

obiecujących i sprawdzonych praktyk oceny i dokumentowania przebiegu procesów i odnoszenia wyników tych ocen do oczekiwań użytkowników. W wyniku takiego postępowania powstają również wytyczne odnośnie zalecanych udoskonaleń [15]. Na ogół przy dokonywaniu oceny przyjmuje się pewien wzorcowy model odniesienia, których może być więcej niż jeden. Właśnie standardy umożliwiają definicję takich modeli odniesienia, co ułatwia ocenę procesów wytwarzania, jak i określenia kierunków zmian. Zalecane modyfikacje powinny być uwzględniane przez wszystkich aktorów (analityków, projektantów, programistów i testerów) biorących udział w wytworzeniu konkretnego przedsięwzięcia informatycznego (systemu czy usługi).

### 5.3. Parametry jakości usług

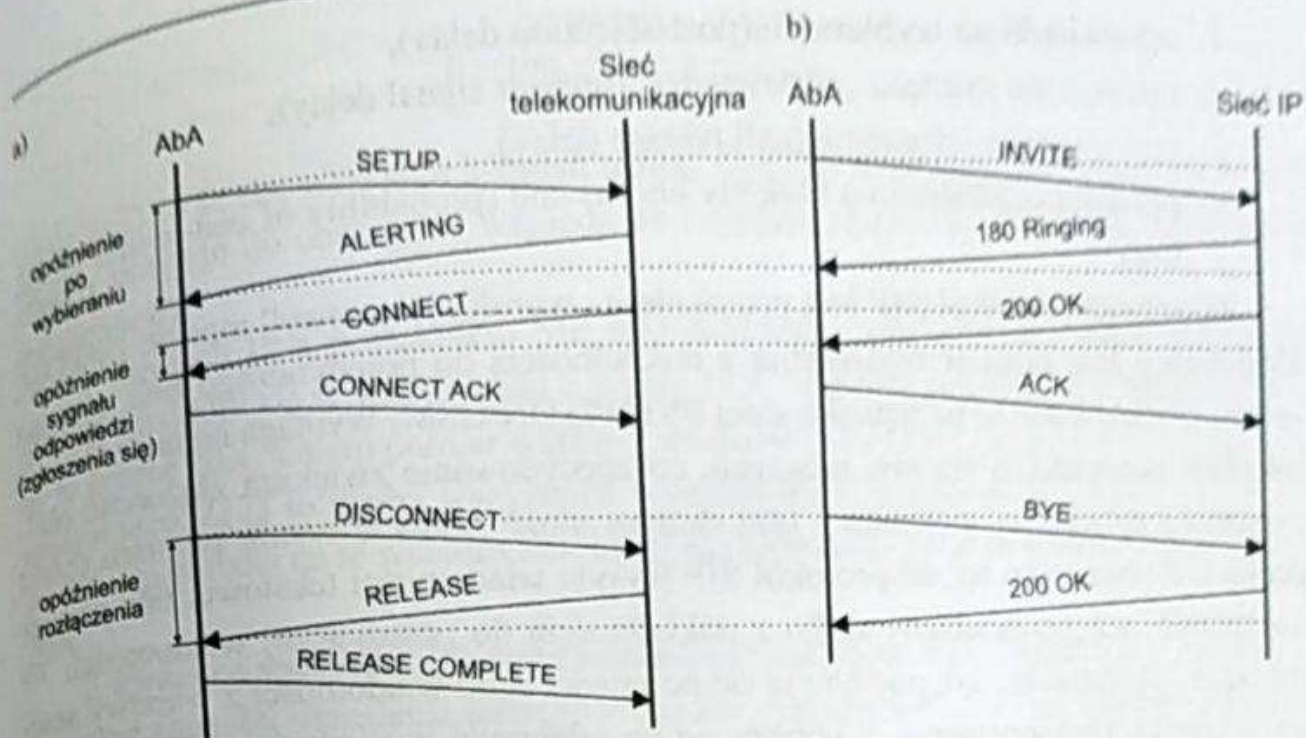
#### 5.3.1. Usługa mowy

Parametry jakości usługi można podzielić na dwie grupy (patrz rozdz. 5.2.1), jedna z nich dotyczy jakości zestawiania połączenia (co nazywane jest poziomem usług, ang. *GoS*, lub wydajnością, ang. *performance*), a druga związana jest z jakością przenoszenia pakietów przez sieć (jakość usługi, ang. *QoS*). W takiej też kolejności zostaną omówione miary wchodzące w skład tych grup parametrów (metryk).

