

Ocena jakości usług telefonii pakietowej

Robert Janowski*

Warszawska Wyższa Szkoła Informatyki, Warszawa

Streszczenie

W artykule kompleksowo opisano podejście do zagadnienia oceny jakości usług telefonii pakietowej. Na wstępie dokonano systematyki pojęć związanych z oceną jakości, wyjaśniając m.in. różnicę pomiędzy używanymi w tym kontekście terminami QoS (*Quality of Service*) oraz QoE (*Quality of Experience*). Następnie omówiono możliwe poziomy oceny jakości, tj. poziom sieci, aplikacji i użytkownika, opisując dla każdego z nich adekwatne metryki jakościowe, stosowne rekomendacje organizacji standaryzujących oraz sugerowane wartości graniczne warunkujące akceptowalną jakość. Dla poziomu użytkownika dokonano krótkiej systematyki stosowanych metod oceny, koncentrując się na metodzie E-model z zalecenia ITU-T G.107 [1]. E-model jest nieinwazyjną, pasywną metodą pozwalającą na oszacowanie subiektywnej oceny użytkownika w skali MOS (*Mean Opinion Score*) na bazie wartości parametrów obiektywnych tzn. w pełni mierzalnych. W pracy zilustrowano praktyczne zastosowanie E-modelu, wyjaśniając sposób pozyskiwania niezbędnych dla niego wartości m.in. z raportów protokołu RTCP (*Real Time Control Protocol*) [2].

Słowa kluczowe: *telefonii pakietowa, ocena jakości, MOS, E-Model, RTCP*

* E-mail: rjanowski@poczta.wwsi.edu.pl.

1 Wprowadzenie

Telefonia pakietowa, popularnie określana skrótem VoIP (*Voice over IP*) staje się coraz powszechniejszą metodą realizacji połączeń głosowych. Aplikacje takie jak Skype [3], Viber [4], będące przykładami realizacji usług VoIP zyskują coraz więcej użytkowników. W segmencie użytkowników końcowych telefonia pakietowa cieszy się sporą popularnością ze względu na stosowany model opłat, tzw. stawek płaskich, w których opłata jest zryczałtowana i nie zależy od częstości korzystania z zasobów. Taki model jest typowy na rynku przewodowego dostępu do sieci Internet (np. technika xDSL), jest także stosowany w przypadku dostępu mobilnego w tzw. pakietach danych z ustalonym wolumenem.

Telefonia pakietowa jest również chętnie używana przez operatorów w szkieletcie sieci, co wynika z wysokiej efektywności wykorzystania zasobów, czyli przepustowości łączy. Dla operatorów opłacalne jest stosowanie telefonii pakietowej w szkieletcie sieci, natomiast pozostawienie w linii abonenckiej dotychczasowej techniki komutacji kanałów, ze względu na wysokie koszty zmian na tym odcinku sieci głównie z powodu cen sprzętu.

Wymagana konwersja pomiędzy obiema technikami wykonywana jest na brzegu sieci szkieletowej. Takie rozwiązania są powszechnie stosowane przez operatorów alternatywnych oferujących tanie telefoniczne połączenia międzynarodowe – użytkownik końcowy nie ma nawet świadomości, że do obsługi jego połączenia głosowego wykorzystywana jest telefonia pakietowa. Wynika to z faktu, że posługuje się on swoim dotychczasowym sprzętem, a konwersja dokonywana wewnątrz sieci jest dla niego niewidoczna. Dlatego skala korzystania z usług telefonii pakietowej jest znacznie większa niż może się to wydawać użytkownikom końcowym, którzy usługi VoIP kojarzą zwykle z aplikacjami typu Skype [3] lub Viber [4].

Powszechność usług VoIP podnosi znaczenie kwestii zapewnienia odpowiedniej jakości. Zagadnienie to jest znacznie bardziej skomplikowane niż w przypadku techniki komutacji kanałów. Sieci pakietowe nie mają, typowych dla techniki komutacji kanałów, wbudowanych mechanizmów gwarantowania zasobów. W efekcie, próbki głosu przenoszone w pakietach IP nie gwarantują dotarcia do odbiorcy ani zachowania reżimu czasowego. Skutkiem straty pakietu lub jego zbyt późnego dostarczenia odbiorcy jest degradacja jakości połączenia głosowego. Zatem w telefonii pakietowej pojawiają się zupełnie nowe zjawiska mające wpływ na jakość tej usługi, które powinniśmy umieć zmierzyć.

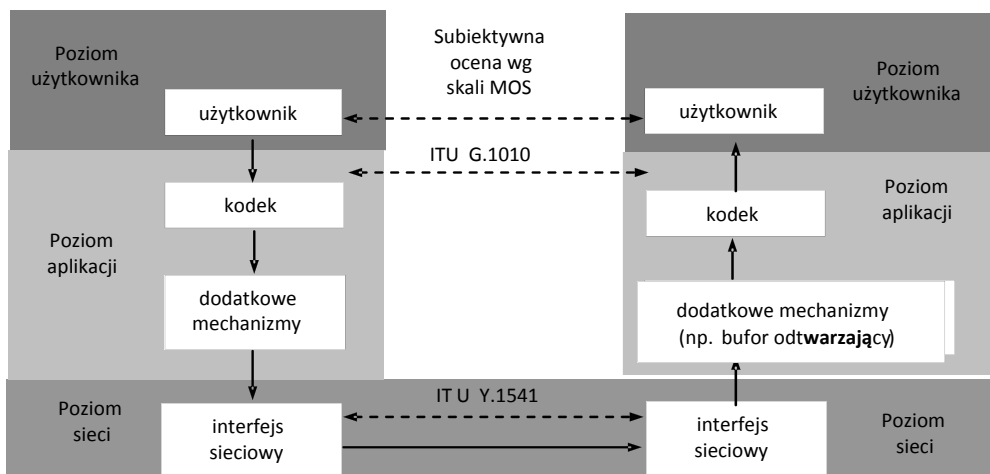
2 Perspektywy, czyli poziomy oceny jakości połączeń telefonii pakietowej

Oceny jakości usług telefonii pakietowej można dokonać na różnych poziomach stosownie do warstwowego modelu przedstawionego na rysunku 1. W tym modelu, analogicznie jak w modelu OSI, warstwa niższa wykonuje pewne funkcje na rzecz warstwy wyższej. Na najwyższym poziomie modelu znajduje się usługa VoIP, która pozwala użytkownikowi na wykonanie rozmowy poprzez sieć IP. Usługa ta korzysta z usług niższego poziomu, które realizują

połączenia VoIP w dedykowanych do tego celu aplikacjach, np. Skype. Jeszcze niżej w modelu umiejscowiony jest poziom sieci, który zapewnia usługi transportowe tzn. przynosi pakiety IP z próbkami głosu.

