Argument impedancji wejściowej obwodu kolektorowego przy częstotliwości podstawowej φ_1 wyraża się wzorem

$$\varphi_1 = \vartheta_1 - \gamma_1 - 180° \cong 35{,}942°, \tag{22}$$

a zatem impedancja ta ma charakter indukcyjny.

Wykres wskazowy prądów i napięć o częstotliwości podstawowej we wzmacniaczu klasy E przedstawiono na rys. 4.

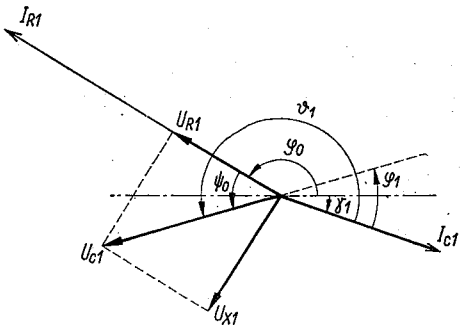
Moc zasilania P_{cc}, moc kolektorową o częstotliwości podstawowej P_{c1} oraz moc wyjściową wzmacniacza P_{wy} można wyznaczyć ze wzorów:

$$P_{cc} = I_{cc} U_{cc}, \tag{23}$$

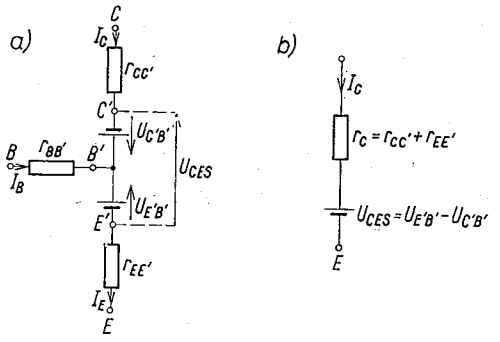
$$P_{c1} = \frac{1}{2} I_{c1} U_{c1} \cos\varphi_1 \cong 0{,}4048 I_{c1} U_{c1}, \tag{24}$$

$$P_{wy} = \frac{U_{R1}^2}{2R} = \frac{8}{\pi^2+4} \frac{U_{cc}^2}{R} \cong 0{,}5768 \frac{U_{cc}^2}{R}. \tag{25}$$

Przy założeniach 4), 5) i 6) moce P_{cc}, P_{c1} i P_{wy} są sobie równe, a sprawność kolektorowa η_c, sprawność obwodu kolektorowego η_o oraz sprawność kolektorowo-obwodowa η_{co} wynoszą 100%.



Rys. 4. Wykres wskazowy składowych podstawowych prądów i napięć we wzmacniaczu klasy E pracującym w warunkach optymalnych



Rys. 5. Model tranzystora w stanie nasycenia: a) model tranzystora w stanie nasycenia uwzględniający zaciski bazy, kolektora i emitera, b) model tranzystora w stanie nasycenia uwzględniający tylko zaciski kolektora i emitera

Współczynnik wydajności energetycznej wzmacniacza wynosi

$$c_p = \frac{P_{c1}}{i_{c\,max} u_{c\,max}} \cong 0{,}0981. \tag{26}$$

Maksymalna moc kolektorowa o częstotliwości podstawowej $P_{c1\,max}$ przy pełnym wykorzystaniu prądowym i napięciowym tranzystora wyraża się wzorem

$$P_{c1\,max} \cong 0{,}0981 I_{c\,max} U_{c\,max}, \tag{27}$$

przy czym:

$I_{c\,max}$ — maksymalna, dopuszczalna wartość prądu kolektora,

$U_{c\,max}$ — maksymalna, dopuszczalna wartość napięcia kolektora (napięcia kolektor-baza).

Maksymalna, dopuszczalna wartość napięcia kolektora U_{cmax} w rozpatrywanym wzmacniaczu równa się w przybliżeniu napięciu przebicia kolektor-baza przy odłączonym emiterze V_{CBO}, mimo że tranzystor pracuje w układzie wspólnego emitera. Maksymalna wartość chwilowa napięcia kolektora u_{cmax} występuje bowiem przy zatkanym złączu baza-emiter.

2.3. Straty mocy w kolektorze i sprawność kolektorowa wzmacniacza

W dotychczasowej analizie pominięto straty mocy w kolektorze związane z niezerowym napięciem kolektora w stanie nasycenia tranzystora oraz z czasem opadania prądu kolektora Δt_f przy wyłączaniu tranzystora.

W celu oszacowania pierwszego z wymienionych składników można przyjąć uproszczony model tranzystora w stanie nasycenia pokazany na rys. 5 [13], [17]. W modelu tym rezystancja nasycenia tranzystora r_c równa się sumie rezystancji rozproszonych obszarów kolektora $r_{CC'}$ i emitera $r_{EE'}$, a napięcie nasycenia tranzystora U_{CES} równa się różnicy napięć złączowych $U_{B'E'}$ i $U_{B'C'}$. Straty mocy w kolektorze powstają wskutek przepływu prądu kolektora i_c przez rezystancję nasycenia tranzystora r_c oraz przez źródło napięciowe U_{CES} w kierunku przeciwnym do wymuszenia tego źródła. W praktycznych układach wzmacniacza reaktancja kondensatora C jest na tyle duża, że w stanie nasycenia tranzystora przez kondensator ten płynie pomijalnie mały prąd, a prawie cała różnica prądów I_{cc} oraz i_{R1} płynie przez tranzystor [19]. Można zatem przyjąć założenie, że prąd kolektora i_c nadal wyraża się wzorem (5). Założenie to prowadzi do nieznacznego zawyżenia mocy strat w kolektorze. Rezystancja nasycenia tranzystora r_c jest w przybliżeniu liniowa. Moc strat $P_{cr\,str}$ wydzielana w tej rezystancji wynosi

$$P_{cr\,str} = \frac{r_c}{2\pi}\int_0^{\pi} i_c^2(\omega t)d(\omega t) = \frac{\pi^2+28}{16} r_c I_{cc}^2 = \frac{\pi^2+28}{2(\pi^2+4)}\frac{r_c}{R} P_{cc} \cong$$

$$\cong 2{,}3669 r_c I_{cc}^2 \cong 1{,}3652 \frac{r_c}{R} P_{cc}. \tag{28}$$

Oszacowanie mocy strat $P_{cu\,str}$ związanej z napięciem nasycenia tranzystora U_{CES} jest bardziej kłopotliwe, ponieważ napięcie to jest nieliniowe. Dokładne wyznaczenie mocy $P_{cu\,str}$ jest więc możliwe przez numeryczne całkowanie chwilowych przebiegów prądu kolektora i_c oraz napięcia nasycenia u_{CES}. Statyczne napięcie U_{CES} wyznacza się na podstawie pomiarów lub na podstawie wielkosygnałowego modelu tranzystora Ebersa-Molla. Do obliczeń inżynierskich można przyjąć stałą wartość napięcia nasycenia U_{CES} [18]. Wówczas moc strat związana z tym napięciem wynosi

$$P_{cu\,str} = I_{cc} U_{CES}. \tag{29}$$

Wyłączanie tranzystora w rozpatrywanym wzmacniaczu, jak już wspomniano, odbywa się poprzez stan aktywny. W czasie Δt_f następuje opadanie prądu kolektora i_c oraz narastanie napięcia kolektora u_c, czemu towarzyszą straty mocy w kolektorze (rys. 2a).

Tranzystor w tym czasie zachowuje się jak źródło prądowe. Model wzmacniacza obowiązujący w czasie Δt_f przedstawiono na rys. 6. W celu wyznaczenia mocy strat w kolektorze $P_{cf\,str}$ związanej z czasem Δt_f należy opisać najpierw przebiegi prądu kolektora i_c oraz napięcia kolektora u_c. Z obserwacji przebiegu prądu kolektora i_c w układzie praktycznym wzmacniacza wynika, że prąd ten w czasie Δt_f jest liniowo opadający. Pomijając czas magazynowania Δt_s oraz obliczając ze wzoru (5) wartość prądu kolektora w chwili rozpoczęcia się stanu aktywnego $i_c(\pi) = 2I_{cc}$ można napisać wzór

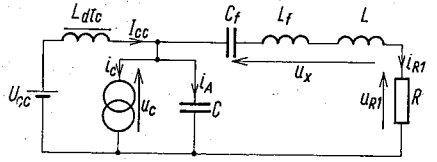
$$i_c(\omega t) = 2I_{cc}\left(1-\frac{\omega t-\pi}{\Theta_f}\right) \quad \text{dla} \quad \pi < \omega t \leqslant \pi+\Theta_f, \tag{30}$$

w którym $\Theta_f = \omega \Delta t_f$.

Jak wynika z rys. 6, prąd i_A płynący przez kondensator C jest wymuszany przez prądy i_c, i_{R1} oraz I_{cc}. Stąd otrzymuje się wzór

$$i_A(\omega t) = I_{cc}-i_{R1}(\omega t)-i_c(\omega t) =$$

$$= I_{cc}\left(\frac{\pi}{2}\sin\omega t-\cos\omega t+\frac{2\omega t}{\Theta_f}-\frac{2\pi}{\Theta_f}-1\right) \quad \text{dla} \quad \pi < \omega t \leqslant \pi+\Theta_f. \tag{31}$$



Rys. 6. Model wzmacniacza klasy E w czasie opadania prądu kolektora Δt_f

Na podstawie prądu i_A można wyznaczyć napięcie na kondensatorze C, które jest jednocześnie napięciem kolektora

$$u_c(\omega t) = \frac{1}{\omega C}\int_{\pi}^{\omega t} i_A(\tau)d\tau =$$

$$= U_{cc}\pi\left[\frac{(\omega t)^2}{\Theta_f}-\left(\frac{2\pi}{\Theta_f}+1\right)\omega t+\frac{\pi^2}{\Theta_f}+\frac{\pi}{2}-\sin\omega t-\frac{\pi}{2}\cos\omega t\right] \tag{32}$$

dla $\pi < \omega t \leqslant \pi+\Theta_f$.

Moc strat w kolektorze $P_{cf\,str}$ związana z czasem opadania prądu kolektora Δt_f wynosi

$$P_{cf\,str} = \frac{1}{2\pi}\int_{\pi}^{\pi+\Theta_f} i_c(\omega t)u_c(\omega t)d(\omega t) =$$

$$= P_{cc}\left[\frac{\pi(1-\cos\Theta_f)}{2\Theta_f}-\frac{\sin\Theta_f}{\Theta_f}-\frac{\Theta_f}{4}\left(\pi+\frac{\Theta_f}{3}\right)+1\right]. \tag{33}$$

Całkowita moc strat w kolektorze $P_{c\,str}$ równa się sumie poszczególnych składników: $P_{cr\,str}$, $P_{cu\,str}$ i $P_{cf\,str}$.

Sprawność kolektorową wzmacniacza η_c po uwzględnieniu wzorów (28), (29) i (33) można przedstawić w postaci

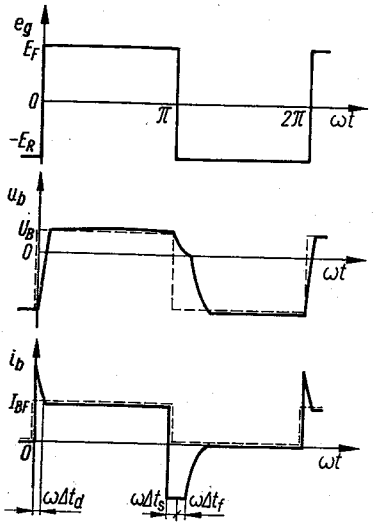
$$\eta_c = 1 - \frac{P_{cstr}}{P_{cc}} =$$

$$= \frac{\sin\Theta_f}{\Theta_f} + \frac{\Theta_f}{4}\left(\pi + \frac{\Theta_f}{2}\right) - \frac{\pi(1-\cos\Theta_f)}{2\Theta_f} - \frac{\pi^2+28}{16}\frac{r_c I_{cc}}{U_{cc}} - \frac{U_{CES}}{U_{cc}}. \tag{34}$$

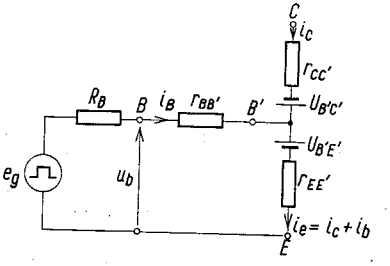
Ze wzoru tego wynika, że w celu uzyskania dużej sprawności kolektorowej wzmacniacza η_c należy wybierać możliwie duże napięcie zasilania U_{cc}. Wówczas przy zadanej mocy wyjściowej P_{wy} uzyskuje się mały prąd zasilania I_{cc}. Wybór taki ma korzystny wpływ na minimalizację mocy strat w kolektorze. Moc P_{crstr}, zgodnie ze wzorem (28), maleje z kwadratem prądu I_{cc}. Moc P_{custr}, zgodnie z przybliżonym wzorem (29), maleje proporcjonalnie ze zmniejszaniem się prądu I_{cc}. W rzeczywistości zależność ta jest dużo silniejsza, ponieważ napięcie nasycenia tranzystora U_{CES} jest nieliniowe i maleje ze zmniejszaniem się prądu kolektora (rys. 15).

2.4. Moc wejściowa i wzmocnienie mocy

W celu oszacowania mocy wejściowej wzmacniacza P_{we} należy określić, przynajmniej w przybliżeniu, chwilowe przebiegi napięcia bazy u_b oraz prądu bazy i_b. Przebiegi te oprócz parametrów tranzystora i obwodu kolektorowego zależą od kształtu napięcia wyjściowego i rezystancji wewnętrznej generatora sterujcego. W praktyce jako stopień sterujący najdogodniej jest zastosować układ aperiodyczny o prostokątnym napięciu wyjściowym, na przykład inwerter napięciowy. Rezystancja wyjściowa tych układów jest zwykle stosunkowo duża w porównaniu z rezystancją wejściową sterowanego tranzystora znajdującego się w stanie nasycenia. Na rys. 7 przedstawiono przybliżone przebiegi występujące w obwodzie bazy we wzmacniaczu klasy E z uwzględnieniem procesów przejściowych związanych z przełączaniem tranzystora. Przebiegi te są analogiczne jak w układzie przełączania tranzystora obciążonego rezystancją [13], [16], z tym jednak, że we wzmacniaczu klasy E czas narastania prądu kolektora Δt_r wynosi zero. Jak widać z przedstawionego rysunku, dokładny opis matematyczny mocy wejściowej P_{we} doprowadzi do dość skomplikowanych wzorów, mało przydatnych w praktyce. Jeżeli jednak czas procesów przejściowych jest dostatecznie mały w stosunku do połowy okresu, wówczas można ograniczyć się do analizy stałoprądowej. Schemat zastępczy obwodu bazy w stanie nasycenia tranzystora pokazano na rys. 8, w którym wykorzystano statyczny model tranzystora z rys. 5a. W stanie nasycenia tranzystora zmienia się napięcie $U_{B'E'}$ (źródło nieliniowe) oraz spadek napięcia na rezystancji $r_{EE'}$, ponieważ płynie przez nie prąd zmienny kolektora i_c opisany wzorem (5). Jest to zasadnicza różnica w stosunku do oporowego układu przełączania tranzystora. Na podstawie rys. 8 można napisać dokładne równania opisujące prąd bazy i_b oraz napięcie bazy u_b. Jednakże w związku z nieliniowością źródła $U_{B'E'}$ moc wejściową P_{we} praktycznie można obliczyć tylko numerycznie. W celu uproszczenia tego problemu na rys. 7 zaznaczono liniami przerywanymi aproksymowane przebiegi i_b oraz u_b. Moc wejściowa wzmacniacza w tym przypadku wyraża się wzorem



Rys. 7. Przebiegi napięć i prądu w obwodzie bazy wzmacniacza klasy E



Rys. 8. Schemat zastępczy obwodu bazy w stanie nasycenia tranzystora

$$P_{we} = \frac{U_B I_{BF}}{2}. \tag{35}$$

Napięcie U_B dla tranzystorów krzemowych jest rzędu $0{,}7 \div 1$ V, a prąd I_{BF} przy dużej ezystancji R_B w przybliżeniu wynosi

$$I_{BF} = \frac{E_F - U_B}{R_B}. \tag{36}$$

W celu zapewnienia pracy tranzystora w stanie nasycenia prąd I_{BF} powinien spełniać arunek

$$I_{BF} > \frac{I_{CM}}{\beta_o}, \tag{37}$$

rzy czym $I_{CM} = i_{cmax}$. Wymaganą wartość prądu I_{BF} można ustalić przez regulację ezystancji R_B.

Wzmocnienie mocy wzmacniacza k_p po uwzględnieniu zależności (25) i (35) wyraża się zorem

$$k_p = \frac{P_{wy}}{P_{we}} = \frac{16}{\pi^2 + 4} \frac{U_{cc}^2}{U_B I_{BF} R} \cong 1{,}1536 \frac{U_{cc}^2}{U_B I_{BF} R}. \tag{38}$$

3. PROJEKTOWANIE WZMACNIACZA KLASY E

Przy projektowaniu wzmacniacza klasy E o zadanych parametrach można dokonać yboru tranzystora oraz obliczyć interesujące wartości prądów, napięć i parametrów ergetycznych na podstawie wzorów zamieszczonych w rozdz. 2. Wprawdzie nie zawsze katalogach są podawane parametry tranzystora konieczne do określenia mocy strat

w kolektorze, takie jak r_c, U_{CES} i Δt_f, jednakże w rozpatrywanym wzmacniaczu nie występuje ograniczenie dopuszczalną mocą strat. Stąd wielkość radiatora może być dobrana na etapie eksperymentalnym.

Wartości elementów obwodu kolektorowego układu podstawowego wzmacniacza z rys. 1 wyznacza się ze wzorów:

$$R = \frac{8}{\pi^2+4}\frac{U_{cc}^2}{P_{wy}} \cong 0{,}5768\frac{U_{cc}^2}{P_{wy}}, \tag{39}$$

$$X_C = \frac{\pi(\pi^2+4)}{8}R \cong 5{,}4466R, \tag{40}$$

$$X_{L_w} = Q_w R, \tag{41}$$

$$X_{C_f} = \left[Q_w - \frac{\pi}{16}(\pi^2-4)\right]R \cong (Q_w - 1{,}1525)R. \tag{42}$$

We wzorze (39) rezystancja strat własnych cewki L_w wchodzi w skład rezystancji R a we wzorze (40) pojemność wyjściowa tranzystora, pojemność własna dławika zasilającego kolektor i pojemność montażu wchodzą w skład pojemności C. Dobroć wypadkowa szeregowego obwodu rezonansowego Q_w, którą należy założyć, powinna mieścić się w granicach od 5 do 15.

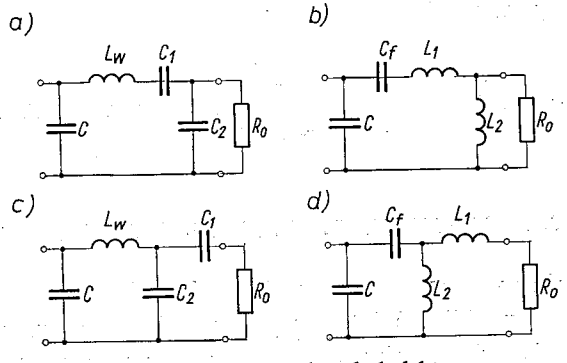
Reaktancja dławika zasilającego kolektor powinna spełniać warunek

$$X_{dł} \geqslant (50 \div 100)R. \tag{43}$$

Obwód kolektorowy w podstawowym układzie wzmacniacza przedstawionym na rys. nie posiada właściwości dopasowujących. Obwody posiadające takie właściwości można otrzymać przez modyfikację szeregowego obwodu rezonansowego. Przykładowe schematy kolektorowych obwodów dopasowujących przedstawiono na rys. 9. Wzory opisujące wartości elementów tych obwodów można wyprowadzić wykorzystując metodę dwójników równoważnych [14], [15].

W przypadku obwodu π1a (rys. 9a) nadal pozostają słuszne wzory (39), (40) i (41). Reaktancje kondensatorów C_1 i C_2 oblicza się ze wzorów:

$$X_{C1} = \left[Q_w - \sqrt{\frac{R_o}{R}-1} - \frac{\pi}{16}(\pi^2-4)\right]R \cong \left(Q_w - \sqrt{\frac{R_o}{R}-1} - 1{,}1525\right)R, \tag{44}$$



Rys. 9. Dopasowujące obwody kolektorowe:
a) obwód π1a, b) obwód π2a, c) obwód π1b, d) obwód π4a

$$X_{C2} = \frac{R_o}{\sqrt{\frac{R_o}{R}-1}}. \tag{45}$$

Za pomocą rozważanego obwodu, jak wynika ze wzorów (44) i (45), można dopasować rezystancje obciążenia wzmacniacza R_o większe od optymalnej rezystancji dopasowania tranzystora R, czyli

$$R_o > R. \tag{46}$$

W obwodzie tym w celu regulacji rezystancji dopasowania R należy zmieniać pojemność C_2, a w celu dostrojenia wzmacniacza jeden z elementów C_1 lub L_w.

W przypadku obwodu $\pi 2a$ nadal pozostają słuszne wzory (39), (40) i (42). Reaktancje cewek L_1 i L_2 oblicza się ze wzorów:

$$X_{L1} = \left(Q_w - \sqrt{\frac{R_o}{R}-1}\right)R, \tag{47}$$

$$X_{L2} = \frac{R_o}{\sqrt{\frac{R_o}{R}-1}}. \tag{48}$$

Zakres rezystancji dopasowania tego obwodu wyraża się nierównością (46). Regulację rezystancji dopasowania R należy przeprowadzić przez zmianę indukcyjności L_2, a dostrojenie wzmacniacza przez zmianę pojemności C_f lub indukcyjności L_1.

W przypadku obwodu $\pi 1b$ pozostają słuszne wzory (39), (40) i (41). Reaktancje kondensatorów C_1 i C_2 oblicza się ze wzorów:

$$X_{C1} = R_o\sqrt{\frac{R[(Q_w-1{,}1525)^2+1]}{R_o}-1}, \tag{49}$$

$$X_{C2} \cong \frac{R[(Q_w-1{,}1525)^2+1]}{Q_w-1{,}1525-\sqrt{\frac{R[(Q_w-1{,}1525)^2+1]}{R_o}-1}}. \tag{50}$$

a pomocą tego obwodu można dopasować rezystancje spełniające nierówność

$$R_o < R[(Q_w-1{,}1525)^2+1]. \tag{51}$$

W przypadku obwodu $\pi 4a$ pozostają słuszne wzory (39), (40) i (42). Reaktancje cewek L_1 i L_2 oblicza się ze wzorów:

$$X_{L1} = R_o\sqrt{\frac{R(Q_w^2+1)}{R_o}-1}, \tag{52}$$

$$X_{L2} = \frac{R(Q_w^2+1)}{Q_w-\sqrt{\frac{R(Q_w^2+1)}{R_o}-1}}. \tag{53}$$

a pomocą tego obwodu można dopasować rezystancje spełniające nierówność

$$R_o < R(Q_w^2+1). \tag{54}$$

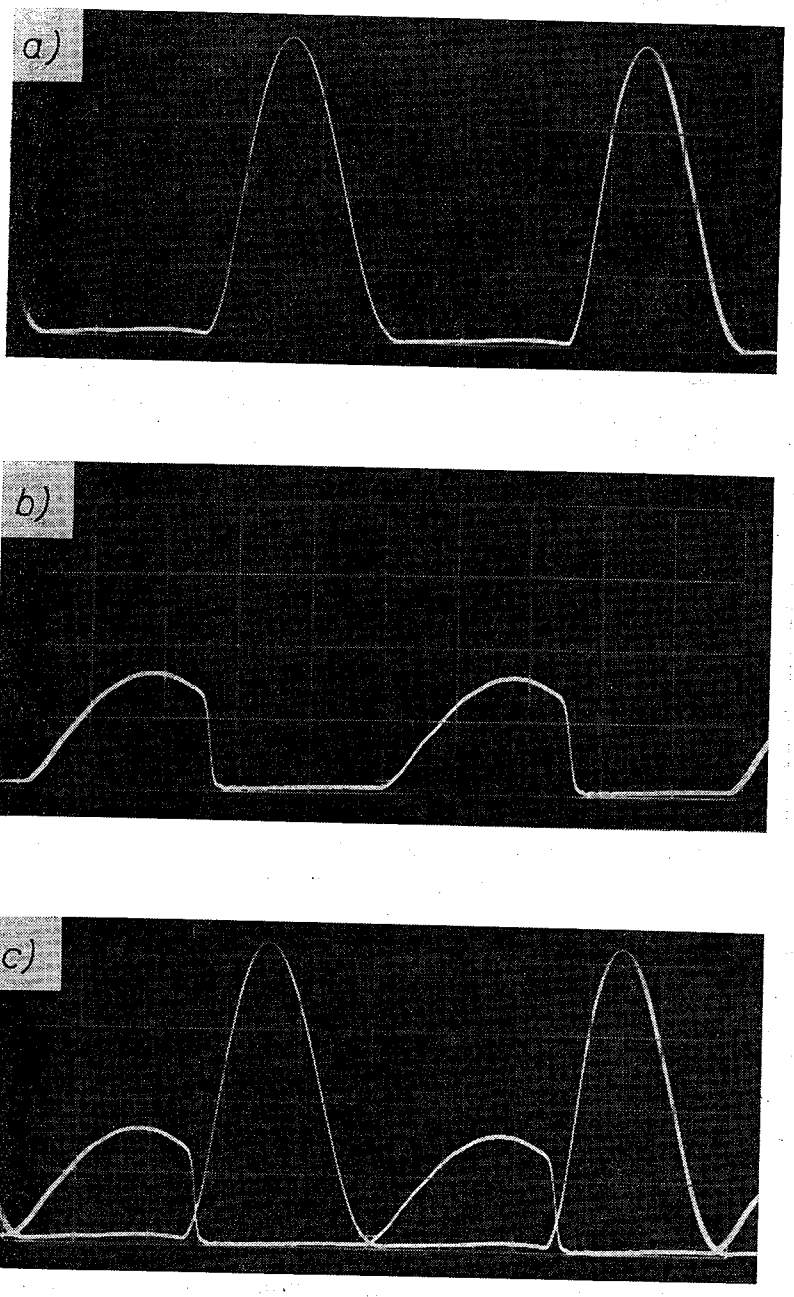
4. BADANIA EKSPERYMENTALNE

W celu sprawdzenia zgodności wyników teoretycznych i eksperymentalnych, takich jak chwilowe przebiegi prądów i napięć, parametry energetyczne oraz optymalne wartości elementów obwodu kolektorowego, zaprojektowano i wykonano układ doświadczalny wzmacniacza klasy E o schemacie pokazanym na rys. 1a. W układzie tym zastosowano tranzystor *BSXP* 60. Częstotliwość robocza f wynosiła 2 MHz. Wzmacniacz sterowano z generatora impulsów prostokątnych o współczynniku wypełnienia 50%. Rezystancja wyjściowa generatora wynosiła 50 Ω. Wyjście generatora sprzężono z bazą tranzystora przez kondensator o pojemności 6,8 nF. Między bazę i emiter włączono dławik o indukcyjności 2,5 mH zamykający obwód składowej stałej prądu bazy. Indukcyjność dławika kolektorowego wynosiła 0,7 mH. Wartości elementów obwodu kolektorowego oraz parametry energetyczne obliczone teoretycznie i zmierzone w układzie doświadczalnym zestawiono w tabl. 1. Wypadkowa rezystancja obciążenia R składała się z rezystancji zewnętrznej równej 25,1 Ω oraz rezystancji własnej cewki L_w równej 1,5 Ω. W wypadkowej

Tablica 1

Wyniki teoretyczne i doświadczalne podstawowego układu wzmacniacza klasy E sterowanego napięciem prostokątnym

		Wyniki teoretyczne	Wyniki doświadczalne	Graniczny błąd względny pomiaru w %	Różnica względna w %
R		26,6	26,6	1,88	0
C	pF	549,5	549,8	1,1	0,06
C_f	pF	511,9	511,3	0,12	0,12
L_w	μH	14,82	14,78	4,12	0,27
U_{cc}	V	12	12,01	0,13	0,08
I_{cc}	mA	260	260,3	0,34	0,12
P_{cc}	W	3,12	3,126	0,36	0,19
$P_{c\,str}$	mW	172	169	4,3	1,74
P_{c1}	W	2,948	2,957	0,46	0,32
P_{wy}	W	2,782	2,788	8,8	0,22
η_c	%	94,45	94,59	5,7	0,15
η_o	%	94,36	94,28	8,9	0,09
η_{co}	%	89,12	89,18	8,9	0,07



ys. 10. Optymalne przebiegi napięcia kolektora u_c oraz prądu kolektora i_c we wzmacniaczu klasy E przy sterowaniu napięciem prostokątnym:

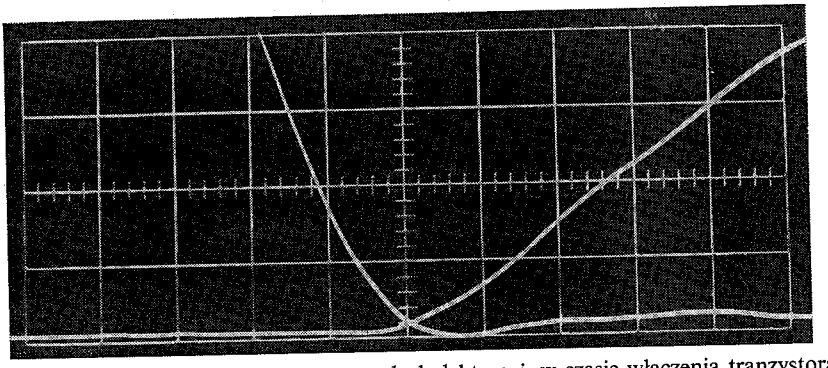
przebieg napięcia kolektora u_c, b) przebieg prądu kolektora i_c, c) współzależności czasowe pomiędzy przebiegami u_c oraz i_c (skala pionowa 10 V/dz i 0,5 A/dz, skala pozioma 0,1 μs/dz)

pojemności C uwzględniono pojemność wyjściową tranzystora 10 pF, pojemność własną dławika kolektorowego 5,3 pF oraz pojemność wejściową sondy napięciowej oscyloskopu 8,2 pF. Dobroć wypadkowa szeregowego obwodu rezonansowego Q_w wynosiła 7.

Pomiar mocy strat w kolektorze P_{cstr} wykonano przy użyciu termistora. Skalowanie termistora oraz pomiar mocy P_{cstr} przeprowadzono przy jednakowej temperaturze otoczenia w specjalnej komorze termicznej. Napięcie U_{cc} oraz prąd I_{cc} mierzono mierni cyfrowymi. Moc wyjściową P_{wy} wyznaczono na podstawie pomiaru napięcia wyjścio woltomierzem wielkiej częstotliwości.

W tabl. 1 zamieszczono graniczne błędy pomiarów oraz różnice względne wyn teoretycznych i eksperymentalnych. Z tablicy tej wynika, że zgodność wspomnianych wyników jest zadowalająca w praktyce.

Obserwacje chwilowych przebiegów napięć i prądów przeprowadzono za pomocą oscyloskopu wyposażonego w sondy napięciowe i prądowe. W rozpatrywanym wzmacniaczu szczególnie istotne są przebiegi napięcia kolektora u_c oraz prądu kolektora i_c. Przebiegi te przedstawiono na rys. 10. Na rys. 11 pokazano powiększony fragment tych prze



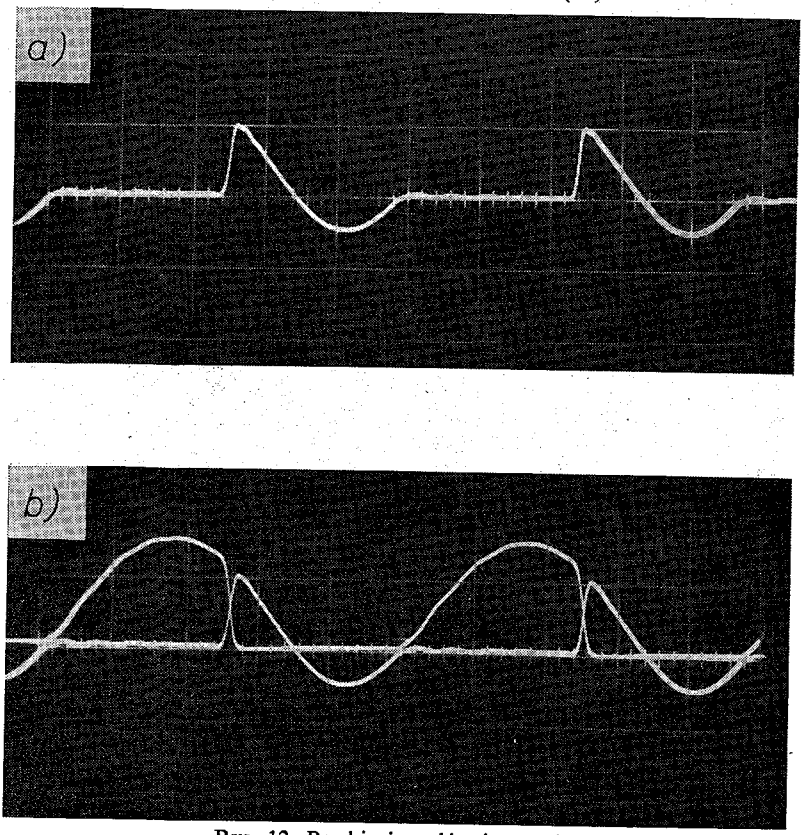
Rys. 11. Przebiegi napięcia kolektora u_c oraz prądu kolektora i_c w czasie włączenia tranzystora (powiększony fragment rys. 10c)

biegów w czasie włączania tranzystora. Z obu tych rysunków wynika, że w praktycznym układzie wzmacniacza można spełnić warunki (3) i (4) zapewniające wyeliminowanie mocy strat w kolektorze związanej z włączaniem tranzystora; napięcie kolektora i pochodna tego napięcia w chwili włączania tranzystora wynoszą bowiem zero. Przy włączaniu tranzystora nie występują szybkie zmiany napięcia i prądu kolektora. Napięcie kolektora u_c przed włączeniem tranzystora łagodnie opada do zera, a prąd kolektora i_c po włączeniu tranzystora łagodnie narasta od zera. Ponadto przebiegi te nie przyjmują wartości ujemnych. Tranzystor przy przejściu ze stanu odcięcia do normalnego stanu nasycenia nie wchodzi więc w inne stany pracy. Ponieważ prąd kolektora i_c oraz napięcie kolektora u_c w chwili włączania tranzystora praktycznie wynoszą zero (rys. 11), stąd moc strat w kolektorze związana z włączaniem tranzystora jest pomijalnie mała. Istotny jest również fakt, że optymalne przebiegi prądu i napięcia kolektora otrzymano przy wartościach elementów obwodu kolektorowego zgodnych z obliczeniami teoretycznymi (tabl. 1).

Na rys. 10 można zaobserwować obie przyczyny występowania mocy strat w kolektorze, a mianowicie w stanie nasycenia tranzystora napięcie kolektora przyjmuje warto

niezerowe oraz przy wyłączaniu tranzystora, w czasie Δt_f, następuje opadanie prądu lektora i narastanie napięcia kolektora. Prąd kolektora w czasie Δt_f opada liniowo, otwierdza słuszność założenia przyjętego w rozdz. 2.3.

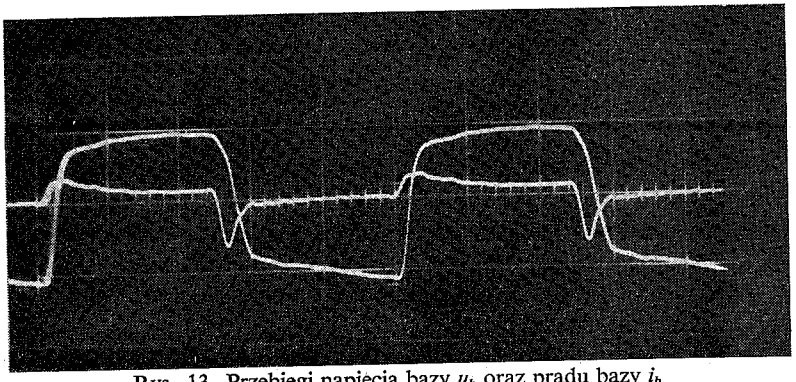
a rys. 12a pokazano przebieg prądu i_A płynącego przez kondensator C, a na rys. 12b egi prądów i_A oraz i_c. Jak widać, opadaniu prądu i_c w czasie Δt_f towarzyszy nae prądu i_A. Ponadto suma prądów i_c oraz i_A ma kształt sinusoidy przesuniętej dową stałą I_{cc}, co potwierdza słuszność równania (31).



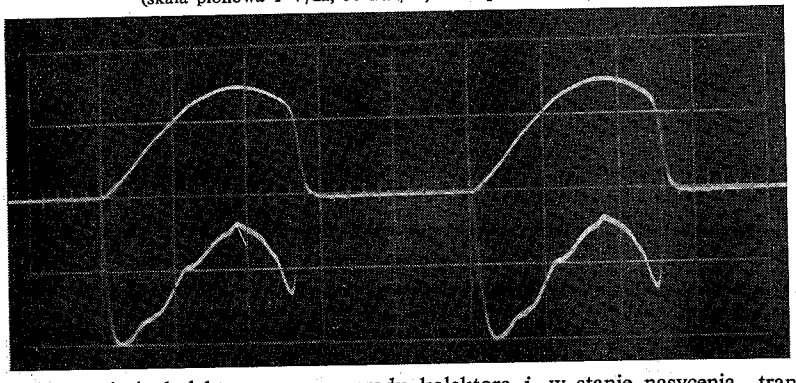
Rys. 12. Przebiegi prądów i_A oraz i_c
przebieg prądu i_A płynącego przez kondensator C, b) przebiegi prądów i_c oraz i_A (skala pionowa 0,5 A/dz, skala pozioma 0,1 μs/dz)

Na rys. 13 przedstawiono przebiegi napięcia bazy u_b oraz prądu bazy i_b. W czasie łączenia tranzystora prąd bazy i_b przyjmuje w przybliżeniu stałą wartość równą 10 mA. rzy wyłączaniu tranzystora następuje usuwanie ładunku z obszaru bazy i wówczas prąd azy przyjmuje wartości ujemne.

Powróćmy do zagadnienia mocy strat w kolektorze. Jak już wspomniano, jeden ze składików tej mocy wynika z niezerowego napięcia kolektora w stanie nasycenia tranzystora. hwilowy przebieg tego napięcia pokazano na rys. 14. W celu oszacowania wspomnianej ocy zmierzono statycznie rezystancję nasycenia tranzystora r_c oraz napięcie nasycenia anzystora U_{CES}. Rezystancja r_c wynosiła 0,51 Ω, a wyniki pomiarów napięcia U_{CES} zedstawiono na rys. 15. Jak widać z tego rysunku, napięcie nasycenia U_{CES} zależy silnie



Rys. 13. Przebiegi napięcia bazy u_b oraz prądu bazy i_b
(skala pionowa 1 V/dz, 50 mA/dz, skala pozioma 0,1 μs/dz)



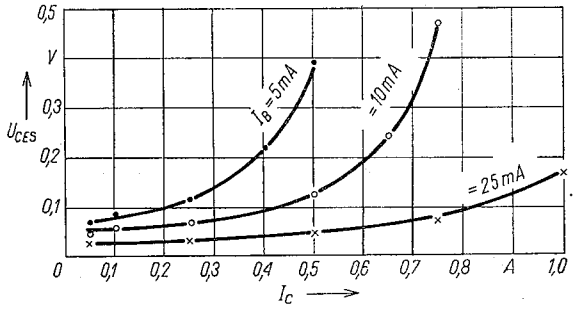
Rys. 14. Przebieg napięcia kolektora u_c oraz prądu kolektora i_c w stanie nasycenia tranzystora
(skala pionowa 0,5 V/dz i 0,5 A/dz, skala pozioma 0,1 μs/dz)

nieliniowo od prądu kolektora I_C oraz prądu bazy I_B. Wykorzystując wzór (28) obliczo no moc strat $P_{cr\,str}$ związaną z rezystancją nasycenia tranzystora r_c

$$P_{cr\,str} = 82 \text{ mW}.$$

Moc strat $P_{cu\,str}$ związaną z napięciem nasycenia tranzystora obliczono metodą numerycz nego całkowania iloczynu chwilowych wartości prądu kolektora i_c oraz napięcia nasyceni u_{CES}, wyznaczonego na podstawie charakterystyki statycznej U_{CES} (I_C) dla $I_B = 10$ m (rys. 15) i równania (5). W wyniku numerycznego całkowania otrzymano

$$P_{cu\,str} = 71 \text{ mW}.$$



Rys. 15. Wyniki pomiarów napięcia nasycenia tranzystora U_{CES}

Na rys. 16 pokazano chwilowe przebiegi prądu kolektora i_c oraz napięcie kolektora u_c w czasie wyłączania tranzystora. Czas opadania prądu kolektora Δt_f odczytany z tego rysunku wynosi 25 ns, co odpowiada kątowi fazowemu $\Theta_f = 18°$. Stąd wykorzystując wzór (33) otrzymano

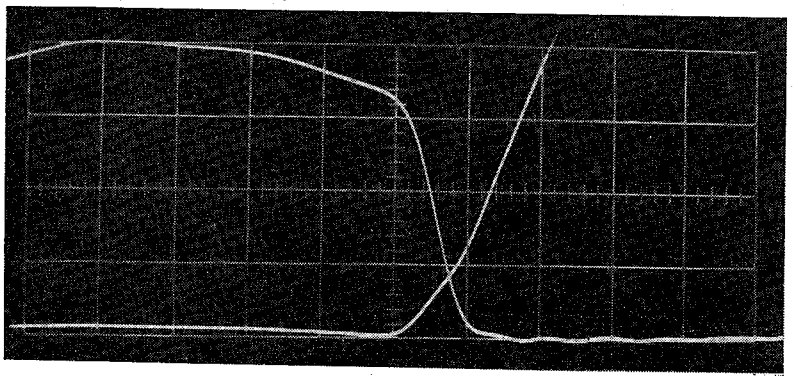
$$P_{cf\,str} = 19 \text{ mW}.$$

Całkowita moc strat w kolektorze $P_{c\,str}$ wynosiła

$$P_{c\,str} = P_{cr\,str} + P_{cu\,str} + P_{cf\,str} = 172 \text{ mW}.$$

Z tabl. 1 wynika, że wartość tej mocy zgadza się z zadowalającą dokładnością z wartością zmierzoną.

Moc wejściową wzmacniacza P_{we} można oszacować na podstawie rys. 13. Przyjmując napięcie bazy w stanie nasycenia tranzystora $U_B = 1$ V oraz prąd bazy $I_{BF} = 10$ mA ze wzoru (35) otrzymano $P_{we} = 5$ mW. Stąd wzmocnienie mocy wzmacniacza k_p wynosiło w przybliżeniu 550 W/W.



Rys. 16. Przebiegi prądu kolektora i_c oraz napięcia kolektora u_c w czasie wyłączania tranzystora (skala pozioma 0,02 μs/dz)

W omawianym dotychczas układzie doświadczalnym przeprowadzono analogiczne adania z kilkudziesięcioma egzemplarzami tranzystorów typu *BSXP* 60, *BSXP* 59 *BSXP* 61. Okazało się, że parametry energetyczne oraz przebiegi prądu i napięcia ko-ktora są mało wrażliwe na parametry tranzystora i mieszczą się w granicach błędów omiarowych. Zauważone różnice dotyczyły głównie napięcia kolektora w stanie nasy-enia tranzystora oraz czasu opadania prądu kolektora Δt_f, a w związku z tym również ocy strat w kolektorze $P_{c\,str}$.

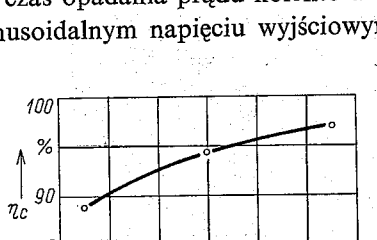
Stwierdzono ponadto, że dobór pojemności C jest mało krytyczny. Stąd tolerancja ojemności kondensatora C nie musi być uwzględniana przy projektowaniu wzmacniacza. totne znaczenie ma natomiast właściwe dostrojenie szeregowego obwodu rezonansowe-o L_w, C_f i R. W celu precyzyjnego dostrojenia wzmacniacza do pożądanej częstotliwości boczej należy regulować w niewielkim zakresie jeden z elementów C_f lub L_w i obserwo-ać na oscyloskopie przebieg napięcia kolektora. Jako kryterium dostrojenia należy zyjąć optymalny przebieg tego napięcia, taki jak przedstawiono na rys. 10a.

Następnie zbadano doświadczalnie zależność sprawności kolektorowej η_c od napięcia silania U_{cc}. W tym celu wykonano trzy modele wzmacniaczy o jednakowej mocy wyj-

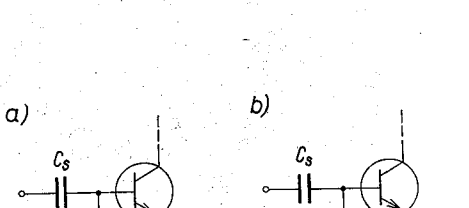
ściowej $P_{wy} = 3$ W i o różnych napięciach zasilania U_{cc}. Wyniki pomiarów sprawności kolektorowej η_c tych wzmacniaczy przedstawiono na rys. 17. Z rysunku tego wynika, że ze wzrostem napięcia zasilania U_{cc} rośnie sprawność kolektorowa wzmacniacza η_c, np. przy $U_{cc} = 17$ V osiągnięto $\eta_c = 97\%$. Wniosek ten jest zgodny z wnioskiem wynikającym z rozważań teoretycznych zamieszczonych w rozdz. 2.3.