Punktem odniesienia dla określenia parametrów poziomu usług są definicje jakie zostały wprowadzone w klasycznej sieci telekomunikacyjnej PSTN/ISDN/GSM, ponieważ w innych rozwiązaniach z punktu widzenia użytkownika (abonenta) sposoby reakcji sieci na żądanie usługi nie ulegają zmianie. Jednakże z praktycznego punktu widzenia dokładne przeniesienie pomiarów, w przypadku pomiaru czasów, z sieci PSTN/ISDN/GSM do sieci IP może okazać się utrudnione z uwagi na stosowanie innych protokołów. Dużym kłopotem jest np. ustalenie tych samych punktów odniesienia w scenariuszu obsługi połączenia. Dlatego w dalszej części monografii przedstawimy dwa podejścia do określenia parametrów poziomu usług. Pierwsze bazujące na bezpośrednim odwzorowaniu scenariuszy i wykorzystaniu tej samej definicji oraz drugie polegające na niezależnym zdefiniowaniu dwóch parametrów dla opóźnień czasowych wprowadzonych dla sieci IP z protokołem SIP, które jednocześnie określają parametry wydajnościowe systemu telekomunikacyjnego.

Rozpatrzmy pierwsze podejście, które odnosi się do uproszczonego schematu (scenariusza) realizacji połączenia w tej sieci w porównaniu do rozbudowanego scenariusza przedstawionego w rozdziale trzecim na rysunku 3.9. W celu udostępnienia usługi abonentowi A konieczne jest zestawienie połączenia pomiędzy terminalami abonentów A i B, które przebiega według określonego scenariusza wymiany wiadomości sygnalizacyjnych. Schemat takiej wymiany w sieci wraz z uwzględnieniem mierzonych wielkości poziomu usług określonych zgodnie z zaleceniem ITU-T E.721, przedstawiono na rysunku 5.7.





**Rysunek 5.7.** Podstawowe scenariusze wymiany wiadomości dla obsługi żądania usługi mowy oraz wskazane parametry poziomu usług dla sieci ISDN (a) i dla sieci IP z protokołem SIP (b)

Wartości parametrów GoS mogą oddziaływać na opinię abonenta i ze względów czysto technicznych muszą być wyrażone poprzez wielkości odpowiadające określonym zdarzeniom zaznaczonym w scenariuszu na rysunku 5.7a. Zgodnie z zaleceniem ITU-T E.721 można zdefiniować następujące parametry poziomu usług (GoS) w sieciach PSTN/ISDN:

1. opóźnienie przed wybieraniem (pre-selection delay),
2. opóźnienie po wybieraniu (post-selection delay),
3. opóźnienie sygnału „odpowiedzi” (answer signal delay),
4. opóźnienie rozłączenia (call release delay),
5. prawdopodobieństwo blokady end-to-end (probability of end-to-end blocking).

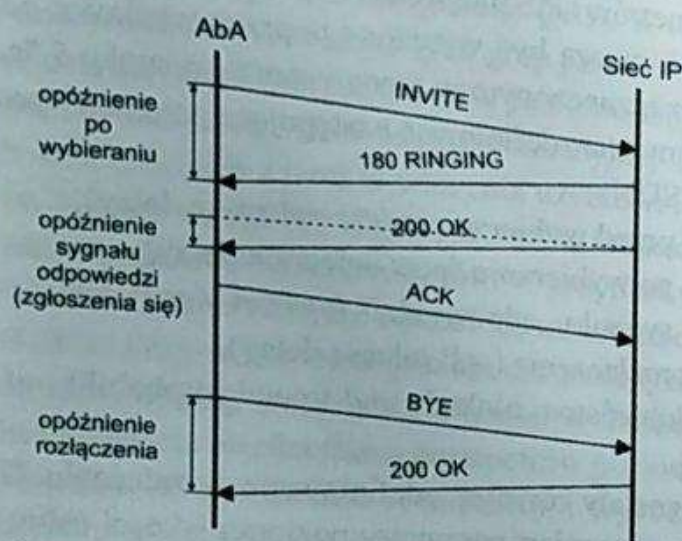
Parametry te zostały również zdefiniowane w zaleceniu ITU-T E.726, gdzie wyróżniono dwie grupy miar: parametry poziomu jakości usług na poziomie połączenia (jednakowe z przedstawionymi dla ISDN) oraz parametry na poziomie komórki w technologii BISDN/ATM, które zostały zaadaptowane dla technologii IP i określają wówczas jakość usług (QoS).

W sieciach IP z sygnalizacją bazującą na protokole SIP, dane potrzebne do zestawienia połączenia są przesyłane w jednej wiadomości INVITE. Odpowiada to, ze względu na sposób zestawienia połączenia, wysyłaniu numeru abonenta „en bloc”. W konsekwencji nie dokonuje się pomiaru opóźnienia przed wybieraniem. Dlatego rozważania ograniczymy przede wszystkim do pomiaru wartości opóźnienia dla scenariusza zgodnego z takim sposobem pracy. W ogólności możliwe jest zdefiniowanie następujących parametrów poziomu usług (GoS) dla sieci IP z protokołem SIP:



1. opóźnienie po wybieraniu (post-selection delay),
2. opóźnienie sygnału „odpowiedzi” (answer signal delay),
3. opóźnienie rozłączenia (call release delay),
4. prawdopodobieństwo blokady end-to-end (probability of end-to-end blocking).

