowaną metodą OJTD z metodą PESQ można zauważyć, że obie metody uwzględniają w swoich ocenach wpływ zniekształceń kodowania, stratę pakietów i przerwy w transmitowanym sygnale dźwięku. Obie metody nie uwzględniają absolutnego opóźnienia, natomiast metoda PESQ potrafi uwzględniać wpływ zjawiska jitter na jakość transmitowanego dźwięku. Jednakże metoda PESQ posiada znacznie większą złożoność obliczeniową oraz przede wszystkim do swojej pracy wymaga sygnału odniesienia, a dostarczenie go do porównania w warunkach rzeczywistych jest praktycznie niewykonalne. Porównanie metody OJTD z E-modelem wymaga zauważenia faktu, iż obie metody uwzględniają w swoich ocenach wpływ zniekształceń kodowania, stratę pakietów i przerwy w transmitowanym sygnale dźwięku. Metoda E-modelu potrafi w swojej ocenie uwzględnić wartość absolutnego opóźnienia. Jednakże metoda ta nie nadaje się do użycia w systemie rzeczywistym ze względu na konieczność dostarczania szeregu parametrów odnośnie stanu pracy sieci jak i parametrów terminali abonentów. Użycie metody w warunkach rzeczywistych jest

praktycznie niewykonalne, podczas gdy proponowana metoda OJTD znakomicie pracuje na jednym końcu systemu nie ingerując w jego strukturę.

Porównanie możliwości zastosowań metod oceny jakości transmitowanego dźwięku P.563 oraz proponowanej metody OJTD zostało przedstawione w tabeli 4.11. Obie metody uwzględniają wpływ zniekształceń kodowania, przy czym podczas testów średni błąd bezwzględny MOS miał tą samą wartość, jednakże odchylenie standardowe błędu bezwzględnego MOS było mniejsze dla metody OJTD. Testy wpływu absolutnego opóźnienia na wyniki wykazały, że obie metody nie uwzględniają tego parametru na końcowy wynik oceny jakości transmisji dźwięku. Podczas badań wpływu starty pakietów, jak i przerw w transmitowanym sygnale dźwięku, metoda OJTD wykazała się mniejszym średnim błędem bezwzględnym MOS w 100% przypadków testowych. Natomiast badania wpływu zjawiska jitter wykazało, że obie metody uzyskują porównywalne wyniki. Badania wpływu echa w transmitowanym dźwięku wykazały, że obie metody nie uwzględniają poprawnie tego parametru.

Tabela 4.11. Porównanie metod oceny jakości transmitowanego dźwięku

Porównywany wskaźnik	P.563	OJTD
Badanie wpływu zniekształceń kodowania	Tak	Tak
Uwzględnienie absolutnego opóźnienia	Nie	Nie
Uwzględnienie straty pakietów – średni błąd bezwzględny metody podczas testów	w 100% przypadków lepsza metoda OJTD	
Uwzględnienie wpływu jitter	Wyniki porównywalne	
Uwzględnienie wpływu przerw w sygnale – średni błąd bezwzględny metody podczas testów	w 100% przypadków lepsza metoda OJTD	
Uwzględnienie echa w sygnale	Nie	Nie
Praca z sygnałem odniesienia (ingerencja w badaną sieć)	Nie	Nie
Wymagania mocy obliczeniowej procesora MFLOPS	Od 175 do 262	W przybliżeniu od 175 do 262
Praca w czasie rzeczywistym	Tak	Tak
Możliwość predykcji wyniku podczas pracy na jednym końcu systemu	Tak	Tak

Metoda OJTD dzięki wykorzystaniu algorytmu metody P.563 posiada zbliżone wymagania na moc obliczeniową procesora, dzięki czemu może być stosowana do pracy w czasie rzeczywistym. Z przeprowadzonych badań wynika, że metoda OJTD najlepiej

nadaje się do monitorowania stanu pracy usługi VoIP ze względu na brak konieczności pracy z sygnałem odniesienia oraz uzyskiwanie wyników predykcji jakości sygnału lepszych niż metoda P.563.

Szczegółowe wyniki badań zostały zamieszczone w Dodatku A do niniejszej pracy.

Wnioski

W wyniku przeprowadzonych prac badawczych i studiów literatury udało się potwierdzić postawione tezy. Do głównych osiągnięć niniejszej rozprawy należą:

- dokonanie analizy parametrów dźwięku mających wpływ na jakość jego transmisji w sieci,
- określenie wpływu parametrów sieci IP na jakość transmisji dźwięku,
- opracowanie algorytmu oceny jakości dźwięku, który cechuje się:
 - dużą szybkością działania,
 - małym zapotrzebowaniem na moc obliczeniową procesora,
 - małym zapotrzebowaniem na pamięć danych,
 - brakiem sygnału referencyjnego do pomiarów,
 - możliwością pracy w miejscu sieci oferującej usługę VoIP, w którym istnieje dostęp do pakietów transmitujących dźwięk jak i protokołów sygnalizacji połączeń,
- zaproponowanie adaptacyjnej sieci oferującej usługę VoIP.

W wyniku prac nad niniejszą rozprawą przeanalizowano parametry, które mają wpływ na jakości transmisji dźwięku w sieci oferującej usługę VoIP. Wyznaczone parametry posłużyły do przygotowania próbek dźwięku wykorzystanych do prac nad algorytmem oceny jakości transmisji dźwięku w telefonii VoIP. Dokonano przeglądu dostępnych metod pomiaru jakości dźwięku w połączeniach głosowych. Po dokonaniu analizy metod oceny jakości dźwięku, ich wad i zalet pod kątem ich użycia w sieci oferującej usługę VoIP, okazało się, że żadna z nich nie spełnia kryteriów związanych z monitorowaniem stanu pracy tej sieci.