W zależności od rozpatrywanego poziomu definicje jakości usług mogą posiadać pewne różnice. Dla usług, których odbiorcami są ludzie (poziom użytkownika na rysunku 1) jakość definiowana jest w kategoriach percepcji ludzkiej. W przypadku mowy jest to określenie jakości słyszanego głosu, a na przykład dla usług telewizji IP jest to określenie jakości i płynności oglądanych obrazów. W celu podkreślenia, że pewne miary jakościowe są ściśle związane z ludzką percepcją w ostatnich latach w literaturze światowej pojawiło się określenie QoE (ang. *Quality of Experience*) [5], które można tłumaczyć jako „jakość odczuwana”. Zatem określenie QoE stosowane jest do oceny jakości w sposób bardziej odwzorowujący percepcję ludzką, czyli w kategoriach subiektywnych. Dla usług VoIP subiektywna ocena jakości wyrażana jest w skali MOS (ang. *Mean Opinion Score*) [1].

Do oceny jakości na poziomach aplikacji i sieci stosuje się metryki mierzalne charakteryzujące opóźnienia i straty pakietów IP. Metryki te mają charakter obiektywny, tzn. są dobrze zdefiniowanymi parametrami technicznymi, niezależnymi od percepcji użytkowników. W odniesieniu do nich stosuje się określenie QoS (ang. *Quality of Service*), odróżniając je w ten sposób od subiektywnych metryk QoE.

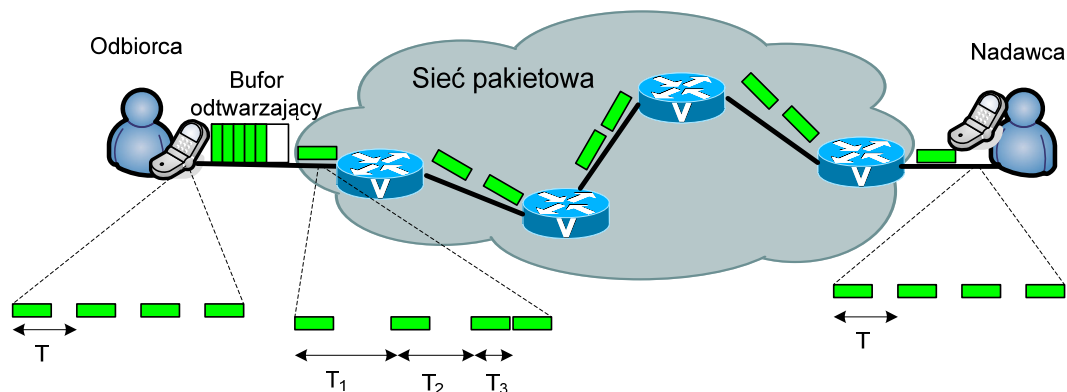


Rysunek 1. Różne poziomy (użytkownika, aplikacji, sieci) oceny jakości usług

Jak widać na rysunku 1, na poziomach aplikacji i sieci jakość definiuje się w oparciu o podobne metryki. Różnice polegają na innych punktach referencyjnych pomiarów i w konsekwencji na operowaniu na innych porcjach danych. Typowo na poziomie sieci podstawową porcją danych

(ang. *Protocol Data Unit*, PDU) jest pakiet, a na poziomie aplikacji, w przypadku usług telefonii pakietowej, jest to ramka kodeka. W jednym pakiecie może być przenoszona więcej niż jedna ramka kodeka, więc strata lub opóźnienie jednego pakietu wpływa na stratę lub opóźnienie więcej niż jednej ramki. W rezultacie wartości parametrów mierzonych na poziomie sieci mogą być inne niż na poziomie aplikacji. Różnice wynikają także z obecności mechanizmów dodatkowych (rysunek 1), mających na celu poprawę parametrów jakościowych uzyskanych na poziomie sieci.

W ogólnym ujęciu, mechanizmy dodatkowe wpływają na zmianę charakterystyk strumienia pakietów mierzonego na poziomie sieci, czyli w pewien sposób modyfikują dane, które podlegają pomiarowi na poziomie sieci. Na przykład dla usług VoIP w celu zapewnienia izochronizmu odbieranego strumienia pakietów przenoszących głos, po stronie odbiorczej stosowany jest tzw. bufor odtwarzający (ang. *playback buffer*) [6]. Idea bufora odtwarzającego polega na kumulowaniu niewielkiej ilości danych (kilku pakietów) w celu kompensacji nierównomiernego napływu pakietów do odbiorcy, który dla poprawnego działania aplikacji wymaga izochronicznego pobierania i odtwarzania próbek z głosem tzn. z takimi charakterystykami, z jakimi został pierwotnie nadany strumień. Efekt nierównomiernego napływu do odbiorcy, czyli zmiana charakterystyk nadawanego strumienia pakietów, jest wynikiem zniekształceń wprowadzanych przez sieć IP (rysunek 2).



Rysunek 2. Idea działania bufora odtwarzającego

Innym mechanizmem stosowanym w celu poprawy jakości, której degradacja wynika z wysokiego poziomu strat pakietów w sieci, jest mechanizm PLC (ang. *Packet Loss Concealment*), czyli mechanizm ukrywania strat pakietów [7]. W jednej z prostszych realizacji może on polegać na replikacji ostatnio odebranego pakietu i podstawianiu go w miejsce straconego. Takie działanie ewidentnie modyfikuje dane otrzymywane z poziomu sieci, co wpływa na wyniki pomiarów dokonywanych na poziomie aplikacji. W ten sposób mechanizmy dodatkowe występujące na

poziomie aplikacje powodują różną ocenę jakości usług mierzonych na poziomach sieci i aplikacji.