W celu sprawdzenia praktycznej przydatności dopasowujących obwodów kolektorowych przedstawionych w rozdz. 4 zaprojektowano i wykonano modele wzmacniaczy klasy E z tymi obwodami. Parametry energetyczne oraz przebiegi prądu i napięcia kolektora nie różniły się od uzyskanych w układzie podstawowym wzmacniacza, co potwierdziło przydatność tych obwodów w praktyce.

Przeprowadzono również badania eksperymentalne wzmacniacza klasy E przy sterowaniu z generatora o sinusoidalnym napięciu wyjściowym. Zmierzone parametry wzmacniacza w tym przypadku były takie same, z dokładnością do błędów pomiarowych, jak parametry wzmacniacza sterowanego z generatora o prostokątnym napięciu wyjściowym (tabl. 1). Jednakże, przebiegi napięcia kolektora w stanie nasycenia tranzystora w obu przypadkach różniły się od siebie (co wynikało głównie z różnych przebiegów prądu bazy) oraz czas opadania prądu kolektora Δt_f był nieco większy przy sterowaniu z generatora o sinusoidalnym napięciu wyjściowym [20].

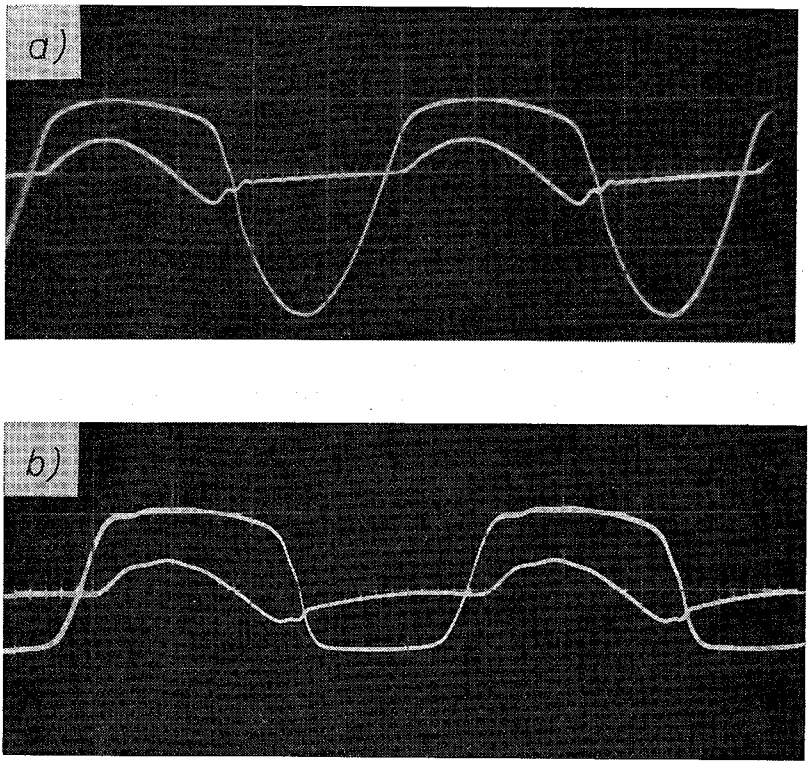


Rys. 17. Zależność sprawności kolektorowej wzmacniacza η_c od napięcia zasilania U_{cc} przy jednakowej mocy wyjściowej $P_{wy} = 3$ W



Rys. 18. Zamknięcie obwodu składowej stałej prądu bazy we wzmacniaczu klasy E
a) przez dławik, b) przez diodę

Jeżeli składowa stała prądu bazy nie zamyka się przez generator sterujący, co ma n przykład miejsce w przypadku pojemnościowego sprzężenia wzmacniacza z generatore sterującym, wówczas obwód składowej stałej prądu bazy można zamknąć przez dławi lub przez diodę impulsową, jak pokazano na rys. 18. Stosowanie diody w obwodzie baz jest korzystne z kilku względów. Po pierwsze, dioda ta zabezpiecza tranzystor przed prz biciem złącza baza-emiter napięciem wejściowym w czasie ujemnej półfali. Przykładow przebiegi napięcia i prądu bazy przy sterowaniu wzmacniacza z generatora o sinusoida nym napięciu wyjściowym, z dławikiem w obwodzie bazy przedstawiono na rys. 19 a z diodą w obwodzie bazy na rys. 19b. Z drugiego z wymienionych rysunków wynik że wejście wzmacniacza w przypadku zastosowania diody w obwodzie bazy zachowu się jak ogranicznik; górna półfala napięcia sterującego ulega spłaszczeniu dzięki działan złącza baza-emiter, a dolna półfala dzięki działaniu diody. Po drugie, wzmacniacz z dio



Rys. 19. Przebiegi napięcia i prądu bazy we wzmacniaczu klasy E przy sterowaniu z generatora o sinusoidalnym napięciu wyjściowym
a) z dławikiem w obwodzie bazy, b) z diodą w obwodzie bazy

w obwodzie bazy jest bardziej stabilny. W przypadku dławika w obwodzie bazy, indukcyjność tego dławika, pojemność kolektor-baza i indukcyjność dławika kolektorowego wraz z reaktancją obwodu kolektorowego tworzą w pewnym zakresie częstotliwości układ o strukturze generatora Hartleya, co może powodować wzbudzenie się wzmacniacza. Ponadto możliwe jest również sprzężenie indukcyjne między wejściem i wyjściem tranzystora. Po trzecie, użycie diody jest korzystniejsze ze względu na mniejsze wymiary, ciężar i koszt układu. Natomiast wadą stosowania diody jest większa moc sterująca i mniejsze wzmocnienie mocy wzmacniacza. Przeprowadzone pomiary wykazały, że pozostałe parametry układu w obu przypadkach były takie same.

5. ZAKOŃCZENIE

Rezultatem niniejszej pracy są zależności matematyczne opisujące podstawowe parametry wzmacniacza klasy E z równoległym kondensatorem pracującego w warunkach optymalnych, wzory opisujące wartości elementów czterech typów kolektorowych obwodów dopasowujących oraz wyniki obszernych badań eksperymentalnych. Z uzyskanych wyników teoretycznych i doświadczalnych wynikają następujące wnioski:

1) w rozpatrywanym wzmacniaczu można uzyskać przebiegi napięcia i prądu kolektora zapewniające wyeliminowanie mocy strat w kolektorze związanej z włączaniem tranzystora,
2) w praktycznych układach wzmacniacza osiąga się bardzo dużą sprawność kolektorową, rzędu 95%,
3) parametry energetyczne wzmacniacza wyznaczone teoretycznie zgadzają się z zadowalającą dokładnością z wynikami eksperymentalnymi,
4) optymalne przebiegi napięcia i prądu kolektora otrzymuje się przy wartościach elementów obwodu kolektorowego zgodnych z obliczeniami teoretycznymi.
5) sprawność kolektorową wzmacniacza można zwiększyć przez wybór możliwie dużego napięcia zasilania kolektora U_{cc},
6) dopasowanie impedancji można uzyskać przez zastosowanie kolektorowych obwodów dopasowujących, które otrzymuje się w wyniku modyfikacji gałęzi szeregowej podstawowego obwodu kolektorowego,
7) wzmacniacz może być sterowany ze stopnia sterującego zarówno o prostokątnym napięciu wyjściowym jak i o sinusoidalnym napięciu wyjściowym,
8) w celu zamknięcia składowej stałej prądu bazy zamiast dławika można zastosować diodę, uzyskując w ten sposób zabezpieczenie złącza baza-emiter przed przebiciem, bardziej stabilną pracę wzmacniacza oraz mniejsze wymiary i ciężar układu.

Dodatek 1

PRZEBIEGI CHWILOWE PRĄDU I NAPIĘCIA KOLEKTORA

Zgodnie z założeniem 2) prąd w szeregowym obwodzie rezonansowym jest sinusoidalny

$$i_{R1}(\omega t) = I_{R1}\sin(\omega t+\varphi), \tag{D1}$$

przy czym:

I_{R1} — amplituda prądu płynącego w szeregowym obwodzie rezonansowym,

φ — faza początkowa tego prądu.

Prąd kolektora i_c oraz prąd i_A płynące przez kondensator C można opisać wzorami:

$$i_c(\omega t) = \begin{cases} I_{cc}-I_{R1}\sin(\omega t+\varphi) & \text{dla} \quad 0 < \omega t \leqslant \pi, \\ 0 & \text{dla} \quad \pi < \omega t \leqslant 2\pi, \end{cases} \tag{D2}$$

$$I_A(\omega t) = \begin{cases} 0 & \text{dla} \quad 0 < \omega t \leqslant \pi \\ I_{cc}-I_{R1}\sin(\omega t+\varphi) & \text{dla} \quad \pi < \omega t \leqslant 2\pi. \end{cases} \tag{D3}$$

Napięcie kolektora u_c można wyznaczyć na podstawie prądu i_A płynącego przez kondensator C

$$u_c(\omega t) = \frac{1}{\omega C}\int_0^{\omega t} i_A(\tau)d\tau =$$

$$= \begin{cases} 0 & \text{dla} \quad 0 < \omega t \leqslant \pi, \\ \dfrac{1}{\omega C}\left[I_{cc}(\omega t-\pi)+I_{R1}\cos(\omega t+\varphi)+I_{R1}\cos\varphi\right] & \text{dla} \quad \pi < \omega t \leqslant 2\pi. \end{cases} \tag{D4}$$

Podstawiając warunek (3) do wyrażenia (D4) otrzymuje się zależność

$$I_{R1} = -I_{cc}\frac{\pi}{2\cos\varphi}, \tag{D5}$$

z której wynika nierówność

$$\frac{\pi}{2} < \varphi < \frac{3\pi}{2}. \tag{D6}$$

Zależność (D4) po uwzględnieniu wyrażenia (D5) można przekształcić do postaci

$$u_c(\omega t) = \begin{cases} 0 & \text{dla} \quad 0 < \omega t \leqslant \pi, \\ \dfrac{I_{cc}}{\omega C}\left[\omega t - \dfrac{3\pi}{2} - \dfrac{\pi}{2\cos\varphi}\cos(\omega t+\varphi)\right] & \text{dla} \quad \pi < \omega t \leqslant 2\pi. \end{cases} \tag{D7}$$

Wykorzystując warunek (4) i zależność (D7) otrzymuje się wzór

$$\operatorname{tg}\varphi_o = -\frac{2}{\pi}. \tag{D8}$$

Ze wzoru tego po uwzględnieniu nierówności (D6) wynika faza początkowa prądu w szeregowym obwodzie rezonansowym φ_o, przy której jednocześnie są spełnione warunki (3) i (4)

$$\varphi_o = \pi - \operatorname{arctg}\frac{2}{\pi} \cong 2{,}57468 \text{ rad} \cong 147{,}518°. \tag{D9}$$

Wykorzystując wzór (D8) zależność (D7) można przekształcić do postaci

$$u_c(\omega t) = \begin{cases} 0 & \text{dla} \quad 0 < \omega t \leqslant \pi, \\ \dfrac{I_{cc}}{\omega C}\left(\omega t - \dfrac{3\pi}{2} - \dfrac{\pi}{2}\cos\omega t - \sin\omega t\right) & \text{dla} \quad \pi < \omega t \leqslant 2\pi. \end{cases} \tag{D10}$$

Po podstawieniu tej zależności do wzoru Fouriera otrzymujemy

$$U_{cc} = \frac{1}{2\pi}\int_0^{2\pi} u_c(\omega t)d(\omega t) = \frac{I_{cc}}{\pi\omega C}. \tag{D11}$$

Z (D10) i (D11) wynika wzór (6), zaś z (D2), (D5) i (D8) wzór (5).

Na podstawie (D8) i (D9) można wykazać, że

$$\cos\varphi_o = -\frac{\pi}{\sqrt{\pi^2+4}}. \tag{D12}$$

Po podstawieniu (D12) do (D5) otrzymuje się wzór (8).

Dodatek 2

WARTOŚCI ELEMENTÓW OBWODU KOLEKTOROWEGO

Na rezystancji obciążenia R występuje sinusoidalne napięcie, ponieważ zgodnie ze wzorem (D1) przez rezystancję tę płynie sinusoidalny prąd i_{R1}. Można zatem napisać

$$u_{R1}(\omega t) = U_{R1}\sin(\omega t+\varphi_o), \tag{D13}$$

przy czym

$$U_{R1} = RI_{R1}. \tag{D14}$$

Przebieg napięcia u_X występującego na elementach C_f, L_f i L jest odkształcony. Składowa podstawowa tego napięcia u_{X1} występuje tylko na cewce L, ponieważ impedancja dwójnika L_f, C_f przy częstotliwości podstawowej równa się zeru. Składową tę można przedstawić w postaci

$$U_{X1}(\omega t) = U_{X1}\cos(\omega t+\varphi_o), \tag{D15}$$

przy czym

$$U_{X1} = \omega L I_{R1}. \tag{D16}$$

Suma napięć u_{R1} i u_{X1} równa się składowej podstawowej napięcia kolektora u_{c1}. Przesunięcie fazowe ψ_o między napięciami u_{R1} i u_{c1} wyraża się wzorem

$$\operatorname{tg}\psi_o = \frac{U_{X1}}{U_{R1}} = \frac{\omega L}{R}. \tag{D17}$$

Amplitudy napięć U_{R1} i U_{X1} otrzymuje się na podstawie wyrażenia (6) i wzorów Fouriera:

$$U_{R1} = \frac{1}{\pi}\int_0^{2\pi} u_c(\omega t)\sin(\omega t+\varphi_o)d(\omega t) = \frac{4}{\sqrt{\pi^2+4}}U_{cc}, \tag{D}$$

$$U_{X1} = \frac{1}{\pi}\int_0^{2\pi} u_c(\omega t)\cos(\omega t+\varphi_o)d(\omega t) = \frac{\pi(\pi^2-4)}{4\sqrt{\pi^2+4}}U_{cc}. \tag{D1}$$

Wyrażenia (D18) i (D19) po uwzględnieniu wzorów (10) i (D11) można przekształci do postaci:

$$U_{R1} = \frac{8}{\pi(\pi^2+4)}\frac{I_{R1}}{\omega C}, \tag{D20}$$

$$U_{X1} = \frac{(\pi^2-4)}{2(\pi^2+4)}\frac{I_{R1}}{\omega C}. \tag{D21}$$

Przyrównując prawe strony (D14) i (D20) otrzymuje się wzór (12), a podstawiając (D20) i (D21) do (D17) wzór (13) oraz przesunięcie fazowe ψ_o między napięciami u_{R1} i u_{c1}. Z zależności (D11) i (12) otrzymuje się wzór (15).

Dodatek

SKŁADOWE PODSTAWOWE PRĄDU I NAPIĘCIA KOLEKTORA

Składową podstawową prądu kolektora i_{c1} można przedstawić w postaci

$$i_{c1}(\omega t) = I_{c1}\sin(\omega t+\gamma_1), \tag{D2}$$

przy czym:

$$I_{c1} = \sqrt{I_{c1A}^2 + I_{c1B}^2}, \tag{D23}$$

$$\mathrm{tg}\gamma_1 = \frac{I_{c1B}}{I_{c1A}}. \tag{D24}$$

Wykorzystując wyrażenie (5) otrzymuje się współczynniki Fouriera:

$$I_{c1A} = \frac{1}{\pi}\int_0^{2\pi} i_c(\omega t)\sin\omega t d(\omega t) = \left(\frac{\pi}{4} + \frac{2}{\pi}\right) I_{cc}, \tag{D25}$$

$$I_{c1B} = \frac{1}{\pi}\int_0^{2\pi} i_c(\omega t)\cos\omega t d(\omega t) = -\frac{I_{cc}}{2}. \tag{D26}$$

Podstawiając (D25) i (D26) do (D23) i (D24) otrzymuje się wyrażenia (16) i (17).

Składową podstawową napięcia kolektora u_{c1} można przedstawić w postaci

$$u_{c1}(\omega t) = U_{R1}\sin(\omega t + \varphi_o) + U_{X1}\cos(\omega t + \varphi_o) = U_{c1}\sin(\omega t + \vartheta_1), \tag{D27}$$

przy czym:

$$U_{c1} = \sqrt{U_{R1}^2 + U_{X1}^2}, \tag{D28}$$

$$\vartheta_1 = \varphi_o + \psi_o. \tag{D29}$$

Podstawiając (D18) i (D19) do (D28) oraz (9) i (14) do (D29) otrzymuje się wyrażenie 3) i (19).

LITERATURA

1. M. Kazimierczuk. *Teoria wzmacniacza mocy wielkiej częstotliwości klasy E*, Rozpr. Elektrot., t. 25, z. 4, 1979, ss. 957—986.
2. M. Kazimierczuk, *Analiza wzmacniacza klasy E przy dowolnym współczynniku wypełnienia prostokątnego napięcia sterującego*, Rozpr. Elektrot., t. 25, z. 4, 1979, ss. 987—1003.
3. J. Ebert, M. Kazimierczuk, *Nowy sposób podwyższenia sprawności rezonansowego wzmacniacza mocy wielkiej częstotliwości*, Raport Instyt. Radioelektr. Polit. Warszawskiej, z. 31, 1977, ss. 1—19.
4. J. Ebert, M. Kazimierczuk, *High efficiency R.F. power transistor amplifier*, Bull. Acad. Polon. Sci., Sér. Sci. Techn, vol. 25, No. 2, 1977, pp. 13—16.
5. J. Ebert, M. Kazimierczuk, *Nowy sposób podwyższenia sprawności rezonansowego wzmacniacza mocy wielkiej częstotliwości*, Prz. Telekom., Nr 6, 1977, ss. 165—168.
6. J. Ebert, M. Kazimierczuk, *Powyszenije kpd moszcznych wysokoczastotnych usilitielej*, Izw. WUZ, Radioelektronika, t. 20, No. 11, 1977, ss. 103—106.
7. M. Kazimierczuk, *Tranzystorowy wzmacniacz mocy wielkiej częstotliwości o podwyższonej sprawności*, Rozprawa doktorska, Inst. Radioelektr. Polit. Warszawskiej, Warszawa 1977.
8. W.B. Kozyriew, *Odnokontaktnyj kluczewoj gienierator s filtrujuszczim konturom, Połuprowodnikowyje pribory w tiechnikie elektroswiazi*, Sbornik statiej, Wyp. 8, 1971, ss. 152—166.
9. N.O. Sokal, A.D. Sokal, *Class E — a new class of high-efficiency tuned single-ended switching power amplifiers*, IEEE J. Solid-State Circuits, vol. SC-10, June 1975, pp. 168—176.
10. N.O. Sokal, *Class E boost the efficiency of RF power amplifiers*, Electron. Design., vol. 25, No. 20, Sept. 27, 1977, pp. 96—102.
11. F.H. Raab, *Idealized operation of the class E tuned power amplifier*, IEEE Trans. Circuits and Systems, vol. CAS-24, Dec. 1977, pp. 725—735.

12. W. J. Chudabiak, D. F. Page, *Frequency and power limitations of Class-D transistor amplifiers*, IEEE J. Solid-State Circuits, vol. SC-4, Feb. 1969, pp. 25—37.
13. J. Baranowski, *Półprzewodnikowe układy implusowe i cyfrowe*, Wyd. 2, WNT, Warszawa 1976.
14. S. Ryżko, J. Ebert, *Wzmacniacze rezonansowe i generatory mocy wielkiej częstotliwości*, Wyd. 2, WNT, Warszawa 1971.
15. J. Ebert, A. Fiok, *Obliczanie wzmacniaczy i generatorów mocy wielkiej częstotliwości*, WNT, Warszawa 1971.
16. A. Filipkowski, *Układy elektroniczne analogowe i cyfrowe*, WNT, Warszawa 1978.
17. J. J. Ebers, J. L. Moll, *Large-signal behavior of junction transistor*, Proc. IRE, vol. 42, Dec., 1954, pp. 1761—1772.
18. F. H. Raab, N. O. Sokal, *Transistor power losses in the class E tuned power amplifier*, IEEE J. Solid-State Circuits (Corresp), vol. SC-13, Dec. 1978, pp. 912—914.
19. M. Kazimierczuk, *Moc strat w kolektorze i sprawność kolektorowa wzmacniacza klasy E z równoległym kondensatorem*, Elektronika, Nr 5, 1980, ss. 24—30.
20. M. Kazimierczuk, *Weryfikacja eksperymentalna metody wyznaczania mocy start w kolektorze i sprawności kolektorowej wzmacniacza klasy E z równoległym kondensatorem*, Elektronika, Nr 4, 1981, ss. 30—35.

M. KAZIMIERCZUK

ANALYSIS, DESIGN AND EXPERIMENTAL RESULTS OF THE *E* CLASS AMPLIFIER FOR OPTIMUM OPERATION

Summary

The basic results of the theoretical analysis, the design and the experimental results of the class *E* tuned power amplifier with a shunt capacitor for optimum operation are presented in the paper. Some parameters of the amplifier are determined on the basis of a simplified model. Next, the collector power losses and the collector efficiency, the driving power and the power amplification are estimated. Four types of the matching output networks and the formulas for evaluating the element values of these networks are reported. Many experimental results are also given.

M. KAZIMIERCZUK

ANALYSE, PROJECTION ET ESSAIS EXPÉRIMENTAUX DE L'AMPLIFICATEUR DE LA CLASSE *E* DANS LES CONDITIONS DE TRAVAIL OPTIMALES

Résumé

L'article présente les résultats les plus importants de l'analyse théorique, le mode de projection et les résultats des essais expérimentaux de l'amplificateur de puissance de haute fréquence de la classe *E* à condensateur parallèle, fonctionnant dans les conditions optimales. En se servant d'un modèle simplifié de l'amplificateur, on a déterminé certains paramètres du système. Ensuite on a évalué les pertes de puissance de collecteur et son rendement, ainsi que la puissance d'entrée et l'amplification de puissance de l'amplificateur. On a montré quatre types de circuits du collecteur aux propriétés adaptables et les formules, décrivant les valeurs des éléments de ces circuits. On y a inclu ausssi les résultats de nombreux essais expérimentaux.

M. KAZIMIERCZUK

ANALYSE, ENTWERFEN UND EXPERIMENTELLE UNTERSUCHUNGEN EINE *E*-KLASSE-VERSTÄRKERS UNTER OPTIMALEN ARBEITS-BEDINGUNGEN

Zusammenfassung

Die wichtigsten Ergebnisse einer theoretischen Analyse wurden dargestellt, auch die Art des Entwerfens sowie Ergebnisse experimenteller Untersuchungen für einen Verstärker mit Hochfrequenzleistun

der *E*-Klasse mit parallelem Kondensator unter optimalen Arbeitsbedingungen. Aufgrund eines vereinfachten Verstärkermodells wurden manche Systemparameter bestimmt. Danach wurden Leistungsverluste im Kollektor und dessen Wirkungsgrad eingeschätzt, auch die Eingangsleistung und die Verstärkung der Versterkerleistung. Es wurden 4 Typen von Kollektorkreisen dargestellt, die Anpassungseigenschaften besitzen sowie Muster, die die Werte dieser Kreiselemente beschreiben. Abschließend wurden Ergebnisse umfangreicher Untersuchungen angeführt.

М. КАЗИМЕРЧУК

АНАЛИЗ, ПРОЕКТИРОВАНИЕ И ЭКСПЕРИМЕНТАЛЬНЫЕ ИССЛЕДОВАНИЯ УСИЛИТЕЛЯ КЛАССА *E* В ОПТИМАЛЬНОМ РЕЖИМЕ

Р е з ю м е

Представлены главные результаты теоретического анализа, способ проектирования и результаты экспериментальных исследований высокочастотного усилителя мощности класса *E* с параллельной емкостью, работающего в оптимальном режиме. Определены некоторые параметры усилителя при использовании его упрощенной эквивалентной схемы. Определены приблизительно потери мощности, кпд, мощность возбуждения и коэффициент усиления мощности. Представлены четыре типа коллекторных, согласованных цепей и формулы для расчета их элементов. Представлены результаты экспериментальных исследований.

ROZPRAWY ELEKTROTECHNICZNE 1981, 27, z. 3, ss. 823-836

621.371.413

Metody wyznaczania zespolonej pulsacji rezonansowej współosiowego rezonatora ze skupioną pojemnością niejednorodnie wypełnionego dielektrykiem

ANDRZEJ KĄCZKOWSKI (WARSZAWA), ANDRZEJ MILEWSKI (WARSZAWA)
Instytut Technologii Elektronowej Politechniki Warszawskiej

Otrzymano 28.12.1979

W pracy przebadano analityczną metodę określania zespolonej pulsacji rezonansowej współosiowego rezonatora ze skupioną pojemnością niejednorodnie wypełnionego dielektrykiem. Zaproponowano metodę przybliżoną opartą na obwodowym modelu rezonatora. Przeanalizowano różne modele obwodu zastępczego rezonatora. Dla dwu z zaproponowanych modeli wykazano, że błędy określenia pulsacji zespolonej tymi metodami są dla większości przypadków mniejsze od błędów pomiarowych. Pozwoliło to na znaczne uproszczenie aparatu matematycznego przy jednoczesnym nieobniżaniu dokładności pomiaru składowych pulsacji rezonansowej.

1. WSTĘP

Dokładna znajomość pulsacji rezonansowej — w ogólności o wartości zespolonej — jest istotna w wielu zastosowaniach wnęk rezonansowych. Między innymi problem ten uwypuklił się ostatnio [3, 9] w rezonatorowym miernictwie materiałów bardzo małostratnych. Sprawa jest w miarę prosta dla rezonatorów pracujących w jednym modzie drgań, dla których do opisu rozkładu pola elektromagnetycznego wystarczą pojedyncze funkcje Bessela lub trygonometryczne. W przypadku rezonatora współosiowego ze skupioną pojemnością — rezonatora często stosowanego w miernictwie materiałowym — rozkład pola jest bardzo skomplikowany. Prowadzi to w efekcie do konieczności opisywania tego rozkładu przy użyciu superpozycji wielu funkcji Bessela. W niniejszej pracy wykazano, że analityczne określanie zesplonej pulsacji rezonansowej dla takiego przypadku pociąga za sobą konieczność rozwiązywania skomplikowanych zależności matematycznych. W związku z tym zaproponowano dowolną metodę przybliżoną traktującą wnękę współosiową ze skupioną pojemnością jako superpozycję elementów o parametrach skupionych i rozłożonych, a następnie przebadano skuteczność zaproponowanych przybliżeń.

2. DOKŁADNA METODA OBLICZANIA CZĘŚCI UROJONEJ PULSACJI REZONANSOWEJ

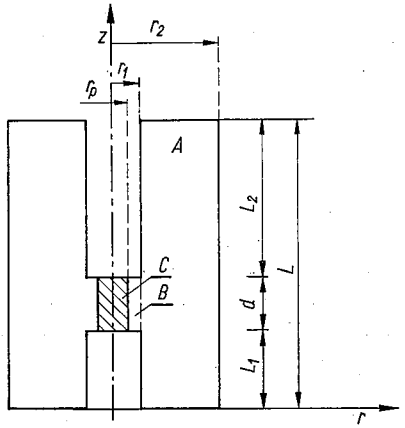
Przedmiotem rozważań jest rezonator współosiowy ze skupioną pojemnością, częściowo wypełniony dielektrykiem, jak na rys. 1. Jak ogólnie wiadomo, własności obwodu

rezonansowego można przedstawić za pomocą zespolonej wartości pulsacji swobodnych drgań rezonatora $\hat{\omega}$

$$\hat{\omega} = \omega' \left(1 + j \frac{1}{2Q}\right). \tag{1}$$

Różni się ona od pulsacji rezonansowej mierzonej metodą drgań wymuszonych. Różnica ta zależy od sposobu pobudzenia drgań [1] i wyraża się zależnością

$$\omega_s = \omega \left(1 + a \frac{1}{Q^2}\right). \tag{2}$$



Rys. 1. Rezonator współosiowy ze skupioną pojemnością częściowo wypełniony dielektrykiem (obszar *C*)

gdzie:

ω_s — pulsacja drgań swobodnych,

ω — pulsacja mierzona rezonansu drgań wymuszonych,

Q — dobroć rezonatora,

przy czym współczynnik a w zależności od charakteru sprzężeń przyjmuje różne wartości $a = \frac{1}{2}$ [2, 3], $a = 5/2$ [1].

Jak wynika z (2) niezależnie od współczynnika a poprawka na wartość częstotliwości jest pomijalna dla dobroci większych od 100. W związku z tym w dalszej części prac przyjęto, że pulsacja drgań swobodnych jest równa mierzonej pulsacji rezonansowej.

Dobroć rezonatora jest z definicji określona zależnością

$$Q = \omega W / P_{st}, \tag{3}$$

gdzie:

W — średnia energia gromadzona w rezonatorze,

P_{st} — średnia moc strat w układzie.

Na wartość mocy strat składają się straty energii pochodzące z prądów przewodzeni w ściankach rezonatora oraz straty w samej próbce dielektrycznej (obszar *C*, rys. 1) Inne mechanizmy rozpraszania energii (jak np. promieniowanie przez szczeliny, strat

w sprzężeniach) można przy starannie prowadzonych pomiarach [3] wyeliminować. Aby otrzymać prawidłowe wartości pulsacji zespolonej (1), należy znaleźć taką dobroć Q_p, jaką miałby rezonator o idelanie przewodzących ściankach z umieszczoną wewnątrz badaną próbką. Jak wynika z zależności (3) dobroć tę można wyliczyć ze wzoru

$$\frac{1}{Q_p} = \frac{1}{Q} - \frac{1}{Q'}, \tag{4}$$

gdzie:

Q — dobroć rzeczywista wnęki z próbką,

Q' — dobroć rezonatora z idealnie bezstratną próbką o przenikalności takiej, aby rozkład pola był taki sam, jak z próbką rzeczywistą.

Dobroć Q' określa się na podstawie znajomości rozkładu pola i rezystywności ścianek rezonatora ϱ_s [3, 4, 5 i in.]. Tę ostatnią należy wyznaczyć eksperymentalnie. Zwykle stosuje się wyliczenie ϱ_s na podstawie pomiaru parametrów rezonatora pustego, tzn. jego częstotliwości rezonansowej f_o oraz dobroci Q_o.

W przypadku, gdy wprowadzenie próbki nieznacznie zaburza rozkład pola, dobroć Q' określona jest równością (5) wynikającą z zależności rezystywności powierzchniowej metalu od częstotliwości

$$Q' = Q_o \sqrt{f_o/f}. \tag{5}$$

Gdy próbka wykazuje dużą przenikalność ε', należy wtedy obliczyć następującą proporcję [7]

$$K = \frac{Q'}{Q_o} = \frac{\int_v \hat{H}_o \hat{H}_o^* dv}{\int_v \hat{H}\hat{H}^* dv} \frac{\int_s \hat{H}\hat{H}^* ds}{\int_s \hat{H}_o \hat{H}_o^* ds} \sqrt{\frac{f_o}{f}} \tag{6}$$

gdzie:

$\hat{H}_o$ — pole magnetyczne w rezonatorze pustym,

H — pole magnetyczne w rezonatorze z próbką,

S — powierzchnia ścianek rezonatora,

V — objętość rezonatora.

Jak wynika z (6), w celu obliczenia współczynnika K należy znaleźć stosunek energii zgromadzonej w rezonatorze do traconej w jego metalowych ściankach dla dwóch przypadków: wnęki pustej oraz z próbką. Dla rozpatrywanego rezonatora w obu sytuacjach zagadnienie sprowadza się do policzenia następujących proporcji

$$\frac{\int |\hat{H}|^2 dv}{\int |\hat{H}|^2 ds} = \frac{\int |\hat{H}_\varphi^A|^2 dv_A + \int |\hat{H}_\varphi^B|^2 dv_B + \int |\hat{H}_\varphi^C|^2 dv_C}{\int |\hat{H}_\varphi^A|^2 ds_A + \int |\hat{H}_\varphi^B|^2 ds_B + \int |\hat{H}_\varphi^C|^2 ds_C}. \tag{7}$$

Korzystając z obliczonych wcześniej [9], dla rezonatora współosiowego ze skupioną pomnością z próbką dielektryczną o konfiguracji jak na rys. 1, współczynników rozwinięa pola magnetycznego w szeregi funkcyjne, poszczególne całki występujące w powyższej leżności przyjmują następującą postać:

$$\frac{\int |\hat{H}_\varphi^A|^2 dv_A}{2\pi(\omega\varepsilon_o)^2} = \frac{L}{2} \sum_{n=0} \frac{\alpha(n)}{(\hat{X}_n^A)^2} \Big\{ A_n^2 [r_2^2 (J_1^2(\hat{X}_n^A r_2) - J_o(\hat{X}_n^A r_2) J_2(\hat{X}_n^A r_2)) -$$

$$-r_1^2(J_1^2(\hat{X}_n^A r_1)-J_o(\hat{X}_n^A r_1)J_2(\hat{X}_n^A r_1))]+\hat{B}_n^2[r_2^2(Y_1^2(\hat{X}_n^A r_2)-Y_o(\hat{X}_n^A r_2)Y_2(\hat{X}_n^A r_2))-$$

$$-r_1^2(Y_1^2(\hat{X}_n^A r_1)-Y_o(\hat{X}_n^A r_1)\,Y_2(\hat{X}_n^A r_1))]+\frac{\hat{A}_n\hat{B}_n}{2}[r_2^2(2J_1(\hat{X}_n^A r_2)Y_1(\hat{X}_n^A r_2)-$$

$$-J_o(\hat{X}_n^A r_2)Y_2(\hat{X}_n^A r_2)-J_2(\hat{X}_n^A r_2)Y_o(\hat{X}_n^A r_2))-r_1^2(2J_1(\hat{X}_n^A r_1)Y_1(\hat{X}_n^A r_1)-$$

$$-J_o(\hat{X}_n^A r_1)Y_2(\hat{X}_n^A r_1)-J_2(\hat{X}_n^A r_1)Y_o(\hat{X}_n^A r_1))]\Big\}, \tag{8}$$

$$\frac{\int|\hat{H}_\varphi^A|^2 ds_A}{2\pi(\omega\varepsilon_o)^2}=\sum_{n=0}^{\Gamma}\Big\{\frac{L\alpha(n)}{(\hat{X}_n^A)^2}\langle\hat{A}_n^2[r_2J_1^2(\hat{X}_n^A r_2)+r_1J_1^2(\hat{X}_n^A r_1)]+$$

$$+\hat{B}_n^2[r_2Y_1^2(\hat{X}_n^A r_2)+r_1\,Y_1^2(\hat{X}_n^A r_1)]+2\hat{A}_n\hat{B}_n[r_2J_1(\hat{X}_n^A r_2)Y_1(\hat{X}_n^A r_2)+r_1J_1(\hat{X}_n^A r_1)Y_1(\hat{X}_n^A r_1)]\rangle+$$

$$+\frac{1}{(\hat{X}_n^A)^2}\langle\hat{A}_n^2[r_2^2(J_1^2(\hat{X}_n^A r_2)-J_o(\hat{X}_n^A r_2)J_2(\hat{X}_n^A r_2))-r_1^2(J_1^2(\hat{X}_n^A r_1)-$$

$$-J_o(\hat{X}_n^A r_1)J_2(\hat{X}_n^A r_1))]+\hat{B}_n^2[r_2^2(Y_1^2(\hat{X}_n^A r_2)-Y_o(\hat{X}_n^A r_2)Y_1(\hat{X}_n^A r_2))-$$

$$-r_1^2(Y_1^2(\hat{X}_n^A r_1)-Y_o(\hat{X}_n^A r_1)Y_2(\hat{X}_n^A r_1))]+$$

$$+\frac{\hat{A}_n\hat{B}_n}{2}[r_2^2(2J_1(\hat{X}_n^A r_2)Y_1(\hat{X}_n^A r_2)-J_o(\hat{X}_n^A r_2)Y_2(\hat{X}_n^A r_2)-$$

$$-J_2(\hat{X}_n^A r_2)Y_o(\hat{X}_n^A r_2))-r_1^2(2J_1(\hat{X}_n^A r_1)Y_1(\hat{X}_n^A r_1)-J_o(\hat{X}_n^A r_1)Y_2(\hat{X}_n^A r_1)-J_2(\hat{X}_n^A r_1)Y_o(\hat{X}_n^A r_1))]\Big\}+$$

$$+\sum_{m=0}^{\Gamma}\sum_{l=0}^{\Gamma}\frac{[1+(-1)^{m+l}][1-\delta(m,l)]}{\hat{X}_m^A\hat{X}_l^A}\Big\{\frac{\hat{A}_m\hat{A}_l}{(\hat{X}_m^A)^2-(\hat{X}_l^A)^2}\langle r_2[\hat{X}_l^A J_o(\hat{X}_l^A r_2)\times$$

$$\times J_1(\hat{X}_m^A r_2)-\hat{X}_m^A J_o(\hat{X}_m^A r_2)J_1(\hat{X}_l^A r_2)]-r_1[\hat{X}_l^A J_o(\hat{X}_l^A r_1)J_1(\hat{X}_m^A r_1)-\hat{X}_m^A J_o(\hat{X}_m^A r_1)J_1(\hat{X}_m^A r_1)]\rangle+$$

$$+\frac{\hat{B}_m\hat{B}_l}{(\hat{X}_m^A)^2-(\hat{X}_l^A)^2}\langle r_2[\hat{X}_l^A Y_o(\hat{X}_l^A r_2)Y_1(\hat{X}_m^A r_2)-\hat{X}_m^A Y_o(\hat{X}_m^A r_2)Y_1(\hat{X}_l^A r_2)]-$$

$$-r_1[\hat{X}_l^A Y_o(\hat{X}_l^A r_1)Y_1(\hat{X}_m^A r_1)-\hat{X}_m^A Y_o(\hat{X}_m^A r_1)Y_1(\hat{X}_l^A r_1)]+$$

$$+\frac{\hat{A}_m\hat{B}_l}{(\hat{X}_m^A)^2-(\hat{X}_l^A)^2}\langle r_2[\hat{X}_l^A Y_o(\hat{X}_l^A r_2)J_1(\hat{X}_m^A r_2)-\hat{X}_m^A J_o(\hat{X}_m^A r_2)Y_1(\hat{X}_l^A r_2)]-$$

$$-r_1[\hat{X}_l^A Y_o(\hat{X}_l^A r_1)J_1(\hat{X}_m^A r_1)-\hat{X}_m^A J_o(\hat{X}_m^A r_1)Y_1(\hat{X}_l^A r_1)]\rangle+$$

$$+\frac{\hat{B}_m\hat{A}_l}{(\hat{X}_m^A)^2-(\hat{X}_l^A)^2}\langle r_2[\hat{X}_l^A J_o(\hat{X}_l^A r_2)Y_1(\hat{X}_m^A r_2)-\hat{X}_m^A Y_o(\hat{X}_m^A r_2)J_1(\hat{X}_l^A r_2)]-$$

$$-r_1[\hat{X}_l^A J_o(\hat{X}_l^A r_1)Y_1(\hat{X}_m^A r_1)-\hat{X}_m^A Y_o(\hat{X}_m^A r_1)J_1(\hat{X}_l^A r_1)]\rangle\Big\}-$$

$$-\frac{dr_1}{(\hat{X}_o^B)^2}[\hat{G}_o^2J_1^2(\hat{X}_o^B r_1)+\hat{F}_o^2Y_1^2(\hat{X}_o^B r_1)+2\hat{G}_o\hat{F}_oJ_1(\hat{X}_o^B r_1)Y_1(\hat{X}_o^B r_1)], \tag{}$$

$$\frac{\int|\hat{H}_{\varphi}^{B}|^2 dv_B}{2\pi(\omega\varepsilon_o)^2} = \frac{d}{2(\hat{X}_o^B)^2}\Big\{\hat{G}_o^2\langle r_1^2[J_1^2(\hat{X}_o^B r_1) - J_o(\hat{X}_o^B r_1)J_2(\hat{X}_o^B r_1) -$$

$$-r_p^2[J_1^2(\hat{X}_o^B r_p) - J_o(\hat{X}_o^B r_p)J_2(\hat{X}_o^B r_p)]\rangle + \hat{F}_o^2\langle r_1^2[Y_1^2(\hat{X}_o^B r_1) - Y_o(\hat{X}_o^B r_1)Y_2(\hat{X}_o^B r_1) -$$

$$-r_p^2[Y_1^2(\hat{X}_o^B r_p) - Y_2(\hat{X}_o^B r_p)Y_o(\hat{X}_o^B r_p)]\rangle + \frac{\hat{G}_o\hat{F}_o}{2}\langle r_1^2[2J_1(\hat{X}_o^B r_1)Y_1(\hat{X}_o^B r_1) -$$

$$-r_p^2[2J_1(\hat{X}_o^B r_p)Y_1(\hat{X}_o^B r_p) - J_o(\hat{X}_o^B r_p)Y_2(\hat{X}_o^B r_p) - J_2(\hat{X}_o^B r_p)Y_o(\hat{X}_o^B r_p)]\rangle\Big\}, \tag{10}$$

$$\frac{\int|\hat{H}_{\varphi}^{B}|^2 ds_B}{2\pi(\omega\varepsilon_0)^2} = \frac{1}{(\hat{X}_o^B)^2}\Big\{\hat{G}_o^2\langle r_1^2[J_1^2(\hat{X}_o^B r_1) - J_o(\hat{X}_o^B r_1)J_2(\hat{X}_o^B r_1)] - r_p^2[J_1^2(\hat{X}_o^B r_p) -$$

$$-J_o(\hat{X}_o^B r_p)J_2(\hat{X}_o^B r_p)]\rangle + \hat{F}_o^2\langle r_1^2[Y_1^2(\hat{X}_o^B r_1) - Y_o(\hat{X}_o^B r_1)Y_2(\hat{X}_o^B r_1)] -$$

$$-r_p^2[Y_1^2(\hat{X}_o^B r_p) - Y_o(\hat{X}_o^B r_p)Y_2(\hat{X}_o^B r_p)]\rangle + \frac{\hat{G}_o\hat{F}_o}{2}\langle r_1^2[2J_1(\hat{X}_o^B r_1)Y_1(\hat{X}_o^B r_1) -$$

$$-J_o(\hat{X}_o^B r_1)Y_2(\hat{X}_o^B r_1) - J_2(\hat{X}_o^B r_1)Y_o(\hat{X}_o^B r_1)] - r_p^2[2J_1(\hat{X}_o^B r_p)Y_1(\hat{X}_o^B r_p) -$$

$$-J_o(\hat{X}_o^B r_p)Y_2(\hat{X}_o^B r_p) - J_2(\hat{X}_o^B r_p)Y_o(\hat{X}_o^B r_p)]\rangle\Big\}, \tag{11}$$

$$\frac{\int|\hat{H}_{\varphi}^{C}|^2 dv_C}{2\pi(\omega\varepsilon_0)^2} = \frac{d(\varepsilon\hat{T}_o r_p)^2}{2(\hat{X}_o^C)^2}[J_1^2(\hat{X}_o^C r_p) - J_o(\hat{X}_o^C r_p)J_2(\hat{X}_o^C r_p)], \tag{12}$$

$$\frac{\int|\hat{H}_{\varphi}^{C}|^2 ds_o}{2\pi(\omega\varepsilon_o)^2} = \frac{(\varepsilon\hat{T}_o r_p)^2}{(\hat{X}_o^C)^2}[J_1^2(\hat{X}_o^C r_p) - J_o(\hat{X}_o^C r_p)J_2(\hat{X}_o^C r_p)]. \tag{13}$$

W wyrażeniach (8)÷(13) oznaczono:

$$\hat{A}_n = \frac{\hat{C}_n Y_o(\hat{X}_n^A r_2)}{J_o(\hat{X}_n^A r_1)Y_o(\hat{X}_n^A r_2) - Y_o(\hat{X}_n^A r_1)J_o(\hat{X}_n^A r_2)},$$

$$\hat{B}_n = \frac{-\hat{C}_n J_o(\hat{X}_n^A r_2)}{J_o(\hat{X}_n^A r_1)Y_o(\hat{X}_n^A r_2) - Y_o(\hat{X}_n^A r_1)J_0(\hat{X}_n^A r_2)},$$

$$\hat{C}_n = \frac{2d}{\pi L}\alpha_n \sum_{m=0}^{\infty}\hat{M}_m \times$$

$$\times \frac{\sin\frac{\pi}{2}\left(n\frac{d}{L} - m\right)}{\frac{nd}{L} - m}\cos\frac{\pi}{2}\left(n\frac{d+2L_1}{L} - m\right) + \frac{\sin\frac{\pi}{2}\left(n\frac{d}{L} + m\right)}{\frac{nd}{L} - m}\cos\frac{\pi}{2}\left(n\frac{d+2L_1}{L} + m\right),$$

$$\alpha_{n,m} = \begin{cases} 1/2 & \text{dla} \quad n, m = 0, \\ 1 & \text{dla} \quad n, m > 0, \end{cases}$$

$$\hat{X}_n^A = [\hat{k}_o^2 - (\beta_n^A)^2]^{1/2}, \quad \beta_n^A = \frac{n\pi}{L}, \quad \hat{k}_o = \frac{\hat{\omega}}{c},$$

$$\hat{X}_m^B = [\hat{k}_o^2 - (\beta_m^B)^2]^{1/2}, \quad \beta_m^B = \frac{m\pi}{d},$$

$$\hat{X}_m^C = [\hat{k}^2 - (\beta_m^C)^2]^{1/2}, \quad \beta_m^C = \beta_m^B, \quad \hat{k} = \hat{k}_o(\hat{\varepsilon})^{1/2},$$

$$\begin{bmatrix} J_o(\hat{X}_m^B r_1) & Y_o(\hat{X}_m^B r_1) & 0 \\ J_o(\hat{X}_m^B r_p) & Y_o(\hat{X}_m^B r_p) & -J_o(\hat{X}_m^C r_p) \\ \frac{1}{\hat{X}_m^B} J_1(\hat{X}_m^B r_p) & \frac{1}{\hat{X}_m^B} Y_1(\hat{X}_m^B r_p) & -\frac{\hat{\varepsilon}}{\hat{X}_m^C} J_1(\hat{X}_m^C r_p) \end{bmatrix} \begin{bmatrix} \hat{G}_m \\ \hat{F}_m \\ \hat{T}_m \end{bmatrix} = \begin{bmatrix} \hat{M}_m \\ 0 \\ 0 \end{bmatrix},$$

J, Y — funkcje Bessela odpowiednio pierwszego i drugiego rodzaju.