W wyniku badań udało się zaproponować algorytm pomiaru jakości rozmów przeprowadzanych w technologii VoIP. Zaproponowany algorytm OJTD bazuje na zmodyfikowanym algorytmie metody P.563. Modyfikacja polegała na wprowadzeniu dodatkowych parametrów pozwalających na usunięcie niezgodności ocen przy zaniku transmitowanego dźwięku. Ponadto metoda OJTD mierzy wpływ takich samych parametrów jak metoda P.563 (rozdział III). Opracowana metoda ma charakter

heurystyczny, jednakże wyniki uzyskiwane proponowanym algorytmem OJTD są lepsze niż uzyskiwane dotychczasowymi metodami możliwymi do zastosowania w środowisku VoIP. W zależności od rodzaju zakłócenia występującego w sieci wyniki uzyskiwane proponowanym algorytmem są lepsze w około 90% przypadków, natomiast w pozostałych są porównywalne. Zaproponowany algorytm OJTD pracuje bez sygnału wzorcowego, dzięki czemu może być wykorzystany w miejscu sieci oferującej usługę VoIP, gdzie istnieje dostęp do protokołu sygnalizacji oraz transmitowanych próbek dźwięku, bez ingerencji w jej strukturę. Przesyłając wynik pomiaru poprzez sieć IP można uzyskać dostęp do wyników ze stacji monitorującej stan pracy całej sieci. *Możliwa jest więc ocena jakości rozmowy realizowanej w telefonii VoIP za pomocą metody nie ingerującej w działającą sieć. Możliwe jest również sprawdzanie stanu jej pracy z punktów zdalnych, co potwierdza postawioną tezę w niniejszej pracy.*

Mała złożoność obliczeniowa zaproponowanego algorytmu OJTD pozwala na zastosowanie metody OJTD bez obciążania urządzeń sieci świadczącej usługę VoIP, co jest zgodne z tezą o *możliwości zastosowania metody w urządzeniach sieci oferującej usługę VoIP wykorzystując ich wolne zasoby sprzętowe.*

Wynik metody OJTD prezentowany jest w skali MOS, co potwierdza tezę o *możliwości zaproponowania metody działającej w sieci oferującej usługę VoIP prezentującej wynik w skali MOS.* Uzyskane rezultaty badań w zaproponowanym podejściu pozwalają znacznie lepiej oceniać jakość transmisji dźwięku w sieci oferującej usługę VoIP niż metody dotychczas stosowane.

Zaproponowana metoda OJTD i algorytm obliczeń (dzięki swoim cechom: szybkości działania i trybie pracy nie ingerującym w badany system) poprzez integrację z protokołem SNMP całkowicie nadaje się do użycia we współczesnych światowych rozwiązaniach technicznych do monitorowania stanu pracy sieci. Metoda OJTD poprzez monitorowanie przeprowadzanych rozmów umożliwia automatyczne informowanie administratorów systemów o wystąpieniu awarii czy jakichkolwiek anomalii w pracy sieci, co znacznie zwiększy poziom usługi VoIP oferowanej w sieci IP. Ponadto pozwala na próbę budowy adaptacyjnej sieci z usługą VoIP reagującą na zmienne warunki swojej pracy.

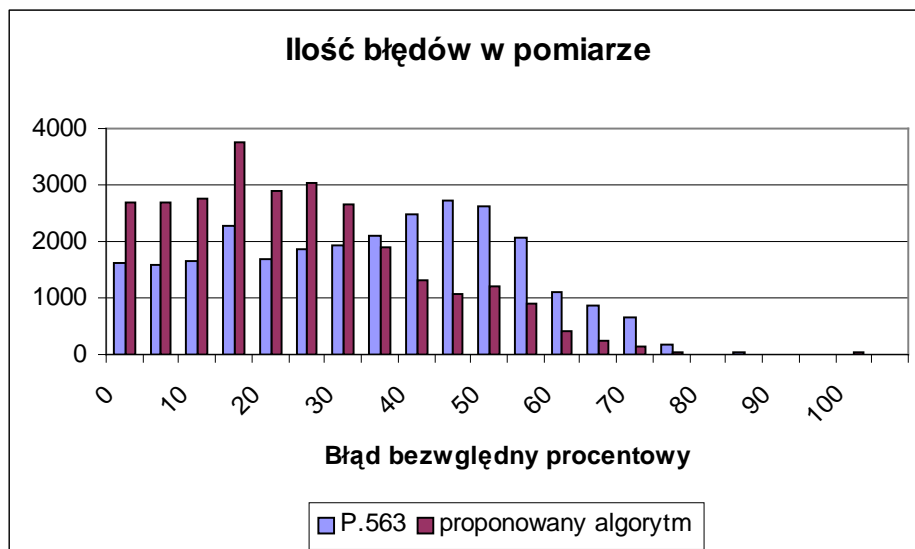
Oczywiście algorytmy muszą być aktualizowane. Obecnie można spotkać już urządzenia pracujące z systemami kodowania dźwięku o paśmie transmisji 8kHz, czy urządzenia łączące systemy VoIP z systemami transmisji radiowej gdzie przesyłane sygnały posiadają szумы o znacznej wartości, a pojęcie jakości sygnału znacznie odbiega od omawianych w niniejszej pracy [89]. Proponowana metoda OJTD w przyszłości może być również stosowana do wykrywania innych aspektów degradacji próbek dźwięku poprzez implementację dodatkowych parametrów pozwalających na ocenę ich wpływu.