W sieciach IP podstawowymi metrykami jakościowymi (w sensie QoS) są opóźnienia i straty. Dla usługi VoIP brak danych lub ich nadmierne opóźnienie (czyli ich brak w określonym momencie, w którym są potrzebne) powoduje niemożność prezentacji próbki głosu co bezpośrednio wpływa na ocenę jakości usługi dokonywaną przez użytkownika.

Rekomendacje wypracowane na bazie analizy percepcji użytkowników jednoznacznie określają akceptowalne poziomy degradacji dla poszczególnych typów usług świadczonych użytkownikom w sieciach IP. Jednym z takich dokumentów jest rekomendacja ITU-T G.1010 [8] podsumowana w tabeli 1, określająca graniczne wartości metryk jakościowych odnoszących się do poziomu aplikacji.

Tabela 1. Podstawowe parametry jakościowe oraz ich wartości graniczne dla usług Audio (na podstawie ITU-T G.1010 Tabela I.1)

Aplikacja	Kierunkowość transmisji	Typowe przepływności [kbit/s]	Podstawowe parametry jakościowe oraz ich wartości graniczne		
			Opóźnienie w jedną stronę	Zmienność opóźnienia	Poziom strat pakietów
Rozmowa	Dwukierunkowa	4–64	< 150 ms < 400 ms (połączenia długodystansowe)	< 1 ms	< 3%
Poczta głosowa	Jednokierunkowa	4–32	< 1 s (odsluchiwanie) < 2 s (nagrywanie)	< 1 ms	< 3%
Strumień audio wysokiej jakości	Jednokierunkowa	16–128	< 10 s	<< 1 ms	< 1%

W nomenklaturze ITU-T wykorzystywanej w rekomendacji G.1010 usługa VoIP odpowiada aplikacji *Conversational Voice*, dla której medium stanowi głos (w oryginale *Audio*). Według tej rekomendacji krytyczne dla odczuwanej jakości usług VoIP są opóźnienia (*One-way delay*), zmienność opóźnienia (*Delay variation*) oraz strata informacji (*Information loss*), czyli strata ramek kodeka VoIP. Graniczne wartości progowe określono na poziomie 150 ms (ewentualnie 400 ms dla połączeń długodystansowych np. międzykontynentalnych), 1 ms oraz 3%, odpowiednio dla opóźnienia, zmienności opóźnienia i poziomu strat informacji. Wartości graniczne dla usługi VoIP nie implikują jednakowych wymagań na poziom QoS dla usług transportowych oferowanych przez poziom sieci. Wynika to z faktu, że pomiędzy poziomem

aplikacji a warstwą sieci istnieją mechanizmy dodatkowe zmieniające charakterystyki strumienia pakietów.

Dla poziomu sieci, który realizuje usługi transportowe, istnieją również rekomendacje precyzujące oczekiwane poziomy jakości. W dokumencie ITU-T Y.1540 [9] zdefiniowano podstawowe metryki poziomu sieci. Należą do nich: opóźnienie pakietów – IPTD (ang. *IP Packet Transfer Delay*), zmienność opóźnienia pakietów – IPDV (ang. *IP Packet Delay Variation*) oraz poziom strat pakietów – IPLR (ang. *IP Packet Loss Ratio*). Parametr IPTD jest mierzony pomiędzy dwoma punktami odniesienia w sieci i zdefiniowany jako czas liczony od wysłania pierwszego bitu pakietu w pierwszym punkcie do momentu odebrania ostatniego bitu pakietu w drugim punkcie sieci. Parametr IPDV jest zdefiniowany jako kwantyl rzędu 10^{-3} parametru IPTD minus jego wartość minimalna, czyli opóźnienie wynikające z czasu transmisji i propagacji. Parametr IPLR jest zdefiniowany jako stosunek liczby pakietów straconych do liczby wszystkich wysłanych pakietów.

W rekomendacji ITU-T Y.1541 [10] określone zostały graniczne wartości metryk z poziomu sieci. Z uwagi na fakt, że na poziomie sieci nie rozważa się poszczególnych aplikacji, ale rodzaje ruchu z uwzględnieniem ich charakterystyk, rekomendacja definiuje tzw. klasy QoS z przypisanymi do nich granicznymi wartościami adekwatnych metryk (tabela 2).

Do przenoszenia ruchu z aplikacji VoIP dedykowana jest klasa 0, ewentualnie klasa 1 w przypadku połączeń długodystansowych. Porównanie granicznych wartości dla poziomów sieci i aplikacji pokazuje, że dla niektórych metryk są one znacząco różne. Na poziomie aplikacji oczekiwana jest bardzo mała zmienność opóźnienia (*Delay variation*) o wartości nieprzekraczającej 1 ms. Natomiast na poziomie sieci graniczna akceptowalna wartość zmienności opóźnienia to 50 ms. Zmniejszenie zmienności opóźnienia z 50 ms do 1 ms jest możliwe do osiągnięcia dzięki wprowadzeniu bufora odtwarzającego, który „wygładzi” odbierany strumień pakietów, czyli spowoduje, że na poziomie aplikacji strumień odbieranych ramek kodeka będzie izochroniczny z dokładnością do 1 ms.