Łącząc wyrażenia (1), (4), (5), (6) otrzymuje się

$$\frac{\hat{\omega}}{\omega} = 1 + j\frac{1}{2}\left(\frac{1}{Q} - \frac{1}{Q_o}\frac{1}{K}\right). \tag{14}$$

W celu dokładnego policzenia zespolonej pulsacji rezonansowej należy oczywiście do (14) podstawić wyrażenia (6)÷(13). Otrzymana zależność na $\hat{\omega}$ jest dość skomplikowana a występujące w niej szeregi mają postać nieskończoną. W praktyce liczbę wyrazów szeregu należy oczywiście doprowadzić do skończonej wartości, w analizowanym przypadku liczbę wyrazów oznaczoną przez N_{AX}, przy której błąd określania rozkładu pola jest pomijalnie mały oraz sposób jej wyznaczania przedstawiono w pracy [9] przy okazji badania możliwości wykorzystania tego typu rezonatora do pomiaru $\hat{\varepsilon}$ dielektryków.

3. UPROSZCZONA METODA OBLICZANIA CZĘŚCI UROJONEJ ZESPOLONEJ PULSACJI REZONANSOWEJ WNĘKI

W poprzednim rozdziale podano ogólną, ścisłą metodę znajdowania składowej Q dobroci rezonatora (4), która służąc do wyliczenia zespolonej pulsacji rezonansowej (14) pozwala wyznaczyć np. straty w dielektryku (obszar C). Relacje (14), (6,) (7) są słuszne dla wszystkich rezonatorów i — jak podkreślono w pracy [3] — liczenie dobroci Q_p powinno odbywać się na podstawie pełnego opisu pola elektromagnetycznego w rezonatorze przed i po wprowadzeniu próbki, zwłaszcza przy badaniu dielektryków o małym współczynniku strat ($\operatorname{tg}\delta$).

Ze względu na fakt, że całki w wyrażeniu (6) mają złożoną postać (8)÷(13), procedura liczenia dobroci Q' a następnie Q_p z pełnego rozwinięcia rozkładu pola w rezonatorze ze skupioną pojemnością jest dość skomplikowana. Korzystanie z niej jest jednak konieczne [3], np. w przypadku badania ε i $\operatorname{tg}\delta$ materiałów o dużej przenikalności rzeczywistej ($\varepsilon > 10$) oraz małym współczynniku strat ($\operatorname{tg}\delta < 1/Q_o$). Taka sytuacja wiąże się ze znaczną zmianą rozkładu pola w rezonatorze, wywołaną wprowadzeniem próbki przy niewielkiej zmianie dobroci. Przy określaniu współczynnika K we wzorze (14) np. z dokładnością 10%, może okazać się, że błąd wyznaczenia części urojonej pulsacji

rezonansowej przekracza 100% [3, 10]. Wspomniana złożoność wyrażeń całkowych (8)÷(13) powoduje, że liczenie Q' zajmuje wiele czasu maszynowego i znacznie rozszerza obszar wymaganej pamięci operacyjnej systemu ETO. Aby pominąć te niedogodności, autorzy postanowili zbadać możliwość zastosowania pewnych przybliżeń mając na uwadze, aby nie odbiło się to w wyraźny sposób na dokładności określenia Q'.

Przybliżone metody stosowane do tej pory przez innych autorów przy okazji badania i tg δ materiałów, opierały się zwykle na jednym z dwóch niżej podanych założeń.

1. $Q_p = Q$. Przybliżenie takie występuje np. w pracy [11]. Błąd względny określenia części urojonej pulsacji wynosi $\delta\omega''/\omega'' = Q/Q'$. Takie założenie jest zatem dopuszczalne jedynie w przypadku, gdy wprowadzenie próbki zmniejszy dobroć przynajmniej kilkaziesiąt razy.

2. $1/Q_p = 1/Q - 1/Q_o$. Jest to często [12÷15 i in.] stosowane przybliżenie wynikające z perturbacyjnego rozwiązania pierwszego rzędu [8, 10]. Z założeń tam poczynionych wynika, że można je stosować dla przypadku małego zaburzenia pola poza próbką. Błąd w określeniu dekrementu tłumienia energii w dielektryku $\delta\omega''/\omega'' = Q(1-K)/[Q_p(1-Q/Q_o)]$. Tak więc gdy $K \to 1$, a ma to miejsce, gdy przestrojenie $(f_o - f)$ jest znikomo małe, błąd popełniany jest niewielki. Wielu autorów zapomina jednak o tym ograniczeniu i stosuje przybliżenie perturbacyjne w niedopuszczalnie szerokim zakresie [14÷16].

Ze względu na warunki pomiarowe w rozpatrywanym rezonatorze (duże przestrojenia) żadne z przedstawionych przybliżeń nie daje dobrych wyników. W celu zbadania możliwości poszczególnych uproszczeń przedstawiono wyrażenie na Q_p wynikające z (14) w następującej postaci

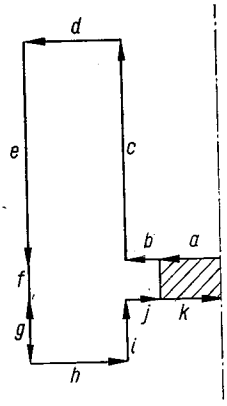
$$\frac{1}{Q_o} = \frac{1}{Q} - \frac{1}{Q_o(K+\delta K)}.$$

Inaczej mówiąc: jeżeli współczynnik $K = Q'/Q_o$ został określony z zależności (6), to $\delta K = 0$. We wszystkich innych przypadkach przy określaniu K z innych, mniej precyzyjnych wzorów otrzymuje się wartość różną od właściwej o δK. Tak zdefiniowane δK jest bezpośrednim wskaźnikiem dokładności stosowanych przybliżeń. Dokładność określenia współczynnika K przenosi się na błąd składowej urojonej pulsacji według następującej zależności

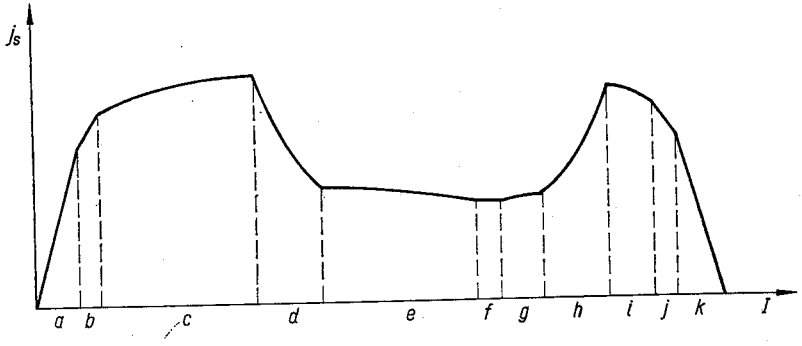
$$\delta\omega''/\omega'' = [(Q/Q_o)/(K - Q/Q_o)]\,\delta K/K.$$

Współczynnik K można określać szybciej niż z pełnego, dokładnego algorytmu (6)÷(13) zastępując szeregi opisujące pole elektromagnetyczne albo stosując przybliżenia obwodowe. Kryterium przydatności jest w każdym z przypadków wartość $\delta K/K$. W celu efektywnego zastosowania przybliżeń obwodowych należało zbadać zmianę rezystancji zastępczej obwodu rezonansowego R_s wywołanego wprowadzeniem próbki do rezonatora. Rezystancja ta reprezentuje straty energii wynikające z rozpływu prądu w metalowych ścianach rozpatrywanego rezonatora (rys. 2 i 3). Przy analizie rozkładu gęstości prądu powierzchniowego j_s oparto się na następujących założeniach.

1. W obrębie skupionej pojemności (obszary B i C, rys. 1) założono jednorodność pola elektrycznego [3, 4, 14 i in.]. Konsekwencją tego założenia jest liniowy charakter zmian gęstości prądu j_s w ściankach ograniczających obszar C (odcinki a i k na rys. 2 i 3) oraz liniowo-hiperboliczny w obszarze B (odcinki b i j).



Rys. 2. Droga prądu w ściankach rezonatora



Rys. 3. Charakter zmian gęstości prądu powierzchniowego j_s w ściankach rezonatora
Literami $a \div k$ opisano odpowiednie obszary ścianek zaznaczone na rys. 2

2. W zwartych odcinkach linii współosiowej rezonatora założono rozkład TEM po
[3, 4, 5 i in.]. Wynikiem tego założenia jest m.in. hiperboliczna zależność amplitudy po
magnetycznego od promienia w obszarze A (odcinki d i h). Charakter tej zależności p
twierdzony został doświadczalnie [18] również dla rozkładów pól odbiegających od
pu TEM.

Opis matematyczny zmian gęstości prądu w ściankach oparty na powyższych zało
niach znajduje się w tablicy 1, w której przytoczono również wyliczone — z rozkładu gęs
ści prądu w ściankach — spadki potencjału na powierzchni metalu pojawiające się wsl
tek niepomijalnej rezystywności powierzchniowej ϱ_s. Na podstawie wyników zamieszc
nych w tablicy 1 można wyliczyć rezystancję zastępczą układu rezonansowego R_s

$$R_s = \frac{\Sigma V_t}{I_o} = \frac{\varrho_s}{2}\left\{\frac{r_1^2+(\varepsilon'-1)\left(1+2\ln\frac{r_1}{r_p}\right)r_p^2}{r_1^2+(\varepsilon'-1)r_p^2} + \right.$$

$$\left.\frac{1}{k}\left(\frac{1}{r_1}+\frac{1}{r_2}\right)(\operatorname{tg} kL_1+\operatorname{tg} kL_2)+\ln\frac{r_2}{r_1}\left(\frac{1}{\cos kL_1}+\frac{1}{\cos kL_2}\right)+\frac{d}{r_2}\right\},$$

gdzie:

ϱ_s — rezystywność powierzchniowa ścianek rezonatora,

ε' — część rzeczywista przenikalności elektrycznej próbki.

Oprócz zaproponowanego sposobu obliczania R_s (15) należało dodatkowo zbadać, który ze sposobów przybliżonego określenia indukcyjności i pojemności zastępczej da najlepsze wyniki określenia k. Ponieważ wybór był duży, po wstępnym przebadaniu większości kombinacji liczenia indukcyjności i pojemności zdecydowano się na dokładną

Tablica 1

Gęstość prądu powierzchniowego j_s oraz spadki potencjału wywołane nim w ściankach rezonatora (rys. 2 i 3)

Odcinek	Zakres zmiennej	$\vec{j_s}$	$\frac{2\pi}{I_0 Q_s} V_t = \frac{2\pi}{I_0 Q_s} \int Q_s \vec{j_s} \vec{dl}$
a	$r \in \langle 0, r_p \rangle$	$\frac{I_0 \varepsilon' r}{2\pi} \frac{1}{r_1^2 + (\varepsilon'-1) r_p^2} (+\vec{i_r})$	$\frac{\varepsilon' r_p^2}{2[r_1^2 + (\varepsilon'-1) r_p^2]}$
b	$r \in \langle r_p, r_1 \rangle$	$\frac{I_0}{2\pi r} \frac{r^2 + (\varepsilon'-1) r_p^2}{r_1^2 + (\varepsilon'-1) r_p^2} (+\vec{i_r})$	$\frac{r_1^2 - r_p^2 + 2(\varepsilon'-1) r_p^2 - \ln \frac{r_1}{r_p}}{2[r_1^2 + (\varepsilon'-1) r_p^2]}$
c	$z \in \langle L - L_2, L \rangle$	$\frac{I_0}{2\pi r_1} \frac{\cos k(L-z)}{\cos k L_2} (+\vec{i_z})$	$\frac{\operatorname{tg} k L_2}{r_1 k}$
d	$r \in \langle r_1, r_2 \rangle$	$\frac{I_0}{2\pi r} \frac{1}{\cos k L_2} (+\vec{i_r})$	$\frac{1}{\cos k L_2} \ln \frac{r_2}{r_1}$
e	$z \in \langle L - L_2, L \rangle$	$\frac{I_0}{2\pi r_2} \frac{\cos k(L-z)}{\cos k L_2} (-\vec{i_z})$	$\frac{\operatorname{tg} k L_2}{r_2 k}$
f	$z \in \langle L_1, L - L_2 \rangle$	$\frac{I_0}{2\pi r_2} (-\vec{i_z})$	$\frac{d}{r_2}$
g	$z \in \langle 0, L_1 \rangle$	$\frac{I_0}{2\pi r_2} \frac{\cos kz}{\cos k L_1} (-\vec{i_z})$	$\frac{\operatorname{tg} k L_1}{r_2 k}$
h	$r \in \langle r_1, r_2 \rangle$	$\frac{I_0}{2\pi r} \frac{1}{\cos k L_1} (-\vec{i_r})$	$\frac{1}{\cos k L_1} \ln \frac{r_2}{r_1}$
i	$z \in \langle 0, L_1 \rangle$	$\frac{I_0}{2\pi r_1} \frac{\cos kz}{\cos k L_1} (+\vec{i_z})$	$\frac{\operatorname{tg} k L_1}{r_1 k}$
j	$r \in \langle r_p, r_1 \rangle$	$\frac{I_0}{2\pi r} \frac{r^2 + (\varepsilon'-1) r_p^2}{r_1^2 + (\varepsilon'-1) r_p^2} (-\vec{i_r})$	$\frac{r_1^2 - r_p^2 + 2(\varepsilon'-1) r_p^2 - \ln \frac{r_1}{r_p}}{2[r_1^2 + (\varepsilon'-1) r_p^2]}$
k	$r \in \langle 0, r_p \rangle$	$\frac{I_0 \varepsilon' r}{2\pi} \frac{1}{r_1^2 + (\varepsilon'-1) r_p^2} (-\vec{i_r})$	$\frac{\varepsilon' r_p^2}{2[r_1^2 + (\varepsilon'-1) r_p^2]}$

Tablica 2

Porównanie współczynników K liczonych z rozwiązania polowego (6) ÷ (13) oraz z zależności obwodowych (18)

Lp.	$\frac{L}{d}$	N_{AX}	f_o	r_p [mm]	ε'	f/f_o	K obwodowe (18) z podstawieniem (16)	K obwodowe (18) z podstawieniem (17)	K całkowe (6)÷(13) $N = 1$	K całkowe (6)÷(13) $N = [N_{AX}/5]$	K całkowe (6)÷(13) $N = N_{AX}$
1	495	1250	151,53 MHz	3,65	0,969	0,969	0,9139	0,9293	0,9331	0,9281	0,9269
2				7	2,92	0,882	0,7335	0,7459	0,7815	0,7504	0,7385
3				1,6	151,7	0,678	0,4352	0,4637	0,5426	0,4818	0,4498
4				2,5	36,7	0,771	0,5654	0,5810	0,6468	0,5941	0,5693
5				6	9,64	0,716	0,4654	0,4901	0,5678	0,5082	0,4775
6				6	3,81	0,875	0,7215	0,7345	0,7724	0,7396	0,7268
7	54,4	267	398,76 MHz	3,65	2,57	0,977	0,9548	0,9596	0,9572	0,9566	0,9565
8				7	2,84	0,910	0,8162	0,8281	0,8276	0,8225	0,8209
9				1,6	142,9	0,729	0,5744	0,6048	0,6089	0,5953	0,5893
10				2,5	36,4	0,813	0,6841	0,7064	0,7081	0,6980	0,6941
11				6	9,64	0,762	0,5726	0,5979	0,6045	0,5907	0,5846
12				6	3,80	0,901	0,8032	0,8160	0,8157	0,8101	0,8083
13	3,91	37	2501 MHz	2,5	2,56	0,966	0,9068	0,9542	0,9310	0,9301	0,9300
14				7,55	2,00	0,836	0,6055	0,6570	0,6450	0,6348	0,6332
15				3,78	9,53	0,722	0,4912	0,5531	0,5421	0,5292	0,5270
16				7,55	9,66	0,459	0,1649	0,2177	0,2143	0,2027	0,2003
17				7,55	3,77	0,675	0,3732	0,4262	0,4231	0,4086	0,4059

analizę dwu z nich dających najlepsze wyniki. Można je zapisać następującymi zależnościami:

$$L_o = \frac{Z_o}{\omega_o}\left[\operatorname{tg}\frac{\omega_o L_1}{C} + \operatorname{tg}\frac{\omega_o L_2}{C}\right], \quad C_o = \frac{1}{\omega_o^2 L_o}.$$
$$L_I = \frac{Z_o}{\omega}\left[\operatorname{tg}\frac{\omega L_1}{C} + \operatorname{tg}\frac{\omega L_2}{C}\right], \quad C_1 = \frac{1}{\omega^2 L_I}. \tag{16}$$

$$L_o = \frac{Z_o}{\omega_o}\left[\operatorname{tg}\frac{\omega_o L_1}{C} + \operatorname{tg}\frac{\omega_o L_2}{C} + \frac{\omega_o d}{C}\right], \quad C_o = \frac{1}{\omega_o^2 L_o},$$
$$L_I = \frac{1}{\omega^2 C_I}, \quad C_I = C_o + \frac{(\varepsilon' - 1)\,\varepsilon_o \pi r_p^2}{d}, \tag{17}$$

gdzie:

L_o, C_o — odpowiednio indukcyjność i pojemność zastępcza rezonatora pustego,

L_I, C_I — odpowiednio idnukcyjność i pojemność zastępcza rezonatora z próbką.

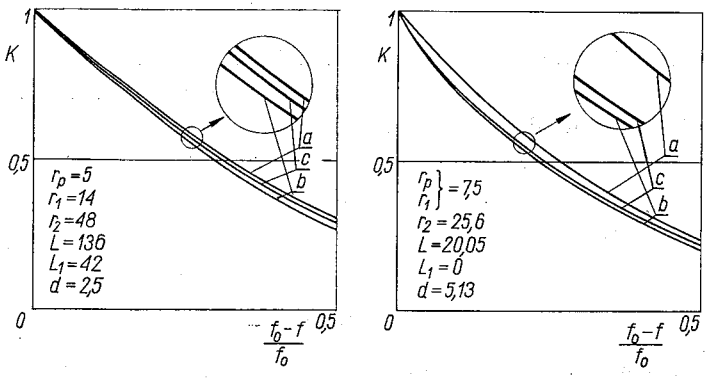
Korzystając z obwodu zastępczego złożonego z elementów dyskretnych interesujący nas współczynnik K można wyrazić następującą zależnością

$$K = \frac{R_{so}}{R_{s1}}\left(\frac{C_o L_I}{C_I L_o}\right)^{1/2} \tag{18}$$

gdzie: R_{so}, R_{s1} — rezystancja zastępcza wyznaczona odpowiednio dla rezonatora pustego i z próbką.

W celu stwierdzenia przydatności w ten sposób określonego współczynnika K, należało przeprowadzić porównanie z wynikami z pełnego algorytmu liczenia energii (6)÷(13). W tablicy 2 przedstawiono porównanie wyników otrzymanych z podanych wyżej dwóch zależności przybliżenia obwodowego (16) i (17) oraz trzech różnych długości szeregu funkcyjnego w obszarze A w przypadku rozwiązania polowego (6)÷(13).

Jak wynika z tablicy 2, zależności (17) podstawione do wzorów (15) i (18) określają współczynnik K z dobrą dokładnością dla szerokiej gamy proporcji rezonatorów i próbek. Postanowiono więc korzystać z nich przy obliczaniu pulsacji zespolonej $\hat{\omega}$. W praktycznie



Rys. 4. Zależność współczynnika K od przestrojenia względnego $f_o - f/f_o$ dla dwu rezonatorów o wymiarach jw., liczona następującymi sposobami: a) z wzorów przybliżonych (17), (18), b) z pełnych zależności polowych (6)÷(13) przy $N = NAX$, c) jak w punkcie b przy $N = NAX/5$

użytecznym zakresie pomiarowym błędy określenia tą metodą ilorazu Q'/Q_o nie przekraczały 5% (tablica 2). Dokładność w ten sposób określonego K poprawia się wraz ze zwiększaniem długości rezonatora. Jak dużą przewagę ma proponowana metoda nad bardzo często stosowanym przybliżeniem $Q = Q_o$ (czyli $K = 1$) pokazuje rysunek 4 wykonany dla dwóch konkretnych proporcji rezonatora i próbki. Krzywe zależności współczynnika K od wartości względnego przestrojenia częstotliwościowego rezonatora leżą w swoim bezpośrednim sąsiedztwie, oddalając się znacznie przy dużych przestrojeniach od linii $Q = Q_o$, czyli $K = 1$.

Podstawiając zależności (15), (17), (18) do (14) otrzymuje się następujące równanie na zredukowaną wartość zespolonej pulsacji

$$\frac{\hat{\omega}}{\omega} = 1+j\frac{1}{2}\left\{\frac{1}{Q}-\frac{1}{Q_o}\left(\frac{f}{f_o}\right)^{3/2}\left[1+\omega_o Z_o \varepsilon_o \pi(\varepsilon'-1)r_p^2\left(\operatorname{tg}\frac{\omega_o L_1}{C}+\right.\right.\right.$$

$$+\operatorname{tg}\frac{\omega_o L_2}{C}+\frac{\omega_o d}{C}\Bigg)\frac{1}{d}\Bigg]\Bigg[\frac{r_1^2+(\varepsilon'-1)\left(1+2\ln\frac{r_1}{r_p}\right)r_p^2}{r_1^2+(\varepsilon'-1)r_p^2}+\frac{C}{\omega}\left(\frac{1}{r_1}+\frac{1}{r_2}\right)\Bigg(\operatorname{tg}\frac{\omega L_1}{C}+$$

$$+\operatorname{tg}\frac{\omega L_2}{C}\Bigg)+\left(\frac{1}{\cos\frac{\omega L_1}{C}}+\frac{1}{\cos\frac{\omega L_2}{C}}\right)\ln\frac{r_2}{r_1}+\frac{d}{r_2}\Bigg]\Bigg[1+\frac{C}{\omega_o}\left(\frac{1}{r_1}+\frac{1}{r_2}\right)\Bigg(\operatorname{tg}\frac{\omega_o L_1}{C}+$$

$$+\operatorname{tg}\frac{\omega_o L_2}{C}\Bigg)+\left(\frac{1}{\cos\frac{\omega_o L_1}{C}}+\frac{1}{\cos\frac{\omega_o L_2}{C}}\right)\ln\frac{r_2}{r_1}+\frac{d}{r_2}\Bigg]^{-1}\Bigg\}. \tag{19}$$

Na zakończenie należy stwierdzić, że właściwości zaproponowanej przez autorów metody określenia Q'/Q_o (rys. 4) przy wykorzystaniu zależności (19) dały dobre wyniki w zastosowaniach praktycznych przy określaniu $\operatorname{tg}\delta$. Wartości błędów w ten sposób określonego $\operatorname{tg}\delta$ zostały przedstawione w [9].

4. PODSUMOWANIE

Opracowana metoda przybliżonego liczenia dobroci Q' ma następujące zalety: jest prosta, szybka i daje wynik Q' z dokładnością lepszą na ogół niż dokładność pomiaru dobroci, a dla rezonatorów długich (tzn. o dużym stosunku L/d) błędy kalkulacji Q' są pomijalnie małe (tablica 2). Jak wynika z (6) i (14) prowadzi to do precyzyjnego określenia pulsacji zespolonej rezonatora $\hat{\omega}$ z dokładnością limitowaną tylko przez techniczne możliwości pomiaru Q_o i Q.

LITERATURA

1. Hsi-Teh Hsieh i inni, *A Resonant Cavity Study of Semiconductors*, J. of Appl. Phys., v. 25, nr 3, pp. 302, 1954.
2. K. Froom, L. Essen, *The velocity of light and radio waves*, Izdatelstwo „Mir", Moskva 1973, s. 66—94.

3. A. Milewski, *Coaxial lumped capacitance resonators for the investigations of dielectrics*, Electron Technology, v. 10, pp. 71—98, 1977.
4. O. V. Karpova, *Ob absolutnom metodie izmerenia dielektrićeskich parametrow twerdowo tela s pomoščiu π-obraznowo rezonatora*, Fizika Twerdowo Tela, 1959, v. 1, s. 246—255.
5. A. A. Brandt, *Isledovanie dielektrikov na sverchvysokich častotach*. Izdatelstvo „Nauka", Moskwa 1963.
6. K. Bro i inni, *Rezonatory dla izmierenia parametrow dielektrykow w oblasti častot 100 Mgc*, AN SSSR, 1960, s. 116—123.
7. L. O. Goldsztain, N. W. Zernov, *Elektromagnitnyje pola i volny*, Mokva 1956.
8. V. V. Nikolskij, *Variacjonnyje metody dla vnutrennych zadač elektrodynamiki*, Izdatelstvo „Nauka", Moskva 1967, rozdział 2 i 5.
9. A. Kączkowski, A. Milewski, *Dokładna metoda określenia szerokiej gamy zespolonej przenikalności elektrycznej* $\hat{\varepsilon}$ *we współosiowym rezonatorze ze skupioną pojemnością*, Arch. Elektr., z. 3, tom 28, 1979, s. 501—522.
10. A. Milewski, *Zastosowanie metody perturbacyjnej do badania dielektryków w pasmie fal centymetrowych*, Praca doktorska, 1969, ITE PW.
11. M. Jaworski, *Teoretyczne podstawy pomiaru zespolonej przenikalności elektrycznej w pasmie mikrofalowym*, Praca doktorska, 1976, IF PAN.
12. H. M. Altschuler, M. Sucher, *Microwave measurements*. Politechnic Institute of Brooklyn 1963, rozdział 3.
13. L. Badian, A. Kączkowski, A. Milewski, *Dielektryczne właściwości ferrytów w pasmie mikrofalowym*, Elektronika, nr 5, s. 188, 1973.
14. L. Badian, A. Milewski, *Badanie dielektryków w pasmie 100—1000 MHz*, Elektronika, nr 7, 1973.
15. I. I. Eldumiati, G. I. Haddad, *Cavity perturbation techniques for measurement of the microwave conductivity and dielectric constant of a bulk semiconductor material*, IEEE Trans. on MTT, nr 2, 1972.
16. H. Sobol, J. J. Hughes, *Measurement of the permittivity of insulating films at microwave frequencies*, IEEE Trans. on MTT, nr 6, s. 378, 1967.

A. KĄCZKOWSKI, A. MILEWSKI

METHODS OF A COMPLEX RESONANT FREQUENCY DETERMINATION OF RE-ENTRENT RESONATORS WITH DIELECTRIC SAMPLE OF OPTIONAL DIAMETER

Summary

The elaboration of an analitic method of a complex resonant frequency determination of re-entrent resonators with dielectric sample of optional diameter is presented in the work. The proposed approximate method is based on the circuit electric model of resonator. For both of the proposed methods it is proved that errors of computing methods are usually less than measurement ones. That conclusion enables the simplification of algorithms without accuracy decreasing.

A. KĄCZKOWSKI, A. MILEWSKI

MÉTHODES DE DÉTERMINATION DE LA PULSATION DE RÉSONANCE COMPLEXE DU RÉSONATEUR COAXIAL À CAPACITÉ RECUEILLIE DU DIÉLECTRIQUE REMPLI NON HÉTÉROGÉNIQUEMENT

Résumé

Dans l'article on a examiné la méthode analytique de détermination de la pulsation de résonance complexe du résonateur coaxial à capacité recueillie du diélectrique rempli non hétérogéniquement. On

a proposé la méthode approximative basée sur le modèle du circuit d'un résonateur. On a analysé les différents modèles du circuit équivalent du résonateur. Sur l'example des deux modèles proposés, on a prouvé que les erreurs de détermination de la pulsation complexe à l'aide de cette méthode sont, pour la plupart, inférieures que les erreurs de mesure. Cela a permis de simplifier considérablement l'appareil mathématique sans abaisser en même temps le précision de mesure des composantes de la pulsation de résonance.

A. KĄCZKOWSKI, A. MILEWSKI

BESTIMMUNGSMETHODEN FÜR DIE RESONANZKOMPLEXPULSATION EINES KOAXIALEN, MIT DIELEKTRISCHEM KÖRPER INHOMOGEN GEFÜLLTEN RESONATORS MIT AGGREGIERTER KAPAZITÄT

Zusammenfassung

Es wurde die analytische Bestimmungsmethode für die Resonanzkomplexpulsation eines koaxialen, mit dielektrischem Körper inhomogen gefüllten Resonators mit aggregierter Kapazität überprüft. Eine angenäherte, auf einem Modell des Kreisresonators basierende Methode wurde dargestellt, auch wurden verschiedene Modelle des Kreises eines Ersatzrezonators analysiert. Für zwei dieser vorgeschlagenen Modelle wurde nachgewiesen, daß die Fehler bei der Bestimmung der komplexierten Pulsation mittels dieser Methoden größtenteils geringer sind, als die Meßfehler. Dieses gestattete, den mathematischen Apparat bedeutend zu vereinfachen unter gleichzeitigem Nichtabsinken der Meßgenauigkeit der Komponenten der Resonanzpulsation.

А. КОНЧКОВСКИ, А. МИЛЕВСКИ

МЕТОДЫ ОПРЕДЕЛЕНИЯ КОМПЛЕКСНОЙ РЕЗОНАНСНОЙ ЧАСТОТЫ СООСНОГО РЕЗОНАТОРА С СОСРЕДОТОЧЕННОЙ ЕМКОСТЬЮ С НЕОДНОРОДНЫМ ДИЭЛЕКТРИКОМ

Резюме

Предложен аналитический метод определения комплексной резонансной частоты п-образного резонатора с неоднородным диелектрическом выполнением. Предложен приближённый метод, базирующий на цепной модели резонатора. Проведен анализ двух моделей цепи резонатора. Для одной и другой доказано, что неточности определения комплексной частоты меньше чем ошибки измерения. Блогодоря этому, программа вычисления частоты является проще без ухудшения точности.

ROZPRAWY ELEKTROTECHNICZNE 1981, 27, z. 3, ss. 837-846

621.394.343: 621.395.3

Multidimensional Birth-Death Processes in the Analysis of Multi-User-Class Systems

JÓZEF LUBACZ (WARSZAWA), WOJCIECH BURAKOWSKI (WARSZAWA)

Instytut Telekomunikacji Politechniki Warszawskiej

The paper received 13.12.1979

The paper considers the solutions for probability of states of multidimensional birth-death processes in equilibrium and the utilization of these processes in the analysis of multi-user-class blocked-calls-cleared and blocked-calls-queued systems.

1. INTRODUCTION

In many of the public data transmission communication networks designed, the principle of circuit (channel) switching is utilized. Channel multiplexors, concentrators and exchanges are the basic components of these networks. For the optimal design of these components careful traffic analysis is necessary. From the point of view of traffic analysis these components are systems offering service to the traffic generated by data transmission terminals. Since terminals operate with various transmission speed rates, they differ from the point of view of service that has to be given to the traffic they generate. This makes it necessary to divide the set of terminals (users) into several subsets, called user-classes. As a consequence multidimensional (the dimension is the number of user-classes) traffic analysis has to be performed.

Several papers, for example [1] ÷ [4], have presented traffic analysis of multi-user-class systems. The papers give solutions for probabilities of states for systems in which blocked calls are not queued (blocked-calls-cleared systems) with the assumptions that:

— user-class calls form Poisson or quasi-Poisson streams,

— service times are negative-exponentially distributed.

It was observed that the solutions have a convenient product form.

In section 2. of this paper we analyse the conditions under which the product form of the solutions holds, using multidimensional birth-death processes and generalize the well known solutions.

In section 3. we present some remarks on the conditions that have to be taken into account if the systems in which blocked calls are queued (blocked-calls queued systems) are to be analysed with the use of multidimensional birth-death processes. Some limitations concerning queueing principles have to be made since it has been shown [5] that

under general conditions blocked-calls-queued systems cannot be modeled with multi-dimensional birth-death processes.

In section 4. we analyse several performance measures of multi-user-class blocked-calls-cleared systems pointing out the relations between them.

2. MULTIDIMENSIONAL BIRTH-DEATH PROCESSES

Let S be a set of k-vectors (states), $J = (j_1, j_2, \ldots, j_k)$, where $j_1, j_2, \ldots, j_k$ are non-negative integers. Let I_i be a k-vector for which $j_i = 1$ and $j_m = 0$ for $m \neq i$. States J and J' are said to be neighbouring states if $J - J'$ equals I_i or $-I_i$ for some i $(0 \leqslant i \leqslant k)$. A set S is said to be connected if for every two states $J, J^* \in S$ there exists a sequence of states $J_1, J_2, \ldots, J_n$ such that: J and J_1, J_j and J_{j+1} for $j = 1, 2, \ldots, n-1$, J_n and J^* are neighbouring states.

Further we consider only connected sets S. We say that a Markov chain taking values from S, is a multidimensional birth-death process if in an infinitesimal time interval (the definition is a straightforward generalization of the one-dimension case):

a) only transitions between neighbouting states are possible;
b) transitions from J to $J+I_i$ take place with the birth rate $\lambda_{j_i}^{(i)}$ (assuming J, $(J+I_i) \in S$);
c) transitions from J to $J-I_i$ take place with death rate $\mu_{j_i}^{(i)}$ (assuming J, $(J-I_i) \in S$)[1].

Let

$$S_J^+ = \{i\colon (J+I_i) \in S\},$$
$$S_J^- = \{i\colon (J-I_i) \in S\}.$$

We see that $\{J+I_i\}_{i \in S_J^+} \bigcup \{J-I_i\}_{i \in S_J^-}$ is a set of all states from S neighbouring with $J \in S$.

It can easily be shown that for every $J \in S$

$$\sum_{i \in S_J^+} [\lambda_{j_i}^{(i)} P(J) - \mu_{j_i+1}^{(i)} P(J+I_i)] = \sum_{i \in S_J^-} [\lambda_{j_i-1}^{(i)} P(J-I_i) - \mu_{j_i}^{(i)} P(J)] \tag{1}$$

where $P(J)$ is the probability of state J (in equilibrium). The problem is to find solutions for $P(J)$, $J \in S$, satisfying the equation above.

Let $J_s = (s_1, s_2, \ldots, s_k)$ be a fixed state from S. It can be verified that

$$P(J) = P(J_s) \prod_{i=1}^{k} P_i(J) \tag{2}$$

where

$$P_i(J) = \begin{cases} \prod\limits_{n=s_i+1}^{j_i} \dfrac{\lambda_{n-1}^{(i)}}{\mu_n^{(i)}} & \text{for} \quad j_i > s_i, \\ \prod\limits_{n=j_i}^{s_i} \dfrac{\mu_n^{(i)}}{\lambda_{n-1}^{(i)}} & \text{for} \quad j_i < s_i, \\ 1 & \text{for} \quad j_i = s_i \end{cases} \tag{3}$$

satisfy equation (1); $P(J_s)$ is determined by the condition $\sum\limits_{J \in S} P(J) = 1$.

[1] It should be noted that this definition is a modification of the usual definition.

The verification becomes straighforward if we notice that (2) implies

$$\lambda_{j_i}^{(i)} P(J) = \mu_{j_i+1}^{(i)} P(J+I_i)$$

and

$$\mu_{j_i}^{(i)} P(J) = \lambda_{j_i-1}^{(i)} P(J-I_i).$$

If $(0, 0, \ldots, 0) \in S$, then it is convenient to put $J_s = (0, 0, \ldots, 0) \triangleq \Theta$, since it leads to the simplification of (3):

$$P_i(J) = \prod_{n=1}^{j_i} \frac{\lambda_{n-1}^{(i)}}{\mu_n^{(i)}}. \tag{4}$$

The presented results are a generalization of the results from [1]÷[4]. It should be emphasized that the only assumption we made is that S is connected which implies that the „coordinate-convexity" of S assumed in [2] (if S is „coordinate-convex" then it is connected too) is not a necessary condition to receive solutions for $P(J)$ in a product form (2). The „coordinate-convexity" of S is often desirable in applications concerning communication systems [2].

3. MULTI-USER-CLASS BLOCKED-CALLS-QUEUED SYSTEMS

Many variants of multi-user-class, blocked-calls-cleared systems have been successfully analysed using multidimensional birth-death processes [1]÷[4]. We present some remarks on the possibility of using multidimensional birth-death processes in the analysis of multi-user-class, blocked-calls-queued systems.

It is obvious that we have to assume that the arriving calls form (for every user-class) a Poisson stream or a quasi-Poisson stream, and that the holding (service) times are (for every user-class) exponentially distributed. This assumption determines $\lambda_{j_i}^{(i)}$ as follows:

$$\lambda_{j_i}^{(i)} = \begin{cases} \lambda_i & \text{for } j_i \text{ such that } (J+I_i) \in S, \\ 0 & \text{otherwise}, \end{cases} \tag{5}$$

when the i-th user-class calls form a Poisson input stream (as usual, λ_i is the parameter of the stream), and

$$\lambda_{j_i}^{(i)} = \begin{cases} (N_i - j_i)\,\lambda_i' & \text{for } j_i \text{ such that } (J+I_i) \in S, \\ 0 & \text{otherwise}, \end{cases} \tag{6}$$

when the i-th user-class calls form a quasi-Poisson input stream, where N_i is the number of users in the i-th user-class; λ_i' is the i-th class individual user calling rate[2].

The assumption about the distribution of holding times is not sufficient to determine $\mu_{j_i}^{(i)}$. Some additional constraints concerning access and queueing have to be considered. These constraints must lead to the possibility of determining $\mu_{j_i}^{(i)}$ explicitly as a function of j_i for all j_i such that $(J-I_i) \in S$ (if $(J-I_i) \notin S$ we put $\mu_j^{(i)} = 0$) independently of all $j_n \neq j_i$. This forces us to consider only systems in which separate queues are formed for

[2] In the text below we write λ_i instead of λ_i' as it will not lead to confusion.

user-classes. It can be shown [5] that a system not sytisfying this condition leads out from processes of the birth-death type.

Next, we notice that the death rates $\mu_{j_i}^{(i)}$ depend on the number, j_i', of users being served. If the i-th user-class queue is not empty then $j_i' \neq j_i$ and j_i is unknown.

We may overcome this problem putting

$$\mu_{j_i}^{(i)} = \begin{cases} j_i\mu_i & \text{for} \quad j_i < m, \\ m\mu_i & \text{for} \quad j_i \geqslant m, \end{cases} \tag{7}$$

(assuming $(J-I_i) \in S$), where m is the number of servers of the analysed system, and μ_i is the parameter of the i-th user-class holding time distribution. We see that (7) is a straightforward generalization of the one-user-class $M/M/m$ system [6]. Unfortunately, for the multi-user-class case, definition (4) leads to systems that are not really interesting in applications. Considering (7) we see that the queue can be formed only for one user-class at a time and only in the cases when all m servers are occupied by users from this user-class. All calls from other user-classes are cleared when a queue is formed.

Definition (7) may be generalized in the following way:

$$\mu_{j_i}^{(i)} = \begin{cases} j_i\mu_i & \text{for} \quad j_i < l_i, \\ l_i\mu_i & \text{for} \quad j_i \geqslant l_i, \end{cases} \tag{8}$$

(assuming $(J-I_i) \in S$), where, for every i, l_i is a fixed number not greater than m.

In systems based on (8), queues for more than one user-class can be formed in the same time (if $l_1, \ldots, l_k$ are properly chosen). Unfortunately, definition (8) leads to the following disadvantages: an arriving call from i-th user-class finding $j_i = l_i$ is queued even if idle servers are available. Moreover, a queued call cannot be served even if idle servers are available, it has to wait until one of the i-th user-class calls being served leaves the system. Hence, the theoretically available throughput of the system is not fully utilized.

Nothing can be done about this if systems with unlimited waiting rooms are considered (and we insist on using birth-death type processes). Luckily, systems with limited waiting rooms are of most interest in applications, especially when quasi-Poisson inputs are considered. If M_i denotes the number of waiting places in the i-th user class waiting room then we may put:

$$\mu_j^{(i)} = \begin{cases} j_i\mu_i & \text{for} \quad j_i < l_i, \\ l_i\mu_i & \text{for} \quad l_i \leqslant j_i \leqslant l_i+M_i, \\ (j_i-M_i)\mu_i & \text{for} \quad j_i > l_i+M_i. \end{cases} \tag{9}$$

In this case, if $j_i = l_i+M_i$ and idle serves are available when a call from i-th user-class arrives, then a waiting call from the i-th user-class will enter service and the arrived call will take its place in the queue. Consequently, the utilization of the availability of the system will be increased (comp. with (8)).

Some rather straighforward generalization of (9) leads to systems in which more than one queue for i-th user-class is formed and the queueing discipline is of a combined FIFO and LIFO type. We feel that not much more can be done if we want to remain in the field of multi-user-class queueing systems analysed with multidimensional birth-death processes. One might expect that using birth-death processes of dimension $2k$ while analysing

k-user-class system may lead to new results. Unfortunately, this is not the case (cf. one-user-class queueing systems).

As we see from the analysis in this section, multidimensional birth-death processes are of limited usefulness in the analysis of blocked-calls-queued systems since they lead to systems which are not realy interesting in technical applications.

4. MULTI-USER-CLASS BLOCKED-CALLS-CLEARED SYSTEMS

4.1. State probabilities

We now present in some detail the analysis of blocked-calls-cleared systems with Poisson or quasi-Poisson input streams and negative-exponentually distributed service times i.e. we define $\lambda_{j_i}^{(i)}$ by (5) or (6) and put

$$\mu_{j_i}^{(i)} = \begin{cases} j_i\mu_i & \text{for } j_i \text{ such that } (J-I_i) \in S, \\ 0 & \text{otherwise.} \end{cases}$$

We consider systems in which i-th user-class calls ($i = 1, 2, \ldots, k$) demand service from m_i servers simultaneously ($m_i \geqslant 1$). Thus the set of states S equals

$$S_p = \left\{J: \sum_{i=1}^{k} j_i m_i \leqslant m\right\}$$

in the case of Poisson input, and

$$S_q = S_p \cap \{J: j_i \leqslant N_i,\ i = 1, 2, \ldots, k\}$$

for quasi-Poisson input (m is the number of servers).

With the assumptions made, eq. (4) gives the following well known results:

$$P(J) = P(\Theta) \prod_{i=1}^{k} \frac{\varrho_i^{j_i}}{j_i!} \tag{10}$$

for the Poisson input case, and

$$P(J) = P(\Theta) \prod_{i=1}^{k} \binom{N_i}{j_i} \varrho_i^{j_i} \tag{11}$$

for the quasi-Poisson input case; $\varrho_i = \lambda_i/\mu_i$. $P(\Theta)$ can by calculated from $\sum_{J \in S} P(J) = 1$ (S equals S_p or S_q respectively).

4.2. Performance measures

Relations between the most important performance measures of systems from 4.1 will be analysed.

The following measures will be considered:

B_i — i-th class call blocking probability (i-th class call congestion probability). i.e. the probability that an arriving i-th class call finds less than m_i idle servers,

B — blocking probability (call congestion probability): the probability that an arriving call finds fewer idle servers than it needs,

B_i' — the probability that the system is in a state in which it is unable to accept a i-th class call (i-th class call time congestion probability), i.e. there are fewer than m_i idle servers,

A_i — the traffic generated by i-th class calls,

A_i^c — the i-th class traffic carried (served) by the system,

U — the servers utilization factor: the average fraction of busy servers.

We begin with the analysis of relations between B_i and B_i'. Let $P_t(J)$ be the probability of state J at moment t, and $P_t^{(i)}(J)$ the probability of state J at moment t on the conditions that an i-th class call has just arrived. If $i_{\Delta t}$ denotes the event that an i-th class call has arrived in time interval $(t, t+\Delta t)$ then

$$P_t^{(i)}(J) = \lim_{\Delta t \to 0} P_t(J|i_{\Delta t}).$$

Using the Bayes formula we receive

$$P_t^{(i)}(J) = \lim_{\Delta t \to 0} \frac{P_t(i_{\Delta t}|J)P_t(J)}{\sum_{J \in S} P_t(i_{\Delta t}|J)P_t(J)} \tag{12}$$

where $P_t(i_{\Delta t}|J)$ is the probability of $i_{\Delta t}$ on the condition the system was in state J at moment t.

From the definition of multidimensional birth-death processes

$$P_t(i_{\Delta t}|J) = \lambda_{j_i}^{(i)} \Delta t + o(\Delta t)$$

thus from (12), after taking the limit we receive

$$P_t^{(i)} = \frac{\lambda_{j_i}^{(i)} P_t(J)}{\sum_{J \in S} \lambda_{j_i}^{(i)} P_t(J)}. \tag{13}$$

We will consider only systems in equilibrium so we may omit index t in $P_t(J)$ and also in $P_t^{(i)}(J)$ since (13) holds. In the text below the index will be omitted.

For systems with Poisson input $\lambda_{j_i}^{(i)} = \lambda_i$ thus from (13)

$$P^{(i)}(J) = P(J), \tag{14}$$

which means that the probability of state J at the moments of i-th class call arrivals equal $P(J)$; this is an intuitively clear and expected result.