Dodatek A

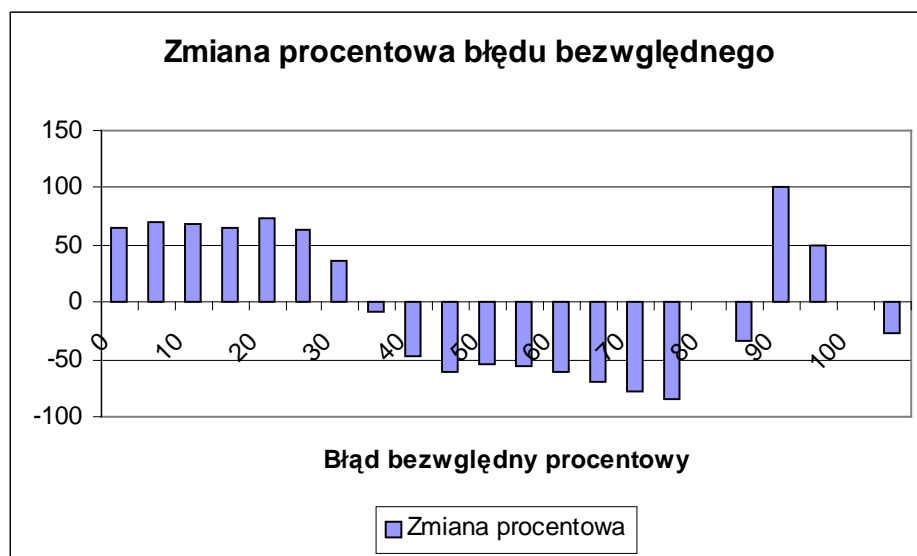
Wyniki pomiarów proponowanej metody oceny jakości transmisji głosowej w telefonii VoIP.

Przedstawione wykresy określają wyniki ocen jakości transmitowanego dźwięku zaproponowanym algorytmem pomiaru w stosunku do metody P563.

Na rysunku 6.1 przedstawiono błąd względny procentowy uzyskiwany metodą OJTD oraz metodą P.563 w badaniach oceny jakości transmitowanego dźwięku przy uwzględnieniu wielkości pakietu IP i modelowaniu ich utraty. Natomiast rysunek 6.2 zawiera zmianę procentową ilości błędów bezwzględnych o określonej wartości uzyskiwanych metodą OJTD w stosunku do metody P.563 w tych samych badaniach. Widać, że ilość błędów o wartościach do 35% zwiększyła się, natomiast ilość błędów większych została zredukowana. Wzrost ilości błędów o wartościach 90 i 95% podyktowany jest bardzo małą ilością wyników pomiarów o tej wartości błędu.

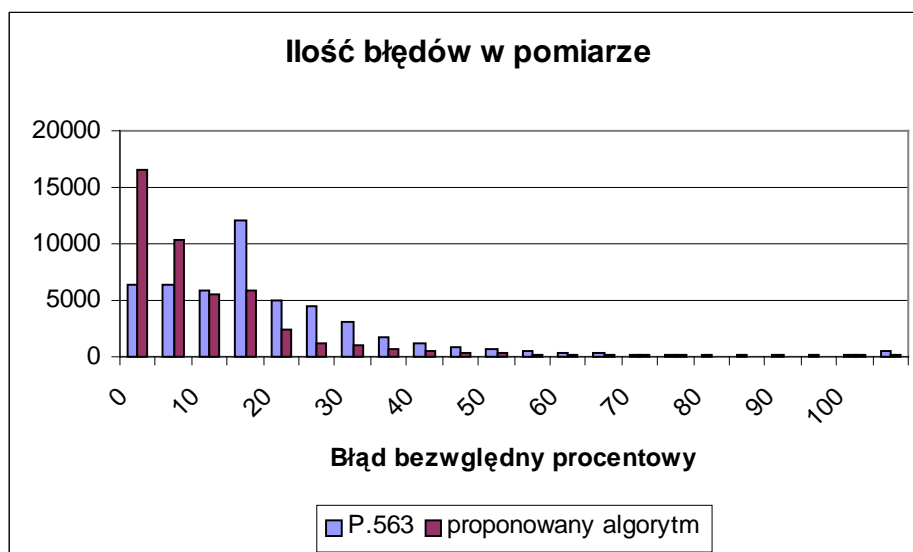


Rysunek 6.1. Błąd względny procentowy uzyskiwany proponowanym algorytmem i metodą P.563 przy badaniu próbek dźwięku z uwzględnieniem wielkości pakietu IP i straty pakietów modelowanej procesami Markowa