Tabela 2. Lista klas QoS według ITU-T Y.1541

Parametry jakościowe	Klasy QoS					
	Klasa 0	Klasa 1	Klasa 2	Klasa 3	Klasa 4	Klasa 5
Opóźnienie przekazu	100 ms	400 ms	100 ms	400 ms	1 s	U
Zmienność opóźnienia	50 ms	50 ms	U	U	U	U
Poziom strat pakietów	10^{-3}	10^{-3}	10^{-3}	10^{-3}	10^{-3}	U
Poziom błędów w pakietach	10^{-4}	10^{-4}	10^{-4}	10^{-4}	10^{-4}	U

*Symbol U oznacza wartość nieokreśloną

Podobnie jest dla dopuszczalnego poziomu strat pakietów – jest on inny dla poziomu sieci niż dla poziomu aplikacji. W tym jednak przypadku, co wydawać się może niezrozumiale,

na poziomie sieci dopuszczane są dużo mniejsze (10^{-3}) straty pakietów niż na poziomie aplikacji (3×10^{-2}). Dopuszczenie wyższych strat na poziomie aplikacji niż na poziomie sieci wynika z zasad percepcji opóźnionych próbek głosu przez ucho ludzkie. Próbki (czyli faktycznie pakiety przynoszące próbki), których zwłoka przekroczyła pewną progową wartość przynoszą taką samą degradację jakości jak próbki stracone. Wobec tego próbki, których opóźnienie przekroczyło graniczną wartość są na poziomie aplikacji odrzucane, co jest uwzględnione poprzez złagodzenie wymagań na poziom strat do wartości 3×10^{-2} .

3 Metody oceny jakości połączeń VoIP na poziomie użytkownika

Według modelu z rysunku 1, na poziomie użytkownika ocena jakości usług telefonii pakietowej dokonywana jest w subiektywnej skali MOS (ang. *Mean Opinion Score*) [11]. Skala MOS jest skalą 5-stopniową, w której wartość 1 oznacza najgorszą jakość, a wartość 5 najlepszą (tabela 3):

Tabela 3. Pięciostopniowa skala MOS z przypisanymi ocenami

Ocena punktowa	Ocena jakości
5	Wspaniała (ang. <i>Excellent</i>)
4	Dobra (ang. <i>Good</i>)
3	Znośna (ang. <i>Fair</i>)
2	Słaba (ang. <i>Poor</i>)
1	Zła (ang. <i>Bad</i>)

Tabela 4. Poziomy subiektywnej oceny jakości połączeń dla grupy testującej

Opinia użytkowników	Wartość MOS (według ITU)
Bardzo zadowoleni (ang. <i>Very satisfied</i>)	4.3–5.0
Zadowoleni (ang. <i>Satisfied</i>)	4.0–4.3
Część użytkowników jest zadowolona (ang. <i>Some users satisfied</i>)	3.6–4.0
Wielu użytkowników jest niezadowolonych (ang. <i>Many users dissatisfied</i>)	3.1–3.6
Prawie wszyscy użytkownicy są niezadowoleni (ang. <i>Nearly all users dissatisfied</i>)	2.6–3.1
Nierekomendowane (ang. <i>Not recommended</i>)	1.0–2.6

*Maksymalna osiągalna wartość dla kodeka G.711 to 4.4

Definityjnie ocena w skali MOS jest oceną subiektywną dokonaną przez uczestnika konwersacji. W praktyce, subiektywnej oceny usług VoIP dokonuje się poprzez uśrednienie ocen dokonanych przez grupę użytkowników biorących udział w teście (tabela 4). Metoda ta określana jest skrótowo ACR (ang. *Absolute Category Rating*) i jest zestandaryzowana przez ITU-T w rekomendacji P.800 [11]. Oddaje ona prawdziwe odczucia użytkowników, natomiast jej wadą jest koszt i czasochłonność, związane z wynajęciem oraz pracą z grupą testerów.

Z tego powodu ITU-T w rekomendacji P.861 zdefiniowało metodę PSQM (*Perceptual Speech Quality Measure*) [12] umożliwiającą ocenę jakości połączenia VoIP bez konieczności posiadania grupy testerów. Idea metody polega na porównaniu sygnału po stronie odbiorczej (tzn. po zniekształceniu przez kanał) z sygnałem oryginalnym. Porównanie odbywa się z wykorzystaniem specjalnego modelu, który ocenia różnice obu sygnałów i ich wpływ na odczuwaną przez użytkownika jakość. W ten sposób obliczany jest wynik wyrażony w skali MOS, który przybliża percepcję człowieka. Metoda wymaga posiadania sygnału odniesienia (tzw. sygnału referencyjnego) i z tego powodu stosowana może być tylko na etapie testowania produktu lub kodeka. Zwykle takie testowanie odbywa się przed wdrożeniem komercyjnym lub okresowo w trakcie komercyjnego świadczenia usługi. Wadą tej metody jest konieczność posiadania sygnału referencyjnego. Warunek ten uniemożliwia ciągły pomiar jakości połączeń VoIP użytkowników, co ma duże znaczenie dla operatorów. Dodatkowym minusem jest kosztowny sprzęt realizujący pomiar. Do grupy tych metod należą również później zaproponowane metody np. PESQ (*Perceptual Evaluation of Speech Quality*) opisana w rekomendacji ITU-T P.862 [13] lub POLQA (*Perceptual Objective Listening Quality Assessment*) opisana w rekomendacji ITU-T P.863 [14].

W celu umożliwienia pomiaru jakości usług VoIP w sposób nieinwazyjny tzn. tylko na podstawie ruchu przekazywanego przez sieć, ITU-T opracowało metodę pozwalającą na przybliżenie subiektywnej oceny jakości połączeń VoIP wyrażonej w skali MOS poprzez parametry obiektywne tzn. mierzalne w sposób ścisły opóźnienia i straty pakietów. Opracowana metoda to E-model opisany w rekomendacji ITU-T G.107 [1].

Podstawą uzyskania wyniku w skali MOS jest obliczenie wartości tzw. wskaźnika R (ang. *R factor*), która następnie jest odwzorowywana na wartość w skali MOS, według następującego wzoru (1) [1]:

$$MOS = 1 + 0.035R + \frac{7R(R-60)(100-R)}{1000000} \quad (1)$$

Wartość wskaźnika R zależy od kilku składników według następującego wzoru (2) [1]:

$$R = R_0 - I_s - I_d - I_e + A \quad (2)$$

gdzie:

R_0 – jest wartością podstawową obliczoną z wartości parametrów analogowych np. głośności sygnału, poziomu szumów,

I_s – jest składnikiem reprezentującym zniekształcenia analogowe występujące równocześnie z sygnałem użytecznym np. zniekształcenia wprowadzone przez kwantyzację lub spowodowane zbyt wysokim poziomem głośności sygnału,

I_d – jest składnikiem modelującym wpływ opóźnienia pakietów na jakość głosu. Według definicji ITU-T jest to opóźnienie „koniec – koniec” tzn. uwzględniające opóźnienia w sieci oraz opóźnienia związane z przetwarzaniem pakietów RTP w odbiorniku np. wpływ bufora odtwarzającego.