Ponieważ protokół SIP jest protokołem „wysokiej warstwy” modelu ISO/OSI, kłopotliwy jest pomiar opóźnienia z dokładnością do pojedynczego bitu, chociaż jest to realizowane w przypadku sieci PSTN/ISDN/GSM. Wymaga on nadzoru protokołów wszystkich warstw niższych, co zwiększa złożoność obliczeniową procedury pomiaru i tym samym może zakłócić wynik pomiaru (efekt próby). Z uwagi na to, że protokół SIP wysyła wiadomości tekstowo, taki pomiar może być przeprowadzany tylko z dokładnością do pojedynczej wiadomości. To oznacza, że pomiary rozpoczynają się po przekazaniu wiadomości z warstwy sesji do warstwy transportowej, a kończą się po odebraniu wiadomości przez warstwę sesji z warstwy transportowej. Uwzględniając wymianę wiadomości sygnalizacyjnych, poziom usług dla sieci IP z protokołem SIP możemy określić w sposób pokazany na rysunku 5.8.



**Rysunek 5.8.** Podstawowy scenariusz wymiany wiadomości dla VoIP z protokołem SIP wraz z zaznaczonymi parametrami poziomu usług

Oprócz trzech wymienionych na rysunku 5.8 parametrów jakościowych (metryk) konieczne jest również zdefiniowanie czwartego tzw. prawdopodobieństwa niezyskania usługi ze względu na zajętość zasobów w sieciach IP z protokołem SIP. Z punktu widzenia pomiaru prawdopodobieństwa blokady end-to-end, sieć IP umożliwia pomiary w sposób analogiczny do stosowanych pomiarów w sieci PSTN/ISDN/GSM. W związku z tym wzór określający prawdopodobieństwo  $B$  blokady end-to-end, inaczej niezyskania usługi w sieciach IP, ma następującą postać:

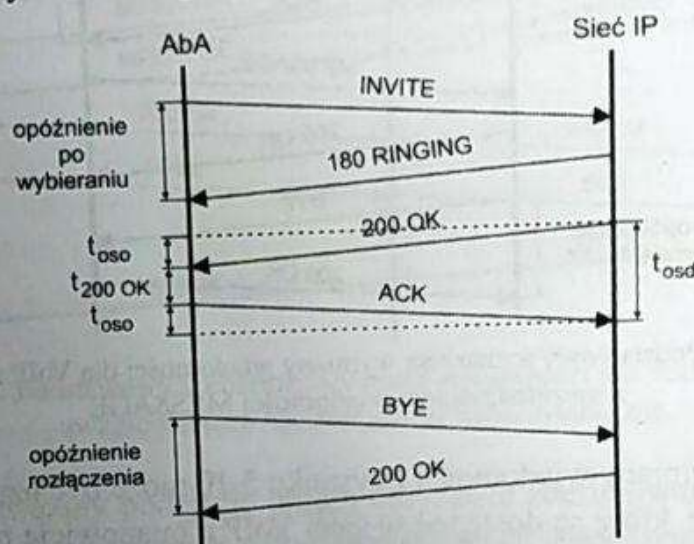


$$B = \lim_{T \rightarrow \infty} \frac{L_s(T)}{L(T)} \cong \frac{L_s(T)}{L(T)}, \quad (5.2)$$

gdzie  $L(T)$  to liczba wszystkich żądań usług,  $L_s(T)$  to liczba żądań usług, które nie zostały przyjęte do obsługi ze względu na zajętość zasobów,  $T$  to czas pomiarów (obserwacji).

W sieciach IP z protokołem SIP fakt dotyczący przyjęcia żądania usługi do obsługi znajduje się w wiadomości z grupy 18X, zaś fakt odpowiedzi drugiej strony w grupie 20X. Zatem pomiar wartości wielkości  $L(T)$  i  $L_s(T)$  można zdefiniować następująco:  $L(T)$  to liczba wszystkich wysyłanych wiadomości INVITE,  $L_s(T)$  to liczba odpowiedzi na te wiadomości, które nie należą do grup odpowiedzi 18X oraz 20X.

Z uwagi na specyfikę działania technologii pakietowej konieczne jest zmodyfikowanie sposobu określania jednego z czasów opóźnienia, tzn. opóźnienia sygnału odpowiedzi (200 OK). W innym przypadku wymagana jest duża dokładność zsynchronizowania zegarów obu terminali abonentów (użytkowników będących AbA czyli UAC i AbB czyli UAS). Aktualnie sieci IP nie gwarantują takich rozwiązań dlatego wykorzystuje się modyfikację analizowanych parametrów, która została przedstawiona na rysunku 5.9 [16].