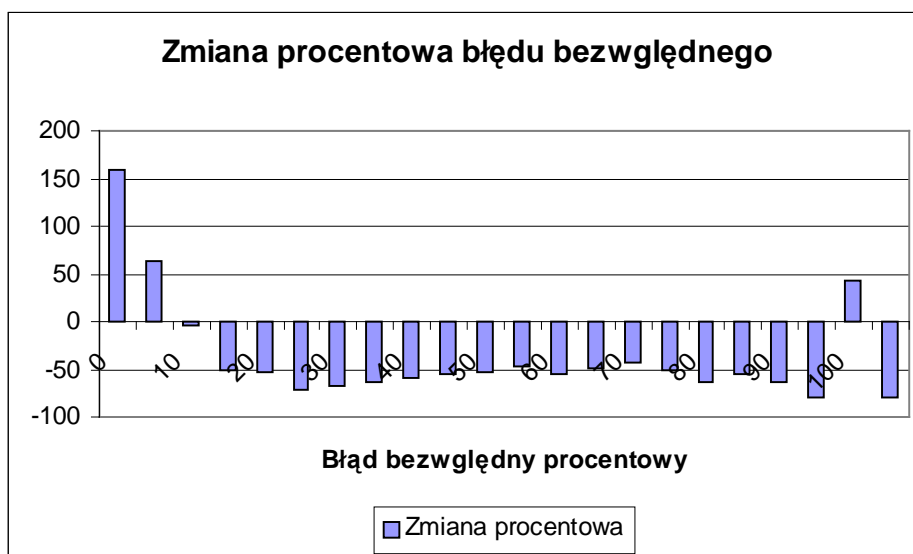


Rysunek 6.2. Zmiana procentowa ilości udziału błędu bezwzględnego oceny jakości transmitowanego dźwięku uzyskiwanej proponowanym algorytmem w stosunku do metody P.563 przy badaniu próbek dźwięku z uwzględnieniem wielkości pakietu IP i straty pakietów modelowanej procesami Markowa

Na rysunku 6.3 przedstawiono błąd względny procentowy uzyskiwany metodą OJTD oraz metodą P.563 w badaniach oceny jakości transmitowanego dźwięku z jego utratą. Natomiast rysunek 6.4 zawiera zmianę procentową ilości błędów bezwzględnych o określonej wartości uzyskiwanych metodą OJTD w stosunku do metody P.563 w tych samych badaniach. Widać, że ilość błędów o wartościach do 10% zwiększyła się, natomiast ilość błędów większych została zredukowana. Wzrost ilości błędów o wartości 105% podyktowany jest bardzo małą ilością wyników pomiarów o tej wartości błędu. Przy tego typu zniekształceniach sygnału metoda OJTD jest wyraźnie lepsza niż metoda P.563.



Rysunek 6.3. Błąd względny procentowy uzyskiwany proponowanym algorytmem i metodą P.563 przy badaniu próbek dźwięku z utratą jego sygnału

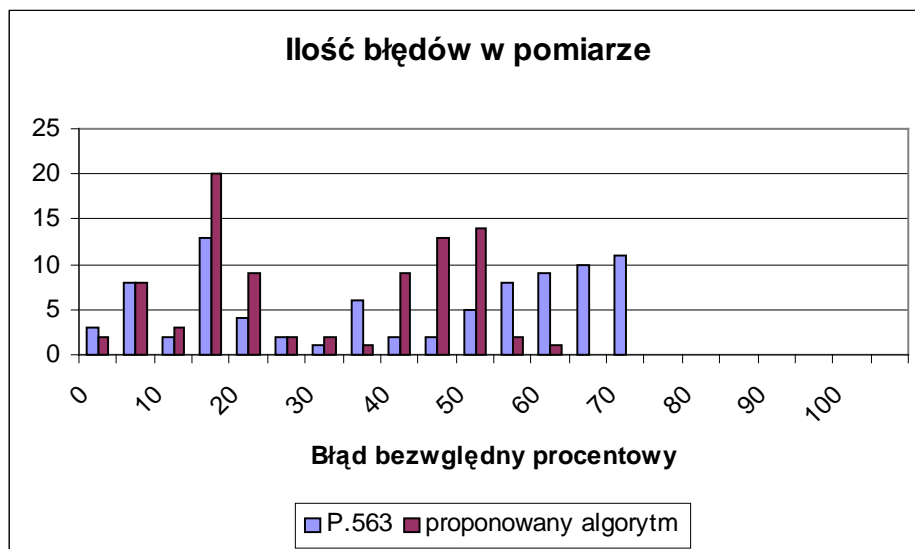


Rysunek 6.4. Zmiana procentowa ilości udziału błędu bezwzględnego oceny jakości transmitowanego dźwięku uzyskiwanej proponowanym algorytmem w stosunku do metody P.563 przy badaniu próbek dźwięku z utratą jego sygnału

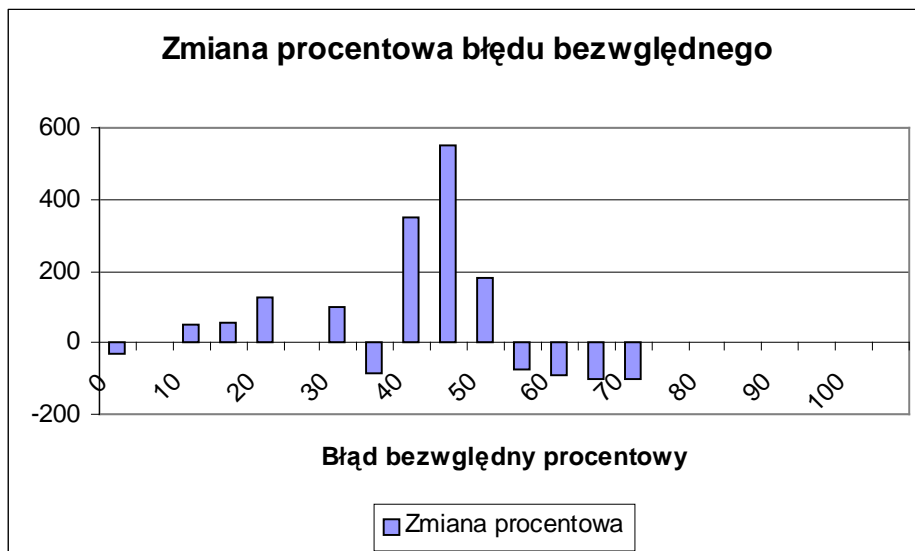
Na rysunku 6.5 przedstawiono błąd względny procentowy uzyskiwany metodą OJTD oraz metodą P.563 w badaniach oceny jakości transmitowanego dźwięku przy wprowadzeniu zniekształceń poprzez zjawisko echa. Natomiast rysunek 6.6 zawiera

zmianę procentową ilości błędów bezwzględnych o określonej wartości uzyskiwanych metodą OJTD w stosunku do metody P.563 w tych samych badaniach. Widać, że ilość błędów o określonych wartościach jest różna dla obydwu metod. W niektórych przypadkach lepsze wyniki uzyskiwano metodą OJTD, w innych za pomocą metody P.563. Ogólnie wartość błędu dla metody OJTD była mniejsza.