Wartość parametru I_d można wyznaczyć z poniższego wzoru (3) [1]:

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

gdzie:

$$X = \frac{\log\left(\frac{d}{100}\right)}{\log 2}$$

A – to tzw. *Advantage factor* – składnik reprezentujący oczekiwania użytkownika uczestniczącego w rozmowie. W sieciach komórkowych użytkownicy skłonni są zaakceptować obiektywnie gorszą jakość połączeń niż w sieciach stacjonarnych. Dla sieci bezprzewodowych wartość A wynosi 5, a dla sieci przewodowych wynosi 0, tzn. podnosi wartość oceny dla połączeń w sieciach bezprzewodowych,

I_e – jest składnikiem modelującym wpływ strat pakietów przy uwzględnieniu specyfiki używanego kodeka.

Wartość składnika I_e obliczana jest według wzoru (4):

$$I_e = a + b \ln \left(1 + c \frac{IPLR}{100} \right) \quad (4)$$

gdzie:

a, b, c – są parametrami charakteryzującymi kodek,

$IPLR$ – jest poziomem strat podanym w procentach.

Wartości parametrów a, b i c dla przykładowych kodeków zostały podane w tabeli 5. Wartości te otrzymywane są w badaniach metodą ACR [11] w wyniku aproksymacji charakterystyk empirycznych krzywymi opisanymi funkcją z równania (4).

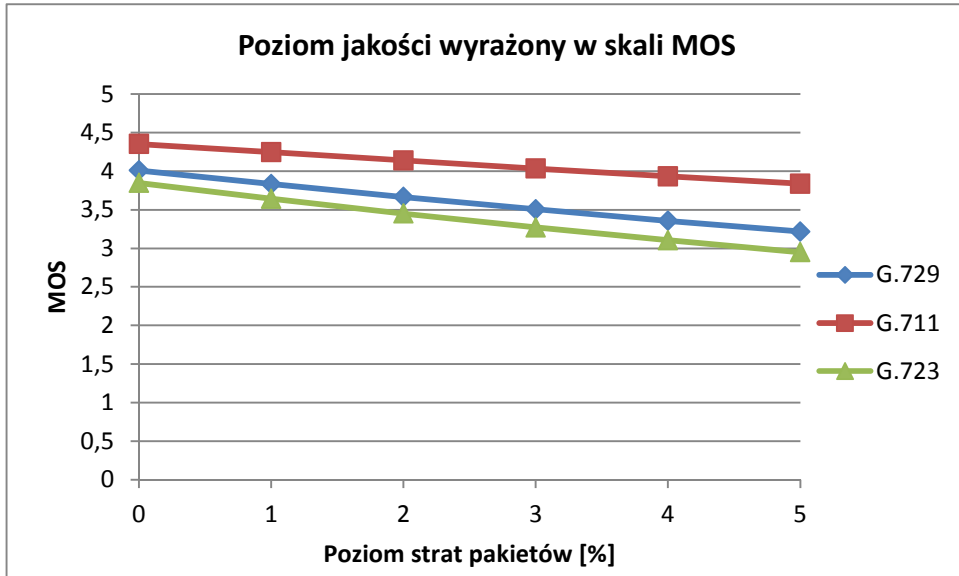
Tabela 5. Wartości parametrów a, b i c dla przykładowych kodeków

Nazwa kodeka	Wartość parametru a	Wartość parametru b	Wartość parametru c
G.711	22	20	0
G.729	31	15	11
G.723	33	15	15

Z definicji E-modelu wynika, że jakość połączenia jest silnie zależna nie tylko od poziomu strat i opóźnień pakietów, ale również od typu użytego kodeka. Niektóre są bardziej, a inne

mniej odporne na określone typy zniekształceń strumienia pakietów IP przenoszącego próbkę głosu.

Na rysunku 3 przedstawiono teoretyczne charakterystyki MOS wybranych kodeków VoIP otrzymane na podstawie wzoru (1) [1], przy założeniu opóźnień „od końca do końca” równych 150 ms i różnego poziomu strat pakietów IP.



Rysunek 3. Teoretyczne charakterystyki MOS wybranych kodeków VoIP otrzymane na podstawie wzoru (1) [1], przy opóźnieniach „od końca do końca” równych 150 ms

W praktycznych zastosowaniach do oceny jakości połączeń wykonywanych w całości w technice pakietowej, tzn. bez transkodowania do techniki komutacji kanałów, wartość wskaźnika R przy domyślnych wartościach parametrów analogowych można wyznaczać ze wzoru (5) [1]:

$$R = 93.2 - I_d - I_e + A \quad (5)$$

E-model jest bardzo atrakcyjny ze względu na możliwość monitorowania jakości usług VoIP bez ingerencji w zachowanie użytkowników. Taki pomiar jest całkowicie przezroczysty dla użytkownika. W praktycznych implementacjach bazuje tylko na raportach protokołu RTCP [2], które i tak są przekazywane pomiędzy terminalami.

Należy jednak podkreślić, że dokładność otrzymanych wyników, rozumiana jako błąd popelnianego oszacowania faktycznej jakości odczuwanej przez użytkowników, zależy od właściwego doboru parametrów modelu, w tym parametrów specyficznych dla kodeków. Doniesienia prac badawczych w tym zakresie wskazują, że popelniany błąd może osiągać wartość

0.5 w skali MOS. Zważywszy jednak na nieinwazyjność tej metody, jej względną prostotę oraz niskie koszty implementacji, można stwierdzić wysoką przydatność tej metody. Dowodem tego są liczne komercyjne programy monitorujące jakość telefonii pakietowej wykorzystujące właśnie tę metodę np. VQMON [15].