**Rysunek 5.9.** Podstawowy scenariusz wymiany wiadomości VoIP z protokołem SIP oraz modyfikacja określania czasu opóźnienia odpowiedzi

Opóźnienie sygnału odebrania  $t_{oso}$  można określić pośrednio poprzez pomiar czasu opóźnienia oznaczonego na rysunku 5.9 przez  $t_{osd}$ :

$$t_{osd} = 2 \cdot t_{oso} + t_{200OK} \quad (5.3)$$

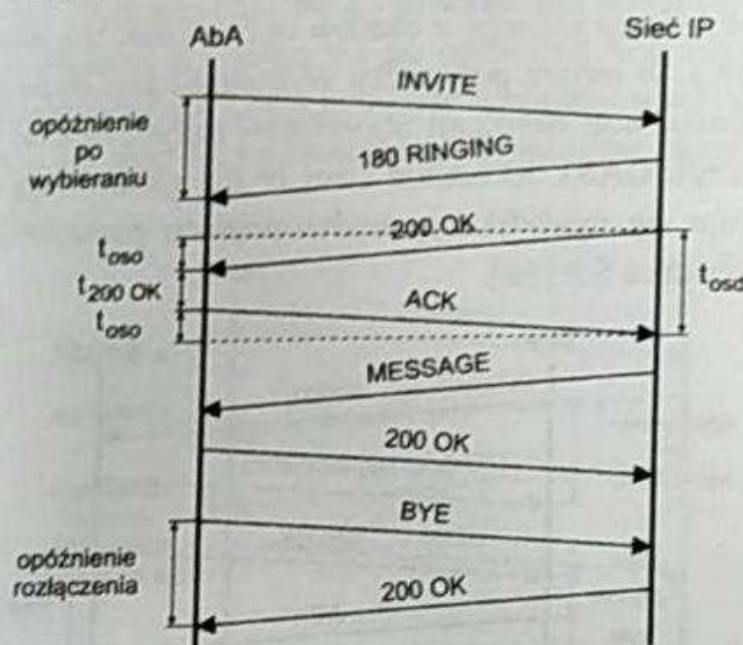
gdzie  $t_{oso}$  to opóźnienie wiadomości 200 OK,  $t_{200OK}$  to czas jaki upływa pomiędzy otrzymaniem odpowiedzi 200 OK a wysłaniem potwierdzenia ACK. Po



przekształceniu wzoru 5.3 otrzymujemy poszukiwaną wielkość  $t_{oso}$  zgodnie ze wzorem 5.4. Zauważmy, że poczyniono założenie dotyczące równości czasów opóźnień przesłania wiadomości w obu kierunkach. Stanowi to potencjalne źródło błędu oszacowania  $t_{oso}$ .

$$t_{oso} = \frac{t_{osd} - t_{200OK}}{2} \quad (5.4)$$

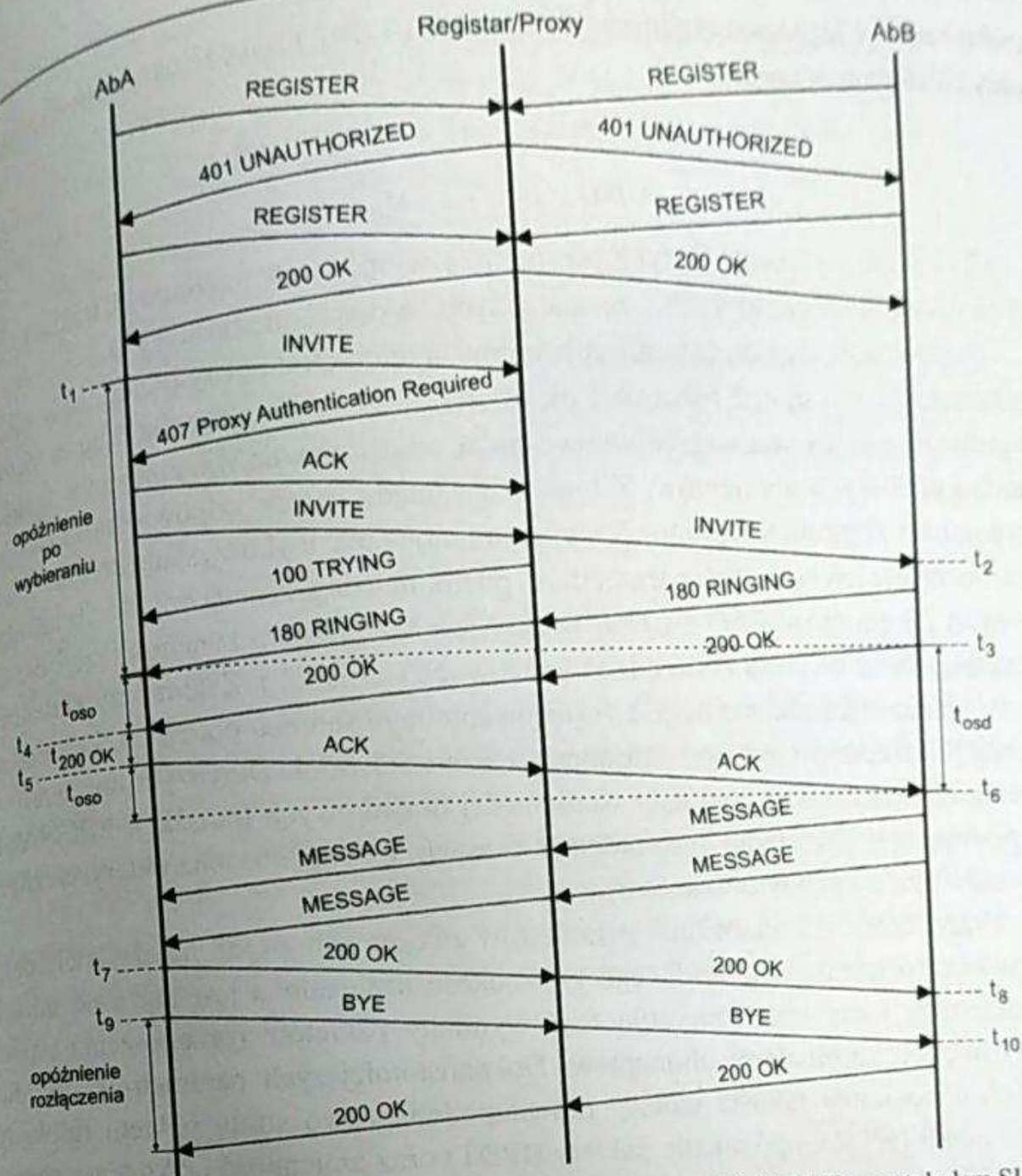
Chcąc zrealizować ten sposób pomiaru  $t_{oso}$  musimy dostarczyć wynik pomiaru czasu  $t_{osd}$  do terminala abonenta AbA (UAC), który jest odpowiedzialny za przebieg pomiaru parametrów GoS. Jest to realizowane poprzez dodanie do scenariusza dodatkowej transakcji wykorzystującej wiadomość MESSAGE (RFC 3261). Umożliwia ona przesyłanie wiadomości tekstowych pomiędzy terminalami abonentów (abonentami Ab czyli użytkownikami UA), co zilustrowano na rysunku 5.10.