For the quasi-Poisson input case $\lambda_{j_i}^{(i)} = (N_i - j_i)\lambda_i$ thus from (13)

$$P^{(i)}(J) = \frac{N_i - j_i}{\sum_{J \in S_q} (N_i - j_i) P(J)} P(J). \tag{15}$$

Equation (14) implies $B_i = B_i'$ for Poisson input systems so we turn to the problem of finding relations between B_i and B_i' for the quasi-Poisson input case. Substituting (11) into (15) and using the relation

$$\binom{N_i}{j_i}(N_i - j_i) = N_i \binom{N_i - 1}{j_i},$$

after simple transformations we receive the following equation:

$$P^{(i)}(J) = \frac{\binom{N_i-1}{j_i} \varrho_i^{j_i} \prod_{\substack{n=1 \\ n=i}}^{k} \binom{N_n}{j_n} \varrho_n^{j_n}}{\sum_{J \in S_q} \binom{N_i-1}{j_i} \varrho_i^{j_i} \prod_{\substack{n=1 \\ n=i}}^{k} \binom{N_n}{j_n} \varrho_n^{j_n}} \tag{16}$$

where

$$S_q' = S_q \setminus \{J \in S_q \colon j_i = N_i\}.$$

Comparing (16) with (11) we see that $P^{(i)}(J)$ for a system with $N_1, \ldots, N_k$ users equals $P(J)$ for a system with $N_1, \ldots, N_{i-1}, N_i - 1, N_{i+1}, \ldots, N_k$ users. This conclusion will be written shortly as follows:

$$P^{(i)}(J)|_{N_i} = P(J)|_{N_i - 1}. \tag{17}$$

Now let

$$S_i = \left\{J \in S_q' \colon m - \sum_{n=1}^{k} j_n m_n < m_i\right\}.$$

Summing both sides of (17) over all $J \in S_i$ we receive

$$B_i|_{N_i} = B_i|_{N_i - 1}$$

which means that the i-th class call congestion probability for a system having $N_1, \ldots, N_k$ users equals the i-th class time congestion probability for a system having $N_1, \ldots, N_{i-1}, N_i - 1, N_{i+1}, \ldots, N_k$ users. This is the result we were looking for. It is a generalisation of an analogous result for the one-dimension ($k = 1$) case [7].

We now turn to the analysis of the relations between B and B_i ($i = 1, 2, \ldots, k$). Using the following general formula:

$$P(C|A_1 \cup \ldots \cup A_k) = \sum_{i=1}^{k} P(C|A_i) P(A_i|A_1 \cup \ldots \cup A_k)$$

for $A_i \cup A_j = \Phi$ $(i \neq j)$, it can be shown that

$$B = \sum_{i=1}^{k} B_i p_i \tag{18}$$

where p_i is the probability that an arriving call is an i-th class call.

The probability p_i can be calculated as follows:

$$p_i = \lim_{\Delta t \to 0} \frac{\sum_{J \in S_q} P(i_{\Delta t}|J) P(J)}{\sum_{i=1}^{k} \sum_{J \in S_q} P(i_{\Delta t}|J) P(J)} = \frac{\sum_{J \in S_q} (N_i - j_i) \lambda_i P(J)}{\sum_{i=1}^{k} \sum_{J \in S_q} (N_i - j_i) \lambda_i P(J)}.$$

The B is in fact an average of the B_i's ($i = 1, \ldots, k$) taken over all user-classes.

It shoud be emphasised that there is no reasonable way of defining the notion of "average time congestion probability" (the average taken over user-classes).

Next we analyse the relation between $B_i(B)$ and A_i, A_i^c. Dividing the nominator and denominator of the left side of (13) by μ_i and summing both sides of the equation over all $J \in S_i^*$, where

$$S_i = \begin{cases} \left\{J \in S_p\colon m - \sum_{n=1}^{k} j_n m_n < m_i\right\} & \text{for the Poisson input case,} \\ S_i & \text{for the quasi-Poisson input case,} \end{cases}$$

we receive

$$B_i = \frac{\sum\limits_{J \in S_i^*} (\lambda_{j_i}^{(i)}/\mu_i) P(J)}{\sum\limits_{J \in S} (\lambda_{j_i}^{(i)}/\mu_i) P(J)} = \frac{\sum\limits_{J \in S} (\lambda_{j_i}^{(i)}/\mu_i) P(J) - \sum\limits_{J \in S \setminus S_i^*} (\lambda_{j_i}^{(i)}/\mu_i) P(J)}{\sum\limits_{J \in S} (\lambda_{j_i}^{(i)}/\mu_i) P(J)}$$

(S equals S_p or S_q depending on the type of input considered), i.e.

$$B_i = \frac{A_i - A_i^c}{A_i}. \tag{19}$$

The result is quite obvious intuitively.

Using (18) and (19), after a simple transformation, another intuitively clear resul can be obtained:

$$B = \frac{A - A^c}{A},$$

where

$$A = \sum_{i=1}^{k} A_i, \qquad A^c = \sum_{i=1}^{k} A_i^c$$

(the total traffic offered and carried respectively).

From definition the servers utilization factor

$$U = \frac{1}{m} \sum_{i=1}^{k} L_i m_i$$

where L_i is the average number of i-th class calls simultneously served. From Little formula applied for the considered systems

$$L_i = A^c{}_i, \tag{2}$$

so

$$U = \frac{1}{m} \sum_{i=1}^{k} A_i^c m_i.$$

We end with an approximate formula for U in the case of quasi-Poisson inp systems. The approximation is made for $B_i \ll 1$ $(i = 1, \ldots, k)$ — the most interesti case in applications. Simple calculations show that

$$A_i = (N_i - L_i)\varrho_i. \tag{}$$

For $B_i \ll 1$ $A_i \cong A_i^c$ as a consequence of (19).

Using this and (20) in (21) we receive

$$A_i^c \cong \frac{\varrho_i}{\varrho_i+1} N_i,$$

thus

$$U \cong \frac{1}{m} \sum_{i=1}^{k} \frac{\varrho_i}{\varrho_i+1} N_i m_i.$$

REFERENCES

1. G. F. W. Fredricson, *Analysis of channel utilization in traffic concentrators*, IEEE Trans. on Com., vol. COM-22, No. 8, 1974.
2. J. M. Aein, *A multi-user-class, blocked-calls-cleared, demand acces model*, IEEE Trans. on Com., vol. COM-26, No. 3, 1978.
3. O. Enomoto, H. Miyamoto; *An analysis of mixtures of multiple band-with traffic on dimension switching networks*, ITC, Stockholm 1973, 635/1—8.
4. L. Katzschner, R. Scheller, *Probability of loss of data traffics width different bit rates hunting one common PCM-channel*, ITC, Melbourne 1976, 525/1—8.
5. E. Wollner, *A queueing problem in data transmission*, ITC, Stockholm 1973, 344/1—6.
6. L. Kleinrock, *Queueing systems: vol. I — Theory*, J. Wiley, 1975.
7. R. B. Cooper, *Introduction to queueing theory*, Macmillan. 1972.

J. LUBACZ, W. BURAKOWSKI

VIELOWYMIAROWE PROCESY URODZIN I ŚMIERCI W ANALIZIE SYSTEMÓW OBSŁUGI Z WIELOMA KLASAMI UŻYTKOWNIKÓW

Streszczenie

W pracy przedstawiono ogólne rozwiązania na prawdopodobieństwa stanów wielowymiarowego rocesu urodzin i śmierci w stanie ustalonym. Przeanalizowano możliwości zastosowania wielowymiaro-ych procesów urodzin i śmierci do analizy systemów obsługi, w których użytkownicy podzieleni są na :ereg klas. Rozważono systemy z kolejką i bez kolejki zgłoszeń.

J. LUBACZ, W. BURAKOWSKI

ROCESSUS POLYDIMENSIONNELS NAISSANCE/MORT À L'ANALYSE DES SYSTÈMES DE SERVICE DES USAGERS PARTAGÉS EN CLASSES

Résumé

Dans le travail on a présenté les solutions générales pour la probabilité des états polydimensionnels ı processus de naissance/mort dans un état stationnaire. Ont été analysées les possibilités de l'appli-tion des processus polydimensionnels naissance/mort aux analyses des systèmes de service, dans lesquels ; usagers sont partagés en plusieurs classes. On a considéré le trafic avec et sans attente.

J. LUBACZ, W. BURAKOWSKI

MEHRDIMENSIONALE GEBURTS- UND TODESPROZESSE BEI DER ANALYSE DER BEDIENUNGSSYSTEME MIT MEHREREN TEILNEHMERKLASSEN

Zusammenfassung

Es wurde eine allgemeine Lösung bezüglich der Zustandswahrscheinlichkeit des mehrdimensionalen Geburts- und Todesprozesses im stationaren Zustand angegeben. Diskutiert wurde die Möglichkeit, mehrdimensionale Geburts- und Todesprozesse für die Analyse der Bedienungssysteme, anzuwenden, wo die Teilnehmer in eine Reihe von Klassen eingeteilt worden sind. Es wurden Warte- und Verlustsysteme erwogen.

Ю. ЛЮБАЧ, В. БУРАКОВСКИ

МНОГОМЕРНЫЕ ПРОЦЕССЫ РОЖДЕНИЯ И ГИБЕЛИ В АНАЛИЗЕ СИСТЕМ МАССОВОГО ОБСЛУЖИВАНИЯ СО МНОГИМИ КЛАССАМИ ПОЛЬЗОВАТЕЛЕЙ

Резюме

Представлены решения для вероятности состояний в многомерном процессе рождения и гибели. Анализируется возможность использования процессов рождения и гибели для анализа систем массового обслуживания со многими классами пользователей.

ROZPRAWY ELEKTROTECHNICZNE 1982, 27, z. 3, ss. 847-862

621.394.343: 621.395.3

Traffic Analysis of a Data Transmission Channel Concentrator

JÓZEF LUBACZ (WARSZAWA), WOJCIECH BURAKOWSKI (WARSZAWA),
MICHAŁ PIÓRO (WARSZAWA)

Instytut Telekomunikacji Politechniki Warszawskiej

The paper received 13.12.1979

The paper presents traffic analysis of a data transmission concentrator utilizing the principle of channel switching in time domain. The concentrator multiplexes traffic generated by data terminals operating at several transmission speed rates: 0.6 kbit/s, 2.4 kbit/s, 4.8 kbit/s, 9.6 kbit/s. A 64 kbit/s digital channel is the output of the concentrator.

1. INTRODUCTION

The paper presents the traffic analysis of a data transmission concentrator working in a network configuration shown in Fig. 1. The basis assumptions in the analysis are as follows. The data terminal equipments (DTE) operate with the following transmission speed rates: $s_1 = 0.6$ kbit/s, $s_2 = 2.4$ kbit/s, $s_3 = 4.8$ kbit/s, $s_4 = 9.6$ kbit/s (comp. CCITT Rec. X.1.) DTEs operating with the same speed rate s_i will be referred to as the i-th DTE class ($i = 1, 2, 3, 4$). The concentrator is a data channel concentrator (DCC) i.e. utilizing the principle of channel switching (in time domain). The DCC is connected to data switching equipment (DSE) via a 64 kbit/s channel organized according to the CCITT Rec.X.51 or X.50. In the DCC—DSE channel only 48 kbit/s is available for the transmission of data (comp. CCITT Rec.X.51 or X.50). The 48 kbit/s channel may be

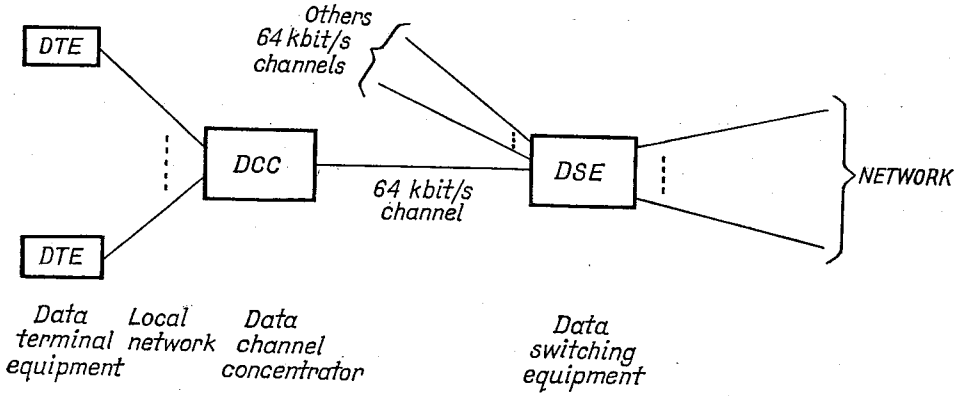


Fig. 1. Network configuration

regarded as a set of 80 fundamental channels of 600 bit/s each ($80 \cdot 600$ bit/s = 48 kbit/s). Thus in order to constitute a DCC—DSE connection for an i-th class DTE call m_i fundamental channels are needed where $m_1 = 1$, $m_2 = 4$, $m_3 = 8$, $m_4 = 16$.

The ratio of the sum of the DTEs transmission speed rates to 48 kbit/s—called the concentration factor (S) — is greater that 1 (i.e. concentration takes place). In a proper design of a DCC the concentration factor should be maximized. The greater the concentration factor, the greater the benefit of using DCC instead of a plain multiplexor (for which $S = 1$). The „price" paid for increasing S is the increase of blocking probabilities for DTE calls and the increase of call delay times (if queueing of blocked calls is considered).

In the paper we analyse relations between S and blocking probability as well as some statistical measures of call delay times for several DCC models. The considered models differ in the way the blocked calls are handled. We start with the analysis of a model in which blocked calls are cleared. Next we analyse several models in which blocked calls are queued. It should be emphasised that two types of DTE calls have to be considered: calls generated by DTE directed through DCC and DSE to the network, and calls coming from the network directed through DSE and DCC to DTE. The types of call will be abbreviated to DTE → DSE or DSE → DTE respectively. The two types of calls are assumed to be equally probable. We also assume that consecutive calls form a random stream (for each DTE class) and since the number of DTEs is finite we arrive at the assumption that each DTE class calls form a quasi-Poissonian process. The distinction becomes important when blocked calls queueing possibility is considered since for technical reasons DTE → DSE calls have to be queued in DCC while DSE → DTE calls have to be queued in DSE [2].

From the trafic analysis point of view the DCC may be seen as a system offering service to the DTE → DSE and DSE → DTE calls. Throughout this paper we assume that the DCC servers (which may be identified with fundamental channels) are chosen according to the „first idle available" principle. This assumption guarantees the maximization of the concentration factor.

In the paper we use traffic parameters forecasts published in [3] and [4].

2. DESCRIPTION OF DCC MODELS

For the purposes of traffic analysis of DCC it is convenient to take some queueing system as a DCC model (section 2.1). For such a system several performance measures (being the natural traffic characteristics of DCC) may be introduced (section 2.2). The traffic analysis is performed in section 3, in order to calculate these measures. The scheme of the analysis is described in section 2.3.

2.1. Queueing systems

DCC models considered here are queueing systems with the following general properties. There are 4 mutually independent quasi-Poissonian streams of calls. The i-th stream („generated" by the i-th DTE class, $i = 1, 2, 3, 4$) has parameters N_i, λ_i where N_i is the

number of DTEs in the i-th class and λ_i is the calling rate of every DTE from this class. N denotes the sum $N_1+N_2+N_3+N_4$.

Calls are served by 80 servers (fundamental channels). The service (i.e. constituting a connection to the DSE) of a i-th stream (i-th DTE class) call demands exactly m_i idle servers where $m_1 = 1$, $m_2 = 4$, $m_3 = 8$, $m_4 = 16$. The cumulated probability distribution of i-th class calls service times is given by $G_i(t)$. The sequence of service times of consecutive calls served (no matter to which class the calls belong) is a sequence of independent random variables. The average service time of i-th class calls will be denoted by $1/\mu_i$. We assume that if a i-th class call is taken into service then the first m_i idle servers are chosen.

In order to simplify the notation a DCC model will be denoted by $C(Q, K, R)$ for arbitrary parameters N_i, λ_i, G_i. Here $Q \subseteq \{1, 2, 3, 4\}$, K is a non-negative integer and $R \in \{0, 1\}$. The interpretation of the parameters Q, K, R is as follows.

If Q is empty no queue of waiting calls is formed, otherwise Q is the list of DTE classes whose calls are queued. If $R = 1$ there is one „waiting room" (organized in DCC) with K places and only DTE → DSE calls of classes from Q are queued. If $R = 0$ there are two waiting rooms having $\lfloor K/2 \rfloor$ waiting places each. In one of them (organized in DCC) DTE → DSE calls are queued and in the other (organized in DSE) DSE → DTE calls are queued (in both cases only calls from classes from Q are queued). The queueing discipline is always FIFO.

For $Q = \Phi$ we will write $C(\Phi)$ instead of $C(\Phi,K,R)$ since the parameters K, R are irrelevant.

2.2. Performance measures

While analysing DCC models we will consider the system in statistical equilibrium. We will be interested in the following performance measures of a DCC model.

1. B — call congestion probability for DTE calls (no matter to which class they belong).
2. B_i — i-th class call congestion probability ($i = 1, 2, 3, 4$).
3. $E(T_i)$ for $i \in Q$.
4. $E(T_i|T_i > 0)$ for $i \in Q$.

Here T_i denotes the random variable of the i-th class call waiting time, $E(\cdot)$ denotes the average, $E(\cdot\,|\,\cdot)$ denotes the conditional average („$T_i > 0$" is the event that an arriving i-th class call joins the queue).

5. $P(T_i > 0)$ for $i \in Q$.
6. t_i for $i \in Q$.

Here t_i is such a real number that $P(T_i > t_i) = \alpha$ (for α the values 0.01, 0.05 and 0.1 will be considered.)

2.3. Analysis scheme

The traffic analysis of the queueing system for each DCC model $C(Q,K,R)$ will be carried out in two main steps.

In the first step we assume that the parameters N_i, λ_i, μ_i, G_i satisfy the following conditions.

C1. For $i = 1, 2, 3, 4$

$$N_i/N = c_i$$

where $c_1 = 0.3$, $c_2 = 0.55$, $c_3 = 0.1$, $c_4 = 0.05$.

C2. The values λ_i, μ_i are given in the following table:

i	1	2	3	4
$1/\lambda_i$ [s]	480	50	600	360
$1/\mu_i$ [s]	120	5	90	90
λ_i/μ_i	0.25	0.1	0.15	0.25

C3. Service times are negative-exponentially distributed i.e.

$$G_i(t) = 1 - e^{-\mu_i t} \quad \text{for} \quad i = 1, 2, 3, 4.$$

For the $C(\Phi)$ model we will also consider systems satisfying C2, C3 and C1′ wher

C1′ ... $c_1 = 0.55$, $c_2 = 0.39$, $c_3 = 0.04$, $c_4 = 0.01$.

Assumptions C1, C2 are taken from [3], C1′ is taken from [4].

The concentration factor of system is the fraction

$$S = \frac{N_1 + 4N_2 + 8N_3 + 16N_4}{80}$$

which in fact is the ratio of the sum of the DTEs transmission speed rates to the throughput of the DCC—DSE channel. Let us notice that for a system satisfying C1 (or C1 $S = NC$ where $C = (c_1 + 4c_2 + 8c_3 + 16c_4)/80$. Hence it is enough to know S in order t determine N_i with satisfactory accuracy since one may put

$$N_i = \lfloor Sc_i/C + 0.5 \rfloor.$$

For a given $C(Q, K, R)$ model we will consider the class of all systems satisfying C1—C N is the only variable of this class or equivalently S is the only variable since $N = S/($

From the analysis we will derive the values of measures M1—M6 as the functio of argument S. Then we will be able to determine a maximal admissible concentratio factor S_{max} of a given DCC model. S_{max} may be informally defined as the greatest co centraction factor S for which the system from considered class has satisfactory valu for M1—M6. The determination of the satisfactory values is not with in the scope this paper. The discussion of this problem may be found in [2].

In the second step of the analysis we will examine the sensibility of S_{max} to the devi tions of assumed traffic parameters λ_i, μ_i and to the type of distributions G_i.

3. DCC ANALYSIS

In this section we assume that service times are negative-exponentially distributed i.e $G_i(t) = 1-e^{-\mu_i t}$ $(i = 1, 2, 3, 4)$ unless stated otherwise. The ratio λ_i/μ_i is denoted by ϱ_i and $\varrho = (\varrho_1, \varrho_2, \varrho_3, \varrho_4)$.

3.1. $C(\Phi)$ DCC model

We start the analysis with the simplest DCC model — $C(\Phi)$ (no queueing). Since $Q = \Phi$ we are interested in measures M1 and M2 exclusively.

System $C(\Phi)$ in statistical equilibrium was considered in [1] by means of multidimensional birth-death processes methods. The analytical relationships between B, B_i $(i = 1, 2, 3, 4)$ and N_i, λ_i, μ_i $(i = 1, 2, 3, 4)$ may be found there. Here we briefly summarize the results applied to the considered DCC model.

Let j_i be the number of i-th class calls simultaneously served $(i = 1, 2, 3, 4)$. A 4-tuple $J = (j_1, j_2, j_3, j_4)$ will be reffered to as a state. Let D denote the set of all admissible states of $C(0)$, i.e.:

$$D = \{J: j_1+4j_2+8j_3+16j_4 \leqslant 80\} \cap \{J: j_i \leqslant N_i \quad \text{for} \quad i = 1, 2, 3, 4\}.$$

Let $P(J)$, $J \in D$ be the probability of a state J. It may be shown that

$$P(J) = P_o \prod_{i=1}^{4} \binom{N_i}{j_i} \varrho_i^{j_i} \tag{3.1}$$

where $P_o = P(0, 0, 0, 0)$. P_o may be calculated from the normalization condition

$$\sum_{J \in D} P(J) = 1.$$

To determine B and B_is we are interested in another probability than $P(J)$ — the probability that an arriving call finds the system in state J.

Let $P_i(J)$ denote the probability that an arriving i-th class call finds the system in state J. One may prove that

$$P_i(J) = P(J) \frac{N_i - j_i}{\sum_{J \in D} (N_i - j_i) P(J)}. \tag{3.2}$$

If D_i is the set of all states such that an arriving i-th class call finds not enough idle servers (the call is cleared) i.e.

$$D_i = \{J \in D: j_1+4j_2+8j_3+16j_4 > 80-m_i\}$$

then

$$B_i = \sum_{J \in D_i} P_i(J) \quad \text{for} \quad i = 1, 2, 3, 4 \tag{3.3}$$

where B_i is the i-th class call congestion probability (M2).

To find the measure M1 i.e. call congestion probability B the equation

$$B = \sum_{i=1}^{4} B_i p_i \tag{3.4}$$

where

$$p_i = \frac{\sum_{J \in D} (N_i - j_i) \lambda_i P(J)}{\sum_{i=1}^{4} \sum_{J \in D} (N_i - j_i) \lambda_i P(J)} \tag{3.5}$$

may be used.

It is worth noticing that B_i depends on ϱ_i exclusively while B depends also on λ_i's.

Now we turn to the first step of the analysis of $C(\Phi)$ (cf. section 2.3). Assuming C1—C3 B_is and B are functions of one argument S — the concentration factor. Fig. 2 shows B as a function of S. It was shown [2] that in the considered data network configuration it is reasonable to take $2.5 \cdot 10^{-3}$ for the upper bound for admissible call congestion probability values. If we choose B to be the criterion of admissible S for $C(\Phi)$ then $S_{max} = 3.17$ and $S_{max} = 3.7$ for conditions C1 and C1′ respectively (cf. fig. 2). Fig. 3, depicting B_i's for C1, shows that the choice of B as the criterion is risky. For $S = 3.17$ the values of B_3 and B_4 are much greater than $2.5 \cdot 10^{-3}$. It seems reasonable to demand $B_i \leqslant 2.5 \cdot 10^{-3}$

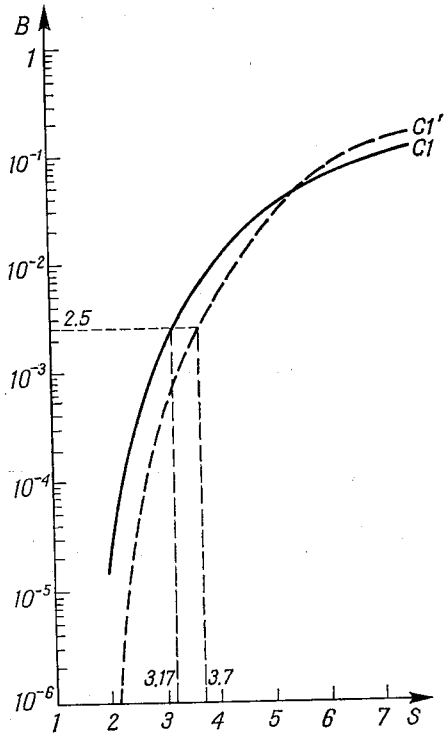


Fig. 2. Call congestion probability (B) for $C(\Phi)$

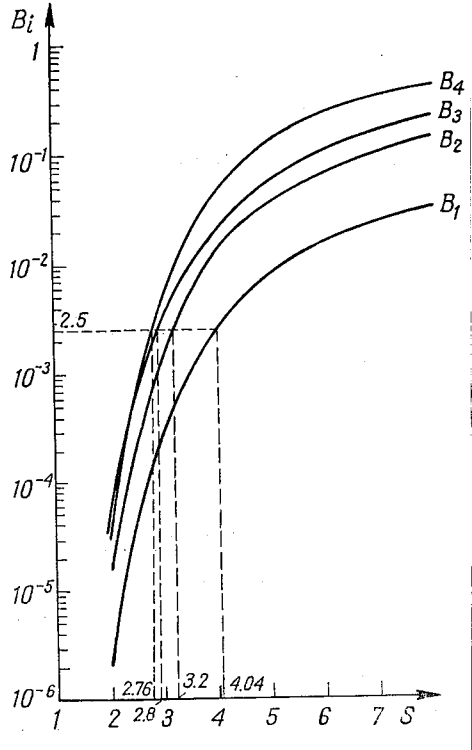


Fig. 3. DTE class call congestion probabilities (B_i) for $C1$ and $C(\Phi)$

($i = 1, 2, 3, 4$) to be fulfilled. In this case $S_{max} = 2.76$ for C1 however B_1 and B_2 are much smaller than $2.5 \cdot 10^{-3}$.

Fig. 4 shows that the situation is similar in the case of C1'.

In the remaining part of this paper we assume $B_i \leqslant 2.5 \cdot 10^{-3}$ ($i = 1, 2, 3, 4$) to be the necessary conditions for admissibility of a DCC model (in the case of $Q = \Phi$ they are also satisfactory conditions).

For $C(\Phi)$ we face the problem that B_1 and B_2 values are unnecessarily small for $S = S_{max}$. It will be shown that introducing the possibility of queueing ($Q \neq \Phi$) is one of the ways to overcome this problem.

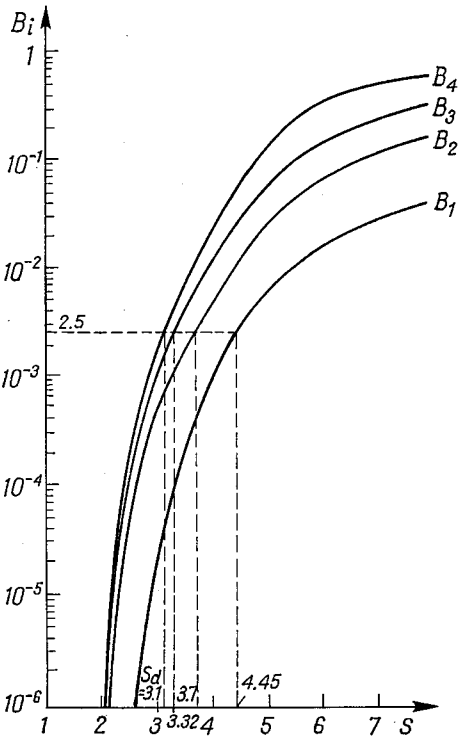


Fig. 4. DTE class call congestion probabilities (B_i) for C2 and $C(\Phi)$

Except from M1 and M2 we have examined another characteristics of $C(\Phi)$ — the average fraction of busy servers U. Fig. 5 depicts U as a function of S. For $S = 2.76$ we have $U = 0.34$.

We proceed to the second step of the analysis of $C(\Phi)$. We start with examining the influence of the deviations of λ_i, μ_i ($i = 1, 2, 3, 4$) values on M1 and M2. As we have noticed B_i depend on ϱ (cf. 3.2, 3.3) while B depends on ϱ and λ_i ($i = 1, 2, 3, 4$) separately (cf. 3.4, 3.5). It came out, however, that in the interesting range of λ_i values one may put $p_i \cong c_i$ ($i = 1, 2, 3, 4$). Hence we may assume that for a fixed S B_i and B are functions of ϱ (for $G_i(t) = 1 - e^{-\mu_i t}$) exclusively.

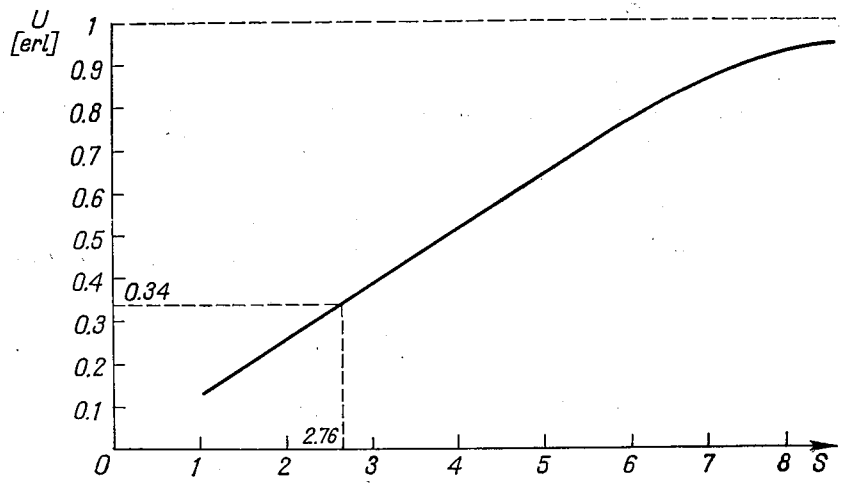


Fig. 5. Utilization factor (U)

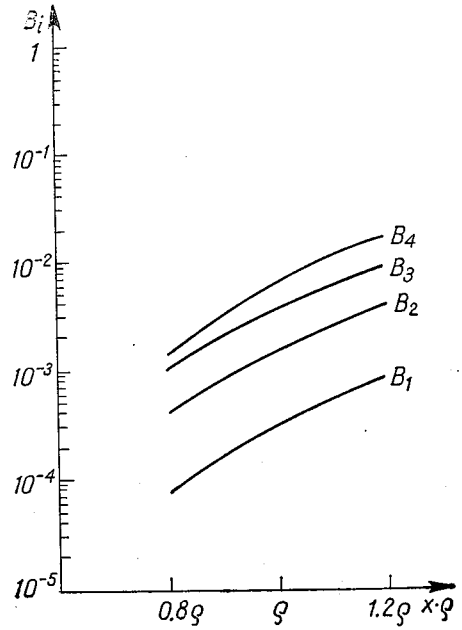


Fig. 6. The B_i's sensibility to ϱ deviations for $S = 3$ and $C(\Phi)$

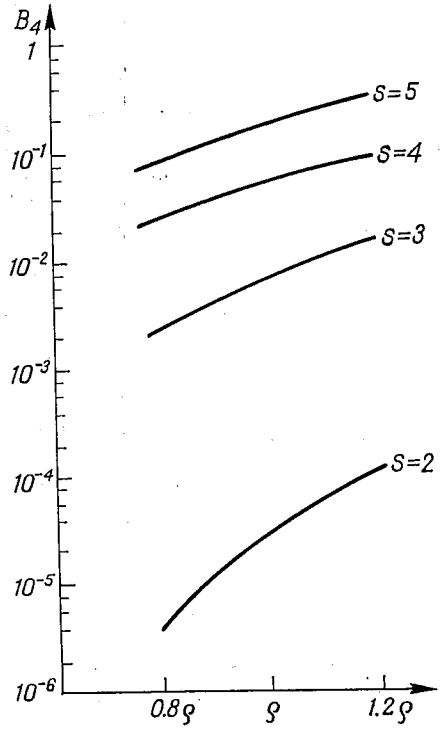


Fig. 7. The sensibility of B_n to ϱ deviations for $C(\Phi)$

The computations were carried out for the case in which all ϱ_i values change with the same ratio. Explicitly: we examined B_i and B for vectors $x\varrho = (x\varrho_1, x\varrho_2, x\varrho_3, x\varrho_4)$ for $0.8 \leqslant x \leqslant 1.2$ where ϱ is given by C2. Fig. 6 shows the sensibility of B_i to ϱ deviations for $S = 3$ (C1 assumed). From Fig. 6 it follows that B_1 is the most and B_4 is the

least sensible to ϱ deviations. The sensibility of B is similar to that of B_4. Since B, B_1, B_2, $B_3 < B_4$, it is enough to know the sensibility of B_4 in order to determine the sensibility of S_{max} to ϱ deviations. Fig. 7 and 8 depict the influence of ϱ on B_4 and S_{max} values.

In order to evaluate the sensibility of B_i's to the type of the G_i distributions we examined, except from the exponential distribution, the two following distruibutions: the one-point distribution

$$G_i(t) = \begin{cases} 0 & \text{for} \quad t < 1/\mu_i \\ 1 & \text{for} \quad t \geqslant 1/\mu_i \end{cases}$$

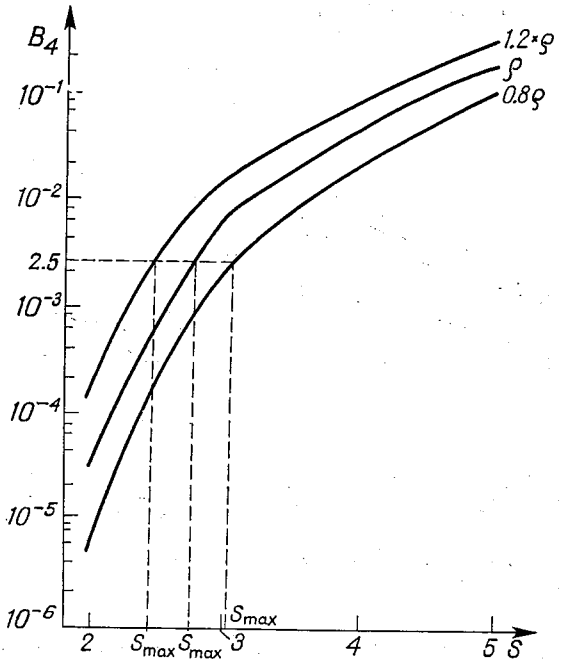


Fig. 8. The sensibility of S_{max} vs. $C(\Phi)$

nd the normal distribution $N(1/\mu_i, \sqrt{2}/\mu_i)$ normalized to $[0, \infty)$. The computations ere carried by means of computer aided simulation (cf. section 3.2). The obtained results em to show that the sensibility of B_i to the type of G_i is irrelevant. Probably the reason that the formulae for B_i are some generalizations of the Erlang loss formula.

3.2. $C(Q, K, R)$ DCC model

The analysis of a DCC model with $Q \neq \Phi$ (queueing possibility) is much more difcult. It comes out that the system with a waiting room cannot be solved by means of nalytical methods.

For the analysis of $C(Q, K, R)$ we applied a simulation program written in FORTRAN. he program (based on the individual source simulation method, cf. [5]) gives the possibi-

lity of analysing the $C(Q, K, R)$ DCC models for arbitrary parameters Q, K, R. Its functional description may be found in [2].

Since the computations are time consuming we restrict the analysis to C1, the condition which is more rigorous than C1′ from the traffic point of view (cf. fig. 3). As we noticed in section 3.1, the maximal admissible concentration factor S_{max} for $C(\Phi)$ is bounded by B_3 and B_4 values. The decrease of these values may be achieved by introducing the possibility of queueing 3rd and 4th class calls.

We start the first step of analysis with the simplest case $C(\{3, 4\}, \infty, 1)$ for which only 3rd and 4th class DTE → DSE calls are queued and there are infinitely many waiting places (in practice $K = N_3 + N_4$ is enough).

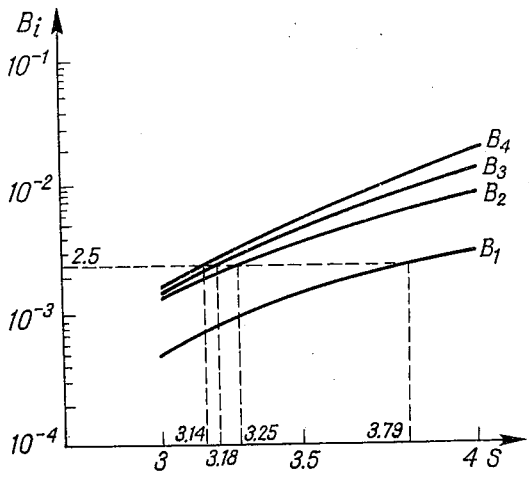


Fig. 9. DTE calls call congestion probabilities (B_i) for $C(\{3, 4\}, \infty, 1)$

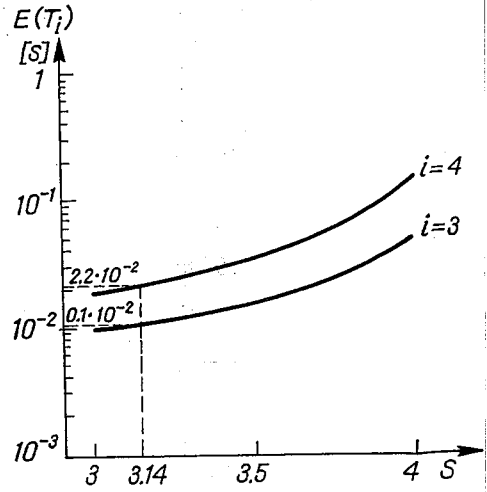


Fig. 10. Average waiting times for $C(\{3, 4\}, \infty,$

Fig. 9 shows B_i as functions of S. The B function is practically the same as B_2 an is not depicted. From Fig. 9 it follows that $S_{max} = 3.14$ (let us remind that $S_{max} = 2.$ for $C(\Phi)$) with the $B_i \leqslant 2.5 \cdot 10^{-3}$ criterion assumed. The „price" that has to be paid f the increase of S_{max} is the increase of waiting times for calls from the 3rd and 4th class. Figur 10, 11 and 12 show this „price".

For $S = 3.14$ if a call waits in the queue then it waits 2.6 seconds in average, $E(T_i | T_i > 0) = 2.6$ for $i = 3, 4$. The average waiting time is, however, very small ($E(T_i)$ $= 0.02$ s for $i = 3, 4$) since the necessity of waiting is quite unlikely ($P(T_3 > 0) = 4.1 \cdot 10^{-}$ $P(T_4 > 0) = 7.2 \cdot 10^{-3}$). It seems that M3, M4 and M5 values are admissible so S_{max} $= 3.14$ for $C(\{3, 4\}, \infty, 1)$.

We have examined also $C(\{3, 4\}, K, 1)$ models for $K < N_3 + N_4$. It came out th such modification of $C(\{3, 4\}, \infty, 1)$ is irrelevant since the M1—M5 values remain pract ally unchanged for $2 \leqslant K < N_3 + N_4$ and for $K = 1$ the S_{max} is smaller than for K $= N_3 + N_4$. $C(\{3, 4\}, \infty, 0)$ is an other modification of $C(\{3, 4\}, \infty, 1)$. In such a D model all 3rd and 4th class calls are queued. It may be verified that this modificatio not reasonable. In this case $B_3 = B_4 = 0$ and B_1, B_2 inrease. Since B_2 increases, t

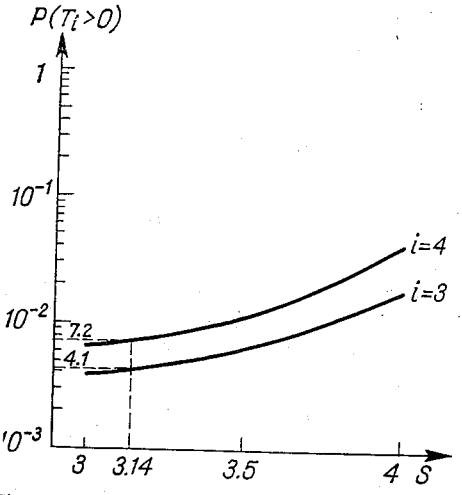


Fig. 11. Queueing probabilities for $C(\{3, 4\}, \infty, 1)$

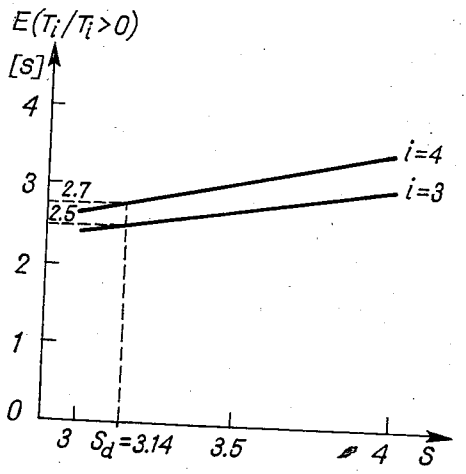


Fig. 12. Conditional average waiting times for $C(\{3, 4\}, \infty, 1)$

S_{max} values cannot be meaningfully greater than S_{max} for $C(\{3, 4\}, \infty, 1)$; 3.25 is the bound for S (cf. fig. 9). The conclusion is that $C(\{3, 4)\}, K, 1)$ is not competitive for $C(\{3, 4\}, \infty, 0)$ and will not be considered in detail.

The above considerations show that introducing the queueing possibility for the 3rd and 4th class makes $B_2 \leqslant 2.5 \cdot 10^{-3}$ the criterion for DCC model admissibility. The decrease of B_2 could be achieved by queueing 2nd class calls (along with the 3rd and 4th class calls) i.e. by introducing $C(\{2, 3, 4\}, K, 1)$ models. In order to increase S_{max} one has to put $K \gg N_3 + N_4$. Otherwise the calls of dominating 2nd class ($N_2 \gg N_3, N_4$; $\lambda_2 > \lambda_3, \lambda_4$) would block the 3rd and 4th class calls causing the inrease of B_3 and B_4. The B_3 and B_4 ncrease makes the improvement of S_{max} impossible.

On the other hand letting $K = \infty$ would cause the increase of the congestion probability

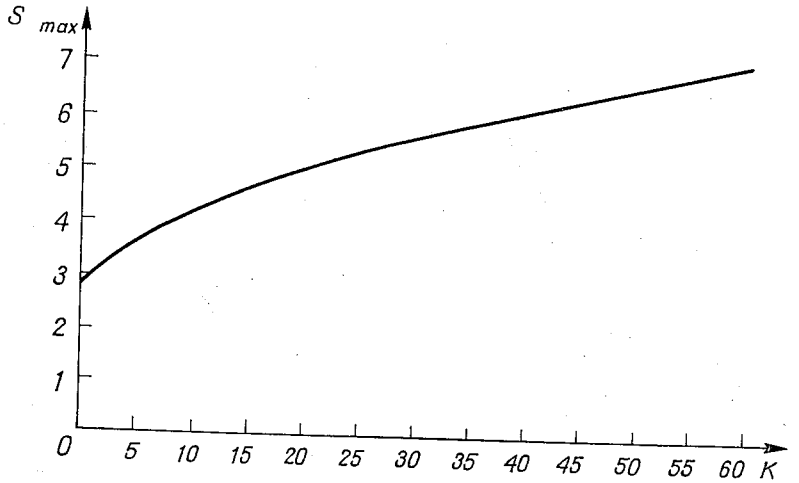


Fig. 13. S_{max} values for $C(\{1, 2, 3, 4\}, K, 0)$

for the DSE → DTE calls (which are not queued). This increase is greater than the decrease of the congestion probability for DTE → DSE calls. It follows that there exists some optimal K value for the $C(\{2, 3, 4\}, K, 1)$ model. The situation is not admissible since the optimal K values depends on traffic parameters of the systems so the $C(\{2, 3, 4\}, K, 1)$ model will not be considered further. Similar argumentations exclude the $C(\{1, 2, 3, 4\}, K, 1)$ model from our considerations.

To complete the first step of analysis the $C(\{1, 2, 3, 4\}, K, 0)$ model (all calls queued in two waiting rooms of $[K/2]$ places each) must be analysed.

From the point of view of M1 and M2 measures there is no limit for S_{max} in this class of models since $B_i \to 0$ when $K \to \infty$ $(i = 1, 2, 3, 4)$. Fig. 13 shows the increase of S_{max} (assuming $B_i \leqslant 2.5 \cdot 10^{-3}$ as the admissibility criterion) with the increament of K. It is obvious that in the case considered measures M3—M6 must be taken into account for the formulation of admissibility criterion since waiting times increase with the increment of K.

Figures 14—20 show the functional relationships between measures M3—M6 and S_{max} (where S_{max} determines K according to Fig. 13). One may observe that starting from $S_{max} = 4$ the values of M3—M6 increase rapidly.

We restrict the second step of the analysis to the most complicated $C(\{1, 2, 3, 4\}, K, 0)$ model. The examination of sensibility was carried in an analogous way to the $C(\Phi)$ case. Computations show that B_i's sensibility to ϱ deviations is greater that it was for the $C(\Phi)$ case. Generally, the sensibility increases with the increment of K. For $S = 4$ $(K = 10)$ B_i are about 1.5 times more sensitive to ϱ deviations in comparison with $C(\Phi)$. It came out that the $E(T_i)$'s sensibility to deviations is similar to the sensibility of B_i. It came out also that the sensibility of the results to the change of the G_i type is irrelevant in comparison with the sensibility to ϱ deviations.