4 Wyznaczanie wartości MOS w oparciu o E-model

Najważniejszą zaletą E-modelu jest łatwość i nieinwazyjność pozyskiwania danych do obliczenia wartości MOS. Próbki głosu przenoszone są w pakietach protokołu RTP (*Real Time Protocol*) enkapsulowanych w pakietach IP. Oprócz strumienia RTP każdemu połączeniu VoIP towarzyszy strumień pakietów protokołu RTCP (*Real Time Control Protocol*) [2]. Zadaniem tego strumienia jest, między innymi, dostarczanie danych do oceny jakości połączeń w ujęciu „koniec – koniec” (ang. *end-to-end*).

Protokół RTCP definiuje 5 rodzajów wiadomości [2]:

- raport nadawcy (ang. *Sender Report*, SR),
- raport odbiorcy (ang. *Receiver Report*, RR),
- opis źródła (ang. *Source Description*, SDES),
- wiadomość Goodbye (*BYE*),
- wiadomość specyficzną dla aplikacji (ang. *Application-defined packet*, APP).

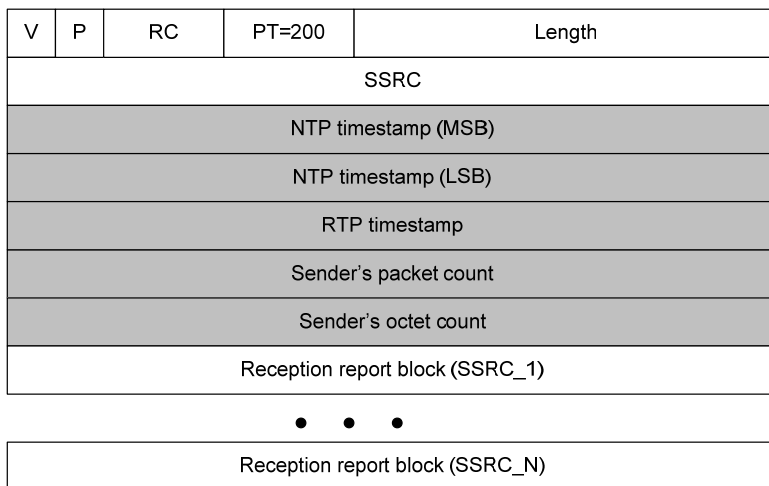
Każda z tych wiadomości ma inny format wynikający z funkcji, których działanie wspiera.

Wiadomości SR i RR zostały zdefiniowane z myślą o przekazywaniu danych związanych z jakością połączeń VoIP.

Różnica pomiędzy wiadomościami SR i RR związana jest z 20 dodatkowymi bajtami, które zawiera wiadomość SR. Bajty te zawierają informację charakteryzującą nadawcę (ang. *Sender info*).

Pozostała część wiadomości SR i RR ma taką samą strukturę i zawiera tzw. bloki raportów (ang. *Report block*). Reguła określająca, który raport (SR czy RR) powinien zostać wygenerowany stanowi, że w przypadku otrzymywania tylko danych odbiorca generuje wiadomości RR. Natomiast gdy odbiorca jednocześnie nadaje dane, wówczas generuje wiadomości SR.

Na rysunku 4 uwidoczniono format wiadomości SR.



Rysunek 4. Format wiadomości z raportem nadawcy (wg RFC 3550); oryginalna terminologia anglojęzyczna

Znaczenie wybranych pól, istotnych z punktu widzenia oceny jakości połączeń, jest następujące:

- pole RC (ang. *Report Count*) określa liczbę raportów przekazywanych w bieżącej wiadomości RTCP,
- pole PT (ang. *Payload Type*) definiuje typ wiadomości RTCP (wartość 200 zdefiniowana jest dla wiadomości SR),
- pole SSRC (ang. *Synchronisation Source*) określa nadawcę raportu.

Pola *NTP timestamp*, *RTP timestamp*, *Sender's packet count* oraz *Sender's octet count* tworzą blok opisu źródła (*Sender Info*). Ich znaczenie jest następujące:

- pole *NTP timestamp* – podaje znacznik czasowy synchronizowany z czasem NTP (*Network Time Protocol*). Wartość zapisana jest na 8 bajtach rozdzielonych na część bardziej (*MSB*) i mniej (*LSB*) znaczącą,
- pole *RTP timestamp* – podaje względny czas wysłania wiadomości,
- pole *Packet count* – podaje liczbę wysłanych pakietów RTP,
- pole *Octet count* – podaje liczbę wysłanych bajtów RTP.

W dalszej części wiadomości SR znajdują się bloki raportów (ang. *Reception report block*) przenoszące dane o parametrach jakościowych związanych z poszczególnymi nadawcami. Nadawcy identyfikowani są przez pole SSRC.

Pełna struktura bloku raportu została uwidoczniiona na rysunku 5 (obszar zaznaczony na szaro), na tle struktury wiadomości RR.

V	P	RC	PT=201	Length
SSRC of the sender				
SSRC_1 (SSRC of the first source)				
Fract. lost		Cumulative number of packets lost		
Ext. highest sequence number received				
Interarrival jitter estimate				
Last sender report timestamp (LSR)				
Delay since last sender report timestamp (DLSR)				
SSRC_2 (SSRC of the second source)				
• • •				

Rysunek 5. Format wiadomości RR (wg RFC 3550); oryginalna terminologia angielszczyzna

Każdy blok raportów rozpoczyna się od pola SSRC wskazującego nadawcę, którego dotyczy raport. Kolejne pola bloku raportu mają następujące znaczenie:

- pole *Fraction lost* – podaje procent straconych pakietów policzony od ostatniego raportu,
- pole *Cumulative number of packets lost* – określa łączną liczbę straconych pakietów od ostatniego raportu,
- pole *Extended highest sequence number received* – podaje najwyższy numer sekwencyjny pakietu RTP odebranego od ostatniego raportu,
- pole *Interarrival jitter estimate* – podaje zmienność opóźnienia pomiędzy pakietami RTP (jest to miara zaburzenia synchronicznie nadanego strumienia pakietów RTP),
- pole *Last sender report timestamp (LSR)* – podaje czas odebrania ostatniego raportu SR z danego źródła,
- pole *Delay since last sender report timestamp (DLSR)* – podaje opóźnienie (wyrażone w jednostkach 1/65535 sekundy) od momentu odebrania ostatniego raportu SR od danego źródła (identyfikowanego polem SSRC) do momentu wysłania do niego bieżącego raportu RR.