**Rysunek 5.10.** Podstawowy scenariusz wymiany wiadomości dla VoIP z protokołem SIP z uwzględnieniem wiadomości MESSAGE

Scenariusz wymiany wiadomości z rysunku 5.10 może być uzupełniony o dwie dodatkowe funkcje, które są dostępne w sieci VoIP a mianowicie rejestrację i autoryzację. Tak rozbudowany scenariusz wymiany wiadomości dla sieci VoIP z protokołem SIP został pokazany na rysunku 5.11. Aby abonent uzyskał dostęp do usługi musi być zarejestrowany w serwerze Registrar co jest realizowane przy pomocy wiadomości REGISTER, na którą serwer nie odpowiada potwierdzeniem 200 OK a 401 UNAUTHORIZED. Jest to prośba o autoryzację ze strony abonenta, który aby ją zrealizować wysyła ponownie wiadomość REGISTER z odpowiednimi danymi dla autoryzacji. Po jej zaakceptowaniu serwer potwierdza odpowiedzią 200 OK.





**Rysunek 5.11.** Podstawowy scenariusz wymiany wiadomości dla VoIP z protokołem SIP uwzględniający rejestrację oraz uwierzytelnienie

Obecnie rozważmy podejście drugie dotyczące zdefiniowania dwóch nowych parametrów opisujących poziom usług, określających jednocześnie miary wydajnościowe systemu telekomunikacyjnego zrealizowanego w technologii VoIP z protokołem SIP. Dla ich opisu wykorzystano ten sam scenariusz obsługi żądania usługi, pokazany na rysunku 5.11, na którym zaznaczono charakterystyczne punkty  $t_i, i=1,2,3, \dots, 10$  na osi czasu a następnie na ich bazie zdefiniowano te dwa nowe parametry. Jeden z tych parametrów dotyczy czasu *CSD* opóźnienia zestawiania połączenia (Call Set-up Delay), natomiast drugi dotyczy czasu *CDD* opóźnienia rozłączenia



połączenia (Call Disengagement Delay) [17, 18, 19]. Czasy te określone są następującymi zależnościami:

$$CSD = (t_2 - t_1) + (t_4 - t_3) + (t_6 - t_5), \quad (5.5)$$

$$CDD = (t_8 - t_7) + (t_{10} - t_9). \quad (5.6)$$

Tak więc wyznaczenie tych parametrów wymaga pomierzenia charakterystycznych momentów czasowych  $t_i$  zaznaczonych na diagramie (rys. 5.11).

Porównanie dwóch ostatnio prezentowanych metod oceny parametrów jakości wskazuje, że nie są one tożsame a określone czasy są mniejsze dla drugiej metody. Wynika to z faktu nieuwzględnienia czasów przetwarzania wiadomości w urządzeniach końcowych abonentów. Z tego też względu są one odpowiednimi miarami wydajności systemu telekomunikacyjnego i mogą być wykorzystane do szacowania czasów opóźnień będących parametrami poziomu usług. Zauważmy, że dla tak określonych parametrów  $CSD$  i  $CDD$ , istotny jest też problem synchronizacji czasów urządzeń sieciowych oraz urządzeń końcowych (terminali). Zatem przy aktualnym rozwiązaniu sieci konieczne jest zastosowanie tego samego podejścia, które odpowiada poprzedniej metodzie opisanej wzorami 5.3 i 5.4. Dotyczyła ona wprowadzenia do scenariusza wymiany wiadomości dodatkowych transakcji MESSAGE. Zapewnią one przesłanie informacji o czasach wystąpienia charakterystycznych punktów czasowych w tym scenariuszu.