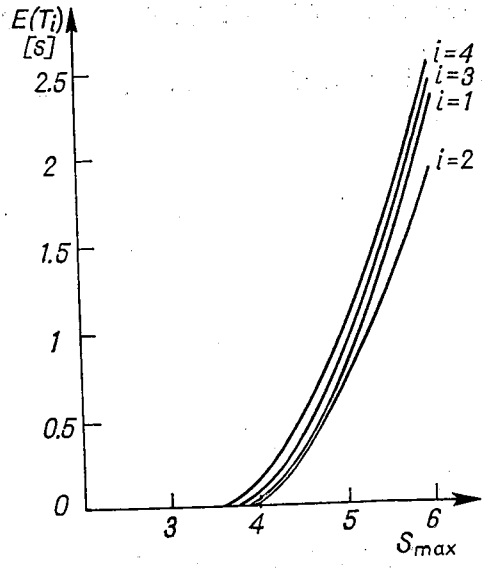


Fig. 14. Average waiting times for $C(\{1, 2, 3, 4\}, K, 0)$

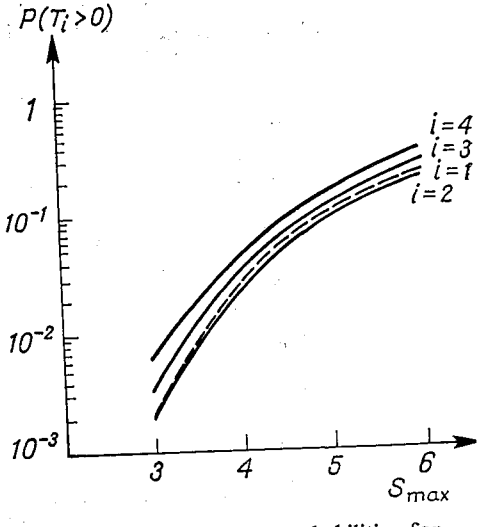
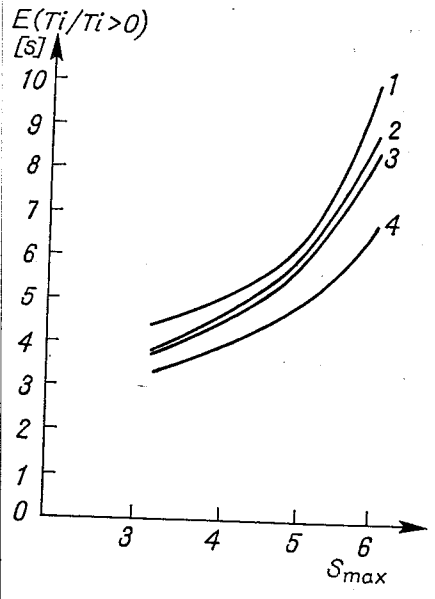


Fig. 15. Queueing probabilities for $C(\{1, 2, 3, 4\}, K, 0)$



ig. 16. Conditional average waiting times for $C(\{1, 2, 3, 4\}, K, 0)$

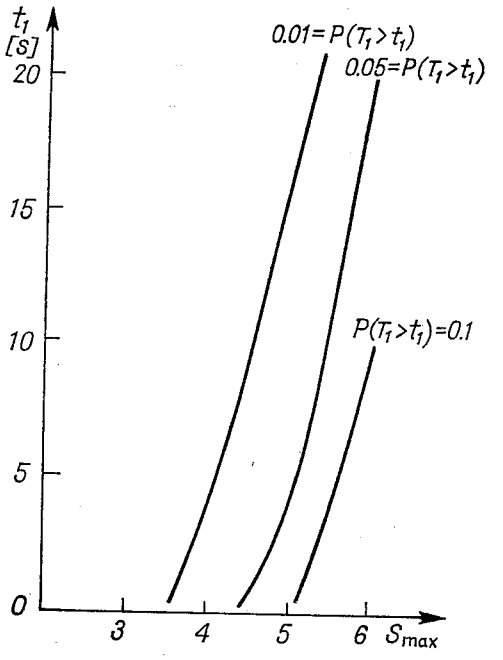
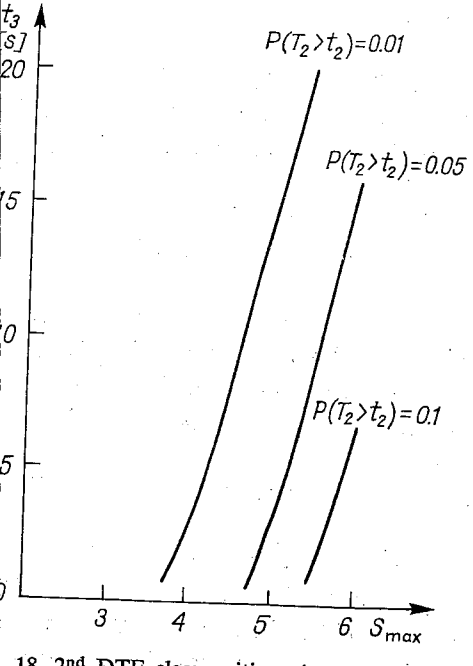


Fig. 17. 1st DTE class waiting time distribution for $C(\{1, 2, 3, 4\}, K, 0)$



. 18. 2nd DTE class waiting time distribution for $C(\{1, 2, 3, 4\}, K, 0)$

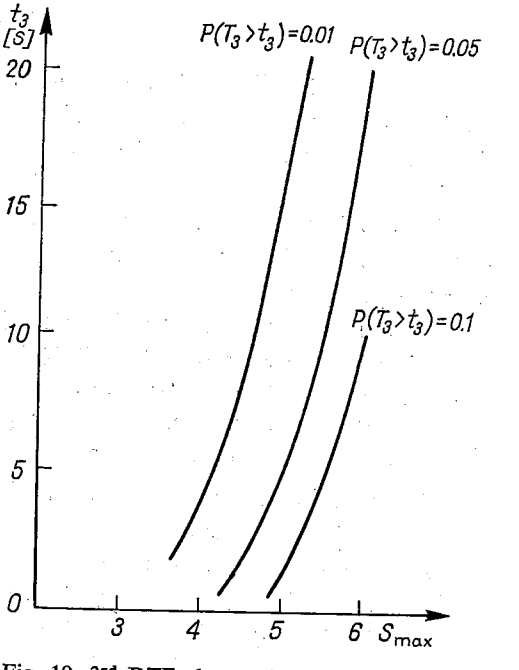


Fig. 19. 3rd DTE class waiting time distribution for $C(\{1, 2, 3, 4\}, K, 0)$

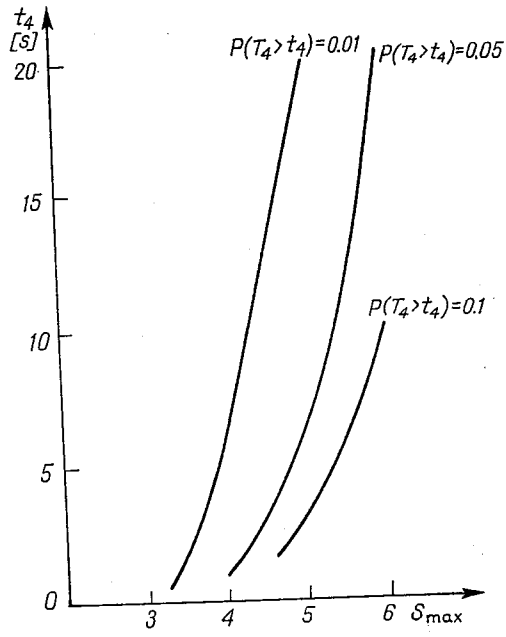


Fig. 20. 4th DTE class waiting time distribution for $C(\{1,2,3,4\}, K, 0)$

4. CONCLUSIONS

Several DCC traffic models were analysed. $C(\Phi)$, $C(\{3,4\}, \infty, 1)$ and $C(\{1,2,3,4\}, K,$ models were analysed in some detail. The remaining models have not been analysed i detail since they happened to be not competitive for three or not admissible for gener traffic reasons.

The problem of choosing the most suitable DCC model depends on the admissibili criterions selected. The question of choosing the admissibility criterions is beyond t scope of this paper.

The results presented are quite general but the comparison of DCC models was pe formed with the assumption that $B_i \leqslant 2.5 \cdot 10^{-3}$ is the admissibility criterion for bloc ing probabilities. For models with queueing the results are given as functions of c queueing time, therefore the choice of a suitable DCC model may be made if waiti time admissibility criterions are fixed.

We briefly summarize the results of the presented analysis for C1—C3 conditio and $B_i \leqslant 2.5 \cdot 10^{-3}$ assumed.

DCC model $C(\Phi)$ allows $S = 2.76$ to be maximal admissible concentration fact For $C(\{3,4\}, \infty, 1)$ $S_{max} = 3.14$ and the average call waiting time is approximately 0.02 $C(\{1,2,3,4\}, K, 0)$ DCC model gives the possibility of the increase of S_{max} as mu as one likes. The „price" for the increase is the increase of call waiting times. Beginn with $S_{max} = 4$ ($K = 10$) the waiting times grow rapidly.

A natural question arises: what is the benefit of using a DCC instead of multiplex

designed according to CCITT Rec.X.50 or X.51 in a data network? The answer may be given by S_M — the number of multiplexors needed to substitute one DCC with a given S_{max}. For $S_{max} = 2.76$ (i.e. $N_1 = 16$, $N_2 = 30$, $N_3 = 6$, $N_4 = 3$) $S_M = 3$, for $S_{max} = 3.14$ (i.e. $N_1 = 18$, $N_2 = 34$, $N_3 = 6$, $N_4 = 3$) $S_M = 4$, for $S_{max} = 4$ (i.e. $N_1 = 23$, $N_2 = 43$, $N_3 = 8$, $N_4 = 4$) $S_M = 5$.

REFERENCES

. J. Lubacz, W. Burakowski, *Multidimensional birth-death processes in the analysis of multi-user-class systems*. Rozprawy Elektroniczne, z. 3, 1981.
. *Modelowanie i optymalizacja algorytmów koncentratora*, Opracowanie zbiorowe Instytutu Telekomunikacji Politechniki Warszawskiej, Warszawa 1979.
. *Nordic Public Network Specification*, NDW 75-1, Doc. No. 91, 1975.
. P. Ostrowski, *Program zaspokajania potrzeb użytkowników krajowych w okresie do 1985 r. oraz realizacja powszechnej sieci teleinformatycznej*, Instytut Łączności, Warszawa 1977.
. H. Weisschuh, *Theoretical aspects of finite source input process simulation*, ITC7, Stockholm 1973, 215/1—4.

J. LUBACZ, W. BURAKOWSKI, M. PIÓRO

ANALIZA RUCHOWA KONCENTRATORA KANAŁÓW TELEINFORMATYCZNYCH

Streszczenie

W pracy przedstawiono wyniki analizy ruchowej koncentratora kanałów teleinformatycznych. Analiwany koncentrator stanowi element pośredniczący pomiędzy zbiorem abonentów transmitujących dane szybkościami 0,6 kbit/s, 2,4 kbit/s, 4,8 kbit/s bądź 9,6 kbit/s a centralą komutacji kanałów. Koncentrator z centralą połączony jest łączem o przepływności 64 kbit/s.

J. LUBACZ, W. BURAKOWSKI, M. PIÓRO

ANALYSE DU TRAFIC DANS LA VOIE DE TRANSMISSION DES DONNÉES PAR CONCENTRATEUR

Résumé

Dans le travail on a présenté les résultats de l'analyse du trafic du concentrateur des canaux téléormatique. Le concentrateur analysé constitue un élément qui sert d'intermédiaire entre l'ensemble abonnés, transmettant les données avec les vitesses 0,6 kbits/s, 2,4 kbits/s, 4,8 kbits/s ou 9,6 kbits/s, le central par une ligne à écoulement de 64 kbits/s.

J. LUBACZ, W. BURAKOWSKI, M. PIÓRO

RKEHRSANALYSE EINES KONZENTRATORS FERNMELDEINFORMATIVER KANÄLE

Zusammenfassung

Im vorliegenden Aufsatz wurde die Verkehrasnalyse eines Konzentrators fernmeldeinformativer näle gebracht. Der analysierte Konzentrator ist ein Vermittlungselement zwischen der Teilnehmermenge, Daten mit Geschwindigkeiten 0,6 kbit/s, 2,4 kbit/s, 4,8 kbit/s bezw. 9,6 kbit/s übertragen werden, und Kanalvermittlungsstelle. Der Konzentrator ist mit der Vermittlungsstelle durch eine Leitung mit Durchrate 64 kbit/s verbunden.

Ю. ЛЮБАЧ, В. БУРАКОВСКИ, М. ПЮРО

АНАЛИЗ ТРАФИКА В КОНЦЕНТРАТОРЕ КАНАЛОВ ПЕРЕДАЧИ ДАННЫХ

Р е з ю м е

В работе представлены результаты анализа трафика в концентраторе каналов передачи данных. Концентратор обсуживает абонентов работающих со скоростями 0,6 кбит/с, 2,4 кбит/с, 4,8 кбит/с и 9,6 кбит/с. Концентратор соединен с центром коммутации каналом 64 кбит/с.

ROZPRAWY ELEKTROTECHNICZNE 1981, 27, z. 3, ss. 863-883

621.39

Zasady tworzenia sieci programów sterujących przesyłaniem informacji cyfrowych w systemach telekomunikacyjnych

ANATOL BADACH (GDAŃSK), RYSZARD SOBCZAK (GDAŃSK),
MARIAN ZIENTALSKI (GDAŃSK)

Instytut Telekomunikacji Politechniki Gdańskiej

Otrzymano 4.1.1980

W pracy przedstawione są zasady tworzenia sieci przedstawiających współdziałanie programów, które są wykonywane przy sterowaniu przesyłaniem informacji cyfrowych pomiędzy procesami zgodnie z przyjętym protokołem komunikacyjnym. W tym celu sformułowano model analityczny protokołu komunikacyjnego, a następnie, na podstawie tego modelu, podano zasadę tworzenia sieci programów, których wykonanie zapewnia realizację przyjętego protokołu. Rozpatrzono zarówno przypadek, gdy kanał przez który przesyłana jest informacja jest kanałem idealnym, jak i przypadek, gdy przesyłana informacja może być w kanale zniekształcona.

Przedstawione tu podejście może być wykorzystywane przy projektowaniu sterowanych programowo systemów telekomunikacyjnych służących do przesyłania informacji cyfrowych, a zwłaszcza węzłów sieci telegraficznych i węzłów sieci transmisji danych budowanych na bazie odpowiednich systemów mini- lub mikrokomputerowych. W szczególności podana w pracy zasada pozwala wprowadzić samoczynne testowanie kanału i ułatwić opracowywanie oprogramowania dla systemów komputerowych wykorzystywanych w telekomunikacji.

1. WSTĘP

Głównym problemem, jaki obecnie powstaje przy projektowaniu programowo sterowanych systemów telekomunikacyjnych dla przesyłania informacji cyfrowych, jest zapewnienie współpracy urządzeń telekomunikacyjnych z systemami mini- lub mikrokomputerowymi. Powstaje przy tym potrzeba opracowania reguł programowego sterowania przesyłaniem informacji cyfrowych w systemach telekomunikacyjnych. Pewne zasady projektowania takich reguł zawarto w niniejszej pracy.

Sterowanie przesyłaniem informacji w nowoczesnych systemach telekomunikacyjnych, budowanych na bazie odpowiednich systemów komputerowych, odbywa się programowo układowo według pewnych z góry ustalonych zasad. Zbiór zasad określających sposób przesyłania informacji cyfrowej pomiędzy dwoma zachodzącymi w systemie telekomunikacyjnym lub systemie nadrzędnym „procesami" nazywać będziemy *protokołem komunikacyjnym* (PK) pomiędzy tymi procesami. Ogólnie mówiąc protokół taki jest zbiorem ustaleń dotyczących formatów przesyłanej informacji cyfrowej oraz wszelkich relacji czasowych przyjętych podczas jej przesyłania. Aby przesyłanie informacji pomiędzy

ustalonymi procesami odbywało się zgodnie z określonym protokołem komunikacyjnym muszą być opracowane odpowiednie programy sterujące tym przesyłaniem, a także zasady współdziałania tych programów. Zbiór wszystkich takich programów i zasady ich współdziałania, które zapewniają przesyłanie informacji zgodnie z określonym protokołem komunikacyjnym, tworzą *regułę sterowania przesyłaniem informacji* według tego protokołu komunikacyjnego. Regułę taką realizowaną z wykorzystaniem odpowiedniego systemu komputerowego możemy przedstawić w postaci prewnej sieci zwanej tu *siecią programów protokołu komunikacyjnego* (SP-PK). Należy na wstępie zaznaczyć, że pojęcie programu rozumiemy tu bardzo ogólnie. Pod tym pojęciem rozumiemy zarówno podprogram (procedurę), jak i dowolny ciąg rozkazów sterujących wykonywaniem ściśle określonych czynności.

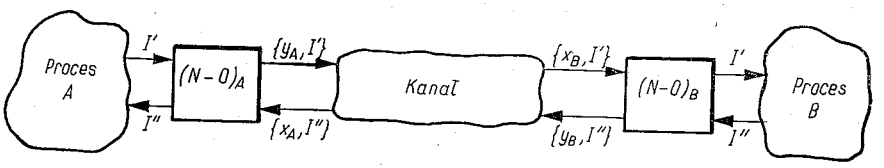
Aby przedstawić zasadę tworzenia musimy przyjąć odpowiedni model protokołu komunikacyjnego. W literaturze znane są modele protokołów komunikacyjnych tworzone głównie w oparciu o automaty skończone [1, 4], tablice logiczne [1], automaty sekwencyjne [1, 4], sieci Petriego [4, 7] i rozszerzone sieci Petriego [7, 10, 11]. Kilka innych metod tworzenia modeli protokołów komunikacyjnych zawiera praca [4]. Opis protokołów z wykorzystaniem modeli znanych z literatury nie jest w pełni dostosowany do realizacji protokołu na bazie systemu komputerowego. W szczególności opisy te nie są dostosowane do sposobu pracy systemów komputerowych czasu rzeczywistego. W szczególności nie uwzględniają one faktu, że realizacja protokołu w takim systemie odbywa się na zasadzie programowej obsługi *zgłoszeń* napływających z kanału, a „wchodzących" do systemu w postaci przerwań. Oczywiście, możliwa jest realizacja w systemie komputerowym protokołu o modelu zbudowanym w oparciu o automat, lecz realizacja ta nie jest efektywna. Modele protokołów zbudowane w oparciu o sieci Petriego to głównie modele symulacyjne utworzone dla celów weryfikacji protokołu [7]. W żadnym ze znanych modeli protokołów nie uwzględnia się możliwości powstania zniekształceń przesyłanej przez kanał informacji sterująco-kontrolnej (*s-k*), na podstawie której przebiega sterowanie przesyłaniem informacji właściwej zgodnie z przyjętym protokołem.

W niniejszej pracy sformułowano najpierw model analityczny protokołu komunikacyjnego, a następnie opierając się na tym modelu pokazano zasadę tworzenia sieci programów, których wykonanie zapewnia realizację przyjętego protokołu. Istotą przedstawionego tu podejścia jest to, że realizację protokołu w systemie komputerowym sprowadzono do programowanej obsługi zgłoszeń kierowanych do systemu w postaci przerwań. W przedstawionej sieci uwzględniono możliwość powstania w kanale zniekształceń przesyłanej informacji *s-k*. Zaproponowana sieć programów zapewnia możliwość kontroli wszystkich parametrów protokołu komunikacyjnego oraz realizację samoczynnej diagnostyki wybranych elementów systemu telekomunikacyjnego.

2. MODEL PROTOKOŁU KOMUNIKACYJNEGO

Weźmy pod uwagę sytuację przedstawioną na rys. 1, gdzie pomiędzy ustalonymi procesami A i B następuje przesyłanie informacji w obu kierunkach za pośrednictwem członów funkcjonalnych oznaczonych symbolicznie $(N\text{-}O)_A$ i $(N\text{-}O)_B$. Człony te nazywa

będziemy dalej *członami nadawczo-odbiorczymi*. Informację przesyłaną od procesu A do procesu B oznaczono symbolem I', natomiast w kierunku przeciwnym — symbolem I''. Przesyłanie informacji pomiędzy procesami A i B odbywa się zgodnie z ustalonym protokołem komunikacyjnym. Sterowanie przebiegiem tego przesyłania odbywa się w członach $(N\text{-}O)_A$ i $(N\text{-}O)_B$. Zgodnie z protokołem do informacji właściwej I' w członie $(N\text{-}O)_A$ dołączona jest informacja y_A i zespół $\{y_A, I'\}$ kierowany jest do kanału. Zespół ten odbierany przez człon $(N\text{-}O)_B$ oznaczono jako $\{x_B, I'\}$. Oczywiście, gdy w kanale nie występują zniekształcenia, to $y_A = x_B$. Ogólnie symbolem x oznaczać będziemy dalej *informację sterująco-kontrolną wejściową* do członu $N\text{-}O$, symbolem y — *informację sterująco-kontrolną wyjściową* z tego członu. W analogiczny sposób przebiega przesyłanie informacji w kierunku od procesu A do procesu B.



Rys. 1. Schemat funkcjonalny systemu przesyłania informacji pomiędzy procesami A i B

W literaturze przy formalnych opisach protokołów komunikacyjnych odpowiednikiem członu $N\text{-}O$ jest człon zwany *interlokutorem* [2, 4, 9]. Nazwy tej nie przyjęliśmy tu z tego względu, gdyż nie przyjmujemy tu w pełni struktury interlokutora, a ponadto ściślej interlokutorami, czyli rozmówcami, wypadałoby nazwać komunikujące się ze sobą procesy A i B.

Należy tutaj podkreślić, że człony $N\text{-}O$ oraz kanał należy traktować jako człony funkcjonalne i w związku z tym operacjom „nadawania" i „odbioru" należy też nadać funkcjonalny charakter. Podamy teraz kilka konkretnych przykładów, które powinny przybliżyć przedstawiony schemat i uwypuklić jego ogólność.

Weźmy pod uwagę dwa sąsiednie węzły sieci telegraficznej zbudowane na bazie systemów komputerowych. Jeżeli procesy A i B przebiegają w tych węzłach, to informacja pomiędzy tymi procesami przesyłana jest zgodnie z ustalonym protokołem komunikacyjnym, zwanym powszechnie *sygnalizacją*. W szczególności może to być sygnalizacja typu D [6], według której informacja przesyłana jest w postaci pojedynczych znaków.

Jeżeli procesy A i B są procesami przebiegającymi w dwóch sąsiednich węzłach sieci teleinformatycznej, w których dokonywana jest komutacja pakietów, wówczas procesy A i B mogą być przykładowo procesami dokonującymi dystrybucji informacji w tych węzłach, zaś człony $N\text{-}O$ zwane są programowanymi stykami [3]. Protokół komunikacyjny w takim przypadku nazywamy *protokołem międzywęzłowym* lub *procedurą wymiany informacji na poziomie węzeł-węzeł*. Spośród różnych protokołów międzywęzłowych najbardziej znanym jest protokół HDLC [5]. Zespoły informacji $\{y, I\}$ w tej sytuacji stanowią pakiety nadawane, zaś zespoły $\{x, I\}$ — pakiety odbierane. Każdy pakiet zawiera nagłówek, w którym zapisywana jest informacja *s-k* oraz pole informacyjne, w którym zapisywana jest informacja właściwa.

Jeżeli procesy A i B są procesami przebiegającymi w dwóch różnych komputerach obliczeniowych połączonych ze sobą poprzez sieć teleinformacyjną, to przedstawiony na

rys. 1 schemat funkcjonalny jest ilustracją realizacji protokołu komputer obliczeniowy-komputer obliczeniowy [3]. W tym przypadku uwidoczniony na rysunku kanał nie jest kanałem fizycznym, lecz tak zwanym *kanałem wirtualnym*, lub niekiedy także zwany jest *kanałem logicznym*. W sytuacji tej protokół komunikacyjny zapewnia przesyłanie informacji pomiędzy parą procesów nie połączonych ze sobą bezpośrednio kanałem fizycznym. Protokół ten jest więc zbiorem zasad sterowania przesyłaniem informacji poprzez kanał, który faktycznie nie istnieje na stałe, a jest tworzony chwilowo, przez odpowiednie połączenia szeregowe i równoległe wielu różnorodnych kanałów fizycznych. Z punktu widzenia procesów A i B tak utworzony kanał jest jak gdyby kanałem jednorodnym. Z tego względu zwany jest on właśnie kanałem wirtualnym. Człony *N-O*, w tej sytuacji, są modułami programowymi zapewniającymi współpracę komputerów obliczeniowych poprzez sieć teleinformatyczną. W tym przypadku zespół informacji tworzy blok informacji, który analogicznie jak pakiet zawiera zespół informacji *s-k* i zespół informacji właściwej.

Na tle tych przykładów widzimy, że ogólnie protokół komunikacyjny pomiędzy parą ustalonych procesów realizowany jest przez dwa człony *N-O*. Przejdźmy teraz do formalnego opisu pojedynczego takiego członu *N-O*. Przyjmujemy co następuje.

1. Informacja właściwa I nie ma wpływu na funkcjonowanie członu *N-O*.

2. Człon *N-O* może się znajdować jedynie w ściśle określonych stanach. Stany takie nazywać będziemy jego *stanami dopuszczalnymi*. Oznaczamy przez[1]

$$\mathscr{S} \equiv \{s_0, s_1, \ldots, s_I\} \tag{2.1}$$

zbiór stanów dopuszczalnych członu *N-O*, natomiast przez $\mathscr{I} = \{0, 1, \ldots, I\}$ zbiór numerów tych stanów.

3. W każdym ze stanów człon *N-O* „oczekuje" na ściśle określone postacie informacji[2]. Oznaczamy przez $\mathscr{X}_i$ zbiór takich postaci x, gdy człon *N-O* jest w stanie $s_i \in \mathscr{S}$. Natomiast zbiór wszystkich dopuszczalnych postaci x oznaczamy przez

$$\mathscr{X} = \bigcup_{i=0}^{I} \mathscr{X}_i. \tag{2.2}$$

4. Postacie zespołów informacji *s-k* wyjściowej y są ściśle określone i zależą także od stanu członu *N-O*. Oznaczamy przez $\mathscr{Y}_i$ zbiór dopuszczalnych postaci zespołów informacji *s-k* gdy człon *N-O* jest w stanie $s_i \in \mathscr{S}$. Zbiorem wszystkich dopuszczalnych postaci tej informacji jest

$$\mathscr{Y} = \bigcup_{i=0}^{I} \mathscr{Y}_i. \tag{2.3}$$

5. Zespoły informacji *s-k* napływają do członu *N-O* asynchronicznie w dyskretnych chwilach czasu. Zespół x odebrany w chwili t_n oznaczać będziemy przez $x(n)$, analogicznie zespół y nadany w chwili t_n — przez $y(n)$. Stan członu *N-O* w przedziale $(t_{n-1}, t_n$ oznaczać będziemy przez $s(n)$.

[1] Symbolem $\equiv$ w całej pracy oznaczamy równość definicyjną.

[2] Informacja *s-k*, w szczególności przy inicjacji komunikacji, może być także przekazywana przez system nadrzędny w stosunku do członu *N-O*, np. system operacyjny węzła sieci w sytuacji, gdy człon *N-O* realizowany jest w węźle. W takiej sytuacji w systemie komputerowym realizującym człon *N-O* z informacją *s-k* pojawia się zgłoszenie, mające sens zgłoszenia wewnętrznego, które przekazywane jest do programowej obsługi w postaci przerwania (wewnętrznego) realizowanego programowo.

6. Przyjmujemy, że w zbiorach $\mathscr{X}$, $\mathscr{S}$, $\mathscr{Y}$ określone są dwie relacje. Mianowicie relacja

$$\mathscr{F} \subset \mathscr{S} \times \mathscr{X} \times \mathscr{S} \tag{2.4}$$

zwana relacją typu *stan-stan* oraz relacja

$$\mathscr{G} \subset \mathscr{S} \times \mathscr{X} \times \mathscr{Y} \tag{2.5}$$

zwana relacją typu *stan-wyjście*. Relacja $\mathscr{F}$ jest zbiorem trójek uporządkowanych postaci (s_i, x, s_j), natomiast relacja $\mathscr{G}$ zbiorem trójek — (s_i, x, y). Zakładamy, że przy uwzględnieniu czasu do relacji $\mathscr{F}$ należy trójka

$$(s(n), x(n), s(n+1)), \tag{2.6}$$

natomiast do relacji $\mathscr{G}$ trójka

$$(s(n), x(n), y(n)). \tag{2.7}$$

Ze względu na postać uporządkowanych trójek (2.6) i (2.7) modelem członu *N-O* jest automat skończony, zwany w literaturze automatem Meale'a.

Dotychczasowe założenia dotyczyły w zasadzie dopuszczalnych postaci informacji *s-k* wejściowej i wyjściowej oraz relacji między nimi. Protokół komunikacyjny narzuca także pewne „wymagania" na dopuszczalne stany członu *N-O*. Sformułujemy teraz założenia dotyczące tych wymagań.

7. Z każdym stanem $s_i \in \mathscr{S}$ związany może być zespół liczb

$$\mathbf{v}_i \equiv (v_{i1}, v_{i2}, \ldots, v_{iK_i}), \tag{2.8}$$

zwanymi dalej *zmiennymi kontrolnymi stanu* s_i. Zbiorem kontrolowanych wartości zmiennej v_{ij}, $j = 1, 2, \ldots, K_i$, jest zbiór $\mathscr{V}_{ij}$.

Oznaczmy

$$\mathscr{V}_i \equiv \{\mathscr{V}_{i1}, \mathscr{V}_{i2}, \ldots, \mathscr{V}_{iK_i}\}. \tag{2.9}$$

Przyjmujemy dalej, że zbiorem kontrolowanych wartości zmiennej v_{i1} jest przedział $\mathscr{V}_{i1} = [V'_{i,1}, V'_{i,1}+V''_{i,1}]$, gdzie: $V'_{i,1}$ jest długością takiego okresu czasu, liczonego od chwili osiągnięcia przez człon *N-O* stanu s_i, że w okresie tym człon nie może osiągnąć innego stanu aniżeli s_i; $V''_{i,1}$ jest długością takiego okresu czasu, po którym człon *N-O* musi osiągnąć inny stan aniżeli s_i.

Okres czasu o długości $V'_{i,1}$, nazywać będziemy *okresem ochrony stanu*, zaś okres o długości $V''_{i,1}$ — *okresem wymuszenia zmiany stanu*. Interpretację tych okresów przedstawiono na rys. 2. W szczególności $V'_{i,1}$ może być równe zeru, wtedy okres ochrony nie istnieje. Ponadto, jeżeli $V''_{i,1} \to +\infty$, to mówić będziemy, że nie następuje wymuszenie zmiany stanu, gdy człon *N-O* znajduje się w stanie s_i. Pozostałe zmienne zespołu v_i mogą mieć różne znaczenie. W szczególności mogą określać ilości zasobów systemu komputerowego, na którego bazie realizowany jest człon *N-O*, rezerwowanych w stanie s_i. Ponadto

Rys. 2. Ilustracja okresu ochrony stanu i okresu wymuszenia zmiany stanu

Człon *N-O* osiąga stan s_i w chwili τ_i, O_i — okres ochrony stanu, W_i — okres wymuszenia zmiany stanu, $\tau'_i = \tau_i + V'_{i,1}$, $\tau''_i = \tau'_i + V''_{i,1}$

może istnieć zmienna określająca ile razy po kolei stan s_i został osiągnięty z tego samego stanu s_i itp.

Niech poszczególne zbiory $\mathscr{V}_i$, $i = 0, 1, \ldots, I$, będą elementami zbioru

$$\mathscr{V} \equiv \{\mathscr{V}_0, \mathscr{V}_1, \ldots, \mathscr{V}_I\}. \tag{2.10}$$

Zespół

$$\mathbf{R} \equiv \langle \mathscr{X}, \mathscr{S}, \mathscr{Y}, \mathscr{V}, \mathscr{F}, \mathscr{G} \rangle \tag{2.11}$$

nazywać będziemy *modelem analitycznym członu N-O*. Oznaczamy przez $\mathbf{R}_A$ i $\mathbf{R}_B$ odpowiednio model analityczny członu $(N\text{-}O)_A$ i $(N\text{-}O)_B$. Jako *model analityczny protokołu komunikacyjnego* ujmującego zasady wymiany informacji pomiędzy ustalonymi procesami A i B przyjmujemy zbiór

$$\mathbf{PK} \equiv \{\mathbf{R}_A, \mathbf{R}_B\}. \tag{2.12}$$

Należy tu podkreślić, że model (2.11) członu *N-O* dotyczy jedynie sytuacji, w której informacja jest przesyłana pomiędzy parą ustalonych procesów przez wydzielony kanał. W wielu przypadkach, a w szczególności w sieciach komputerowych, mamy taką sytuację, że przez pojedynczy kanał komunikuje się równocześnie wiele procesów. Ponadto protokoły komunikacyjne w sieciach komputerowych mają strukturę wielopoziomową [4]. W przypadku takich protokołów modele analityczne członów *N-O* są modelami bardziej złożonymi od modelu (2.11).

Podamy teraz zasady tworzenia sieci programów, których wykonywanie zapewnia realizację protokołu komunikacyjnego o modelu analitycznym (2.12).

3. SIEĆ PROGRAMÓW PROTOKOŁU W PRZYPADKU KANAŁU IDEALNEGO

Zakładamy, że człon *N-O* realizowany jest na bazie odpowiedniego systemu komputerowego (mini- lub mikro-). Opracowanie reguły sterowania takim członem sprowadza się do opracowania zasad współdziałania programów, zgodnie z którymi przetwarzane są poszczególne zespoły informacji *s-k* przekazywanej do systemu komputerowego w postaci tzw. przerwań. Takie przerwanie interpretować będziemy dalej jako napływające z kanału, a przy inicjacji komunikacji z systemu nadrzędnego w stosunku do członu *N-O*, zgłoszenie wymagające programowej obsługi. Ponieważ pomiędzy poszczególnymi zgłoszeniami istnieją pewne zależności czasowe i poszczególne zgłoszenia pomiędzy sobą są ściśle powiązane, przeto pojawia się problem opracowania zasad współdziałania programów, których wykonywanie traktować można jako procesy obsługi tych zgłoszeń. Podamy teraz ogólne zasady tworzenia takich programów.

Niech s_i będzie dowolnym stanem ze zbioru $\mathscr{S}$. Oznaczmy przez Ω_i zbiór wszystkich możliwych postaci informacji *s-k* wejściowej x, jakie mogą się pojawić na wejściu członu *N-O*, gdy jest on w stanie s_i. Zbiór

$$\mathring{\mathscr{X}}_i \equiv \{x \in \Omega_i : x \notin \mathscr{X}_i\} \tag{3.1}$$

nazywać będziemy *zbiorem niepożądanach postaci* x w stanie $s_i \in \mathscr{S}$. Oznaczmy przez Ω_{ij} zbiór wszystkich możliwych wartości zmiennej kontrolnej v_{ij} stanu s_i. Zbiór

$$\mathring{\mathscr{V}}_{ij} \equiv \{v_{ij} \in \Omega_{ij} : v_{ij} \notin \mathscr{V}_{ij}\} \tag{3.2}$$

nazywać będziemy *zbiorem niepożądanych wartości zmiennej kontrolnej* v_{ij}.

Mówić będziemy, że pomiędzy członami funkcjonalnymi $(N\text{-}O)_A$ i $(N\text{-}O)_B$ istnieje *kanał idealny*, gdy w każdym ze stanów obu członów *N-O* na wejściu pojawiają się oczekiwane postacie x oraz zmienne kontrolne poszczególnych stanów członu *N-O* osiągają jedynie wartości kontrolowane. Innymi słowy, w przypadku kanału idealnego w żadnym ze stanów $s_i \in \mathcal{S}$ na wejściu członu *N-O* nie może się pojawić $x \in \mathring{\mathcal{X}}_i$ oraz nie istnieje takie $j = 1, \ldots, K_i$, że $v_{ij} \in \mathring{\mathcal{V}}_{ij}$.

Przyjmować będziemy dalej w tym punkcie, że kanał pomiędzy parą członów *N-O* jest kanałem idealnym. Jeżeli ustalony człon *N-O* jest w stanie s_i, to po napłynięciu zgłoszenia z kanału może on w zależności od postaci zespołu informacji *s-k* wejściowej x osiągnąć jeden ze stanów należących do podzbioru

$$\mathcal{S}_i \equiv \{s \in \mathcal{S} : \bigwedge_{x \in \mathcal{X}_i} (s_i, x, s) \in \mathcal{F}\}. \tag{3.3}$$

Podzbiór $\mathcal{S}_i$ nazywać będziemy *podzbiorem stanów bezpośrednio osiągalnych ze stanu* $s_i \in \mathcal{S}$.

Załóżmy teraz, że człon *N-O* jest w stanie $s_i \in \mathcal{S}$. Aby zapewnić programową obsługę wszystkich zgłoszeń jakie mogą napłynąć w tym stanie z kanału, opracować należy odpowiedni podzbiór programów i określić zasady ich współdziałania. Zajmiemy się teraz tym problemem. W tym celu oznaczamy

$$\bar{\mathcal{S}}_i \equiv \{s_i\} \cup \mathcal{S}_i \tag{3.4}$$

i wprowadzamy związaną ze stanem s_i relację typu stan-stan postaci $\mathcal{F}_i \subset \bar{\mathcal{S}}_i \times \mathcal{X}_i \times \bar{\mathcal{S}}_i$, którą określamy jako

$$\mathcal{F}_i \equiv \{(s_i, x, s_j) : x \in \mathcal{X}_i \wedge s_j \in \mathcal{S}_i\}. \tag{3.5}$$

Oznaczmy przez $\mathcal{X}_{ij} \subseteq \mathcal{X}_i$ podzbiór takich postaci informacji x, przy których następuje w członie *N-O* przejście ze stanu s_i do stanu s_j; w szczególności może być $i = j$. Podzbiór ten określamy jako

$$\mathcal{X}_{ij} \equiv \{x \in \mathcal{X}_i : (s_i, x, s_j) \in \mathcal{F}_i\}. \tag{3.6}$$

Poszczególne podzbiory postaci (3.6) spełniają następujące warunki:

$$\mathcal{X}_i = \bigcup_{j \in \mathcal{J}_i} \mathcal{X}_{ij}, \tag{3.7}$$

$$\bigwedge_{j \in \mathcal{J}_i} \bigwedge_{j' \in \mathcal{J}_i} \mathcal{X}_{ij} \cap \mathcal{X}_{ij'} = \emptyset, \tag{3.8}$$

gdzie

$$\mathcal{J}_i \equiv \{j : s_j \in \mathcal{S}_i\} \tag{3.9}$$

jest zbiorem numerów wszystkich stanów bezpośrednio osiągalnych ze stanu s_i.

Przy założeniu, że zachodzą warunki (3.7) i (3.8) relację $\mathcal{F}_i$ zapisujemy jako

$$\mathcal{F}_i = \bigcup_{j \in \mathcal{J}_i} \mathcal{F}_{ij}, \tag{3.10}$$

gdzie

$$\mathcal{F}_{ij} \equiv \{(s_i, x, s_j) : x \in \mathcal{X}_{ij}\}. \tag{3.11}$$

Relacji $\mathcal{F}_{ij}$ przyporządkowujemy program $P'_{ij|x}$ zawierający zespół rozkazów wykonywanych przy zmianie stanu członu *N-O* ze stanu s_i na stan s_j. Program ten wykonywany

jest wtedy, gdy człon *N-O* jest w stanie s_i oraz pojawia się takie zgłoszenie, że zawiera ono $x \in \mathscr{X}_{ij}$.

Wprowadźmy teraz związaną ze stanem $s_i \in \mathscr{S}$ relację typu stan-wyjście $\mathscr{G}_i \subset \overline{\mathscr{S}}_i \times \times \mathscr{X}_i \times \mathscr{Y}_i$, którą określamy jako

$$\mathscr{G}_i \equiv \{(s_i, x, y): x \in \mathscr{X}_i \wedge y \in \mathscr{Y}_i\}. \tag{3.12}$$

Przy spełnieniu warunków (3.7) i (3.8) relację tę zapisujemy jako

$$\mathscr{G}_i = \bigcup_{j \in \mathscr{J}_i} \mathscr{G}_{ij}, \tag{3.13}$$

gdzie

$$\mathscr{G}_{ij} \equiv \{(s_i, x, y): x \in \mathscr{X}_{ij} \wedge y \in \mathscr{Y}_i\}. \tag{3.14}$$

Relacji $\mathscr{G}_{ij}$ przyporządkowujemy program $P''_{ij|x}$ zawierający zbiór rozkazów wykonywanych przy określaniu postaci zespołu y wyjściowej informacji *s-k* i podczas jej wysyłania do kanału. Procedura $P''_{ij|x}$ jest wykonywana, gdy określona jest postać x i następuje w członie *N-O* przejście ze stanu s_i do stanu s_j. W szczególności do relacji $\mathscr{G}_{ij}$ należeć może trójka (s_i, x, Θ), w której symbol Θ oznacza, że nie jest generowana żadna wyjściowa informacja *s-k*.

W realizacji praktycznej niekiedy parę programów $P'_{ij|x}$ i $P''_{ij|x}$ wygodniej jest potraktować jako jeden program wykonywany przy obsłudze zgłoszenia pojawiającego się wtedy, gdy człon *N-O* jest w stanie s_i i zawierającego zespół $x \in \mathscr{X}_{ij}$. Z tego więc względu oznaczamy

$$P_{ij|x} \equiv (P'_{ij|x}, P''_{ij|x}). \tag{3.15}$$

Symbol x przy indeksach w programie $P_{ij|x}$ oznacza, że ustalona jest już postać x. Z drugiej zaś strony zapis taki ma sugerować, że wykonanie programu zależy od postaci x. Oczywiście w chwili pojawienia się zgłoszenia z żądaniem obsługi postać x nie jest jeszcze określona. Oznaczmy zatem przez P_i *program wstępnej obsługi zgłoszenia* napływającego do członu *N-O*, gdy jest on w stanie s_i. Podczas wykonywania programu P_i ustalona jest postać zespołu x i modyfikowane są odpowiednie zmienne kontrolne. Po ustaleniu postaci x, gdy $x \in \mathscr{X}_{ij}$, następuje przejście do wykonywania programu $P_{ij|x}$.

Dotychczas rozpatrywaliśmy przejście ze stanu s_i do stanu s_j. Ponieważ ze stanu s_i osiągalne są stany o numerach ze zbioru $\mathscr{J}_i$, przeto przy przejściu ze stanu s_i do stanu s_j, $j \in \mathscr{J}_i$, wykonywany jest program ze zbioru

$$\mathscr{P}_{i|x} \equiv \{P_{ij|x}: j \in \mathscr{J}_i\}. \tag{3.16}$$

Taki zbiór programów zapewnia istnienie relacji $\mathscr{F}_i$ postaci (3.10) oraz relacji $\mathscr{G}_j$ postaci (3.13). Wykonywanie programu ze zbioru $\mathscr{P}_{i|x}$ możemy interpretować jako kontynuację procesu obsługi zgłoszenia, przybyłego do członu *N-O*, gdy był on w stanie s_i, po ustaleniu postaci zespołu informacji *s-k* wejściowej x.

Zbiór

$$\mathscr{P}_i \equiv \{P_i\} \cup \mathscr{P}_{i|x} \tag{3.17}$$

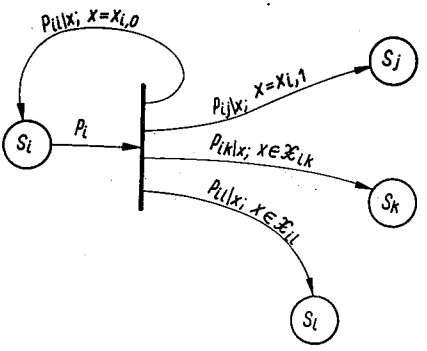
jest zbiorem wszystkich takich programów, które są związane ze stanem $s_i \in \mathscr{S}$. Oznaczmy

$$\overline{\mathscr{X}}_i \equiv \{\mathscr{X}_{ij}: j \in \mathscr{J}_i\}. \tag{3.18}$$

Podsiecią programów w stanie $s_i \in \mathscr{S}$ nazywać będziemy zespół

$$\mathscr{N}_i \equiv \langle \overline{\mathscr{S}}_i, \mathscr{P}_i, \mathscr{V}_i, \overline{\mathscr{X}}_i \rangle, \tag{3.19}$$

gdzie zbiór $\mathscr{V}_i$ określony jest wzorem (2.9). Podsieć $\mathscr{N}_i$ przedstawiać będziemy graficznie. Przykładową postać takiej podsieci przedstawiono na rys. 3.



Rys. 3. Przykładowa struktura podsieci programów
$\overline{\mathscr{S}}_i = \{s_i, s_j, s_k, s_l\}$, $\mathscr{X}_{ii} = \{x_{i,0}\}$, $\mathscr{X}_{ij} = \{x_{i,1}\}$, $\mathscr{X}_{ik} = \{x_{i,3}, x_{i,4}\}$, $\mathscr{X}_{il} = \{x_{i,4}, x_{i,5}\}$. Zastąpiono tu $x \in \mathscr{X}_{ii}$, przez $x = x_{i,0}$; analogicznie $x \in \mathscr{X}_{ij}$, przez $x = x_{i,1}$

Dotychczas rozpatrywaliśmy sytuację, gdy człon *N-O* był w ustalonym stanie $s_i \in \mathscr{S}$. Przeprowadzając podobne rozumowanie dla pozostałych stanów dopuszczalnych i tworząc dla nich podsieci, a następnie „łącząc" te podsieci otrzymujemy w wyniku sieć programów członu *N-O*, którą określamy jako

$$\mathscr{N} \equiv \langle \mathscr{S}, \mathscr{P}, \mathscr{V}, \overline{\mathscr{X}} \rangle, \tag{3.20}$$

gdzie:

$$\mathscr{P} \equiv \bigcup_{i=0}^{I} \mathscr{P}_i, \quad \mathscr{V} \equiv \{\mathscr{V}_i : i = 0, 1, \ldots, I\},$$
$$\overline{\mathscr{X}} \equiv \{\overline{\mathscr{X}}_i : i = 0, 1, \ldots, I\}. \tag{3.21}$$

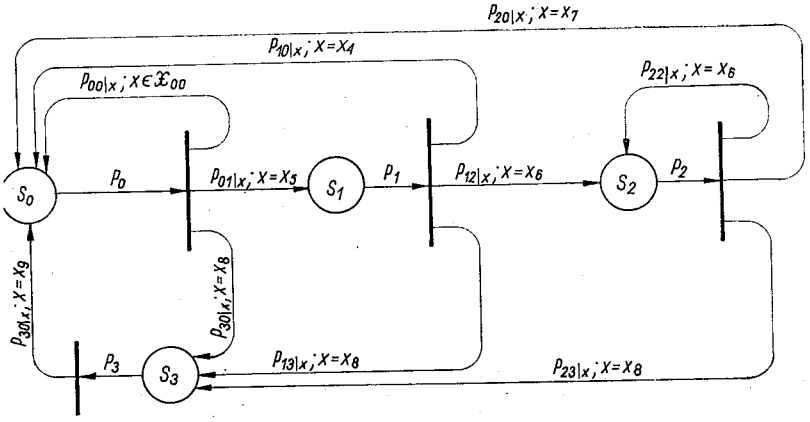
W szczególności mogą istnieć takie stany s_i oraz s_j, że $\mathscr{P}_i \cap \mathscr{P}_j \neq \emptyset$.