Dane z powyżej opisanych wiadomości SR lub RR pozwalają na obliczenie parametrów jakościowych połączenia VoIP tzn. poziomu strat pakietów RTP, ich opóźnień i zmienności tego opóźnienia.

Informacje o poziomie strat (metryka IPLR) są dostępne bezpośrednio z pola *Fraction lost*. Analogicznie, informacje o zmienności opóźnienia (ang. *Jitter*) dostępne są bezpośrednio z pola *Interarrival jitter estimate*.

Należy zwrócić uwagę, że przekazywana w tym polu wartość J wyliczana jest jako średnia krocząca według następującej formuły (6):

$$J(i) = J(i-1) + \frac{(|D(i-1,i)| - J(i-1))}{16} \quad (6)$$

gdzie:

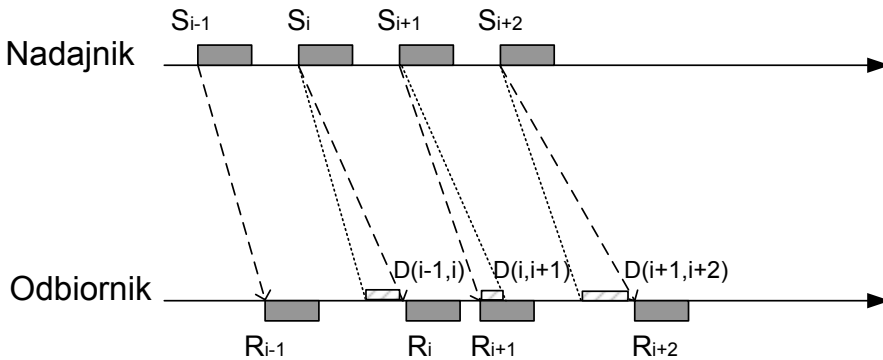
$J(i)$ oraz $J(i-1)$ są wartościami obliczonymi dla dwóch kolejnych pakietów RTP, i -tego oraz jego poprzednika,

$|D(i-1,i)|$ – jest wartością bezwzględną różnicy czasów przekazu dwóch kolejnych pakietów RTP, tj. i -tego pakietu oraz jego poprzednika.

Wartość $|D(i-1,i)|$ obliczana jest na podstawie czasów odbioru (R) i nadania (S) pakietów RTP według wzoru (7):

$$D(i-1, i) = (R_i - R_{i-1}) - (S_i - S_{i-1}) = (R_i - S_i) - (R_{i-1} - S_{i-1}) \quad (7)$$

Zmienność opóźnienia obliczana według wzoru (6) jest faktycznie średnim odchyleniem wielkości D , czyli różnic odstępów między pakietami mierzonych w odbiorniku i w nadajniku. Innymi słowy, jest to średnie odchylenie różnicy czasów przekazu od nadajnika do odbiornika, mierzonych dla dwóch kolejnych pakietów RTP. Definicję wielkości D zilustrowano graficznie na rysunku 6.



Rysunek 6. Ilustracja zjawiska zmian odstępów między pakietami obserwowanego w odbiorniku

Zmienność opóźnienia (*Interarrival jitter*) powinna być obliczana w sposób ciągły, dla każdego odebranego pakietu RTP, według wzoru (6). Po każdym przedziale pomiarowym, wyznaczonym przez wysłanie kolejnych raportów SR lub RR, zmienność opóźnienia jest zerowana.

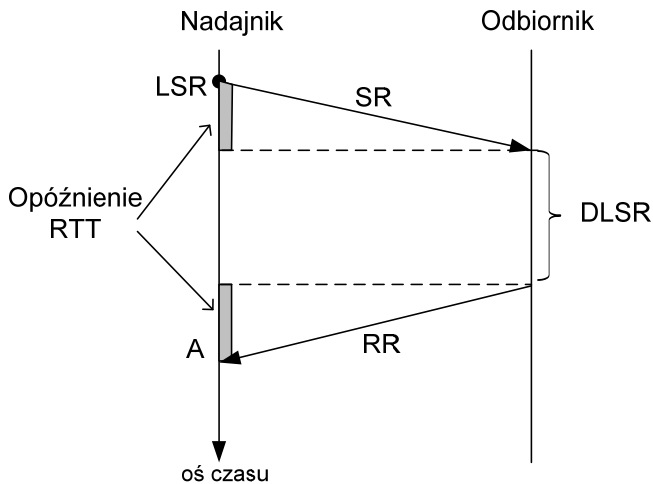
W przeciwieństwie do informacji o poziomie strat i zmienności opóźnienia, informacje o opóźnieniu pakietów RTP nie są dostępne bezpośrednio z raportów SR lub RR. Format

tych raportów nie definiuje pól do przenoszenia takiej informacji. Istnieje jednak sposób obliczenia wartości opóźnienia na podstawie danych z pól LSR i DLSR oraz dodatkowych pomiarów wykonywanych przez odbiornik.

Z definicji czasów LSR i DLSR wynika, że odbiorca wiadomości RR może wyznaczyć opóźnienie, którego doświadczają pakiety RTCP na drodze od nadajnika do odbiornika i z powrotem, czyli opóźnienie w pętli (*Round Trip Time; RTT*). Wartość tego opóźnienia wyznaczana jest jako różnica czasu odbioru wiadomości RR (moment A), czasu nadania wiadomości SR (wartość pola LSR) oraz dodatkowego opóźnienia (wartość pola DLSR) wprowadzonego przez nadawcę wiadomości RR, według wzoru (8):

$$\text{Opóźnienie RTT} = A - \text{LSR} - \text{DLSR} \quad (8)$$

Na rysunku 7 przedstawiono ideę pomiaru opóźnienia RTT z wykorzystaniem wiadomości protokołu RTCP.