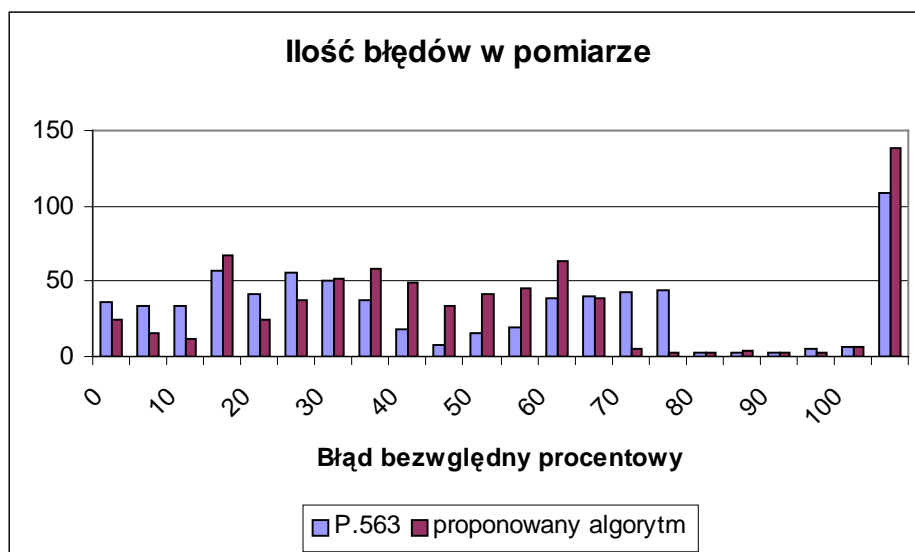
Na rysunku 6.7 przedstawiono błąd względny procentowy uzyskiwany metodą OJTD oraz metodą P.563 w badaniach oceny jakości transmitowanego dźwięku przy wprowadzeniu zmiennego opóźnienia – jitter'a. Natomiast rysunek 6.8 zawiera zmianę procentową ilości błędów bezwzględnych o określonej wartości uzyskiwanych metodą OJTD w stosunku do metody P.563 w tych samych badaniach. Widać, że ilość błędów o określonych wartościach jest różna dla obydwu metod. W niektórych przypadkach lepsze wyniki uzyskiwano metodą OJTD, w innych za pomocą metody P.563.



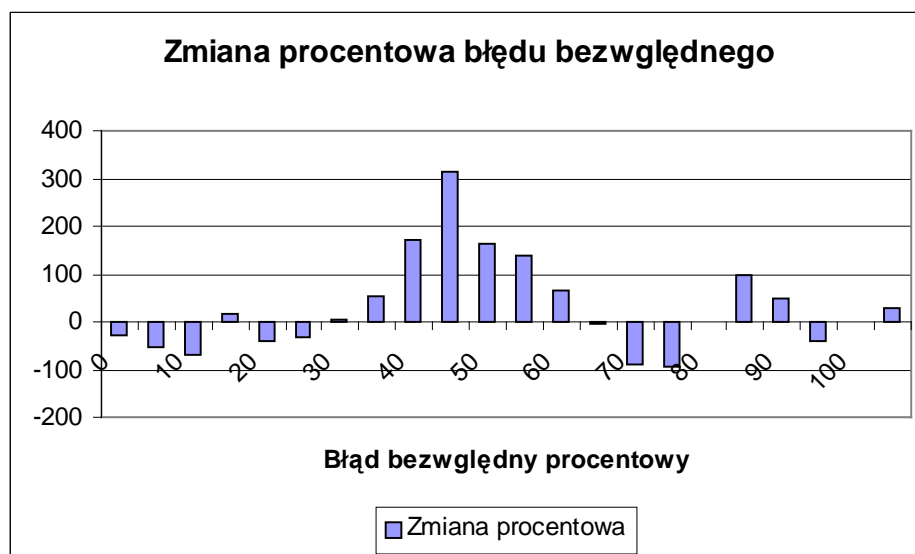
Rysunek 6.5. Błąd względny procentowy uzyskiwany proponowanym algorytmem i metodą P.563 przy badaniu próbek dźwięku z degradacją echem w sygnale



Rysunek 6.6. Zmiana procentowa ilości udziału błędu bezwzględnego oceny jakości transmitowanego dźwięku uzyskiwanej proponowanym algorytmem w stosunku do metody P.563 przy badaniu próbek dźwięku z degradacją echem w sygnale



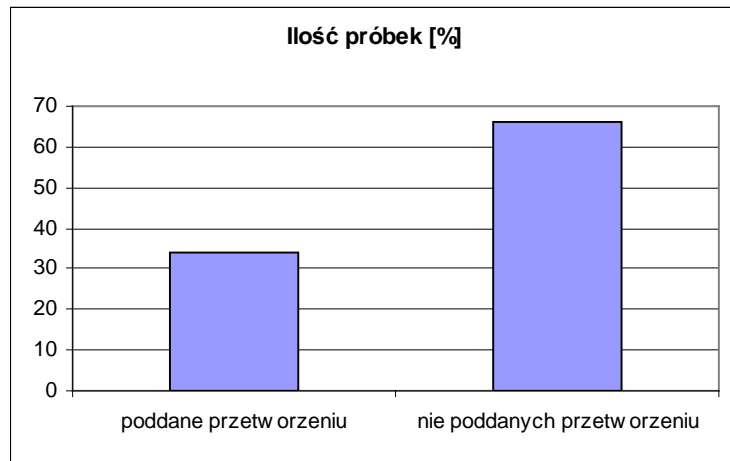
Rysunek 6.7. Błąd względny procentowy uzyskiwany proponowanym algorytmem i metodą P.563 przy badaniu próbek z degradacją jitter'em