Przykład 1. Przedstawimy teraz dla członu *N-O*, zwanego w literaturze członem DCE [5], sieć programów zapewniającą, przy tzw. *trwałym połączeniu* (*permanent virtual call*) tego członu z innym członem *N-O*, zwanym DTE, przesyłanie informacji w postaci pakietów zgodnie z protokołem X.25. Zbiorem stanów dopuszczalnych członu DCE jest $\mathscr{S} = \{s_0, s_1, s_2, s_3\}$, gdzie: s_0 stan gotowości do przesyłania informacji, s_1 — stan oczekiwania na wznowienie przesyłania informacji, s_2 — stan oczekiwania na potwierdzenie, s_3 — stan rozłączenia. Sieć programów dla tego członu przedstawiono na rys. 4.

Zbiorem dopuszczalnych postaci zespołów informacji *s-k* wejściowej, w tej szczególnej sytuacji typów pakietów, jest zbiór $\mathscr{X} = \{x_1, x_2, \ldots, x_9\}$, gdzie: x_1 — dane od DTE, x_2 — krótka wiadomość od DTE, x_3 — DTE potwierdza krótką wiadomość, x_4 — człon DTE jest gotowy do odbioru, x_5 — człon DTE nie jest gotowy do odbioru, x_6 — żądanie wznowienia, x_7 — człon DTE potwierdza wznowienie, x_8 — żądanie rozłączenia, x_9 — człon DTE potwierdza rozłączenie.

Zbiorem postaci zespołów informacji *s-k* wyjściowej jest $\mathscr{Y} = \{y_1, y_2, \ldots, y_9\}$, gdzie y_1 — dane od DCE, y_2 — krótka wiadomość od DCE, y_3 — potwierdza krótką wiadomość, y_4 — człon DCE jest gotowy do odbioru, y_5 — człon DCE nie jest gotowy do odbioru, y_6 — żądanie wznowienia, y_7 — człon DCE potwierdza wznowienie, y_8 — żądanie rozłączenia, y_9 — człon DCE potwierdza rozłączenie.

Zbiór $\mathscr{V}$ określony wzorem (2.10), zawierający podzbiory dopuszczalnych wartości zmiennych kontrolnych poszczególnych stanów dopuszczalnych członu DCE, nie jest w pełni sprecyzowany w zaleceniu X.25. W szczególności w zaleceniu tym nie są podane dopuszczalne wartości zmiennych kontrolnych v_{i1}, $i = 0, 1, 2, 3$. Jest natomiast wprowadzona zmienna kontrolna, którą oznaczamy tu jako v_{02}, określająca ile razy kolejno wykonywana jest procedura $P_{00|x}$. Zbiorem dopuszczalnych wartości v_{02}, może być zbiór $\mathscr{V}_{02} = \{0, 1, \ldots, 7\}$ albo $\mathscr{V}_{02} = \{0, 1, \ldots, 127\}$.



Rys. 4. Sieć programu dla funcjonalnego członu DCE określonego zgodnie z X.25; $\mathscr{X}_{00} = \{x_1, x_2, x_3, x_4\}$

Oznaczamy przez $\mathscr{N}_A$ i $\mathscr{N}_B$ sieć programów odpowiednio członu $(N\text{-}O)_A$ i $(N\text{-}O)_B$. Protokół komunikacyjny, zgodnie z którym odbywa się przesyłanie informacji pomiędzy procesami przedstawionymi na rys. 1, realizowany jest w oparciu o zbiór sieci programów

$$\mathscr{N}_{PK} = \{\mathscr{N}_A, \mathscr{N}_B\}, \tag{3.22}$$

zwany dalej *zbiorem sieci programów protokołu komunikacyjnego.*

Ze względu na to, że zasady tworzenia sieci programów dla obu członów $(N\text{-}O)_A$ i $(N\text{-}O)_B$ są identyczne, przeto dalej rozpatrywać będziemy wyłącznie sieci programów dla pojedynczego członu *N-O*.

Podstawowym problemem, jaki powstaje przy tworzeniu sieci programów jest ustalenie zbioru dopuszczalnych stanów $\mathscr{S}$. W zależności od postaci tego zbioru otrzymujemy różne struktury sieci. Pojawiają się więc dwa problemy. Pierwszy to problem takiego wyboru sieci, by była ona obrazem prawidłowo funkcjonującej reguły sterującej przebiegiem przesyłania informacji; a drugi, to problem takiej optymalizacji tej sieci, by realizacja reguły, którą ona opisuje, była jak najefektywniejsza. W tej pracy zajmiemy się wyłącznie pierwszym problemem, gdyż rozwiązanie problemu drugiego jest zbyt złożone. Na to by sieć przedstawiała prawidłowo funkcjonującą regułę, musi ona posiadać pewne własności. O sieci programów, która jest obrazem prawidłowo funkcjonującej reguły, mówić będziemy, że jest ona *siecią dopuszczalną.*

4. WŁASNOŚCI DOPUSZCZALNEJ SIECI PROGRAMÓW

Aby podać własności, jakie musi mieć dopuszczalna sieć programów członu *N-O* wprowadzimy kilka pojęć. W zbiorze $\mathscr{S}$ stanów dopuszczalnych członu *N-O* wyróżniam

jeden stan s_0, zwany dalej *stanem początkowym*. Jako stan początkowy przyjmujemy taki stan, że człon *N-O* znajduje się w tym stanie, gdy z kanału nie pojawiło się żadne zgłoszenie z żądaniem obsługi.

Mówić będziemy, że w sieci *stan s_k jest osiągalny ze stanu s_i w L krokach*, gdy w sieci tej może być wykonany następujący ciąg

$$\mathbf{Q}_{i,k}(L) \equiv (\mathbf{Q}_{i,j_1}, \mathbf{Q}_{j_1,j_2}, \ldots \mathbf{Q}_{j_{L-1},k}) \tag{4.1}$$

L programów o postaci

$$\mathbf{Q}_{j,k} \equiv P_j, P_{jk|x}).$$

W szczególności w sieci przedstawionej na rys. 4 stan s_2 jest osiągalny ze stanu s_0 w dwóch krokach.

Jeżeli w sieci istnieje taki stan $s_i \in \mathcal{S}$, że może być wykonany ciąg programów $\mathbf{Q}_{i,i}(L)$, to mówić będziemy, że w sieci istnieje *cykl* o długości L i o początku w stanie s_i. Cykl taki oznaczać będziemy symbolem $\mathbf{C}_i(L)$. Oznaczamy przez $\mathcal{J}[\mathbf{C}_i(L)]$ zbiór numerów wszystkich stanów, które osiąga człon *N-O* przy wykonywaniu programów z cyklu $\mathbf{C}_i(L)$. O stanach s_j, $j \in \mathcal{J}[\mathbf{C}_i(L)]$, mówić będziemy, że są one *powtarzalne co L kroków*. W sieci przedstawionej na rys. 4 istnieje między innymi cykl $\mathbf{C}_0(3)$, w którym $\mathcal{J}[C_0(3)] = \{0, 1, 2\}$.

Jeżeli istnieje cykl $\mathbf{C}_i(1)$, stan s_i nazywać będziemy *stanem wymuszającym się*. W szczególności w sieci przedstawionej na rys. 4 stanami wymuszającymi się są stany s_0 i s_2.

O stanie s_i mówić będziemy, że jest *stanem pochłaniającym*, gdy podzbiorem stanów bezpośrednio osiągalnych z tego stanu jest zbiór jednoelementowy $\mathcal{S}_i = \{s_i\}$.

Cyklem pracy członu N-O długości L nazywać będziemy wykonanie programów z cyklu $\mathbf{C}_0(L)$, $L \geqslant 2$. Oznaczamy przez $\tau[\mathbf{C}_0(L)]$ czas trwania cyklu $\mathbf{C}_0(L)$. Zachodzi nierówność

$$\tau[\mathbf{C}_0(L)] \leqslant \sum_{j \in \mathcal{J}[\mathbf{C}_0(L)]} (V'_{j,1} + V''_{j,1}). \tag{4.2}$$

Z określenia $V''_{j,1}$ wynika, że na to by czas trwania każdego cyklu pracy był skończony musi następować wymuszenie zmiany stanu członu *N-O* w każdym jego stanie $s \in \mathcal{S}$.

Wykorzystując wprowadzenie pojęcia podamy teraz własności, jakie posiadać musi sieć programów, by mogła być ona obrazem prawidłowo działającej reguły.

Własność 1. *Do zbioru stanów dopuszczalnych nie może należeć stan pochłaniający*. Gdyby istniał stan pochłaniający, to człon *N-O* osiągnąwszy taki stan na stałe pozostałby w tym stanie. Oczywiste jest więc, że taka sytuacja jest niedopuszczalna.

Własność 2. *Każdy stan $s \in \mathcal{S} - \{s_0\}$ musi być osiągalny ze stanu początkowego. Stan początkowy s_0 nie musi być stanem wymuszającym się*. Gdyby istniał taki stan, który nie byłby osiągalny ze stanu początkowego, to człon *N-O* nie znalazłby się w tym stanie w żadnym cyklu pracy. Zatem do zbioru stanów dopuszczalnych powinny należeć tylko stany osiągalne w określonej liczbie kroków ze stanu początkowego. Z określenia stanu początkowego wynika zaś, że stan ten nie musi, ale może być stanem wymuszającym się.

Własność 3. *Stan początkowy s_0 musi być osiągalny z każdego stanu $s \in \mathcal{S} - \{s_0\}$*. Gdyby istniał taki stan $s \in \mathcal{S} - \{s_0\}$, że z tego stanu nie byłby osiągalny stan początko-

wy s_0, człon *N-O* po wyjściu ze stanu s_0 mógłby nie osiągnąć z powrotem stanu początkowego. W takiej sytuacji czas trwania cyklu pracy, w którym człon *N-O* osiągnąłby stan s, byłby nieograniczony.

Z własności 2 i 3 wynika następująca własność.

Własność 4. *Dla każdego stanu* $s \in \mathcal{S} - \{s_0\}$ *istnieje takie* L, *że stan* s *jest osiągalny w cyklu pracy* $\mathbf{C}_0(L)$. Przez wprowadzenie zmiennych kontrolnych stanów dopuszczalnych w sieci programów dopuszczalne są inne cykle $\mathbf{C}_i(L)$ $(i \neq 0, L \geqslant 1)$, które są wykonywane w cyklu pracy członu *N-O*. Cykle takie muszą być jednak kontrolowane, by czas trwania cyklu pracy zawsze był skończony. Osiągnąć to można poprzez wykorzystanie zmiennych kontrolnych stanów, kontrolując na przykład ile razy kolejno wykonywany jest określony cykl, nie stanowiący jednak cyklu pracy.

5. SIEĆ PROGRAMÓW W PRZYPADKU KANAŁU NIEIDEALNEGO

Dotychczas omawiając sieci programów zakładaliśmy, że kanał pomiędzy członami *N-O* był kanałem idealnym. Konsekwencją tego założenia było to, że informacja *s-k* wejściowa mogła przyjmować w każdym stanie członu *N-O* tylko postacie dopuszczalne zgodne z przyjętym protokołem komunikacyjnym. W przypadku kanału nieidealnego może zaistnieć taka sytuacja, że informacja *s-k* napływająca do członu *N-O* w ustalonym jego stanie może mieć taką postać, która nie jest zgodna z przyjętym protokołem. Ponadto w przypadku kanału nieidealnego zmienne kontrolne poszczególnych stanów osiągać mogą wartości nie należące do zbioru z góry ustalonych ich wartości dopuszczalnych. W takiej sytuacji muszą być wprowadzone pewne dodatkowe działania, których ujęcie wymaga rozbudowania omawianej sieci programów o nowe elementy. Nową sieć, jaką otrzymujemy przez dodanie do omawianej sieci programów nowych elementów, nazywać będziemy *rozszerzoną siecią programów*. W punkcie tym omówimy zasadę tworzenia takiej sieci.

Aby utworzyć sieć programów, będącą obrazem prawidłowo funkcjonującej reguły sterowania przesyłaniem informacji w przypadku kanału nieidelanego, wprowadzić musimy szereg dodatkowych relacji. W tym celu weźmy pod uwagę definicję (3.1) i (3.2).

Dla każdego stanu $s_i \in \mathcal{S}$, gdy $\mathring{\mathcal{X}}_i \neq \emptyset$, tworzymy relację

$$\mathcal{F}_i^* \equiv \{(s_i, x, s_i') : x \in \mathring{\mathcal{X}}_i\} \tag{5.1}$$

oraz zespół relacji

$$\bigwedge_{j \in \mathcal{J}_i^*} \mathcal{F}_{ij}^* \equiv \{(s_i, v_{ij}, s_{ij}'') : v_{ij} \in \mathring{\mathcal{V}}_{ij}\}, \tag{5.2}$$

przy czym

$$\mathcal{J}_i^* \equiv \{j\text{: Może wystąpić taka sytuacja, że } v_{ij} \in \mathring{\mathcal{V}}_{ij}\}.$$

Ze względu na istnienie relacji (5.1) i (5.2) podzbiorem stanów bezpośrednio osiągalnych ze stanem dopuszczalnego $s_i \in \mathcal{S}$ jest podzbiór

$$\mathcal{S}_i^* = \mathcal{S}_i \cup \{s_i'\} \cup \mathcal{S}_i'', \tag{5.3}$$

gdzie

$$\mathscr{S}_i'' \equiv \{s_{ij}'' : j \in \mathscr{J}_i^*\}.$$

Z postaci relacji (5.1) wynika, że stan s_i' osiągany jest ze stanu s_i, gdy w stanie s_i napły-nie informacja x o postaci niepożądanej. Z postaci relacji (5.2) wynika natomiast, że ze stanu s_i osiągany jest stan s_{ij}'' w sytuacji, gdy z kolei zmienna kontrolna v_{ij} stanu s_i osiągnie wartość niepożądaną. Z tego względu podzbiory stanów

$$\mathscr{S}' \equiv \{s_i' : i \in \mathscr{I}'\}, \quad \mathscr{S}'' \equiv \bigcup_{j \in \mathscr{I}''} \mathscr{S}_i'', \tag{5.4}$$

gdzie:

$$\mathscr{I}' \equiv \{i : i \in \mathscr{I} \wedge \mathring{\mathscr{X}}_i \neq \emptyset\}, \quad \mathscr{I}'' \equiv \{i : i \in \mathscr{I} \wedge \mathscr{J}_i^* \neq \emptyset\} \tag{5.5}$$

mają sens podzbiorów stanów niepożądanych osiągalnych ze stanów dopuszczalnych. Zbiór $\mathscr{I}'$ jest więc zbiorem numerów wszystkich takich stanów $s_i \in \mathscr{S}$, z których osiągane są stany niepożądane należące do podzbioru $\mathscr{S}'$. Analogicznie $\mathscr{I}''$ jest zbiorem numerów wszystkich takich stanów $s_i \in \mathscr{S}$, z których osiągane są stany niepożądane z podzbioru $\mathscr{S}''$. Zatem, jeśli $\mathscr{I}' \cap \mathscr{I}'' \neq \emptyset$, to istnieją stany $s_i \in \mathscr{S}$, z których osiągalne są zarówno stany niepożądane z podzbioru $\mathscr{S}'$ jak i z podzbioru $\mathscr{S}''$.

Zakładamy, że w sytuacji, gdy człon *N-O* osiąga stan niepożądany $s' \in \mathscr{S}'$ informacje *s-k* wyjściowa y generowana jest zgodnie z relacją

$$\mathscr{G}^* = \bigcup_{i \in \mathscr{I}'} \mathscr{G}_i^*, \quad \text{gdzie} \quad \mathscr{G}_i^* \equiv \{(s_i, x, y_i^*) : x \in \mathring{\mathscr{X}}_i\}. \tag{5.6}$$

Jeżeli zaś człon *N-O* osiąga stan niepożądany $s'' \in \mathscr{S}''$, informacja y generowana jest zgodnie z relacją

$$\mathscr{G}^{**} = \bigcup_{i \in \mathscr{I}''} \mathscr{G}_i^{**}, \quad \text{gdzie} \quad \mathscr{G}_i^{**} = \bigcup_{j \in \mathscr{J}_i^*} \mathscr{G}_{ij}^{**}, \tag{5.7}$$

przy czym

$$\bigwedge_{j \in \mathscr{J}_i^*} \mathscr{G}_{ij}^{**} \equiv \{(s_i, v_{ij}, y_{ij}^*) : v_{ij} \in \mathring{\mathscr{V}}_{ij}\}. \tag{5.8}$$

Ze wzorów (5.6)÷(5.8) wynika, że jeśli człon *N-O* osiąga stan niepożądany, to generowana jest informacja *s-k* wyjściowa y ze zbioru $\mathscr{Y}^* = \mathscr{Y}_1^* \cup \mathscr{Y}_2^*$, gdzie

$$\mathscr{Y}_1^* \equiv \{y_i^* : i \in \mathscr{I}'\}, \quad \mathscr{Y}_2^* \equiv \{y_{ij}^* : i \in \mathscr{I}'' \wedge j \in \mathscr{J}_i^*\}.$$

Postępując podobnie jak w punkcie 3-cim, wprowadzamy teraz programy odpowiadające poszczególnym wprowadzonym tu dodatkowym relacjom typu stan-stan. Mianowicie, relacji $\mathscr{F}_i^*$ określonej wzorem (5.1) przyporządkowujemy program $P_{i|x}^*$. Wykonywany jest on wtedy, gdy człon *N-O* jest w stanie s_i oraz pojawia się takie zgłoszenie, które zawiera $x \in \mathring{\mathscr{X}}_i$. Przyjmujemy, że program $P_{i|x}^*$ zgodnie z relacją (5.6) generuje także informację *s-k* wyjściową y_i^*. Relacji $\mathscr{F}_{ij}^*$ określonej wzorem (5.2) przyporządkowujemy program P_{ij}^*, który wykonywany jest wtedy, gdy człon *N-O* był w stanie s_i i zmienna kontrolna v_{ij} osiągnęła wartość niepożądaną. Przyjmujemy, że program ten zgodnie z relacją (5.8) generuje także informację *s-k* wyjściową y_{ij}^*.

Zbiorem programów związanych z każdym stanem $s_i \in \mathscr{S}$ w przypadku kanału nieidealnego jest zbiór

$$\mathscr{P}_i^* = \begin{cases} \mathscr{P}_i \cup \{P_{i|x}^*\} \cup \mathscr{P}_i^{**}, & \text{dla } i \in \mathscr{I}' \cap \mathscr{I}'', \\ \mathscr{P}_i \cup \{P_{i|x}^*\}, & \text{dla } i \in \mathscr{I}' \text{ oraz } i \notin \mathscr{I}'', \\ \mathscr{P}_i \cup \mathscr{P}_i^{**}, & \text{dla } i \notin \mathscr{I}'' \text{ oraz } i \in \mathscr{I}', \\ \mathscr{P}_i, & \text{dla } i \notin \mathscr{I}' \text{ oraz } i \notin \mathscr{I}'', \end{cases} \tag{5.9}$$

gdzie zbiór $\mathscr{P}_i$ określony jest wzorem (3.17), zaś

$$\mathscr{P}_i^{**} \equiv \{P_{ij}^{**} : j \in \mathscr{J}_i^*\} \tag{5.10}$$

Rozszerzoną podsiecią programów w każdym stanie $s_i \in \mathscr{S}$ nazywać będziemy zespół

$$\mathscr{N}_i^* \equiv \langle \overline{\mathscr{S}}_i^*, \mathscr{P}_i^*, \mathscr{V}_i^*, \overline{\mathscr{X}}_i^* \rangle, \tag{5.11}$$

gdzie

$$\overline{\mathscr{S}}_i^* \equiv \mathscr{S}_i^* \cup \{s_i\},$$

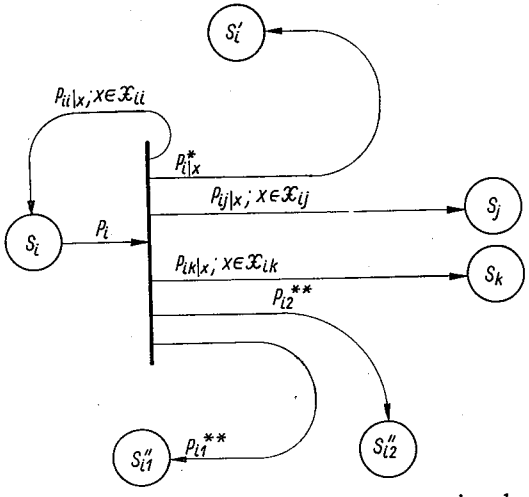
natomiast:

$$\mathscr{V}_i^* \equiv \{\mathscr{V}_i, \mathring{\mathscr{V}}_i\}, \quad \overline{\mathscr{X}}_i^* \equiv \{\overline{\mathscr{X}}_i, \mathring{\mathscr{X}}_i\}$$

przy czym zbiór $\mathscr{V}_i$ określony jest wzorem (2.9), zaś

$$\mathring{\mathscr{V}}_i \equiv \{\mathring{\mathscr{V}}_{i1}, \mathring{\mathscr{V}}_{i2}, \ldots, \mathring{\mathscr{V}}_{i,Ki}\}.$$

Podsieć taką, podobnie jak podsieć $\mathscr{N}_i$, przedstawiać będziemy graficznie. Przykładową postać podsieci $\mathscr{N}_i^*$ przedstawiono na rys. 5.



Rys. 5. Przykładowa struktura rozszerzonej podsieci programów

Zakładamy, że jeżeli człon *N-O* znajduje się w stanie niepożądanym, następuje odpowiednie testowanie kanału. Oczywiście wynik takiego testowania może być „pozytywny" bądź „negatywny". Niech x_i' oznacza informację o wyniku testowania kanału, gdy człon *N-O* jest w stanie $s_i' \in \mathscr{S}'$. Pozytywny wynik testowania oznaczmy jako $x_{i|+}'$, natomiast wynik negatywny — jako $x_{i|-}'$. Analogicznie, niech x_{ij}'' oznacza informację o wyniku testowania kanału, gdy człon *N-O* jest w stanie $s_{ij}'' \in \mathscr{S}''$. Pozytywny wynik testowania oznaczmy jako $x_{ij|+}''$, zaś wynik negatywny — jako $x_{ij|-}''$.

Przyjmijmy, że jeśli wynik testowania kanału nie był pozytywny, człon *N-O* osiąga stan oznaczony symbolem s_{I+1}. Tak więc w przypadku kanału nieidealnego człon *N-O* w dowolnie wybranej chwili może się znajdować w stanie należącym do zbioru

$$\mathscr{S}_R = \mathscr{S} \cup \mathscr{S}' \cup \mathscr{S}'' \cup \{s_{I+1}\}. \tag{5.12}$$

Dla każdego stanu niepożądanego $s'_i \in \mathscr{S}'$ tworzymy relacje:

$$\mathscr{F}'_{i,0} \equiv \{(s'_i, x'_i, s_0) : x'_i = x'_{i|+}\}, \tag{5.13a}$$

$$\mathscr{F}'_{i,I+1} \equiv \{(s'_i, x'_i, s_{I+1}) : x'_i = x'_{i|-}\}. \tag{5.13b}$$

Wobec istnienia takich relacji, podzbiorem stanów bezpośrednio osiągalnych z każdego stanu niepożądanego $s'_i \in \mathscr{S}'$ jest zbiór

$$\mathscr{S}'_i \equiv \{s_0, s_{I+1}\}. \tag{5.14}$$

Analogicznie jak w przypadku stanu $s'_i \in \mathscr{S}'$, także dla każdego stanu niepożądanego $s''_{ij} \in \mathscr{S}''$ tworzymy relacje:

$$\mathscr{F}''_{ij,0} \equiv \{(s''_{ij}, x''_{ij}, s_0) : x''_{ij} = x''_{ij|+}\}, \tag{5.15a}$$

$$\mathscr{F}''_{ij,I+1} \equiv \{(s''_{ij}, x''_{ij}, s_{I+1}) : x''_{ij} = x''_{ij|-}\}. \tag{5.15b}$$

Zatem podzbiorem stanów bezpośrednio osiągalnych z każdego stanu niepożądanego $s''_{ij} \in \mathscr{S}''$ jest podzbiór $\mathscr{S}''_{ij} = \mathscr{S}'_i$.

Rozpatrzmy teraz stan niepożądany $s'_i \in \mathscr{S}'$. Relacji $\mathscr{F}'_{i,0}$ przyporządkowujemy program $P'_{i,0}$, natomiast relacji $\mathscr{F}'_{i,I+1}$ program $P'_{i,I+1}$. Program $P'_{i,0}$ wykonywany jest wtedy, gdy człon *N-O* znajduje się w stanie s'_i i wynik testowania kanału jest pozytywny, natomiast program $P'_{i,I+1}$ — w przypadku, gdy wynik testowania okazał się negatywny. Oznaczmy przez P'_i program testowania kanału, gdy człon *N-O* znajduje się w stanie niepożądanym $s'_i \in \mathscr{S}'$. Zbiorem programów, jakie mogą być realizowane w członie *N-O*, gdy osiągnął on stan niepożądany s'_i jest zbiór

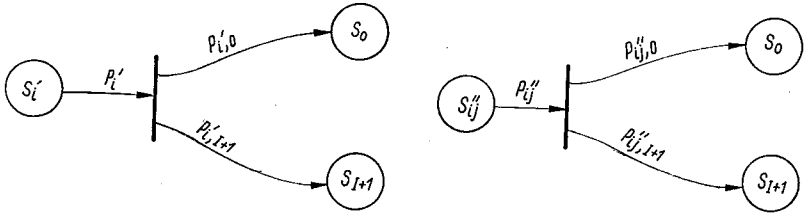
$$\mathscr{P}'_i \equiv \{P'_i, P'_{i,0}, P'_{i,I+1}\}. \tag{5.16}$$

Oznaczmy $\mathscr{X}'_i \equiv \{x'_{i|+}, x'_{i|-}\}$. Podsiecią programów w stanie niepożądanym $s'_i \in \mathscr{S}'$ nazywamy zespół

$$\mathscr{N}'_i \equiv \langle \overline{\mathscr{S}}'_i, \mathscr{P}'_i, \mathscr{X}'_i \rangle, \tag{5.17}$$

gdzie $\overline{\mathscr{S}}'_i \equiv \{s'_i, s_0, s_{I+1}\}$. Podsieć taką przedstawiono na rys. 6a.

Weźmy wreszcie pod uwagę stan niepożądany $s''_{ij} \in \mathscr{S}''$. Postąpimy teraz analogicznie, jak w przypadku stanu niepożądanego s'_i. Mianowicie relacji $\mathscr{F}''_{ij,0}$ przyporządkowujemy program $P''_{ij,0}$ wykonany w przypadku, gdy wynik testowania był pozytywny, nato-



Rys. 6. Przykładowa struktura podsieci programów w stanie niepożądanym
a) stan $s'_i \in \mathscr{S}'$, b) stan $s''_{ij} \in \mathscr{S}''$

miast relacji $\mathscr{F}''_{ij,I+1}$ — program $P''_{ij,I+1}$ wykonywany przy negatywnym wyniku testowania kanału. Zbiorem programów, które mogą być realizowane w członie *N-O*, gdy osiągnął stan niepożądany s''_{ij} jest zbiór

$$\mathscr{P}''_{ij} \equiv \{P''_{ij}, P''_{ij,0}, P''_{ij,I+1}\}. \tag{5.18}$$

Oznaczając $\mathscr{X}''_{ij} \equiv \{x''_{ij|+}, x''_{ij|-}\}$, jako podsieć programów w stanie niepożądanym s''_{ij} przyjmujemy zespół

$$\mathscr{N}''_{ij} \equiv \langle \bar{\mathscr{S}}''_{ij}, \mathscr{P}''_{ij}, \mathscr{X}''_{ij} \rangle, \tag{5.19}$$

gdzie $\bar{\mathscr{S}}''_{ij} \equiv \{s''_{ij}, s_0, s_{I+1}\}$.

Podsieć tę przedstawiono na rys. 6b.

Oznaczmy przez $P_{I+1,0}$ czynność powodującą zmianę stanu kanału ze stanu s_{I+1} na stan s_0. Czynność tę interpretować będziemy jako czynność sprowadzającą się do naprawy kanału, który miał uszkodzenie o charakterze trwałym. Czynność tę przedstawiać będziemy w postaci przerywanego łuku skierowanego od stanu s_{I+1} do stanu s_0.

Rozszerzoną siecią programów dla ustalonego członu *N-O* nazywamy zespół

$$\mathscr{N}_R \equiv \langle \mathscr{S}_R, \mathscr{P}_R, \mathscr{V}_R, \bar{\mathscr{X}}_R \rangle, \tag{5.20}$$

gdzie:

$$\mathscr{P}_R \equiv (\bigcup_{i\in\mathscr{S}} \mathscr{P}^*_i) \cup (\bigcup_{i\in\mathscr{S}'} \mathscr{P}'_i) \cup (\bigcup_{i\in\mathscr{S}''} \mathscr{P}''_i) \cup \{P_{I+1,0}\},$$

$$\mathscr{P}''_i \equiv \bigcup_{j\in\mathscr{J}^*_i} \mathscr{P}''_{ij},$$

natomiast:

$$\mathscr{V}_R \equiv \{\mathscr{V}^*_i : i \in \mathscr{I}\}, \quad \bar{\mathscr{X}}_R \equiv \{\bar{\mathscr{X}}^*, \mathscr{X}', \mathscr{X}''\},$$

przy czym:

$$\bar{\mathscr{X}}^* \equiv \{\bar{\mathscr{X}}^*_i : i \in \mathscr{I}\}, \quad \mathscr{X}' \equiv \{\mathscr{X}'_i : i \in \mathscr{I}'\},$$
$$\mathscr{X}'' \equiv \{\mathscr{X}''_i : i \in \mathscr{I}''\}, \quad \mathscr{X}''_i \equiv \{\mathscr{X}''_{ij} : j \in \mathscr{J}^*_i\}.$$

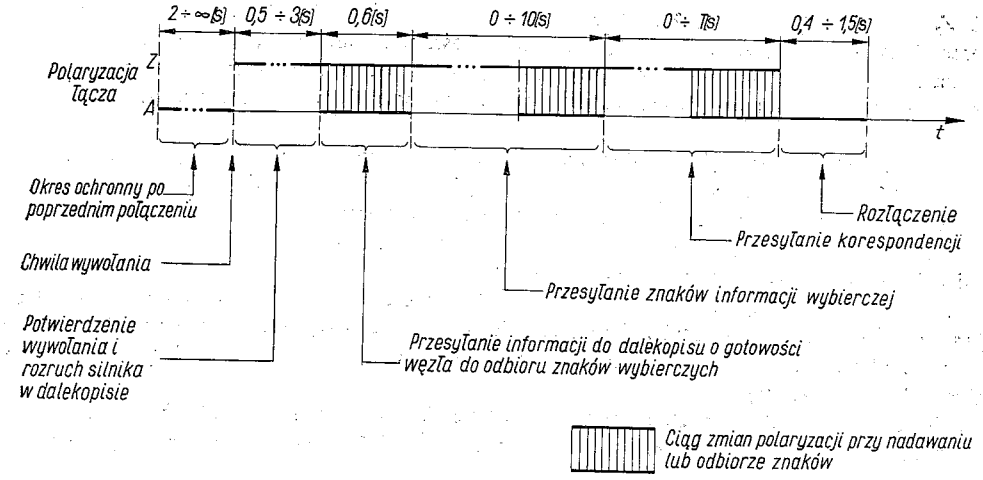
Rozszerzoną sieć programów otrzymujemy więc w wyniku połączenia rozszerzonych podsieci programów dla wszystkich stanów dopuszczalnych ze zbioru $\mathscr{S}$ i wszystkich stanów niepożądanych s' i s'' odpowiednio ze zbiorów $\mathscr{S}'$ i $\mathscr{S}''$.

Sposób graficznego przedstawiania rozszerzonej sieci programów zilustrujemy teraz na bardzo uproszczonym konkretnym przykładzie.

Przykład 2. Weźmy pod uwagę węzeł sieci, np. telegraficznej, zbudowany na bazie odpowiedniego systemu komputerowego. Przyjmijmy, że zgodnie z rys. 1 proces *A* przebiega w tym węźle, natomiast człon $(N\text{-}O)_B$ oraz proces *B* reprezentuje dalekopis podłączony do tego węzła. W pewnym uproszczonym protokół komunikacyjny, zgodnie z którym informacja przesyłana jest pomiędzy węzłem a dalekopisem przedstawiono na rys. 7. Przedstawimy teraz rozszerzoną sieć programów członu $(N\text{-}O)_A$ funkcjonującej w węźle i zapewniającego realizację tego protokołu.

Przyjmijmy, że zbiorem stanów dopuszczalnych członu $(N\text{-}O)_A$ jest: $\mathscr{S} = \{s_0, s_1, \ldots, s_5\}$, gdz s_0 — stan spoczynkowy, s_1 — stan oczekiwania na rozruch silnika w dalekopisie, s_2 — stan gotowo do odbioru informacji wybierczej z dalekopisu, s_3 — stan odbioru informacji wybierczej, s_4 — stan k respondencji, s_5 — stan oczekiwania na potwierdzenie rozłączenia.

Zbiorem stanów niepożądanych s' jest $\mathscr{S}' = \{s'_3, s'_5\}$, gdzie s'_3 — stan po odebraniu niedopuszczalnej informacji wybierczej, s'_5 — stan po niewłaściwym potwierdzeniu rozłączenia.



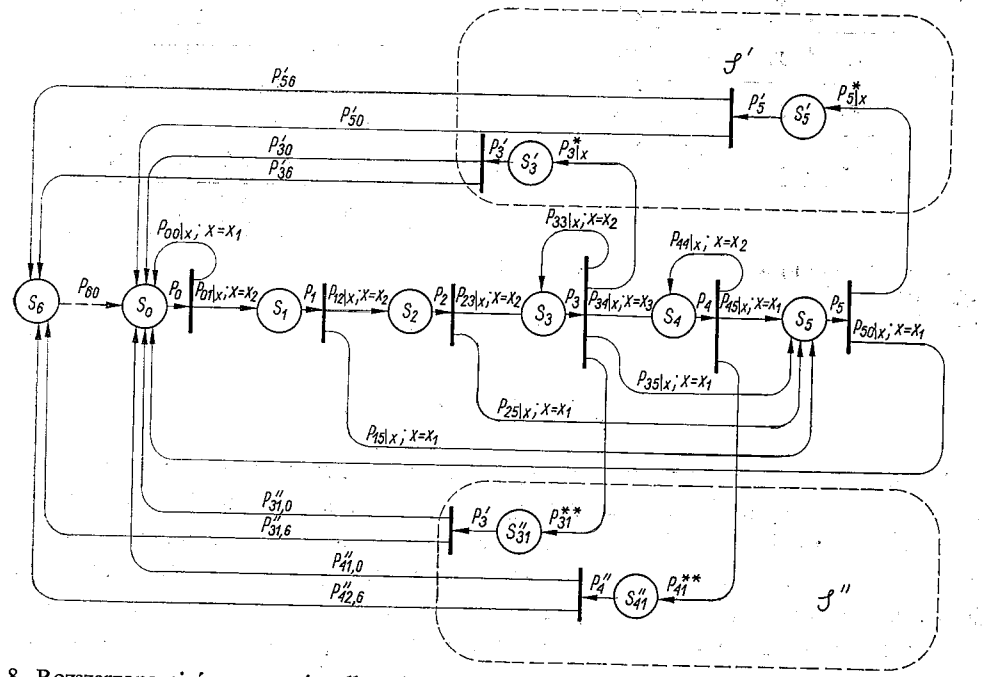
Rys. 7. Uproszczony protokół komunikacyjny: dalekopis-węzeł

Zbiorem stanów niepożądanych s'' jest $\mathcal{S}'' = \{s''_{31}, s''_{41}\}$, gdzie s''_{31} — stan po przekroczeniu czasu oczekiwania na informację wybierczą, s''_{41} — stan po przekroczeniu czasu trwania korespondencji.

Zbiorem dopuszczalnych postaci x jest $\mathcal{X} = \{x_1, x_2, x_3\}$, gdzie: x_1 — informacja o tym, że w łączu do dalekopisu jest polaryzacja A, x_2 — informacja o tym, że w łączu do dalekopisu jest polaryzacja Z, x_3 — dopuszczalna postać informacji wybierczej przesyłanej z dalekopisu.

Wszystkim stanom $s_i \in \mathcal{S}$ przyporządkowana jest tylko jedna zmienna kontrolna v_{i1} związana z upływem czasu. Elementy zbiorów $\mathcal{V}_i$ określonych wzorem (2.9) mają postać

$$\mathcal{V}_{01} = [2; +\infty); \quad \mathcal{V}_{11} = [0{,}15;\ 3]; \quad \mathcal{V}_{21} = \{0; 6\},$$



Rys. 8. Rozszerzona sieć programów dla członu *N-O* zapewniającego przesyłanie informacji pomiędzy dalekopisem a węzłem

gdyż $V'_{21} = 0{,}6$ oraz $V''_{21} = 0$; $\mathscr{V}_{31} = [0; 10]$; $\mathscr{V}_{41} = [0; T]$, gdzie T dopuszczalny czas trwania korespondencji; $\mathscr{V}_{51} = [0{,}4; 1{,}5]$. Występujące tu wartości wyrażone są w sekundach.

Rozszerzoną sieć programów sterujących funkcjonowaniem w węźle członu *N-O* zapewniającego przesyłanie informacji pomiędzy dalekopisem a węzłem, zgodnie z protokołem zilustrowanym na rys. 7. przedstawiono na rys. 8.

Zbiorem postaci informacji *s-k* wyjściowej w sytuacji, gdy człon *N-O* znajduje się w stanach dopuszczalnych $s \in \mathscr{S}$, jest zbiór $\mathscr{Y} = \{y_1, y_2, y_3\}$, gdzie y_1 — informacja o gotowości do odbioru, y_2 — wymuszenie w łączu do dalekopisu polaryzacji A, y_3 — wymuszenie w łączu do dalekopisu polaryzacji Z. Informacje *s-k* wyjściowe $y \in \mathscr{Y}$ generowane są zgodnie z relacją $\mathscr{G}$ określoną wzorem (2.5). Uporządkowane trójki należące do tej relacji wyszczególniono w tablicy 1.

Jeżeli człon *N-O* osiąga stan niepożądany $s' \in \mathscr{S}'$, to zbiorem informacji *s-k* wyjściowej y jest zbiór $\mathscr{Y}_1^* = \{y_1^*, y_2^*\}$, gdzie: y_1^* — znaki z informacją dalekopisu, y_2^* — wymuszenie polaryzacji A w łączu. Informacje *s-k* ze zbioru $\mathscr{Y}_1^*$ generowane są zgodnie z relacją $\mathscr{G}^*$ określoną wzorem (5.6), której elementy wyszczególniono w tablicy 2.

Jeżeli zaś człon *N-O* osiąga stan niepożądany $s'' \in \mathscr{S}''$, to zbiorem informacji *s-k* wyjściowej generowanej zgodnie z relacją $\mathscr{G}^{**}$ daną wzorem (5.7) jest zbiór $\mathscr{Y}_2^* = \{y_2^*\}$. Elementy relacji $\mathscr{G}^{**}$ wyszczególniono w tablicy 3.

Tablica 1

s \ x	x_1	x_2	x_3
s_0	y_2	y_3	×
s_1	y_2	y_1	×
s_2	y_2	y_3	×
s_3	y_2	⊖	y_3
s_4	y_2	y_3	×
s_5	y_2	×	×

Tablica 2

s \ $x \in \mathring{\mathscr{X}}$	$x \in \mathring{\mathscr{X}}_3$	$x \in \mathring{\mathscr{X}}_5$
s_3	y_1^*	×
s_5	×	y_2^*

Tablica 3

s \ $v \in \mathring{\mathscr{V}}$	$v \in \mathring{\mathscr{V}}_{31}$	$v \in \mathring{\mathscr{V}}_{41}$
s_3	y_2^*	×
s_1	×	y_2^*

× — odpowiadający tej pozycji element nie należy do relacji,
$\mathring{\mathscr{X}}_3 = \{x: x \neq x_3\}$,
$\mathring{\mathscr{X}}_5 = \{x_2\}$.

6. WNIOSKI

W niniejszej pracy ograniczyliśmy się do problemu tworzenia sieci programów proto kołów komunikacyjnych przy założeniu, że zadane są stany współpracujących ze sob funkcjonalnych członów *N-O*. Jak już wcześniej wspomniano, problemem, który należ przed przystąpieniem do budowy sieci programów rozwiązać jest ustalenie zbioru stanó dopuszczalnych członu *N-O*. Nie są znane ogólne zasady doboru takich stanów. Postępo wanie przy ich wyznaczaniu w większości przypadków jest heurystyczne. Problem ten wy maga jednak podejścia bardziej formalnego, dzięki któremu można tworzyć bardzi efektywnie funkcjonujące sieci programów.

Omawiając w pracy sieci programów nie wchodziliśmy zbytnio w szczegóły ich praktycznej realizacji, gdyż silnie uzależnione jest to od wyboru konkretnego systemu komputerowego, na którego bazie realizowany jest przyjęty protokół komunikacyjny. Z podanej tu struktury sieci wynikają jednak pewne ogólne zasady realizacji. Ponieważ przyjęto, że protokół realizowany jest na zasadzie programowej obsługi zgłoszeń z kanału przekazywanych do systemu komputerowego w postaci przerwań, przeto przyjmujemy, że ustalony jest zbiór $\mathscr{A}$ adresów zapisanych w pamięci programów P_i, $i \in \mathscr{I}$, których zadaniem jest wstępna obsługa zgłoszeń. Ze względu na to, że program P_i jest jednoznacznie związany ze stanem s_i, przeto tworząc wzajemnie jednoznacznie przyporządkowanie $\mathbf{A}: \mathscr{S} \xrightarrow{\text{na}} \mathscr{A}$ relację $\mathscr{F}$ typu stan-stan możemy zastąpić przy praktycznej realizacji protokołu realizacją wtórną typu adres-adres. Tak więc „aktualny stan" członu *N-O* możemy zastąpić pojęciem „aktualny adres programu" wstępnej obsługi zgłoszenia. Jeżeli zarezerwujemy ustaloną komórkę pamięci, np. komórkę K, na zapis adresu takiego programu, to zmianę stanu członu *N-O* możemy interpretować jako zmianę zawartości komórki K. Przy pojawieniu się zgłoszenia z kanału odczytywana jest zawartość komórki K; zawartość ta stanowi adres programu, zgodnie z którym to zgłoszenie jest wstępnie obsługiwane. Po ustaleniu postaci informacji sterująco-kontrolnej x wykonywany jest kolejny program $P_{ij|x}$, który kończy się ustawieniem odpowiedniej zawartości komórki K.

Jeżeli $\mathscr{A}'$ jest zbiorem adresów programów testujących kanał w poszczególnych stanach niepożądanych $s' \in \mathscr{S}'$, to wprowadzając wzajemnie jednoznaczne przyporządkowanie $\mathbf{A}': \mathscr{S} \xrightarrow{\text{na}} \mathscr{A}'$ możemy stosować samoczynne testowanie kanału, gdy człon *N-O* znajduje się w dowolnym stanie niepożądanym. W tym celu, jeżeli człon *N-O* osiąga stan $s'_i \in \mathscr{S}'$, to na listę programów oczekujących na wykonanie zapisywany jest program testujący P'_i o adresie $a'_i = \mathbf{A}'(s_i)$. Podobnie przeprowadzone może być samoczynne testowanie kanału w sytuacji, gdy człon *N-O* osiągnął stan niepożądany $s'' \in \mathscr{S}''$. W tym celu należy także wprowadzić wzajemnie jednoznaczne przyporządkowanie $\mathbf{A}'': \mathscr{S}'' \xrightarrow{\text{na}} \mathscr{A}''$, gdzie $\mathscr{A}''$ jest zbiorem adresów programów testujących kanał w poszczególnych stanach niepożądanych ze zbioru $\mathscr{S}''$.

Poprzez wprowadzenie testowania kanału w zależności od niepożądanego stanu członu *N-O* możliwa staje się kontrola jakości kanału w poszczególnych fazach procesu przesyłania informacji. Z drugiej zaś strony w przypadku uszkodzenia się kanału, czyli gdy człon *N-O* osiąga stan s_{I+1}, osiągalna jest informacja o tym, w jakim stanie członu *N-O* kanału uległ uszkodzeniu.

Wprowadzając sieć programów protokołu komunikacyjnego dążyliśmy do tego, aby sieć taka była obrazem realizacji protokołu w systemie komputerowym. Dzięki temu przy tworzeniu oprogramowania zapewniającego przesyłanie informacji zgodnie z przyjętym protokołem wprowadzona tu sieć programów spełnia podobną funkcję, jaką spełnia tzw. sieć czynności przy realizacji złożonych przedsięwzięć [8]. Z tego względu zawarte w niniejszej pracy koncepcje powinny przyczynić się do bardziej usystematyzowanego podejścia do oprogramowania systemów komputerowych wykorzystywanych w telekomunikacji przy sterowaniu procesem przesyłania informacji oraz ułatwić projektowanie programowo sterowanych systemów telekomunikacyjnych.