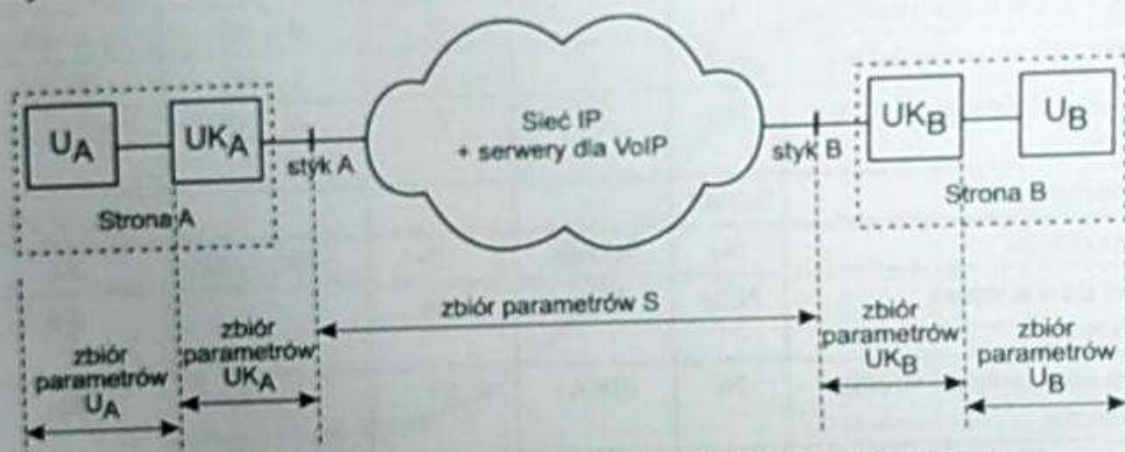
Przechodząc do omawiania parametrów związanych z QoS należy stwierdzić, że w sieci o technologii VoIP tego typu jakość uzależniona jest także od innych parametrów, które charakteryzują fazę wymiany pakietów (przenoszenia pakietów) między terminalami abonentów. Do najistotniejszych parametrów decydujących o poziomie jakości należą: prawdopodobieństwo straty pakietu (blokady end-to-end) (IPLR), opóźnienie pakietu (IPDT) oraz zmienność opóźnienia pakietu (IPDV). Są to metryki, które można bezpośrednio określić przez parametry techniczne sieci IP. Oczywiście na jakość usługi odczuwanej przez abonenta mają także wpływ inne czynniki, takie jak np. typ kodeka, szumy pomieszczenia, itp. Stąd też dla usługi mowa został opracowany specjalny ogólny model zapewniający kompleksowe uwzględnienie różnych ale istotnych parametrów. Jest to tzw. model E opisany w zaleceniu ITU-T G.107, który może być wykorzystany do oceny jakościowej połączenia VoIP utworzonego dla realizacji usługi mowa. Dzięki temu mamy możliwość określenia parametru  $R$  (patrz podrozdział 5.2.1), który bezpośrednio opisuje jakość usługi i może on przyjmować wartości od 0 do 100. Parametr ten można także przeliczyć na parametr  $MOS$  (Mean Opinion Score), obecnie szeroko wykorzystywany do opisu jakości usługi mowy.



Model E określania jakości jest najpopularniejszym i najdokładniejszym spośród tego typu dostępnych modeli. Model ten bierze pod uwagę wszystkie aspekty, które mają wpływ na jakość usługi mowa, począwszy od parametrów jakościowych sieci przenoszącej strumień pakietów, poprzez rodzaj kodeka sygnału mowy, parametry sygnału mowy, czy cechy urządzenia końcowego (terminala), aż po cechy otoczenia w którym przeprowadzana jest rozmowa. Na rysunku 5.12 pokazano schemat tej usługi uwzględniający tego typu zbiory parametrów skojarzone z cechami sieci ( $S$ ), cechami urządzenia końcowego ( $UK$ ) oraz cechami użytkownika i jego otoczenia ( $U$ ). W modelu E parametr  $R$  przedstawiony jest jako liniowa kombinacja kilku współczynników z których każdy jest wyrażony przez wzór uwzględniający wybrane cechy z wymienionych zbiorów  $S$ ,  $UK$  i  $U$ .

$$R = f(S, UK, U) = R_o - I_s - I_d - I_e - eff + A, \quad (5.7)$$

gdzie:  $R_o$  uwzględnia stosunek sygnału do szumu pochodzącego z różnych źródeł w tym szum otoczenia oraz szum wykorzystanych układów elektronicznych,  $I_s$  ujmuje wpływ wszystkich czynników, które mają miejsce podczas transmisji sygnału mowy,  $I_d$  reprezentuje wpływ wszystkich zjawisk wynikających z opóźnienia sygnału w tym także echa,  $I_e - eff$  reprezentuje wpływ wyposażenia i stosowanych kodeków a w tym także zjawiska strat pakietów, współczynnik  $A$  uwzględnia wszystkie czynniki, które wpływają korzystnie na jakość sygnału. W zaleceniu (G.107) można znaleźć dokładny opis każdej z tych wielkości i odpowiadające im zależności analityczne. Zostały one zestawione w tabeli 5.5 wraz z podanymi wartościami dopuszczalnymi i domyślnymi.