Rysunek 6.8. Zmiana procentowa ilości udziału błędu bezwzględnego oceny jakości transmitowanego dźwięku uzyskiwanej proponowanym algorytmem w stosunku do metody P.563 przy badaniu próbek dźwięku z degradacją jitter'em

Ważnym testem oceny metody OJTD był pomiar jakości próbek dźwięku, które nie podlegały przetworzeniu polegającemu na jego degradacji, a jedynie na transmisji z użyciem różnych sposobów kodowania dźwięku.. Przebadano 4176 próbek dźwięku przygotowanych poprzez nagranie przeprowadzonych rozmów z użyciem czterech sposobów kodowania dźwięku. Analizując wyniki przedstawione przez metodę P.563 oraz przez zaproponowany algorytm OJTD okazało się, że:

- 66% próbek nie zostało poddanych korekcje przez proponowany algorytm, wyniki uzyskane były jednakowe jak za pomocą metody P.563,
- 34% próbek zostało poddanych analizie przez proponowaną metodę OJTD i wyniki oceny w większości przypadków zostały poprawione.



Rysunek 6.9. Procentowy wykres liczby próbek dźwięku poddanych przetworzeniu poprzez algorytm proponowanej metody

Osiągnięte wyniki badań, świadczą, że zaproponowany algorytm OJTD nadaje się do wykrywania pogorszenia jakości w sieci oferującej usługę VoIP. Więcej wyników badań znajduje się na dołączonej płycie CD.

Bibliografia

- [1] P. Szkudlarek, *Konwergencja jako czynnik wzrostu konkurencji na rynku usług telekomunikacyjnych*, Przedsiębiorstwo i państwo – wybrane problemy konkurencyjności, Katedra mikroekonomii Uniwersytetu Szczecińskiego, Szczecin 2007
- [2] P. Argyroudis, R. McAdoo, S. Toner, L. Doyle, D. O'Mahony, *Analysing the Security Threats against Network Convergence Architectures*, Proceedings of 3rd International Symposium on Information Assurance and Security (IAS'07), Manchester 2007
- [3] J. Kucharski, P. Sienkiewicz, *Analiza systemowa rozwoju teleedukacji*, Materiały konferencyjne Społeczeństwo informacyjne – wizja czy rzeczywistość, Kraków 2003
- [4] RFC791, *Internet Protocol Darpa Internet Program Protocol Specification*, Information Science Institute University of South California, California 1981
- [5] RFC1883, *Internet Protocol version 6 Specification*, Network Working Group, USA 1995
- [6] J. M. Wojciechowski, *Sygnały i systemy*, Wydawnictwo Komunikacji i Łączności, Warszawa 2008
- [7] R. Tadeusiewicz, *Sygnal mowy*, Wydawnictwo Komunikacji i Łączności, Warszawa 1988
- [8] Ł. Apiecionek, *Speech coding in Voice over Internet Protocol Telephony*, Konferencja System Modelling Control-2005, Zakopane 2005
- [9] Ł. Apiecionek, *Telefonia Voice over IP*, Infotel, maj 2005, str. 40-42
- [10] ITU-T RECOMMENDATION H.323 v 1: *VISUAL TELEPHONE SYSTEMS AND EQUIPMENT FOR LOCAL AREA NETWORKS WHICH PROVIDE A NON-GUARANTEED QUALITY OF SERVICE*, 1996
- [11] ITU-T RECOMMENDATION H.323 v 2: *PROPOSED REVISION OF RECOMMENDATION H.323 - PACKET-BASED MULTIMEDIA COMMUNICATIONS SYSTEMS*, 1998
- [12] ITU-T RECOMMENDATION H.323 v 3: *H323*, 1999

- [13] ITU-T RECOMMENDATION H.323 v 4 :*Draft H.323v4 (Including Editorial Corrections - February 2001)*, 2001
- [14] ITU-T RECOMMENDATION H.323 v 5: *Draft revised Recommendation H.323 V5 (for Consent)* , 2003
- [15] J.I. Moreno, I. Soto, D. Larrabeiti, *Signalling in Voice over IP Networks*, PISCIS and MobyDick, The Pennsylvania State University CiteSeer Archives, 2001
- [16] M. Arango, *Media Gateway Control Protocol (MGCP)*, RFC 2705, 1999
- [17] F. Cuervo, N. Greene, A. Rayhan, *Megaco Protocol Version 1.0*, RFC 3015, 2000
- [18] ITU-T H.248, *Gateway Control Protocol*, 2000
- [19] M. Spencer, *IAX: Inter Asterix eXchange version 2*, RFC5456, 2009
- [20] A. Czyżewski, *Dźwięk cyfrowy Wybrane Zagadnienia teoretyczne, technologia, zastosowania*, Akademicka Oficyna Wydawnicza EXIT, Warszawa 2001
- [21] O. Hagsand, K. Hanson, I. Marsh, *Measuring Internet telephony quality: where are we today?*, Global Telecommunications Conference, 1999. GLOBECOM apos;99 Volume 3, Issue , 1999 Page(s):1838 - 1842 vol.3
- [22] O. Avellaneda, B. Pettitt, *IP Telephony Quality-of-Service Aspects IP Protocol - Changing the Paradigm*, Connect Word ICT, styczeń 2001
- [23] NEC Unified Solutions, *Ensuring IP Telephony performance through remote monitoring*, White paper, 2006
- [24] M.J. Trzaskowska, *Matematyczny model do oceny jakości mowy przesyłanej w sieciach telekomunikacyjnych*, Telekomunikacja i techniki informacyjne, nr 1, 2001
- [25] ITU-T RECOMMENDATION P.800, METHODS FOR SUBJECTIVE DETERMINATION OF TRANSMISSION QUALITY, 1996
- [26] P. Pacyna, *Przegląd wybranych standardów kodowania dźwięku i obrazu*, Krajowa Konferencja Radiodifuzji i Radiokomunikacji 1998
- [27] J. Cała, *Organizacja transmisji głosu w sieci IP*, Studia Informatica Volume 22 Number 2(44), str. 99-112, 2001
- [28] D. Miras, *A Survey of Network QoS Needs of Advanced Internet Applications*, Internet2 QoS Working Group, Working Document, 2002
- [29] M. Gartkiewicz, *Usługa VoIP – czynniki wpływające na jakość transmitowanego głosu*, Telekomunikacja i techniki informacyjne, nr 1-2, 2004