LITERATURA

1. S. Alfonzetti, S. Casale, A. Faro, S. Palazzo, *Algorithm to Pass Between Various Protocol Formalizations: from Logical Matrices to Variable Structure Sequential Machine and Vice-versa*, Alta Frequenza, vol. XLVIII, nr 2, 1979, s. 68—77.
2. S. Alfonzetti, S. Casale, A. Faro, *A formal description of the DTE packet level in the X.25 Recommendation*, Alta Frequenza, vol. XLVIII, nr 8, 1979, s. 513—522.
3. A. Badach, *Sterowanie siecią teleinformatyczną. Rozdz. 5 „Wprowadzenie do teleinformatyki"*, WKŁ, Warszawa 1979.
4. M. Bazewicz, *Modelowanie protokółów sieci komputerowych. Rozdział w „Własności i funkcje sieci komputerowych"*, Biblioteka WASC, Wrocław 1979.
5. *CCiTT, Recommendation X.25*, Orange Book, vol. VII-2, 1977, s. 70—108.
6. *CCiTT, Zalecenie U.12*, Zielona Księga, T. VII, 1972.
7. A. Danthine, *Petri nets for protocol modeling and verification*, Materiały z konferencji COMNET, Budapeszt 1976, s. 663—685.
8. T. Kasprzak (red); *Badania operacyjne w nowoczesnym zarządzaniu*, PWE, Warszawa 1974.
9. G. Le Moli, *A Theory of Colloquies*, Alta Frequenza, vol. XVLII, nr 10, 1973, s. 493—500.
10. P. M. Merlin, *A Methodology for the Design and Implementation of Communication Protocols*, IEEE Transactions on Communication, vol. COM-24, nr 6, 1976, s. 614—621.
11. P. M. Merlin, D. J. Farber, *Recoverability of Communication Protocols — Implementations of a Theoretical Study*, IEEE Transactions on Communication, vol. COM-24, nr 8, 1976, s. 1036—1043.

A. BADACH, R. SOBCZAK, M. ZIENTALSKI

PRINCIPLES OF SETTING UP NETWORKS OF COMPUTER PROGRAMS FOR CONTROL OF DIGITAL INFORMATION TRANSMISSION IN TELECOMMUNICATIONS SYSTEMS

Summary

The paper analyzes the principles of setting up the networks describing the cooperation of programs executed to control the digital information exchange between processes. The information transmission is performed in compliance with communications protocol. An analytical model of communications protocol is formed. On the basis of this model the rule is given for creating the network of programs, whose execution ensures the protocol realization. Two cases are considered, namely the case in which the information is transmitted by an ideal channel and the other one, when the information can be distorted in the channel.

The approach presented in the paper may be helpful in designing of computer-controlled telecommunications systems for digital information transmission, especially for the telegraphic network nodes and the nodes of data transmission network, which are based on mini- or microcomputers. The rule given in the paper allows to introduce the self-acting channel testing and to simplify the software building for the telecommunications computer systems.

A. BADACH, R. SOBCZAK, M. ZIENTALSKI

PRINCIPES DE LA FORMATION DU RÉSEAU DES PROGRAMMES COMMANDANT LA TRANSMISSION DES INFORMATIONS NUMÉRIQUES DANS LES SYSTÈMES DE TÉLÉCOMMUNICATIONS

Résumé

Dans cet ouvrage on a présenté les principes de la formation des réseaux, représentants la collaboration des programmes, réalisés pendant la commande de la transmission des informations numériques entre les processus établis, selon le protocole de communication accepté. Dans ce but on a formulé un modèle

analytique du protocole de communication et ensuite, en se servant de ce modèle, on a présenté le principe de la formation des réseaux des programmes, dont l'exécution permet la réalisation du protocole accepté. On a examiné le cas, ou la voie, par laquelle on transmet l'information, est idéale, ainsi que le cas ou l'information transmise est déformée dans la voie.

La méthode présentée peut être utilisée, pendant établissement du projet des systèmes de télécommunication, commandés par les programmes servant à la transmission des informations numériques, surtout des noeuds des réseaux télégraphiques et des noeuds des réseaux de transmission de données, conçus sur la base des systèmes convenables de mini- ou microordinateur. La méthode décrite permet, en particulier, d'introduire le contrôle automatique de la voie et de faciliter l'élaboration des programmes pour les systèmes d'ordinateur utilisés en télécommunications.

A. BADACH, R. SOBCZAK, M. ZIENTALSKI

PRINZIPIEN DER NETZBILDUNG VON PROGRAMMEN ZUR STEUERUNG DER ÜBERTRAGUNG VON DIGITALEN INFORMATIONEN IN NACHRICHTENSYSTEMEN

Zusammenfassung

In der Arbeit werden die Prinzipien der Bildung von Netzen dargestellt, welche das Zusammenwirken von Programmen beschreiben, die die digitale Informationsübertragung zwischen bestimmten Prozessen in Übereinstimmung mit dem Kommunikationsprotokoll steuern. Zu diesem Zweck wird ein analythisches Modell des Kommunikationsprotokolls geschaffen, auf dessen Grundlage die Prinzipien der Programmnetzbildung aufgeführt werden. Ein solches Programmnetz sichert die Realisierung des entsprechenden Protokolls. Es wird sowohl ein idealer Übertragungskanal betrachtet, als auch der Fall der Informationsverzerrung im nichtidealen Kanal. Die hier dargestellte Konzeption kann zur Projektierung von programmgesteuerten Nachrichtensystemen zur Übertragung digitaler Informationen benutzt werden, insbesondere zur Projektierung von Fernschreib- und Datenübertragungsnetzknoten, die auf der Grundlage entsprechender Mini- oder Mikrocomputersysteme arbeiten. Die dargestellte Methode erlaubt einen selbständigen Kanaltest und erleichtert Programmerarbeitung für die in der Nachrichtentechnik angewandten Computersysteme.

А. БАДАХ, Р. СОБЧАК, М. ЗЕНТАЛЬСКИ

ПРИНЦИПЫ ПОСТРОЕНИЯ СЕТЕЙ ПРОГРАММ УПРАВЯЮЩИХ ПЕРЕДАЧЕЙ ДИСКРЕТНЫХ СООБЩЕНИЙ

Резюме

Представлены принципы построения сетей взаимодействия программ управляющих передачей дискретных сообщений между установленными процессами согласно с принятым коммуникационным протоколом. С этой целью сформулирована аналитическая модель коммуникационного протокола, на основе этой модели представлены принципы построения сетей программов, исполнение которых гарантирует правильную реализацию принятого протокола. Рассматривается так случай, когда канал по которому передается сообщение является идеальным, как и случай, когда, передаваемое сообщение может быть в канале искажено.

Представленный подход может быть использован при проектировании программно управляемых систем предзначенных для передачи дискретных сообщений, а особенно узлов телеграфических и других сетей передачи дискретных сообщений создаваемых на основе соответствующих мини- или микрокомпьютерных систем. Предложенные принципы дают возможность исполнять автоматическое испытание канала и облегчить разработку программного обеспечения компьютеров используемых в сетях связи.

ROZPRAWY ELEKTROTECHNICZNE 1981, *27*, z. 3, ss. 885-922

621.394.18

Zastosowanie przekształceń kryptograficznych do ochrony informacji w bazach danych

JÓZEF PIEPRZYK (BYDGOSZCZ)

Instytut Telekomunikacji i Elektrotechniki
Akademii Technicznej w Bydgoszczy

DOMINIK RUTKOWSKI (SOPOT)

Instytut Telekomunikacji Politechniki Gdańskiej

Otrzymano 5.11.1979

W pracy przedstawiono zastosowanie przekształceń kryptograficznych do ochrony elementów bazy danych. Problem ten został rozwiązany w oparciu o wielopoziomowy model organizacji bazy danych. Szczególną uwagę zwrócono na wpływ związków semantycznych między poszczególnymi elementami bazy danych na stosowane przekształcenie kryptograficzne. Podano też reguły stosowania przekształceń kryptograficznych, przy z góry podanych warunkach określających rodzaj chronionych elementów bazy danych przed działalnością nieautoryzowanego użytkownika. Z kolei oszacowano prawdopodobieństwo złamania szyfru dla poszczególnych reguł realizacji przekształceń kryptograficznych przy warunku, że znane są związki semantyczne zachodzące pomiędzy elementami bazy danych. Na końcu pracy przedstawiono zastosowanie przekształceń kryptograficznych do szyfrowania elementów przykładowej bazy danych.

1. WSTĘP

Wraz z rozwojem zastosowań techniki komputerowej powstają nowe problemy związane z upowszechnianiem jej zastosowań. W większości przypadków wiążą się one ze wspólnym wykorzystywaniem zasobów komputerowych przez wielu użytkowników [46]. Wspólne użytkowanie zasobów komputerowych występuje w wielodostępnych systemach przetwarzania zarówno ze skupioną jak również rozproszoną mocą obliczeniową (sieci komputerowe). Z samej zasady wspólnego użytkowania systemów komputerowych wynika, że programy różnych użytkowników mogą wzajemnie na siebie oddziaływać wszędzie tam, gdzie w grę wchodzi prawo dostępu a zwłaszcza modyfikacji i kasowania wspólnie użytkowanych zbiorów informacji i programów. Bez odpowiedniej organizacji dostępu oraz określenia zakresu i uprawnień dostępu dla poszczególnych użytkowników, wspólne użytkowanie zasobów komputerowych może prowadzić do straty informacji. Znacznie groźniejsze może być niekiedy wykorzystanie dostępu w celu odczytu, modyfikacji bądź zniszczenia zbiorów informacji określonego użytkownika bez uzyskania jego zgody [30, 36].

Problem ochrony różnego rodzaju zbiorów informacji przed dostępem do nich nieautoryzowanego użytkownika jest więc ważnym problemem organizacji przetwarzania. Straty ponoszone przez użytkowników systemu komputerowego, których zbiory informacji stały się obiektami przestępstwa, są znaczne i liczba takich przestępstw nieustannie rośnie [28, 29, 35, 43]. O tym, że problem jest poważny, szczególnie w tych krajach, w których technika komputerowa jest wysoko rozwinięta, świadczy fakt, że np. w Stanach Zjednoczonych w 1974 roku wydano specjalną ustawę normującą ochronę informacji przechowywanych w bazach danych i ich zgodne z przeznaczeniem użytkowanie [17, 22, 31, 33]. Na podstawie dotychczasowych wyników badań przestępstw komputerowych można wyróżnić kilkadziesiąt różnych rodzajów przestępstw (patrz np. [3, 10, 12, 26, 27, 34]). Jednakże wszystkie te przestępstwa można podzielić na cztery podstawowe kategorie, które wiążą się z nieautoryzowanym

— odczytaniem informacji,
— modyfikacją informacji,
— zniszczeniem informacji,
— zablokowaniem dostępu dla autoryzowanego użytkownika.

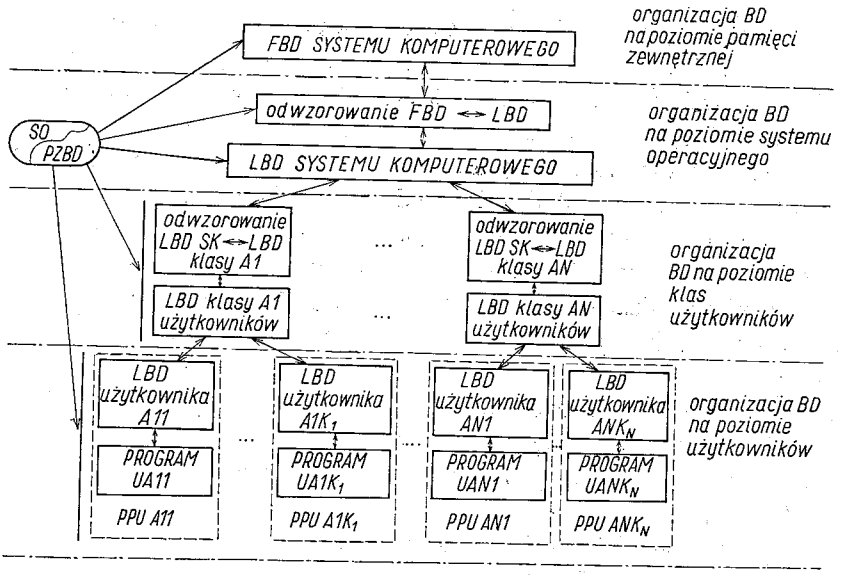
W nowoczesnych systemach komputerowych standardowe programy przetwarzania i ogólnie dostępne dla użytkowników zbiory danych oraz własne programy i zbiory danych poszczególnych użytkowników przechowywane są w pamięci zewnętrznej. Te zbiory informacji wraz z organizacją ich zapisu i odczytu tworzą tzw. bazę danych (BD) [4, 40, 42]. W wielu systemach komputerowych ich użytkownicy mają ograniczony w różnym stopniu dostęp do zbiorów informacji przechowywanych w BD. Wynika on z potrzeby zapewnienia wyłączności lub ograniczenia dostępu do tych zbiorów informacji dla określonych użytkownilów, co wymaga odpowiedniej ochrony zbiorów informacji. Znane są trzy podstawowe metody ochrony systemu komputerowego przed zamierzonym lub przypadkowym naruszeniem jego integralności, a mianowicie:

— ochrona fizyczna [12],
— ochrona operacyjna [13, 24],
— ochrona kryptograficzna [23, 32].

Metoda fizycznej ochrony systemu komputerowego wiąże się z jego zabezpieczeniem przed skutkami działania sił przyrody. Natomiast organizacja zarządzania zasobam komputera przez system operacyjny określa metodę operacyjnej ochrony systemu komputerowego [13, 21, 24]. Z kolei metoda ochrony kryptograficznej jest stosowana do ochrony zarówno kanałów łączących końcówki z komputerem, jak również do ochrony informacj przechowywanych w pamięci zewnętrznej komputera [19, 20, 32]. Znacznie lepszą ochron uzyskuje się przez powiązanie ochrony operacyjnej z ochroną kryptograficzną informacj przechowywanej w BD, przy czym ze względów praktycznych ochronę operacyjną wy godniej jest stosować do informacji znajdującej się w pamięci operacyjnej, a ochron kryptograficzną do informacji znajdującej się w BD. Przy takim rozwiązaniu ochron kryptograficzna stwarza zarówno ochronę informacji przekazywanej między pamięci operacyjną a pamięcią zewnętrzną, jak też ochronę informacji zapisanej na odpowiednicl nośnikach informacji bazy danych.

2. STRUKTURA BAZY DANYCH

Zanim przejdziemy do przedstawienia sposobu zastosowania metod kryptograficznych do ochrony informacji w bazie danych, omówimy najpierw wielopoziomową organizację i zarządzanie bazą danych przedstawione na rys. 1 [9, 15, 16, 21, 23, 40, 41]. Mimo faktu, że w systemie komputerowym istnieje jedna baza danych, to z punktu widzenia organizacji można wyróżnić jej strukturę: fizyczną oraz logiczną. Zatem można mówić o fizycznej i logicznej bazie danych. Fizyczną bazę danych (FBD) tworzy pamięć zewnętrzna[1] wraz z zawartą w niej informacją oraz organizacją zapisu i odczytu tej informacj.



Rys. 1. Schemat organizacji i zarządzania bazą danych przez system operacyjny
SO — system operacyjny, PZBD — program zarządzający bazą danych, PPU — przestrzeń pracy użytkownika, LBDSK — logiczna baza danych systemu komputerowego

Logiczną bazę danych (LBD) tworzy informacja zawarta w pamięci zewnętrznej i organizacja jej zapisu i odczytu uzależniona od zastosowań, które wyznaczają naturalne związki między przechowywaną w BD informacją, a zatem logiczna baza danych jest określona przez użytkowników bazy danych. Na ogół użytkownicy w zależności od rodzaju wykorzystywania zbiorów informacji podzieleni są na klasy, z czego wynika, że bazę danych systemu komputerowego podzieloną w sposób logiczny na podzbiory można rozpatrywać jako bazę danych podzieloną w sposób logiczny na podzbiory odpowiednich klas użytkowników. Natomiast programy użytkowników wraz z ich własnymi logicznymi bazami danych tworzą odpowiednie przestrzenie pracy użytkowników (PPU). Taka dualna organizacja bazy danych umożliwia uniezależnienie LBD od zmian organizacji FBD[2] nosi w literaturze nazwę organizacji z niezależnymi danymi[3] [14, 15].

[1] Np. pamięć dyskowa lub taśmowa.

[2] Chodzi tutaj zarówno o zmianę organizacji zapisu i odczytu jak też sprzętu fizycznej bazy danych.

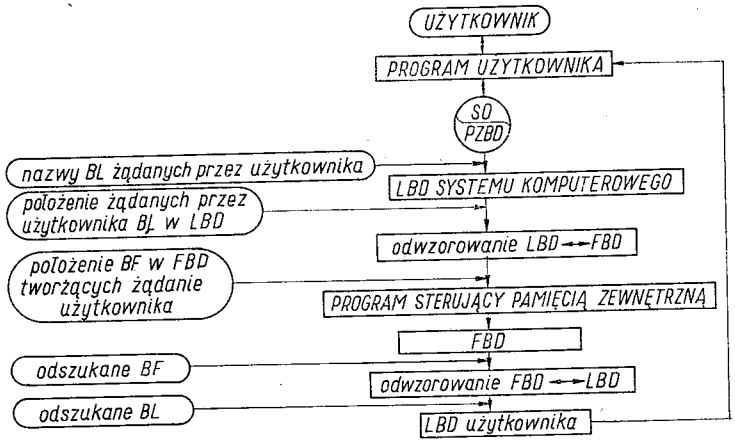
[3] Ang. *data independence*.

Informacje w FBD i LBD zapisywane są w postaci bloków[4)], przy czym można wyróżnić bloki fizyczne (BF) i bloki logiczne (BL). Powiązania pomiędzy blokami fizycznymi i logicznymi określa odpowiednie odwzorowanie [15]. Mówi ono, w jaki sposób bloki są zapisywane oraz w jaki sposób są adresowane. Gdy ulegnie zmianie organizacja FBD i sposób dostępu do niej, to musi zmienić się odwzorowanie pomiędzy blokami FBD a blokami LBD, natomiast organizacja LBD nie musi ulegać zmianie [15]. Podobnie LBD klasy użytkowników określa bloki należące do tej bazy danych. Odwzorowanie pomiędzy LBD systemu komputerowego a LBD określonej klasy użytkowników wyznacza powiązania pomiędzy tymi BD. W ogólności LBD poszczególnych klas użytkowników nie są rozłączne.

Program zarządzający bazą danych (PZBD) jest programem, który wspólnie z systemem operacyjnym (SO) kontroluje dostęp do BD i nadzoruje wykonywanie działań na zawartości BD. Każde żądanie dostępu użytkownika, po sprawdzeniu jego autoryzacji przez SO, jest przekazywane do PZBD i program ten określa, które BL są niezbędne do spełnienia tego żądania (rys. 2). Następnie PZBD uruchamia program dostępu do pamięci zewnętrznej. Program ten wyszukuje odpowiednie bloki fizyczne, które po odwzorowaniu w bloki logiczne są wprowadzane do pamięci operacyjnej. Zatem program określonego użytkownika może wykonywać działania na zawartości tych bloków.

Nad zasadami organizacji BD, autoryzacji dostępu do niej i realizacją programową tych zasad czuwa administrator BD. Jest on odpowiedzialny za:

— zawartość bazy danych,
— wybór organizacji bazy danych [42, 44],
— zapewnienie komunikacji między użytkownikami a bazą danych,
— określenie zasad autoryzacji poszczególnych użytkowników do korzystania z bazy danych [44, 45, 48, 49],
— strategię powrotu bazy danych do normalnego zakresu działania po uszkodzeniu systemu komputerowego [11, 43, 47],



Rys. 2. Schemat reakcji systemu komputerowego na żądanie przesłania określonych zbiorów informac z bazy danych

[4)] Istnieją bloki typu znakowego i bloki typu binarnego.

— nadzór nad działaniem BD i dokonywaniem modyfikacji w oprogramowaniu wynikających ze zmian żądań użytkowników bądź też sprzętu stosowanego w bazie danych.

Użytkownicy komunikują się z BD za pomocą określonego języka, który zawiera pewne dodatkowe rozkazy[5] dostosowane do korzystania z BD. Tak zmodyfikowane języki zwane są językami manipulacyjnymi [6, 7, 38, 39]. Zasady komunikacji użytkownika z BD zawarte w określonym języku manipulacyjnym zależą od organizacji LBD, przy czym można wyróżnić trzy podstawowe rodzaje organizacji [15, 42]:

— hierarchiczną,
— sieciową,
— relacyjną.

Okazuje się [14, 15], że organizacja relacyjna wydaje się być najbardziej obiecująca przy tworzeniu bazy danych w nowych systemach komputerowych i jej w dalszym ciągu poświęcimy szczególną uwagę.

2.1. Wielopoziomowy model bazy danych

Niezależnie od tego czy baza danych ma organizację hierarchiczną, sieciową, czy też relacyjną ma ona w ogólnym przypadku strukturę wielopoziomową z punktu widzenia użytkownika, systemu operacyjnego i pamięci zewnętrznej [15, 23, 39] (patrz rys. 1). Z rys. 1 wynika, że organizację na poziomie użytkowników tworzą LBD poszczególnych użytkowników wraz z odpowiadającymi im programami tych użytkowników (PPU). Ogniwem pośrednim organizacji BD są LBD poszczególnych klas użytkowników. Organizację BD na poziomie SO tworzy LBD całego systemu komputerowego. Dualną formą organizacji LBD jest organizacja FBD. Pomiędzy poszczególnymi poziomami organizacji istnieją powiązania i aby je określić, wprowadzamy formalny opis bazy danych. Opis ten zostanie oparty na pojęciu jednostki danych, która jest najmniejszym i podstawowym elementem BD i którą oznaczać będziemy literą *d*. Jednostki danych mają następujące cechy:

W1. nazwa (interpretacja),
W2. sposób przedstawienia (kodowanie),
W3. długość,
W4. wartość,
W5. adres.

Te same jednostki danych występujące na różnych poziomach organizacji BD mogą mieć różny zapis[6], mimo że dostarczać będą tych samych informacji. Dlatego każdą jednostkę danych będziemy rozpatrywali w powiązaniu z określonym poziomem organizacji, na którym jest ona rozpatrywana. Zatem *j*-tą jednostkę danych na *i*-tym poziomie organizacji będziemy oznaczać przez d_j^i. Oznaczenie to będziemy upraszczać do postaci d_j wtedy, gdy poziom, na którym występuje określona jednostka, nie będzie miał znaczenia. Jednostka danych stanowi podstawę dla zdefiniowania bloku fizycznego (BF).

[5] Chodzi tutaj w szczególności o rozkazy zapisu, odczytu, modyfikacji i kasowania danych [5, 37, 38].

[6] Odmienność w zapisie tej samej jednostki danych na różnych poziomach organizacji BD wynikać będzie, jak to później okaże się, ze stosowania różnych przekształceń kryptograficznych (patrz punkt 3).

Definicja 1. Blokiem fizycznym na i-tym poziomie organizacji BD nazywamy uporządkowany zbiór n jednostek danych

$$BF_j^i = (d_1^i, d_2^i, \ldots, d_n^i). \tag{1}$$

Każdy BF w bazie danych ma określony adres A_j. Adresy poszczególnych jednostek danych wchodzących w skład bloku fizycznego są określane na podstawie adresu A_j tego bloku oraz długości jednostek danych wchodzących w skład bloku. W oparciu o pojęcie jednostki danych i bloku fizycznego możemy zdefiniować fizyczną bazę danych (FBD).

Definicja 2. Fizyczną bazą danych na i-tym poziomie organizacji BD nazywamy zbiór wszystkich BF na tym poziomie, zatem

$$FBD^i = \{BF_1^i, BF_2^i, \ldots, BF_m^i\}, \tag{2}$$

gdzie m — liczba bloków fizycznych.

W celu określenia pojęcia logicznej bazy danych (LBD) wprowadzamy relację podobieństwa pomiędzy jednostkami danych występującymi na tym poziomie organizacji bazy danych.

Definicja 3. Dwie jednostki danych d_1 oraz d_2 są podobne, gdy mają tę samą nazwę, co zapisujemy symbolicznie $d_1 \sim d_2$.

Można wykazać, że relacja podobieństwa pomiędzy jednostkami danych w bazie danych jest relacją równoważności [32]. Czyli w zbiorze wszystkich jednostek danych na określonym poziomie organizacji bazy danych można utworzyć klasy równoważności według relacji podobieństwa.

Definicja 4. Polem P^i na i-tym poziomie organizacji BD nazywamy klasę równoważności według relacji podobieństwa jednostek danych określonych na tym poziomie. Jeżeli nazwa ma postać x, to pole będziemy oznaczać przez $[x]^i$ (stąd $P^i = [x]^i$). Oznaczenie pola będziemy zapisywali w uproszczonej postaci $[x]$, jeżeli poziom organizacji BD nie będzie odgrywał roli.

Definicja 5. Blokiem logicznym BL^i na i-tym poziomie organizacji BD, określonym na polach $P_1^i, P_2^i, \ldots, P_k^i$ nazywamy podzbiór iloczynu kartezjańskiego $P_1 \times P_2 \times \ldots \times P$ z wyróżnioną nazwą (identyfikatorem). Oznaczać go będziemy przez

$$BL^i = I^i(P_1, P_2, \ldots, P_k), \tag{3}$$

gdzie I^i jest nazwą bloku.

Przykład 1, Załóżmy, że polami P_1 i P_2 są odpowiednio zbiory P_1 = [imię] = {Jan, Tadeusz}, P_2 = [nazwisko] = {Adamski, Mikołajczyk, Nowak}. W tablicy podano przykład bloku logicznego określonego na polach P_1 oraz P_2.

Tablica 1

Przykład bloku logicznego

P_1 = [imię]	P_2 = [nazwisko]
Jan	Adamski
Tadeusz	Mikołajczyk
Jan	Nowak

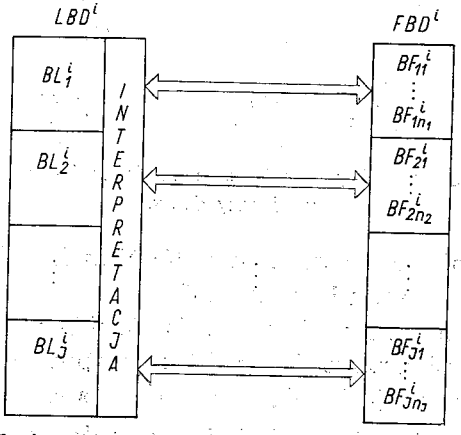
Mając zdefiniowane pojęcie bloku logicznego możemy na jego podstawie określić pojęcie logicznej bazy danych.

Definicja 6. Logiczną bazę danych LBD^i na i-tym poziomie organizacji bazy danych jest zbiór bloków logicznych

$$\{BL_1^i, BL_2^i, \ldots, BL_n^i\} \tag{4}$$

wraz z ich interpretacją.

Z powyższych określeń wynika, że baza danych na i-tym poziomie występuje w dualnej postaci fizycznej oraz logicznej BD, przy czym wzajemnie jednoznaczne odwzorowanie między nimi jest określone przez interpretację (patrz rys. 3).



Rys. 3. Powiązania pomiędzy logiczną bazą danych a fizyczną bazą danych na i-tym poziomie organizacji bazy danych

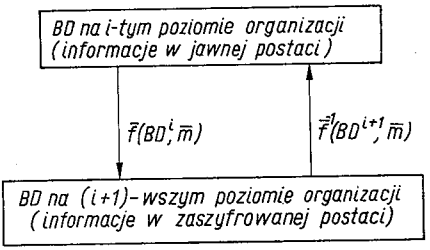
3. ZASTOSOWANIE PRZEKSZTAŁCEŃ KRYPTOGRAFICZNYCH W BAZACH DANYCH

Jak stwierdziliśmy we wstępie, przekształcenia kryptograficzne są użytecznym sposoem ochrony informacji zawartej w bazie danych zarówno przed nieautoryzowanym żytkownikiem (NU), który chce korzystać z zasobów systemu komputerowego mając ostęp do jednej z końcówek systemu, jak również przed NU, który chce uzyskać interesuce go zbiory z BD wykorzystując podsłuch, przechwytywanie nośników informacji, itp. / dalszym ciągu uwagę skupimy na zastosowaniu przekształceń kryptograficznych do chrony informacji w bazie danych.

1. Przekształcenie kryptograficzne zachowujące strukturę bazy danych

Zakładamy, że będziemy dalej rozważać tylko dwa poziomy organizacji bazy danych:
- poziom i-ty, na którym znajdują się zbiory informacji w postaci nie zaszyfrowanej,
- poziom $(i+1)$-wszy, na którym znajdują się zbiory informacji w postaci zaszyfrowanej (rys. 4).

Na rys. 4 zapis $\bar{f}(BD^i, \bar{m}) = BD^{i+1}$ oznacza, że baza danych na $(i+1)$-wszym poziomie jest przekształconą bazą z poziomu i-tego, przy czym $\bar{f}(\cdot, *)$ jest funkcją kryptograficzną (tutaj rozumianą jako zbiór przekształceń kryptograficznych poszczególnych elementów BD), a wektor $\bar{m}$ jest kluczem (zbiorem kluczy dla poszczególnych przekształceń kryptograficznych). W punkcie tym ograniczymy się do przekształceń kryptograficznych szczególnego rodzaju, a mianowicie do przekształceń zachowujących strukturę bazy danych [23], z uwagi na ich znaczenie praktyczne.



Rys. 4. Zastosowanie przekształcenia kryptograficznego pomiędzy dwoma kolejnymi poziomami organizacji bazy danych

Definicja 7. Przekształcenie kryptograficzne $f(\cdot, *)$ zachowuje strukturę bazy danych gdy kryptogram[7] jest elementem bazy danych[8].

Zajmiemy się teraz podstawowymi przekształceniami kryptograficznymi zachowującymi strukturę BD. Do tych przekształceń należą: podstawienie, przestawienie, redukcja jednostek danych, zwielokrotnianie jednostek danych i szyfrowanie adresu dostępu.

P1. Podstawienie jednostek danych

W tym przekształceniu kryptograficznym elementy przed i po przekształceniu są jednostkami danych, przy czym

$$d^{i+1} = f(d^i, m). \tag{5}$$

Przekształcenie odwrotne przyjmuje postać

$$d^i = f^{-1}(d^{i+1}, m), \tag{6}$$

gdzie $d^i \in P^i$, $d^{i+1} \in P^{i+1}$ oraz $|P^i| = |P^{i+1}|$[9].

Tak określone przekształcenie kryptograficzne oddziaływuje także na pola, do których należą jednostki danych. Można tu wyróżnić dwa przypadki:

— gdy jednostki danych przed i po zaszyfrowaniu należą do tego samego pola, tzn. $P^i = P^{i+1}$,

— gdy jednostki danych przed i po zaszyfrowaniu należą do różnych pól, a więc $P^i \neq P^{i+1}$.

Podstawienie jednostek danych zmienia postać BF oraz BL, przy czym dla tych ostatnich można tę zmianę przedstawić w następujący sposób:

$$BL^i = I^i(P_1^i, P_2^i, \ldots, P_n^i) \xrightarrow{\bar{f}(\cdot, \bar{m})} BL^{i+1} = I^{i+1}(P_1^{i+1}, P_2^{i+1}, \ldots, P_n^{i+1}),$$

[7] Zaszyfrowaną wiadomość nazywać będziemy kryptogramem.

[8] Elementami BD są jednostki danych, pola, BF oraz BL.

[9] $|P^i|$, $|P^{i+1}|$ oznacza liczbę jednostek danych w polach P^i oraz P^{i+1}.

gdzie przekształcenie kryptograficzne zostało zastosowane w odniesieniu do pól P_r^i $(1 \leqslant r \leqslant n)$.

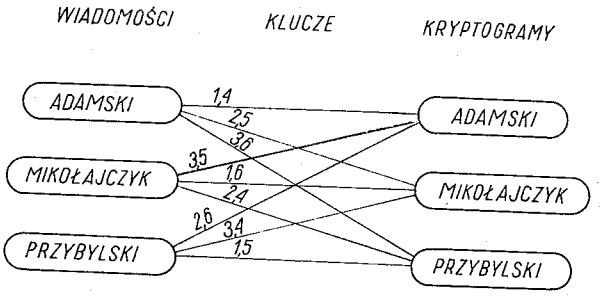
Przykład 2. Dany jest blok logiczny $BL^i = I^i(P_1^i, P_2^i, P_3^i)$ w postaci tablicy 2.

Tablica 2

Przykład bloku logicznego

P_1^i = [imię]	P_2^i = [nazwisko]	P_3^i = [rok urodzenia]
Jan	Adamski	1960
Tadeusz	Mikołajczyk	1951
Jerzy	Przybylski	1946

Załóżmy, że interesuje nas przekształcenie kryptograficzne zachowujące strukturę BD określone dla pola P_2^i, które jest zilustrowane na rys. 5. Jeżeli przyjmiemy, że wartość lucza kryptograficznego $m = 3$, to zaszyfrowany blok logiczny $BL^{i+1} = I^{i+1}(P_1^i, P_2^i, P_3^i)$ ia postać określoną w tablicy 3, przy czym f (Adamski, 3) = Przybylski, f (Mikołajczyk,) = Adamski, f (Przybylski, 3) = Mikołajczyk.



Rys. 5. Ilustracja przekształcenia kryptograficznego $f(\cdot, *)$ omawianego w przykładzie 2

Tablica 3

Blok logiczny w postaci zaszyfrowanej

P_1^i = [imię]	P_2^i = [nazwisko]	P_3^i = [rok urodzenia]
Jan	Przybylski	1960
Tadeusz	Adamski	1951
Jerzy	Mikołajczyk	1946

Rozważane przekształcenie kryptograficzne można realizować w oparciu o podany iżej algorytm:

Każdej jednostce danych $d \in P^i$ przyporządkowujemy w sposób wzajemnie jednoznaczny liczbę $l \in P_l^i$. To przyporządkowanie można zapisać w BD jako blok logiczny $BL^i = I_l^i(P^i, P_l^i)$.

A2. Liczbę $l \in P_l^i$ poddajemy przekształceniu kryptograficznemu dla wybranego klucza m (można tutaj zastosować odpowiedni program szyfrujący [32]). Otrzymujemy w ten sposób kryptogram, który jest liczbą $f(l, m)$,

A3. Stosujemy odwrotne przyporządkowanie do określonego w punkcie A1; określamy więc jednostkę danych.

Jak wynika z powyższego algorytmu dla określonego klucza i jednostki danych otrzymujemy kryptogram, który jest także jednostką danych. Taki sposób szyfrowania jest elastyczny, można go stosować do przekształceń różnych pól i jednocześnie odpowiedni program szyfrujący nie zajmuje zbyt dużo miejsca w pamięci.

Do ważnej klasy przekształceń podstawieniowych należą przekształcenia kryptograficzne zachowujące operację, które mogą być wykonywane na jednostkach danych. Prowadzi to do definicji przekształceń kryptograficznych zachowujących operację.

Definicja 8. Przekształcenie kryptograficzne $f(\cdot, *)$ zachowuje operację F wtedy i tylko wtedy, gdy dla jednostek danych d_1^i, d_2^i przed szyfrowaniem oraz po szyfrowaniu d_1^{i+1}, d_2^{i+1} jest słuszna zależność

$$F(d_1^i, d_2^i) \Leftrightarrow \bigwedge_{m \in M} F(d_1^{i+1}, d_2^{i+1}),$$

gdzie M jest zbiorem kluczy kryptograficznych.

Dla przykładu funkcja kryptograficzna $f(\cdot, *)$ zachowuje operację porównania, gdy $(d_1^i = d_2^i) \Leftrightarrow \bigwedge_{m \in M} f(d_1^i, m) = f(d_2^i, m)$. Z przedstawionych powyżej rozważań wynika że podstawienie jednostek danych jest prostym przekształceniem kryptograficznym i ma tę zasadniczą wadę, że przy znajomości dodatkowych elementów BD można je stosunkowo łatwo złamać, szczególnie wówczas gdy liczba szyfrowanych jednostek danych w określonym polu jest niewielka.

P2. Przestawienie jednostek danych w bloku fizycznym

Rozważać teraz będziemy blok logiczny $BL^i = I^i(P_1^i, P_2^i, \ldots, P_k^i)$ z i-tego poziomu organizacji BD. Elementami tego bloku są bloki fizyczne $BL^i = \{BF_1^i, BF_2^i, \ldots, BF_n^i\}$. Każdy blok fizyczny jest ciągiem k jednostek danych. czyli $BF_r^i = (d_{r,1}^i, d_{r,2}^i, \ldots, d_{r,k}^i)$ gdzie $d_{r,1}^i \in P_1^i$, $d_{r,2}^i \in P_2^i$, $\ldots$, $d_{r,k}^i \in P_k^i$ oraz $1 \leqslant r \leqslant n$. Omawiane przestawienie jednostek danych jest przekształceniem kryptograficznym określonym następująco

$$d_{r,l}^{i+1} = d_{r,f_r(l,m)}^i,$$

przy czym $1 \leqslant r \leqslant n$, $1 \leqslant l \leqslant k$, a $f_r(\cdot, *)$ jest funkcją kryptograficzną określoną liczbie l oraz kluczu kryptograficznym m. Funkcja kryptograficzna $f_r(\cdot, *)$ ma inną postać dla każdego r.

Przykład 3. W celu zilustrowania omawianego przekształcenia kryptograficznego rozważmy przykład. W bazie danych istnieje blok logiczny $BL^i = I^i(P_1^i, P_2^i, P_3^i)$ oraz bloki fizyczne BF_1^i, BF_2^i, BF_3^i przy czym

$$BF_1^i = (d_{1,0}^i, d_{1,1}^i, d_{1,2}^i),$$
$$BF_2^i = (d_{2,0}^i, d_{2,1}^i, d_{2,2}^i),$$
$$BF_3^i = (d_{3,0}^i, d_{3,1}^i, d_{3,2}^i)$$

oraz $BL^i = \{BF_1^i, BF_2^i, BF_3^i\}$. Załóżmy, że funkcja kryptograficzna $f_r(\cdot,*)$ jest określona wzorem $f_r(l, m) = (m+r+1) \bmod 3$, $1 \leqslant r \leqslant 3$, $0 \leqslant 1 \leqslant 2$. Oznacza to, że blok fizyczny z i-tego poziomu przechodzi w blok fizyczny o zmienionym porządku jednostek danych

$$BF_r^i = (d_{r,0}, d_{r,1}, d_{r,2}) \xrightarrow{f_r(\cdot, m)} BF_r^{i+1} = (d^i_{r,f_r(0,m)}, d^i_{r,f_r(1,m)}, d^i_{r,f_r(2,m)}).$$

W szczególności dla wartości klucza $m = 2$ ($m \in \{0, 1, 2\}$) otrzymujemy:

$$BF_1^i = (d^i_{1,0}, d^i_{1,1}, d^i_{1,2}) \to BF_1^{i+1} = (d^i_{1,0}, d^i_{1,1}, d^i_{1,2}),$$
$$BF_2^i = (d^i_{2,0}, d^i_{2,1}, d^i_{2,2}) \to BF_2^{i+1} = (d^i_{2,1}, d^i_{2,2}, d^i_{2,0}),$$
$$BF_3^i = (d^i_{3,0}, d^i_{3,1}, d^i_{3,2}) \to BF_3^{i+1} = (d^i_{3,2}, d^i_{3,0}, d^i_{3,1}),$$

przy czym $d^i_{r,0} \in P_1^i$, $d_{r,1} \in P_2^i$, $d_{r,2} \in P_3^i$ dla $1 \leqslant r \leqslant 3$.

Modyfikacją tego rodzaju przekształcenia kryptograficznego jest przekształcenie opisane równościami

$$\begin{aligned} d^{i+1}_{j,1} &= d^i_{j,f(l,m_j)}, \\ d^i_{j,k} &= d^{i+1}_{j,f^{-1}(k,m_j)}, \end{aligned} \tag{8}$$

gdzie $k = f(l, m_j)$.

Przekształca ono j-ty blok fizyczny z i-tego poziomu na poziom $(i+1)$-wszy zgodnie z zależnościami.

$$BF_j^i = (d^i_{j,0}, \ldots, d^i_{j,k-1}) \to BF_j^{i+1} = (d^i_{j,f(0,m_j)}, \ldots, d^i_{j,f(k-1,m_j)}).$$

Ponieważ BL jest zbiorem n bloków fizycznych, więc zależność pomiędzy blokami logicznymi nie zaszyfrowanym i zaszyfrowanym ma postać

$$BL^i = \begin{Bmatrix} BF_1^i \\ BF_2^i \\ \vdots \\ BF_n^i \end{Bmatrix} \xrightarrow{f(\cdot, m)} \begin{Bmatrix} BF_1^{i+1} \\ BF_2^{i+1} \\ \vdots \\ BF_n^{i+1} \end{Bmatrix} = BL^{i+1},$$

gdzie $f(\cdot, m) = \{f_1(\cdot, m_1), f_2(\cdot, m_2), \ldots, f_n(\cdot, m_n)\}$.

Natomiast klucz dla całego przekształcenia kryptograficznego pomiędzy blokami logicznymi można napisać jako ciąg kluczy $m = (m_1, m_2, \ldots, m_n)$, gdzie $m_k \in M$ dla $1 \leqslant k \leqslant n$. Ten rodzaj przekształcenia kryptograficznego ma tę zaletę, że poszczególne BF wchodzące w skład BL szyfruje się kluczami wybieranymi niezależnie od pozostałych. Dzięki temu złamanie któregokolwiek z kluczy m_k ($1 \leqslant k \leqslant n$) nie ułatwia NU złamania pozostałych kluczy. Warto zauważyć, że nie szyfrowany BL ma na ustalonych pozycjach jednostki danych z poszczególnych pól, natomiast w zaszyfrowanym BL takie uporządkowanie nie występuje. Kolejne miejsca w zaszyfrowanym BF,w których występują jednostki danych z poszczególnych pól, zmieniają się według klucza m. Dla każdego BF po zaszyfrowaniu istnieje odmienna kolejność występowania jednostek danych. Ten rodzaj przekształcenia chroni bazę danych przed tzw. „przeglądaniem" jej zawartości — NU nie zna adresów jednostek danych wchodzących w skład bloku fizycznego.

P3. Redukcja jednostek danych

Przekształcenie kryptograficzne zwane redukcją jednostek danych jest określone przez wzór

$$d_1^{i+1} = f((d_1^i, d_2^i, \ldots, d_p^i), m), \tag{9}$$

przy czym odwrotne przekształcenie kryptograficzne ma postać

$$(d_1^i, d_2^i, \ldots, d_p^i) = f^{-1}(d_1^{i+1}, m). \tag{10}$$

Redukcja p jednostek danych na i-tym poziomie organizacji BD wyznacza pojedynczą jednostkę danych na $(i+1)$-wszym poziomie organizacji BD zgodnie z określonym kluczem m. Załóżmy dla przykładu, że należy zaszyfrować blok $BL^i = I^i(P_1^i, \ldots, P_p^i, P_{p+1}^i, \ldots, P_k^i)$ z i-tego poziomu organizacji BD. Po zaszyfrowaniu pierwszych p pól otrzymujemy blok logiczny o postaci $BL^{i+1} = f(BL^i, m) = I^{i+1}(P_1^{i+1}, P_{p+1}^i, \ldots, P_k^i)$, gdzie jednostki danych należące do pola P_1^{i+1} są kryptogramami ciągu jednostek danych należących do pól $P_1^i, \ldots, P_p^i$. Podstawową wadą przekształcenia kryptograficznego polegającego na redukcji jednostek danych jest niemożliwość wywołania niektórych jednostek danych, a co za tym idzie niemożliwość wykonywania niektórych operacji na zaszyfrowanych elementach. W szczególności nie można bezpośrednio żądać dostępu do zawartości pól $P_1^i, \ldots, P_p^i$, gdyż nie są one zdefiniowane.

P4. Zwielokrotnianie jednostek danych

Przekształcenie kryptograficzne zwane zwielokrotnianiem jednostek danych jest ściśle związane z omawianą poprzednio redukcją jednostek danych i jest opisane za pomocą wzorów

$$\begin{aligned}(d_1^{i+1}, d_2^{i+1}, \ldots, d_r^{i+1}) &= f(d^i, m),\\ d^i &= f^{-1}\big((d_1^{i+1}, d_2^{i+1}, \ldots, d_r^{i+1}), m\big).\end{aligned} \tag{11}$$

Zatem pojedynczej jednostce danych w jawnej postaci odpowiada kryptogram, który jest ciągiem r jednostek danych. W szczególnym przypadku przy zwielokrotnianiu pojedynczego pola pomiędzy blokami logicznymi znajdującymi się na i-tym oraz $(i+1)$-wszym poziomie organizacji BD występuje zależność

$$BL^i = I^i(P_1^i, P_2^i, \ldots, P_k^i) \xrightarrow{f(\cdot, m)} BL^{i+1} = I^{i+1}(P_1^{i+1}, P_2^{i+1}, \ldots, P_n^{i+1}),$$

przy czym:

$$P_1^i \xrightarrow{f(\cdot, m)} P_1^{i+1} \times P_2^{i+1} \times \ldots \times P_r^{i+1};$$

$$P_2^i = P_{r+1}^{i+1},\ P_3^i = P_{r+2}^{i+1}, \ldots, P_k^i = P_n^{i+1}, \quad \text{gdzie} \quad n = (r+k-1).$$

Jak już wcześniej wspomniano, przekształcenie kryptograficzne polegające na zwielokrotnieniu jest stosowane wspólnie z przekształceniem kryptograficznym zwanym redukcją danych i ze względu na wymaganą pojemność pamięci najczęściej stosuje się najpierw redukcję jednostek danych a następnie zwielokrotnianie.