Oznaczenia:  $U_x$  - Użytkownik  $x$ ,  $UK_x$  - Urządzenie końcowe  $x$  (Terminal  $x$ )

**Rysunek 5.12.** Ilustracja zbiorów parametrów uwzględnianych w modelu E



**Tabela 5.5.** Wartości domyślne parametrów modelu E (G.107)

Parametr	Skrót	Jednostka	Wartość domyślna	Dopuszczalne wartości	Współczynnik
Głośności strony nadawczej	SLR	dB	+8	0 ... +18	Ro
Głośności strony odbiorczej	RLR	dB	+2	-5 ... +14	Ro, Is, Id
Maskowanie efektu lokalnego	STMR	dB	15	10 ... 20	Is
Efekty lokalne strony odbiorczej	LSTR	dB	18	13 ... 23	Ro
Wartość D telefonu strony nadawczej	Ds	–	3	-3 ... +3	Ro
Wartość D telefonu strony odbiorczej	Dr	–	3	-3 ... +3	Ro
Głośność echa rozmówcy	TELR	dB	65	5 ... 65	Is
Tłumienie echa	WEPL	dB	110	5 ... 110	Id
Średnie opóźnienie	T	ms	0	0 ... 500	Is, Id
Opóźnienie w obie strony	Tr	ms	0	0 ... 1000	Id
Opóźnienie dla połączeń bez echa	Ta	ms	0	0 ... 500	Id
Czułość na opóźnienie	sT	-	1	0.4-1	Id
Akceptowalne opóźnienie	mT	ms	100	20-150	Id
Liczba jednostek zniekształcenia kwantyzacji	qdu	–	1	1 ... 14	Is
Wpływ wyposażenia (kodeka)	Ie	–	0	0 ... 40	Ie-eff
Odporność na stratę pakietu	Bpl	–	1	1 ... 40	Ie-eff
Prawdopodobieństwo straty pakietu	Ppl	%	0	0 ... 20	Ie, Ie-eff
Wybuchowość	BurstR	–	1	1 ... 2	Ie, Ie-eff
Szum układu	Nc	dBm0p	-70	-80 ... -40	Ro
Szum tła dla strony odbiorczej	Nfor	dBmp	-64	–	Ro
Szum otoczenia po stronie nadawczej	Ps	dB(A)	35	35 ... 85	Ro
Szum otoczenia po stronie odbiorczej	Pr	dB(A)	35	35 ... 85	Ro
Zysk	A	–	0	0 ... 20	-



Bazując na parametrach zawartych w tabeli 5.5 możemy wyznaczyć wszystkie metryki występujące we wzorze 5.7 i tym samym parametr  $R$ . Na podstawie  $R$  można określić wielkości określające jakości połączenia, takie jak: Good or Better ( $GoB$ ), Poor or Worse ( $PoW$ ), Mean Opinion Score ( $MOS$ ) (G.107).

$$GoB = 100 \cdot F\left(\frac{R-60}{16}\right) \quad [\%], \quad (5.8)$$

$$PoW = 100 \cdot F\left(\frac{45-R}{16}\right) \quad [\%], \quad (5.9)$$

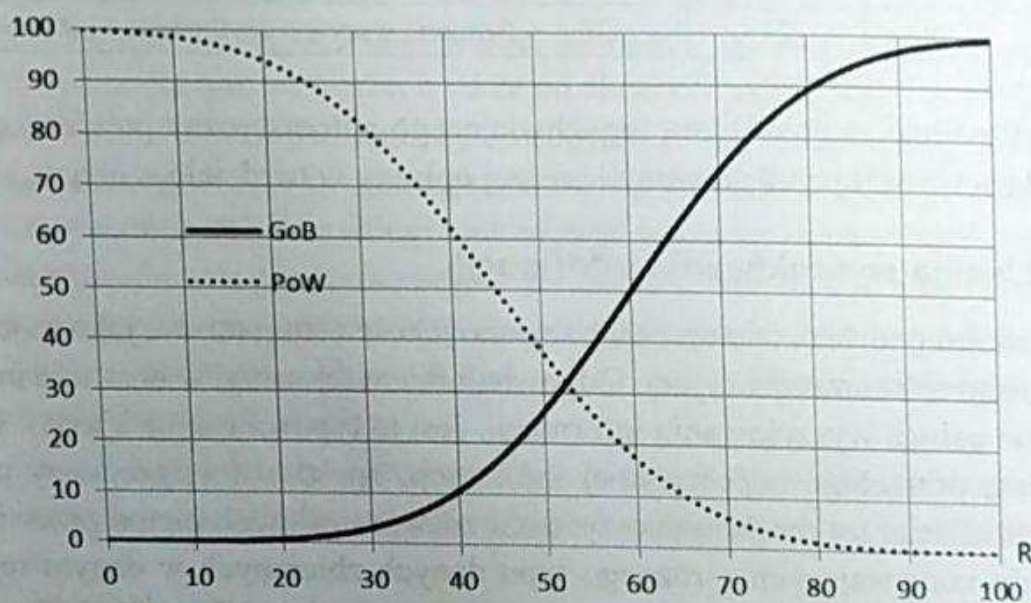
gdzie  $F(x)$  jest dystrybuantą rozkładu Gaussowskiego, tj.