Rysunek 7. Pomiar opóźnienia RTT z wykorzystaniem wiadomości SR i RR protokołu RTCP

Czasy nadania wiadomości SR i dodatkowego opóźnienia jej przetwarzania odczytywane są z odpowiednich pól odebranej wiadomości RR (pola LSR i DLSR), a czas A jest znany na podstawie pomiaru. Wyznaczone w ten sposób opóźnienie RTT można wykorzystać do określenia wartości opóźnienia w jedną stronę (*one-way delay*), czyli wartości parametru d koniecznego do obliczenia wartości parametru Id według wzoru (3). Ze względu na fakt, że drogi w obu kierunkach nie muszą być symetryczne, wartość opóźnienia d wyznaczona jako połowa wartości opóźnienia RTT jest tylko przybliżeniem:

$$d = 0.5 \text{ Opóźnienie RTT} + \text{Jitter} \quad (9)$$

Postać wzoru (9) wynika z faktu, że parametr 'd' opisuje opóźnienie „od końca do końca” więc musi uwzględniać oprócz opóźnienia wnoszonego przez sieć (połowa opóźnienia RTT) również opóźnienie wnoszone w terminalu. W odniesieniu do opóźnienia w terminalu przyjmuje się, że jest ono głównie spowodowane kolejkowaniem w buforze odtwarzającym. Bufor odtwarzający kolejkuje tyle ramek, żeby skompensować zmienność opóźnienia wprowadzaną przez sieć, zatem opóźnienie kolejkowania w przybliżeniu równe jest wartości parametru *jitter*.

5 Podsumowanie

O jakości połączeń telefonii pakietowej można mówić na kilku poziomach oceny. W wielu, szczególnie mniej specjalistycznych publikacjach popularno-naukowych, zapomina się o tym, mieszając pojęcia oceny subiektywnej (wyrażanej w skali MOS) z parametrami obiektywnymi takimi jak poziom strat, opóźnienia pakietów lub ich zmienność (*jitter*). Zrozumienie różnic w poziomach oceny oraz właściwe stosowanie metryk przyjętych przez organizacje standaryzujące, jest konieczne dla właściwej interpretacji wyników. Równie ważne jest wybranie właściwej metody. W pracy wskazano obszary zastosowań, w których wykorzystywane są odpowiednie typy metod. W tym kontekście jeszcze raz warto podkreślić znaczenie metody E-model, której poświęcona została znaczna część tego opracowania.

Pomimo niedokładności oszacowania wartości MOS, E-model jest metodą szczególnie atrakcyjną dla operatorów, gdyż umożliwia ciągle i nieinwazyjne monitorowanie jakości połączeń telefonii pakietowej realizowanych przez abonentów. Jest ona też atrakcyjna dla programistów rozwijających aplikacje terminalowe służące ocenie jakości połączeń VoIP. W oparciu o dostępne w terminalu raporty RTCP można w przybliżony sposób wyznaczyć jakość połączeń w skali MOS. W niniejszym tekście szczegółowo wyjaśniono sposób pozyskiwania wartości parametrów koniecznych do wyznaczenia wartości MOS w oparciu o E-model. Wiedza ta ma istotne znaczenie dla implementacji algorytmu wyznaczania wartości MOS bazującego na danych przekazywanych w raportach RTCP.

Bibliografia

- [1] Rekomendacja ITU-T G.107, *The E-model, a computational model for use in transmission planning*, <http://www.itu.org>, marzec 2003
- [2] Schulzrinne H. inni, *RTP: A Transport Protocol for Real-Time Applications*, IETF RFC 3500 dostępne pod adresem: <http://rfc.net/rfc3261.html>, lipiec 2003
- [3] <http://www.skype.com>

- [4] <http://viber.com>
 - [5] Rekomendacja ITU-T G.1080, *Quality of experience requirements for IPTV services*, <http://www.itu.org>, grudzień 2008
 - [6] Ramjee R., Kurose J., Towsley D., Shulzrinne H., *Adaptive playout mechanisms for packetized audio applications in wide-area networks*, materiały konferencyjne IEEE Infocom, (1994)
 - [7] Gündüzhan E., Momtahan K., *A linear prediction based packet loss concealment algorithm for PCM coded speech*, IEEE Trans. Speech Audio Process., Vol. 9, No. 8, listopad 2001
 - [8] Rekomendacja ITU-Y G.1010, *End-user multimedia QoS categories*, <http://www.itu.org>, listopad 2001
 - [9] Rekomendacja ITU-T Y.1540, *Internet protocol data communication service – IP packet transfer and availability performance parameters*, <http://www.itu.org>, marzec 2011
 - [10] Rekomendacja ITU-T Y.1541, *Network performance objectives for IP-based services*, <http://www.itu.org>, grudzień 2011
 - [11] Rekomendacja ITU-T P.800, *Methods for subjective determination of transmission quality*, <http://www.itu.org>, sierpień 1996
 - [12] Rekomendacja ITU-T P.861, *Objective quality measurement of telephoneband (300-3400 Hz) speech codecs*, <http://www.itu.org>, sierpień 1996
 - [13] Rekomendacja ITU-T P.861, *Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs*, <http://www.itu.org>, luty 2001
 - [14] Rekomendacja ITU-T P.863, *Perceptual Objective Listening Quality Analysis*, <http://www.itu.org>, styczeń 2011
 - [15] www.telchemy.com
-

Quality evaluation of the IP packet telephony

Abstract

This paper summarizes quality evaluation of the IP packet telephony service. At the beginning the taxonomy of the quality related concepts is introduced and the fundamental differences between QoS (Quality of Service) and QoE (Quality of Experience) approaches are explained. Furthermore the possible quality evaluation levels including user, application and network perspective are proposed and for each quality evaluation level the relevant recommendations of

standardization bodies, the adequate metrics and associated thresholds necessary for the acceptable service quality are discussed. For the user level, the brief review of the quality evaluation methods is presented with the special focus on E-model as defined in ITU-T G.107 [1] recommendation. E-model is a not invasive passive method that let to estimate the user subjective opinion in MOS (Mean Opinion Score) scale based on the values of the measurable objective parameters. In this paper the practical application of the E-model is described including details related to acquisition of the necessary values from RTCP (Real Time Control Protocol) reports [2].

Keywords: *voice over IP, quality evaluation, MOS, E-Model, RTCP*