- [30] M. Gajewski, K. Sienkiewicz, W. Latoszek, *Przewodnik w zakresie usług dostępu do internetu*, Urząd Komunikacji Elektronicznej – Instytut Łączności – PIB, Warszawa 2007
- [31] M. Flanagan, *Administring Cisco QoS in IP Networks*, Syngress, 2001
- [32] Audio-Video Transport Working Group, H. Schulzrinne, S. Casner, R. Frederick, V. Jacobson, *Request for Comments: 1889, RTP: A Transport Protocol for Real-Time Applications*, styczeń 1996
- [33] M. Gartkiewicz, *Usługa VoIP – mechanizmy poprawy jakości*, Telekomunikacja i techniki informacyjne, nr 1-2, 2004
- [34] D. Sisalem, H. Schulzrinne, *The Loss-Delay Based Adjustment Algorithm: A TCP-Friendly Adaptation Scheme*, Proceedings of Eight International Workshop on Network and Operating Systems Support for Digital Audio and Video, 1998
- [35] W. Jiang, H. Schulzrinne, *Modeling of Packet Loss and Delay and Their Effect on Real-Time Multimedia Service Quality*, Proceedings of NOVSDAV'2000, 2000
- [36] T. Hunt, *Voice Quality in Next Generation Networks*, Proceedings of Pacific Telecommunications Council Annual Conference'2001, 2001
- [37] J. Tihelka, P. Sovka, *Implementation Effective One-Channel Noise Reductuon System*, Proceedings of Eurospeech 2001 Scandinavia 7th European Conference on Speech Communication and Technology, 2001
- [38] J. Faneuff, D.R. Brown III, *Noise reduction and increased VAD accuracy using spectral subtraction*, Proceedings of the 2003 International Signal Processing Conference, 2003
- [39] P.A.A. Esquef, *Model-Based Analysis of Noisy Musical Recordings with Applications to Audio Restoration*, rozprawa doktorska, Helsinki Universitet of Technology, 2004
- [40] R. A. Khalil, *Adaptive Filter Application in Echo Cancellation System and Implementation using FPGA*, Al-Rafidain Engineering, Vol.16, No. 5, str. 20-32, Grudzień 2008,
- [41] The Test Plan Primer, *Assessing VoIP Call Quality Using the E-model*, IXIA 2005
- [42] J. Davidson, J. Peters, *VOICE OVER IP PODSTAWY*, Mikom, 2005

- [43] 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
- [44] S. Brachmański, *AUTOMATYZACJA METOD OCENY JAKOŚCI OBRAZÓW I MOWY*, Konferencja Poznańskie Warsztaty Telekomunikacyjne, Poznań, 2003
- [45] E. Bernex, A. Gaineau, *Quality of service in VoIP environments*, White Paper, www.neotip.com
- [46] M.P. Hollier, *PAMS: Measuring speech quality over networks, as the customers hear it*, Psytechnics white paper, 2001
- [47] J. Anderson, *Methods for measuring perceptual speech quality*, Agilent Technologies white paper, 2000
- [48] R. Dai, *A Technical White Paper on Sage's PSQM Test*, White paper, 2000
- [49] ITU-T Recommendation G.107, *The E-model, a computational model for use in transmission planning*, grudzień 1998
- [50] ITU-T Recommendation G.109, *Definition of categories of speech transmission quality*, wrzesień 1999
- [51] A.D. Clark, *Modeling the Effects of Burst Packet Loss and Recency on Subjective Voice Quality*, 2nd IPTel Workshop, 2001, str.123 – 127
- [52] ITU-T Recommendation P.563, *Single-ended method for objective speech quality assessment in narrow-band telephony applications*, maj 2004
- [53] Opticom GmbH, *P.563 – Single Sided Speech Quality Measure*, www.opticom.de, 2006
- [54] T.H. Falk, Y. Hua, C. Wai-Yip, *Single-Ended Quality Measurement of Noise Suppressed Speech Based on Kullback-Leibler Distances*, JOURNAL OF MULTIMEDIA, VOL. 2, NO. 5, str. 19-26, wrzesień 2007
- [55] P. Ordas, B. Fox, *Perceptual Evaluation of Speech Quality (PESQ)*, Microtronix Systems Ltd, 2004, <http://www.microtronix.ca/pesq-disc.html>
- [56] ITU-T Recommendation, P-Sup 23 02/98
- [57] J. Willamowius, *The GNU Gatakeeper – User manual*, www.gnugk.org, 2009