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

Zależność między  $MOS$  a parametrem  $R$  jest następująca:

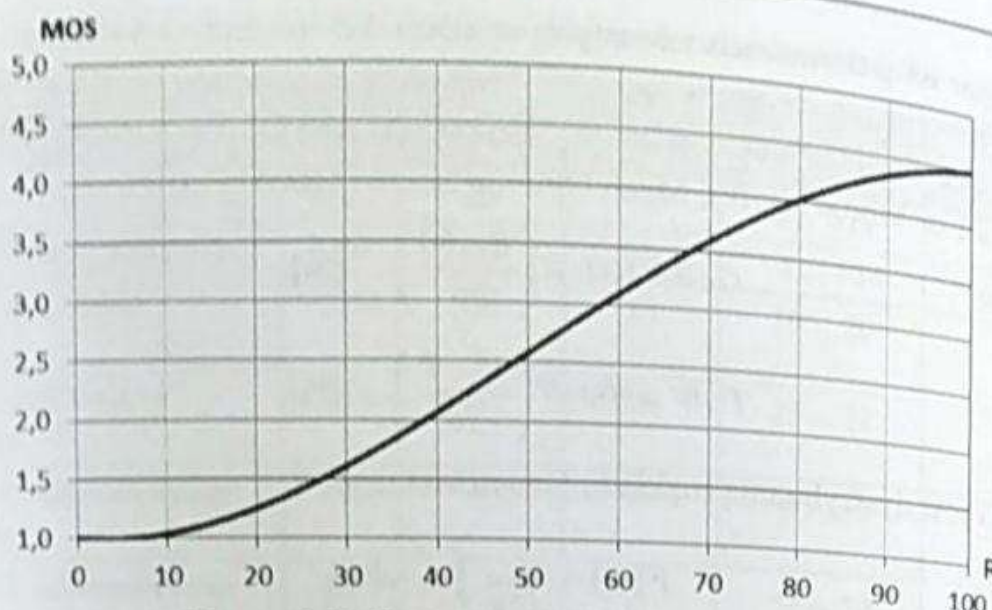
$$MOS = \begin{cases} 1 & \text{dla } R \leq 0 \\ 1 + 0.035 \cdot R + R \cdot (R-60) \cdot (100-R) \cdot 7 \cdot 10^{-6} & \text{dla } 0 < R < 100 \\ 4.5 & \text{dla } R = 100 \end{cases} \quad (5.10)$$

Na rysunkach 5.13 i 5.14 przedstawiono zależności (5.8), (5.9) i (5.10) w postaci graficznej, natomiast w tabeli 5.6 związek między odczuciami użytkownika (abonenta) a tymi parametrami usługi mowa.



**Rysunek 5.13.** Parametry  $GoB$  oraz  $PoW$  w funkcji parametru  $R$





Rysunek 5.14. Zależności MOS od parametru  $R$

Tabela 5.6. Zależność między parametrem  $R$  a innymi ocenami jakości usług

R	MOS <sub>cor</sub>	GoB [%]	PoW [%]	Odczucia użytkownika
90	4.34	97	~0	bardzo zadowolony
80	4.03	89	~0	zadowolony
70	3.60	73	6	niektórzy użytkownicy niezadowoleni
60	3.10	50	17	wielu użytkowników niezadowolonych
50	2.58	27	38	prawie wszyscy użytkownicy niezadowoleni

Przykład realizacji systemu umożliwiającego automatyczny pomiar tego parametru jakości został przedstawiony i szerzej opisany w rozdziale szóstym.

### 5.3.2. Usługa wyszukiwania informacji

W poprzednim podrozdziale przeanalizowano jakość usługi mowa, jako reprezentatywnej usługi telekomunikacyjnej. Odpowiednikiem tej usługi w obszarze informatyki będzie usługa wyszukiwania informacji. Jest to typowa usługa klient – serwer, gdzie klient potrzebuje odpowiedniej informacji, np. aktualnej prognozy pogody a serwer dostarcza jej dzięki wykorzystaniu zaawansowanych metod przewidywania pogody korzystających z różnego typu danych zbieranych w danym regionie geograficznym poprzez odpowiednie systemy monitoringu. Pogodę można przewidywać w okresie krótko, średnio i długoterminowym. Tak więc zadaniem serwera jest przygotowanie danego typu prognozy w formie strony internetowej i jej udostępnienie na żądanie. Ponieważ wielu użytkowników może korzystać z takiego portalu, dlatego typowy model obliczeń w tym przypadku dotyczy wielu klientów