- [58] A. Raja, R. Muhammad Atif Azad, C. Flanagan, C. Ryan, *Real-Time, Non-intrusive Evaluation of VoIP*, EuroGP'2007, strony 217-228, 2007
- [59] V. Martin, M. Ian, G. Björn, *A Systematic Study of PESQ's Performance (from a Networking Perspective)*. 5th International Conference MESAQIN 2006: MEASUREMENT OF AUDIO AND VIDEO QUALITY IN NETWORKS, Praga, Czechy, 2006
- [60] W. Jiang, H. Schulzrinne, *QoS Measurement of Internet Real-Time Multimedia Services*, Technical Report CU-CS-015-99, Department of Computer Science Columbia University, Grudzień 1999
- [61] C. Boutremans, J.Y. LeBoudec, *Adaptive Delay Aware Error Control for Internet Telephony*, EPFL, Lausanne, Switzerland, 2nd IP Telephony Workshop, Columbia University, New York, Kwiecień 2001
- [62] K. Park, Y. Oh, K. Lim, K.R. Cho, *A Dynamic Packet Recovery Mechanism for Realtime Service in Mobile Computing Environments*, ETRI Journal, Volume 25, Number 5, Październik 2003, strony 356 – 368
- [63] J. Bolot, S. Fosse-Parisis, D. Towsley. *Adaptive FEC-based Error Control for Interactive Audio in the Internet*, Proceedings of IEEE INFOCOM, 1998, strony 1453-1460
- [64] J. Bolot, S. Fosse-Parisis, D. Towsley, *Adaptive FEC-based error control for Internet telephony*, in Proc. IEEE INFOCOM'99, Marzec 1999, strony 1453-1460
- [65] Network Working Group, *RFC 1157 – Simple Network Management Protocol*, 1990
- [66] W. Stallings, *Protokoły SNMP i RMON. Vademecum profesjonalisty*, Helion, 2007
- [67] F. Dressler, G. Carle, *HISTORY - High Speed Network Monitoring and Analysis*, 24th IEEE Conference on Computer Communications (IEEE INFOCOM 2005), Miami, FL, USA. 2005
- [68] B. Claise, *Cisco Systems NetFlow Services Export Version 9*, RFC 3954, 2004
- [69] A. C. Davenhall, M. J. Leese, *An Introduction to Computer Network Monitoring and Performance*, Astrophysical Virtual Observatory (AVO) Project, <http://www.euro-vo.org/>, 2005

- [70] G. Münz, G. Carle, *Real-Time Analysis of Flow Data for Network Attack Detection*, Materiały konferencyjne IFIP/IEEE Symposium on Integrated Management 2007 (IM2007), Monachium, RFN, 2007
- [71] F. Dressler, I. Dietrich. *Simulative Analysis of Adaptive Network Monitoring Methodologies for Attack Detection*, Materiały konferencyjne IEEE EUROCON 2005 - The International Conference on "Computer as a Tool", strony 624-627. Belgrad, Serbia i Czarnogóra, 2005
- [72] ITU-T Recommendation G.168, *Digital network echo cancellers*, 2004
- [73] ITU-T Recommendation G.168, *Digital network echo cancellers*, 2007
- [74] ITU-T Recommendation G.168, *Digital network echo cancellers*, Implementators guide, 2003
- [75] The International Engineering Consortium, *Echo cancellation*, www.iec.org
- [76] Z. Zhang, G. Schmer, *Performance Analysis of Line Echo Cancellation Implementation Using TMS320C6201*, Digital Signal Processing Solution, Texas Instruments, 1998
- [77] S. B. Weinstein, *A Passband Data-Driven Echo Canceller for Full-Duplex Transmission on Two-Wire Circuits*, IEEE Transactions on Communications, July 1977
- [78] Network Working Group, RFC 2328, *OSPF Version 2*, 1998
- [79] Christian Huitema, *Routing in the Internet*, Prentice Hall PTR, 2000
- [80] T. Lammle, *CCNA™:Cisco® Certified Network Associate Study Guide, 5th Edition*, SYBEX, 2005
- [81] A. S. Tanenbaum, *Sieci komputerowe*, Helion, 2004
- [82] M. Sportack, *Sieci komputerowe. Księga eksperta. Wydanie II poprawione i uzupełnione*, Helion, 2004
- [83] J. S. Haugdahl, *Diagnozowanie i utrzymanie sieci*. Księga eksperta, Helion, 2001
- [84] I. Cohen, B. Berdugo, *Noise estimation by minima controlled recursive averaging for robust speech enhancement*, IEEE Signal Processing Letters, 2002
- [85] S. Ahn, Y. Lee, *AN IMPROVED STATISTICAL MODEL-BASED VAD ALGORITHM WITH AN ADAPTIVE THRESHOLD*, Journal of the Chinese Institute of Engineers, Vol. 29, No. 5, strony 783-789, 2006

- [86] T. Ravichandran, K. D. Samy, *Performance Enhancement on Voice using VAD Algorithm and Cepstral Analysis*, Journal of Computer Science 2 (11): 835-840, 2006
- [87] Spirit Corp, *Voice Activity Detector (VAD) Algorithm User's Guide*, Texas Instruments, 2003
- [88] Y. Liu, C. Bao, *A New Voice Activity Detection Algorithm Based On SOM & LVQ*, International Journal of Information Technology Vol. 12 No.6. 2006
- [89] H. Kruszyński, *Zastosowanie systemu Jaśmin*, Nowa Technika Wojskowa, Nr 9, str. 82-84, 2006