P5. Szyfrowanie adresu dostępu

Ten rodzaj przekształcenia kryptograficznego odnosi się do bloków logicznych. Jeżeli blokiem logicznym na poziomie nie zaszyfrowanych jednostek danych będzie blok $BL^i = I^i(P_1^i, P_2^i, \ldots, P_j^i, \ldots, P_n^i)$, to po zastosowaniu przekształcenia kryptograficznego polegającego na szyfrowaniu adresu dostępu będziemy mieli w najprostszym przypadku dwa bloki logiczne:

$$\begin{aligned}BL_1^{i+1} &= I_1^{i+1}\big(P_1^i, P_2^i, \ldots, P_j^i, f(A, m)\big),\\ BL_2^{i+1} &= I_2^{i+1}(P_{j+1}^i, \ldots, P_n^i),\end{aligned}$$

gdzie A jest polem adresów drugiej części bloku fizycznego (bloku logicznego BL_2^{i+1}), natomiast przez $f(A, m)$ rozumiemy pole zaszyfrowanych adresów drugiej części bloku fizycznego.

Opisany sposób szyfrowania adresu dostępu ilustruje następujący przykład.

Przykład 4. Niech blok logiczny ma postać podaną w tablicy 4.

Tablica 4

Blok logiczny BL^i

P_1^i = [imię]	P_2^i = [nazwisko]	P_3^i = [rok urodzenia]	P_4^i = [zawód]
Jan	Adamski	1960	mechanik
Tadeusz	Mikołajczyk	1951	mechanik
Jerzy	Przybylski	1946	elektronik

Załóżmy, że chcemy zaszyfrować lata urodzenia oraz zawody poszczególnych osób, możemy zatem zastosować szyfrowanie adresu i w rezultacie otrzymamy dwa bloki logiczne BL_1^{i+1} oraz BL_2^{i+1} w formie pokazanej w tablicy 5, gdzie a_1, a_2, a_3 są adresami pozostałej części bloku fizycznego ($A = a_1, a_2, a_3$). Omawiane przekształcenie kryptograficzne jest stosunkowo proste i łatwe w realizacji praktycznej.

Tablica 5

Postać bloków logicznych po szyfracji adresu

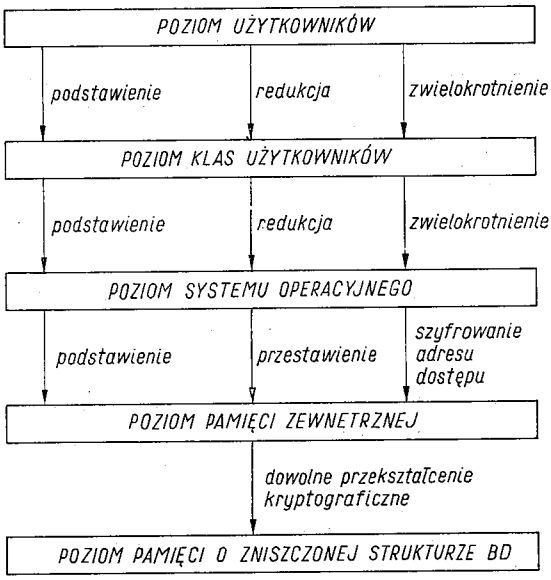
P_1^{i+1} = [imię]	P_2^{i+1} = [nazwisko]	$f(A, m)$	P_3^{i+1} = [rok. ur.]	P_4^{i+1} = [zawód]
Jan	Adamski	$f(a_1, m)$	1946	elektronik
Tadeusz	Mikołajczyk	$f(a_2, m)$	1960	mechanik
Jerzy	Przybylski	$f(a_3, m)$	1951	mechanik

3.2. Przekształcenia kryptograficzne a poziomy organizacji BD

W punkcie tym rozważymy przydatność omówionych poprzednio przekształceń kryptograficznych do ochrony informacji znajdującej się na różnych poziomach organizacji BD. Najpierw zajmiemy się zastosowaniem przekształcenia kryptograficznego przy przekazywaniu jednostek danych z poziomu użytkowników do poziomu klas użytkowników. Jak wiadomo z punktu 2 użytkownik w swojej przestrzeni pracy wykorzystuje wyłącznie te elementy BD, które występują na poziomie klas użytkowników i do których ma autoryzowany dostęp. Jeżeli przyjmiemy, że elementy BD z poziomu klas użytkowników są zaszyfrowane, to określeni użytkownicy będą mogli z nich korzystać wówczas, gdy program zarządzający bazą danych dostarczy im kluczy, przy czym zależeć to będzie od przysługującego tym użytkownikom prawa dostępu.

Przy przekazywaniu jednostek danych między poziomem użytkownika a poziomem klas użytkowników mogą być zastosowane trzy podstawowe przekształcenia kryptograficzne, a mianowicie (patrz rys. 6):

— podstawienie jednostek danych,
— redukcja jednostek danych,
— zwielokrotnianie jednostek danych.



Rys. 6. Rodzaje przekształceń kryptograficznych stosowanych przy przekazywaniu informacji pomiędzy różnymi poziomami organizacji bazy danych

Wszystkie bloki logiczne BD z poziomu klas użytkowników tworzone są w oparciu o bloki logiczne z poziomu SO, ponadto bloki z obydwu poziomów znajdują się w pamięci operacyjnej komputera w czasie wykonywania operacji na zawartości BD. Zatem przy przekazywaniu informacji pomiędzy tymi poziomami można stosować przekształcenie kryptograficzne takie samo, jak pomiędzy poziomem użytkowników a poziomem klas użytkowników. Elementy na poziomie systemu operacyjnego mogą więc być szyfrowane „podwójnie". Rozważmy, jakie przekształcenia kryptograficzne mogą być stosowane pomiędzy poziomem systemu operacyjnego a poziomem pamięci zewnętrznej. Przypomnijmy, że poziom pamięci zewnętrznej związany jest z organizacją BD w pamięci zewnętrznej, tzn. z FBD. Przy przekazywaniu informacji pomiędzy poziomem SO a poziomem pamięci zewnętrznej mogą być stosowane wszystkie wymienione w punkcie 3.1 przekształcenia kryptograficzne. W praktyce jednakże zalecane są tylko trzy spośród nich [2, 23]:

— podstawienie jednostek danych,
— przestawienie jednostek danych,
— szyfrowanie adresu dostępu.

Jak wynika z powyższych rozważań, jednostki danych z poziomu pamięci zewnętrznej mogą być wielokrotnie szyfrowane, przy czym wymaga się, żeby stosowane szyfry zacho-

wywały strukturę BD. Jeśli weźmiemy pod uwagę rozważania przeprowadzone w punkcie 1, to okaże się, że podstawowym wymaganiem przy ochronie informacji w BD na wypadek uszkodzenia systemu komputerowego jest posiadanie duplikatów FBD przechowywanych w odpowiednio chronionym miejscu [43, 47]. Aby zabezpieczyć informacje zawarte w duplikatach przed ich odczytaniem przez NU, korzystamy z przekształcenia kryptograficznego, które w ogólnym przypadku nie musi spełniać postulatu zachowania struktury BD. Tym przekształceniem może być np. przekształcenie kryptograficzne oparte o koncepcję wytwarzania długich ciągów klucza z tzw. „nasienia" (patrz [18, 19, 32]). Zaletą takiego przekształcenia kryptograficznego jest duży koszt lub równoważny mu czas niezbędny do złamania klucza przy odpowiednio długim nasieniu. Mankamentem stosowania tego rodzaju przekształcenia jest konieczność całkowitego rozszyfrowania BD, jeśli żądana jest określona jednostka danych w niej zawarta.

W stosowanych dotychczas rozwiązaniach BD informacje zawarte w pamięci operacyjnej przechowywane są w postaci źródłowej, co oczywiście znacznie zwiększa ryzyko dostępu NU do tych informacji. Wprawdzie omawiane powyżej przekształcenia kryptograficzne są kontrolowane przez PZBD, jednakże do programu tego może mieć dostęp personel ośrodka obliczeniowego, co może ułatwić NU dostęp do informacji przechowywanej w BD. Korzystniejszym rozwiązaniem wydaje się ochrona, w której przekształcenie kryptograficzne jest kontrolowane przez użytkownika [25], który wykorzystuje klucz do szyfrowania informacji przechowywanej w BD. Wówczas odpowiedni program szyfrujący przekształca dostarczoną informację, a PZBD umieszcza zaszyfrowaną już informację w BD. W tym przypadku użytkownik może otrzymać informację w postaci źródłowej tylko wtedy, gdy posiada stosowny klucz i uruchomi program deszyfrujący.

4. WYKORZYSTANIE ZWIĄZKÓW SEMANTYCZNYCH W PRZEKSZTAŁCENIACH KRYPTOGRAFICZNYCH INFORMACJI PRZECHOWYWANYCH W BAZIE DANYCH

Pomiędzy różnymi informacjami przechowywanymi w BD zachodzą bardzo często powiązania semantyczne. Powiązania te umożliwiają NU znaczne zwiększenie prawdopodobieństwa złamania klucza użytego do szyfrowania określonego zbioru informacji. W krańcowym przypadku NU, który nie ma bezpośredniego dostępu do interesującego go zbioru informacji, może go uzyskać pośrednio wykorzytując związki semantyczne między informacjami zawartymi w tym zbiorze. W punkcie tym będziemy zajmować się zastosowaniem przekształceń kryptograficznych do szyfrowania zbiorów informacji przechowywanych w BD w taki sposób, aby pomimo istnienia powiązań semantycznych prawdopodobieństwo załamania klucza kryptograficznego było możliwie małe.

4.1. Pojęcie relacji semantycznej

Dobór przekształceń kryptograficznych do ochrony informacji przechowywanych w BD zależy od związków semantycznych istniejących pomiędzy tymi informacjami. W dalszym ciągu będziemy przyjmować, że w pamięci operacyjnej istnieją zbiory informacji należące do trzech poziomów organizacji BD: użytkownika, klas użytkowników

i systemu operacyjnego. Z tego założenia wynika, że nie będziemy wyróżniać powyżej wymienionych poziomów, czyli zrezygnujemy z górnego indeksu przy oznaczeniach BD. Warto też zauważyć, że informacje zawarte w pamięci operacyjnej systemu komputerowego mogą być dostępne dla użytkownika w dwojaki sposób:

— w jawnej postaci, gdy użytkownik nie przekracza swoich uprawnień (działa w swojej przestrzeni pracy),

— w postaci zaszyfrowanej, gdy użytkownik działa bez autoryzacji (ma dostęp do zbiorów informacji z poziomów klas użytkowników i systemu operacyjnego).

W celu sprecyzowania rozważań zdefiniujemy najpierw pojęcie relacji semantycznej.

Definicja 9. Dane są dwa pola $[x_1]$, $[x_2]$ o nazwach x_1 oraz x_2. Relacją semantyczną $R(x_1, x_2)$ nazywamy podzbiór iloczynu kartezjańskiego $[x_1] \times [x_2]$, czyli

$$R(x_1, x_2) \subset [x_1] \times [x_2]. \tag{12}$$

Zauważmy, że relacja semantyczna jest uogólnieniem pojęcia bloku logicznego, przy czym blok logiczny jest ściśle związany z określonym poziomem organizacji BD, a relacja semantyczna jest związana ze zbiorami informacji przechowywanymi w pamięci operacyjnej. Definicja relacji umożliwia określenie złożenia dwóch relacji semantycznych.

Definicja 10. Złożeniem dwóch relacji semantycznych $R_1(x_1, x_2)$ oraz $R_2(x_3, x_2)$ nazywamy relację semantyczną $R_3(x_1, x_3)$ o postaci

$$R_3(x_1, x_3) = \{(d_i, d_j);\ \bigvee_{d_k \in [x_2]} (d_i, d_k) \in R_1(x_1, x_2) \wedge (d_j, d_k) \in R_2(x_3, x_2)\}, \tag{13}$$

gdzie jednostki danych $d_i \in [x_1]$, $d_j \in [x_3]$.

Określenie relacji semantycznej stanowi podstawę do wprowadzenia rzutu relacji na określone pole.

Definicja 11. Rzutem relacji semantycznej $R(x_1, x_2)$ na pole $[x_1]$ jest zbiór określony następująco

$$R(x_1, x_2)\langle x_1\rangle = \{d_i \in [x_1];\ \bigvee_{d_j \in [x_2]} (d_i, d_j) \in R(x_1, x_2)\}. \tag{14}$$

Podobnie określamy rzut relacji semantycznej na pole $[x_2]$, jako zbiór

$$R(x_1, x_2)\langle x_2\rangle = \{d_j \in [x_2];\ \bigvee_{d_i \in [x_1]} (d_i, d_j) \in R(x_1, x_2)\}. \tag{15}$$

Relacja semantyczna staje się w szczególnym przypadku funkcją semantyczną, gdy spełniony jest warunek zawarty w następującej definicji.

Definicja 12. Niech $[x_1]$ i $[x_2]$ są dowolnymi polami. Jeżeli relacja semantyczna $R(x_1, x_2)$ spełnia warunek, że dla każdej jednostki danych $d_i \in [x_1]$ istnieje dokładnie jeden element $d_j \in [x_2]$ taki, że $(d_i, d_j) \in R(x_1, x_2)$, to relację nazywamy funkcją semantyczną jednej zmiennej x_1, a pole $[x_1]$ nazywamy polem podstawowym relacji R. Wprowadzone pojęcia zilustrujemy na przykładzie.

Przykład 5. Załóżmy, że mamy daną relację o postaci podanej w tablicy 6. Rzut relacji R_1 na pole [Nazwisko] × [Zawód] ma formę, którą przedstawia tablica 7, czyli $R_2 = R_1\langle$Nazwisko, Zawód$\rangle$. W tej samej tablicy umieszczono relację $R_3 = R_1\langle$Zarobek, Zawód$\rangle$.

Tablica 6

Przykładowa relacja semantyczna R_1

[Nazwisko]	[Imię]	[Zawód]	[Zarobek]
Błaszczak	Andrzej	elektronik	5000
Chmara	Krzysztof	elektronik	6000
Michlewski	Krzysztof	fizyk	6000
Molski	Marian	informatyk	6000
Przybylski	Jerzy	informatyk	5000

Tablica 7

Relacje R_2 oraz R_3

[Nazwisko]	[Zawód]
Błaszczak	elektronik
Chmara	elektronik
Michlewski	fizyk
Molski	informatyk
Przybylski	informatyk

[Zarobek]	[Zawód]
5000	elektronik
6000	elektronik
6000	fizyk
6000	informatyk
5000	informatyk

Tablica 8

Relacja R_4

[Nazwisko]	[Zarobek]
Błaszczak	5000
Błaszczak	6000
Chmara	5000
Chmara	6000
Michlewski	6000
Molski	5000
Molski	6000
Przybylski	5000
Przybylski	6000

Na podstawie relacji R_2 oraz R_3 można wykonać złożenie tych relacji według pola [Zawód] (relacja R_4), która została przedstawiona w tablicy 8. Warto zauważyć, że $R_4 \supset R_1\langle$Nazwisko, Zarobek$\rangle$.

4.2. Zastosowanie metod kryptograficznych do ochrony powiązanych semantycznie elementów bazy danych

Na ogół w spotykanych bazach danych chronione są przed NU tylko ich wybrane elementy. Zajmiemy się teraz takim doborem elementów BD, które zostaną zaszyfrowane, aby znajomość przez NU powiązań semantycznych między elementami szyfrowanymi

i nie szyfrowanymi nie pozwoliła na określenie wartości szyfrowanych elementów[10]. Zastanowimy się także nad zastosowaniem metod kryptograficznych do ochrony, w szczególności

— relacji semantycznej,
— bloku fizycznego,
— pola.

Załóżmy, że E oznacza regułę kryptograficzną stosowaną do ochrony informacji w BD. Reguła ta dzieli BD na dwa rozłączne podzbiory:

$A(E)$ — podzbiór zawierający elementy, które mogą być dostępne dla NU w postaci jawnej,

$B(E)$ — podzbiór zawierający elementy, które są chronione przed nieautoryzowanym dostępem (szyfrowane elementy).

W praktyce możliwe są do stosowania reguły kryptograficzne, które poniżej przedstawimy, przy czym przyjmować będziemy, że baza danych zawiera skończoną liczbę pól.

Reguła E1. Reguła całkowitego zaszyfrowania relacji semantycznej

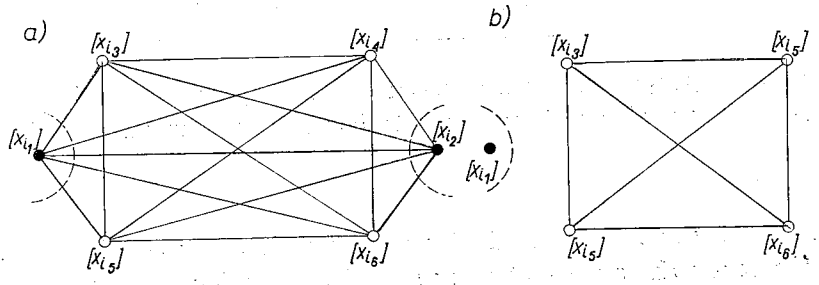
Rozważana reguła poddaje przekształceniu kryptograficznemu każdą relację semantyczną o postaci $R(x_{i_1}, x_{i_k})$ oraz $(R(x_{i_2}, x_{i_j})$, gdzie $[x_{i_k}], [x_{i_j}] \in \mathrm{BD}$, $i_k \neq i_1$, $i_j \neq i_2$. Oznacza to, że reguła zaszyfrowuje wszystkie te relacje, z których można byłoby skonstruować złożenie relacji semantycznych. Zatem podzbiór $B(E1)$ ma postać

$$B(E1) = \bigcup_{\substack{i_k \\ i_k \neq i_1}} R(x_{i_1}, x_{i_k}) \cup \bigcup_{\substack{i_j \\ i_j \neq i_2}} R(x_{i_2}, x_{i_j}).$$

Regułę tę zilustrujemy na rysunku, przy czym bazę danych przedstawimy jako graf, którego wierzchołkami są pola, a krawędziami relacje semantyczne. Krawędzie będą zaznaczane cienką linią, gdy odpowiadające im relacje nie podlegają szyfracji, natomiast linią pogrubioną, gdy odpowiednie relacje będą szyfrowane (patrz rys. 7).

Reguła E2. Reguła selektywnego szyfrowania relacji semantycznej

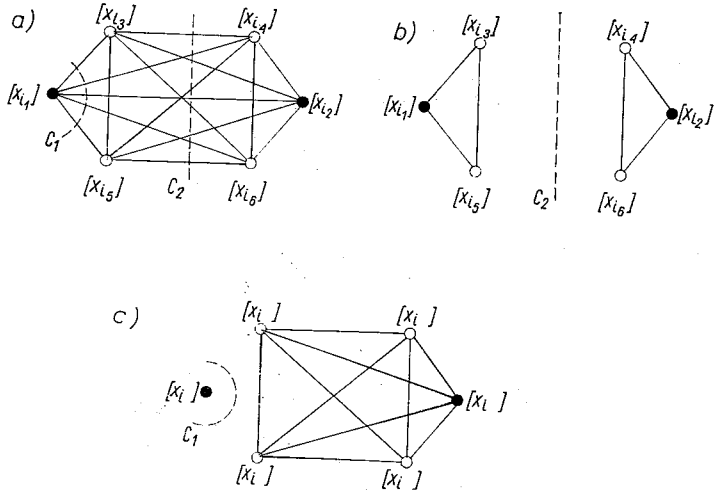
Omawiana reguła udostępnia NU wszystkie te relacje, na podstawie których nie można ustalić złożenia relacji semantycznej $R_z(x_{i_1}, x_{i_2})$ $(R_z \supset R(x_{i_1}, x_{i_2}))$. Reguła ta ma prostą interpretację graficzną. Dla rozważanej bazy danych tworzymy graf niezorientowany



Rys. 7. Schemat działania reguły E1 w odniesieniu do relacji $R(x_{i_1}, x_{i_2})$
a) baza danych widziana przez NU przed zastosowaniem reguły E1, b) baza danych widziana przez NU po zastosowaniu reguły E1 (krawędzie odpowiadające zaszyfrowanym relacjom zostały pominięte)

[10] Taki przypadek zachodzi wtedy, gdy NU jest projektant systemu komputerowego.

(wszystkie krawędzie są zaznaczone cienką linią). Jeżeli chcemy zaszyfrować relację semantyczną $R(x_{i_1}, x_{i_2})$, to dokonujemy przekroju grafu w taki sposób, aby wierzchołki $[x_{i_1}]$ i $[x_{i_2}]$ znalazły się w różnych podgrafach. Dla każdego przekroju można ustalić zbiór usuwanych krawędzi odpowiadających zbiorowi relacji semantycznych, które będą podlegały szyfrowaniu. Zbiór ten oznaczamy przez C. Przekrojów rozdzielających wierzchołki $[x_{i_1}]$ i $[x_{i_2}]$ może być wiele. Oczywiście interesować nas będzie taki przekrój, dla którego zbiór C jest najmniej liczny. Niech $|C|$ oznacza liczbę relacji zawartych w zbiorze C oraz niech $C_1, C_2, \ldots, C_n$ jest układem zbiorów relacji semantycznych dla wszystkich możliwych przekrojów grafu; wtedy $B(E2) = C_i$, gdzie $|C_i| = \min_j |C_j|$ (rys. 8). Warto zauważyć, że optymalny przekrój można wybierać także w inny sposób uwzględniając inne kryteria.



Rys. 8. Zastosowanie reguły E2 do szyfrowania relacji $R(x_{i_1}, x_{i_2})$
a) baza danych widziana przez NU przed zastosowaniem reguły E2, b), c) baza danych widziana przez NU po zastosowaniu reguły E2 (dwie równorzędne wersje)

Reguła E3. Reguła całkowitego szyfrowania pary jednostek danych (d_j, d_k)

Załóżmy, że blok fizyczny $BF(d_j, d_k)$ należy do relacji semantycznej $R(x_{i_1}, x_{i_2})$, $d_j \in [x_{i_1}]$, $d_k \in [x_{i_2}]$ i chcemy go zaszyfrować, przy czym nie chcemy szyfrować całej relacji R. Postępujemy tu podobnie jak w przypadku reguły E1, z tym że w odniesieniu do jednostek danych. Szyfrujemy więc wszystkie bloki fizyczne, w których występują jednostki danych d_j lub d_k, tzn. chronimy te bloki, które mogą nieść jakąkolwiek informację o parze jednostek danych (d_j, d_k). Stąd

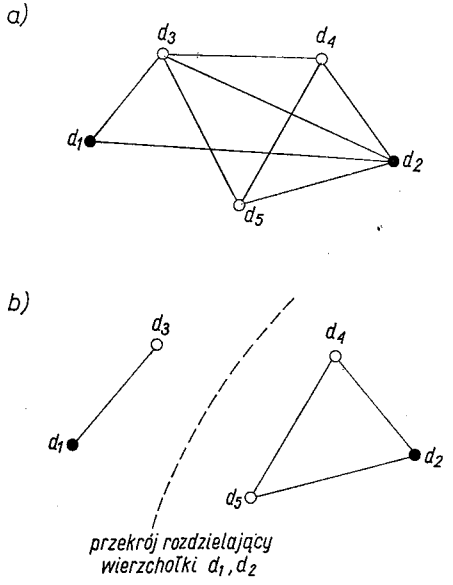
$$B(E3) = \bigcup_{p,r} BF(d_p, d_r) \cup \bigcup_{p,r} BF(d_r, d_p),$$

gdzie $(p = j) \wedge (r \neq j) \vee (p = k) \wedge (r \neq k) \vee (p = j) \wedge (r = k)$.

Reguła E4. Reguła selektywnego szyfrowania pary jednostek danych (d_j, d_k)

Spośród relacji semantycznych o postaci $R(x_{i_1}, x_{i_l})$ oraz $R(x_{i_2}, x_{i_l})$ wybieramy wszystkie elementy, w których występują dane d_j lub d_k i otrzymujemy zbiór D wszystkich

par, które są semantycznie powiązane z jednostkami danych d_j lub d_k. Jednostkom danych występującym w zbiorze D przyporządkowujemy wierzchołki grafu. Wierzchołki d_j i d_k łączymy krawędzią, jeżeli istnieje para (d_p, d_r) w zbiorze D. Podobnie jak w przypadku reguły E3 dokonujemy przekroju otrzymanego grafu na dwa podgrafy, które zawierają odpowiednio jednostkę danych d_j i jednostkę danych d_k. Każdemu przekrojowi przyporządkowujemy zbiór krawędzi, które muszą być usunięte z grafu, aby powstały dwa rozłączne podgrafy. Niech c_i, $i = 1, 2, \ldots, n$, jest zbiorem wszystkich takich krawędzi. Jeżeli dla każdego istnieje n różnych przekrojów rozdzielających wierzchołki (d_j, d_k) rozważanego grafu, wtedy $B(E4) = c_i$. Omawianą regułę ilustruje rys. 9.



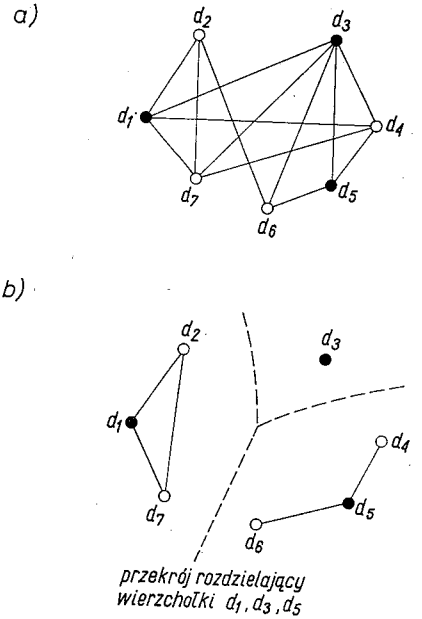
Rys. 9. Zastosowanie reguły E4 do szyfrowania pary jednostek danych d_1, d_2
a) zbiór D widziany przez NU przed zastosowaniem reguły E4, b) zbiór D widziany przez NU po zastosowaniu reguły E4 (krawędzie odpowiadające zaszyfrowanym parom są pominięte)

Reguła E5. Reguła selektywnego szyfrowania jednostek danych $(d_1, \ldots, d_r)$

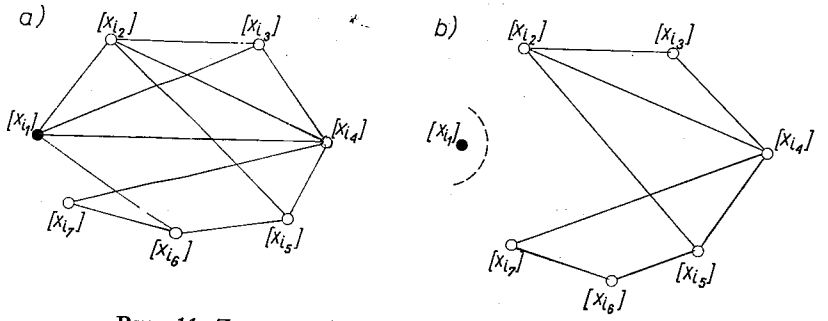
Tworzymy graf dla omawianej bazy danych wybierając tyle wierzchołków, ile jest jednostek danych występujących w związku semantycznym z jednostkami danych $d_1, \ldots, d_r$. Każdą parę wierzchołków łączymy krawędzią, jeżeli para ta pozostaje w odpowiedniej relacji semantycznej. Dokonujemy przekroju grafu na r rozłącznych podgrafów, z których każdy zawiera jedną jednostkę danych ze zbioru $\{d_1, \ldots, d_r\}$. Ilustracja tej reguły jest podana na rys. 10.

Reguła E6. Reguła szyfrowania pola $[x_{i_1}]$

Aby NU podczas penetracji w BD nie mógł odtworzyć pola $[x_{i_1}]$, należy szyfrować wszystkie relacje semantyczne, które są określone na tym polu, czyli te które mają postać $R(x_{i_1}, x_{i_j})$ lub $R(x_{i_j}, x_{i_1})$ $(j \neq 1)$.



Rys. 10. Zastosowanie reguły E5 do szyfrowania bloku fizycznego (d_1, d_3, d_5)
a) powiązania semantyczne widziane przez NU przed zastosowaniem reguły E5, b) powiązania semantyczne widziane przez NU po zastosowaniu reguły E5



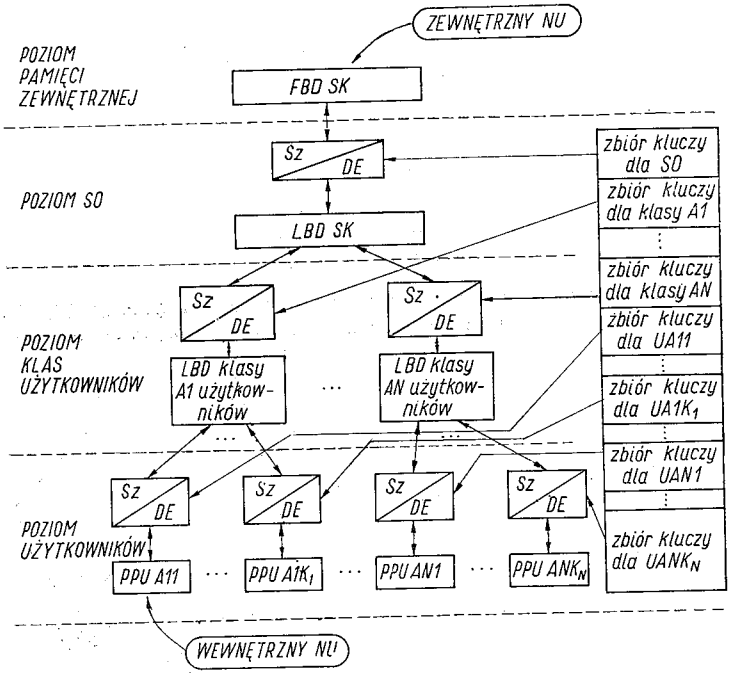
Rys. 11. Zastosowanie reguły E6 do ochrony pola $[x_{i_1}]$
a) powiązania semantyczne przed wprowadzeniem reguły E6, b) powiązania semantyczne po wprowadzeniu reguły E6

Dla rozważanej przykładowej bazy danych otrzymujemy wyniki, które są przedstawione na rys. 11.

4.3. Ocena jakości reguł kryptograficznej ochrony bazy danych z punktu widzenia prawdopodobieństwa złamania klucza

Podsumowanie dotychczasowych rozważań dotyczących kryptograficznej ochrony BD jest przedstawione schematycznie na rys. 12. Na rysunku tym odróżniamy wewnętrzne NU od zewnętrznego NU. Różnica pomiędzy nimi polega na tym, że ten pierwszy jest użytkownikiem BD usiłującym otrzymać dostęp do chronionych zbiorów informacji,

mimo że nie jest do tego autoryzowany. Drugi natomiast nie jest użytkownikiem systemu, a ponieważ zakładamy, że nie ma dostępu do końcówek, więc może uzyskać interesujące go informacje tylko poprzez obserwację zawartości pamięci zewnętrznej. Zagrożenie stwarzane przez wewnętrznego NU jest na ogół poważniejsze niż zagrożenie wywoływane przez zewnętrznego NU, a to dlatego, że dla wewnętrznego NU dostępne są niektóre dodatkowe informacje. W szczególności znając procedury wywołania, modyfikacji i kasowania danych może on określić przekształcenie kryptograficzne stosowane pomiędzy poszczególnymi poziomami organizacji BD lub niektóre relacje semantyczne. W tej sytuacji interesująca nas ocena jakości reguł kryptograficznej ochrony BD będzie zależna od rodzaju NU.



Rys. 12. Schemat bazy danych z możliwymi do zastosowania przekształceniami kryptograficznymi (PZBI został pominięty)

W celu dokonania takiej oceny przyjmujemy następujące założenia:

Z1. każdy użytkownik korzystając z BD jest poprawnie identyfikowany przez PZBI (korzysta z odpowiedniej przestrzeni pracy),

Z2. NU zna rodzaje przekształceń kryptograficznych stosowanych do ochrony poszczególnych elementów BD[11],

Z3. NU zna w jawnej postaci elementy BD w przydzielonej mu przestrzeni, pozostałe zaś elementy z poziomów klas użytkowników i systemu operacyjnego zna w postaci zaszyfrowanej, oraz wie jakie powiązania semantyczne zachodzą pomiędzy poszczególnymi elementami,

[11] Taki przypadek ma miejsce, gdy NU jest projektant systemu komputerowego.

Z4. wartości kluczy, które są użyte do zaszyfrowania bloków logicznych, są dla użytkowników nieznane i użytkownicy nie mają do nich dostępu,

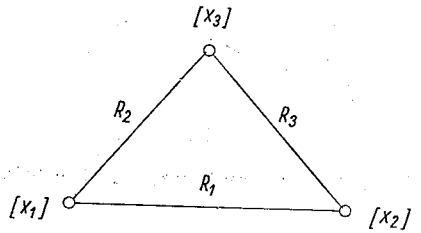
Z5. PZBD przy stosowaniu przekształcenia kryptograficznego wybiera klucz ze zbioru jednakowo prawodopodobnych kluczy.

Przy powyższych założeniach wyznaczymy prawdopodobieństwo złamania klucza przez NU. Można je obliczyć na podstawie sformułowanych poniżej twierdzeń.

Twierdzenie 1. Jeżeli baza danych składa się z trzech relacji: $R_1 = R_1(x_1, x_2)$, $R_2 = R_2(x_1, x_3)$, $R_3 = R_3(x_2, x_3)$ określonych na polach $[x_1]$, $[x_2]$. $[x_3]$ (rys. 13), oraz NU zna relacje R_2 i R_3 w jawnej postaci, a relację R_1 w postaci zaszyfrowanej (tzn. zna relację $R_1' = f_1(R_1, m_{1g})$, gdzie $m_{1g} \in M_1$, $1 \leqslant g \leqslant K_1 = |M_1|$) i jeżeli spełnione są założenia Z1—Z5, wtedy prawdopodobieństwo złamania klucza jest równe $\frac{1}{K_p}$, gdzie K_p est liczbą kluczy należących do zbioru M_{1p} oraz

$$M_{1p} = \{m_{1j_1} \in M_1; \quad f_1(R_{23}, m_{1j_1}) \cap R_1' = R_1'\}, \tag{16}$$

a relacja R_{23} jest złożeniem relacji R_2 i R_3.



Rys. 13. Schemat bazy danych rozważanej w twierdzeniu 1

Dowód. NU w celu rozszyfrowania relacji R_1 będzie chciał wykorzystać związki semantyczne istniejące w BD. Dlatego dokona najpierw złożenia relacji R_2 i R_3. Ponieważ NU wie, że zachodzi inkluzja $R_1 \subset R_{23}$, więc wie, że inkluzja taka zachodzi także dla relacji zaszyfrowanych tym samym kluczem

$$R_1' \subset R_{23}'. \tag{17}$$

W tej sytuacji NU wykreśla ze zbioru wszystkich kluczy M_1 te klucze które nie spełniają inkluzji (17), ponieważ nie mogą być one zastosowane do zaszyfrowania relacji R_1. Jako klucze, które mogłyby wchodzić w grę, zostaną wybrane te klucze, które spełniają warunek:

$$(f_1(R_{23}, m_{1j_1}) \supset R_1') \Leftrightarrow (f_1(R_{23}, m_{1j_1}) \cap R_1' = R_1'). \tag{18}$$

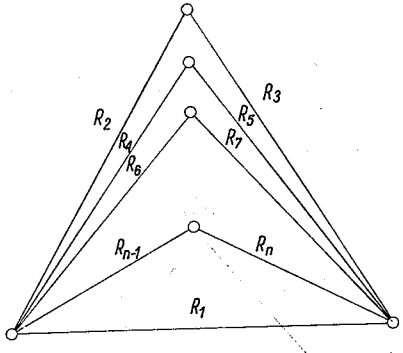
Stąd zbiór kluczy, wśród których jest poszukiwany klucz, ma postać określoną wzorem (16). Korzystając z założenia Z5 możemy stwierdzić, że NU chcąc określić wartość klucza kryptograficznego będzie musiał wybrać jeden spośród $|M_{1p}| = K_p$ jednakowo prawdopodobnych kluczy spełniających warunek (18).

Uogólnienie sformułowanego twierdzenia, gdy istnieje więcej relacji semantycznych powiązanych z relacją R_1, jest podane w poniższym twierdzeniu.

Twierdzenie 2. Jeżeli BD składa się z relacji $R_1, R_2, \ldots, R_n$, gdzie n jest liczbą nieparzystą, określonych na polach $[x_1], [x_2], \ldots, [x_t]$ (rys. 14), oraz NU zna wszystkie relacje w postaci jawnej oprócz relacji R_1 i spełnione są założenia Z1—Z5, wtedy prawdopodobieństwo złamania klucza jest równe $\frac{1}{K_p}$, gdzie K_p jest liczbą kluczy należących do zbioru M_{1p}

$$M_{1p} = \{m_{1j_1} \in M_1; \quad \bigwedge_{g=2,4,\ldots,n-1} f_1(R_{g(g+1)}, m_{1j_1}) \cap R'_1 = R'_1\}, \tag{19}$$

przy czym $R_{g(g+1)}$ jest złożeniem relacji R_g oraz R_{g+1}, a relacja R'_1 jest zaszyfrowaną relacją R_1.



Rys. 14. Schemat bazy danych rozważanej w twierdzeniu 2

Dowód. Jak wynika z rys. 14 zachodzą następujące inkluzje

$$R_{23} \supset R_1, \; R_{45} \supset R_1, \; R_{67} \supset R_1, \ldots, R_{(n-1)n} \supset R_1. \tag{20}$$

Dla klucza, który został zastosowany do szyfrowania relacji R_1, zachodzi więc inkluzja

$$\bigcap_{g=2,4,\ldots,n-1} f_1(R_{g(g+1)}, m_{1k}) \supset R'_1. \tag{21}$$

Ponieważ NU wie, że spełnione jest założenie Z5, będzie więc wybierał jeden z jednakowo prawdopodobnych kluczy należących do zbioru M_{1p}, przy czym

$$M_{1p} = \{m_{1j_1} \in M_1; \quad \bigwedge_{g=2,4,\ldots,n-1} f_1(R_{g(g+1)}, m_{1j_1}) \supset R'_1\} \Leftrightarrow$$
$$\{m_{1j_1} \in M_1; \quad \bigwedge_{g=2,4,\ldots,n-1} f_1(R_{g(g+1)}, m_{1j_1}) \cap R'_1 = R'_1\}. \tag{22}$$

Jeżeli $|M_{1p}| = K_p$, wtedy prawdopodobieństwo wyznaczenia prawidłowej wartości klucz będzie równe $\frac{1}{K_p}$.

W powyżej sformułowanych twierdzeniach rozważaliśmy prawdopodobieństwo wy znaczenia klucza przez NU, gdy szyfracji podlega pojedyncza relacja. W bardziej ogólny przypadku szyfracji może podlegać więcej relacji, jak ma to miejsce, gdy w grę wchod reguła E2. Prawdopodobieństwo wyznaczenia klucza przez NU może być teraz określo na podstawie następującego twierdzenia.

Twierdzenie 3. Jeżeli BD składa się z trzech relacji semantycznych: $R_1 = R_1(x_1, x_2)$, $R_2 = R_2(x_1, x_3)$, $R_3 = R_3(x_2, x_3)$ (rys. 13) oraz NU ma dostęp do relacji R_1 i R_2 w postaci zaszyfrowanej[12] a do relacji R_3 w postaci jawnej i spełnione są założenia Z1—Z5, to prawdopodobieństwo złamania klucza zastosowanego do ochrony relacji R_1 jest równe $\frac{1}{K_r}$, przy czym K_r jest liczbą kluczy m_{1j_1} należących do zbioru

$$M_{1r} = \{m_{1j_1} \in M_1; \quad \bigvee_{1 \leqslant k \leqslant K_2} R_1' \cap f_1(R_{23}(k), m_{1j_1}) = R_1'\}, \tag{23}$$

przy czym relacja $R_{23}(k)$ jest złożeniem relacji $f_2^{-1}(R_2', m_{2k})$ z relacją R_3; $f_1(\cdot, *), f_2(\cdot, *)$ są funkcjami kryptograficznymi stosowanymi do szyfrowania odpowiednio relacji R_1 i R_2, a m_{2k} jest kluczem używanym do szyfrowania relacji R_2.

Dowód. Ponieważ NU zna funkcje kryptograficzne $f_1(\cdot, *)$ i $f_2(\cdot, *)$ oraz relacje R_1', R_2' i R_3, może więc określić dla każdego klucza $m_{2k} \in M_2$ ($|M_2| = K_2$) relacje

$$f_2^{-1}(R_2', m_{2k}) = R_2(k); \quad 1 \leqslant k \leqslant K_2. \tag{24}$$

W zbiorze relacji $\{R_2(1), \ldots, R_2(K_2)\}$ znajduje się relacja R_2. Jeżeli dokonamy złożenia relacji $R_2(k)$ ($1 \leqslant k \leqslant K_2$) z relacją R_3, to otrzymamy relację $R_{23}(k)$. W rezultacie dostaniemy zbiór $\{R_{23}(1), \ldots, R_{23}(K_2)\}$ i jeżeli do każdego elementu tego zbioru zastosujemy przekształcenie kryptograficzne $f_1(\cdot, m_{1j_1})$, $m_{1j_1} \in M_1$, to otrzymamy zbiór

$$\{f_1(R_{23}(1), m_{1j_1}), \ldots, f_1(R_{23}(K_2), m_{1j_1})\}.$$

Zauważmy, że NU odrzuci klucz $m_{1j_1} \in M_1$, gdy

$$\bigwedge_{1 \leqslant k \leqslant K_2} R_1' \cap f_1(R_{23}(k), m_{1j_1}) \neq R_1'. \tag{25}$$

Wynika stąd, że NU chcąc wyznaczać klucz zastosowany do ochrony relacji R_1 musi wybrać go ze zbioru M_{1r} jednakowo prawdopodobnych kluczy (zał. Z5), przy czym

$$M_{1r} = \{m_{1j_1} \in M_1; \quad \bigvee_{1 \leqslant k \leqslant K_2} R_1' \cap f_1(R_{23}(k), m_{1j_1}) = R_1'\}. \tag{26}$$

Oczywiście prawdopodobieństwo poprawnego określenia klucza jest równe

$$\frac{1}{K_r} = \frac{1}{|M_{1r}|}.$$

Twierdzenie 3 można uogólnić w następującej postaci.

Twierdzenie 4. Jeżeli BD składa się z n relacji semantycznych (rys. 14), gdzie n jest liczbą nieparzystą, oraz NU zna relacje $R_1, R_2, R_4, \ldots, R_{n-1}$ w postaci zaszyfrowanej, relacje z pozostałymi indeksami $R_3, R_5, \ldots, R_n$ w jawnej postaci i spełnione są założenia Z1—Z5, to prawdopodobieństwo wyznaczenia klucza zastosowanego do relacji R_1 wynosi $\frac{1}{K_r}$, gdzie K_r jest liczbą elementów zbioru M_{1r}

$$M_{1r} = \{m_{1j_1} \in M_1; \quad \bigwedge_{g=2,4,\ldots,n-1} \bigvee_{1 \leqslant k_g \leqslant K_g} R_1' \cap f_1(R_{g(g+1)}(k_g), m_{1j_1}) = R_1'\} \tag{27}$$

[12] Tzn. ma dostęp do relacji R_1' i R_2'.

a relacja $R_{g(g+1)}(k_g)$ jest złożeniem relacji $R_g(k_g) = f^{-1}(R'_g, m_{gk_g})$ z relacją R_{g+1} ($g = 2, 4, \ldots, n-1$), natomiast funkcje $f_1(\cdot, *)$, $f_2(\cdot, *)$, $f_4(\cdot, *)$, $f_6(\cdot, *)$, $\ldots, f_{n-1}(\cdot, *)$ są funkcjami kryptograficznymi stosowanymi do szyfrowania relacji $R_1, R_2, R_4, R_6, \ldots, R_{n-1}$.

Dowód. W celu udowodnienia twierdzenia przeprowadzamy podobne rozumowanie jak przy dowodzie twierdzenia 3 i utworzymy w tym celu zbiory relacji

$$\begin{gathered}\{R_2(k_2) = f_2^{-1}(R'_2, m_{2k_2}); \quad 1 \leqslant k_2 \leqslant K_2\}, \\ \vdots \\ \{R_{n-1}(k_{n-1}) = f_{n-1}^{-1}(R'_{n-1}, m_{(n-1)k_{(n-1)}}); \quad 1 \leqslant k_{n-1} \leqslant K_{n-1}\}.\end{gathered} \tag{28}$$

Na tej podstawie otrzymujemy złożenia relacji dla elementów powyższych zbiorów, a więc

$$\begin{gathered}\{R_{23}(k_2); \quad 1 \leqslant k_2 \leqslant K_2\}, \\ \vdots \\ \{R_{(n-1)n}(k_{n-1}); \quad 1 \leqslant k_{n-1} \leqslant K_{n-1}\}.\end{gathered} \tag{29}$$

NU usiłujący określić relację R_1 odrzuci każdy klucz $m_{1j_1} \in M_1$, gdy spełniona jest zależność

$$\bigvee_{g=2,4,\ldots,n-1} \bigwedge_{1 \leqslant k_g \leqslant K_g} R'_1 \cap f_1(R_{g(g+1)}(k_g), m_{1j_1}) \neq R'_1. \tag{30}$$

Zatem zbiór kluczy, wśród których znajduje się poszukiwany przez NU klucz, będzie miał postać

$$M_{1r} = \left\{m_{1j_1} \in M_1; \quad \bigwedge_{g=2,4,\ldots,n-1} \bigvee_{1 \leqslant k_g \leqslant K_g} R'_1 \cap f_1(R_{g(g+1)}(k_g), m_{1j_1}) = R'_1\right\}.$$

Prawdopodobieństwo poprawnego określenia klucza zastosowanego do zaszyfrowania relacji R_1 jest, przy uwzględnieniu założenia Z5, równe

$$\frac{1}{K_r}; \quad \text{gdzie} \quad K_r = |M_{1r}|. \tag{31}$$

Przedstawione powyżej twierdzenia 3 i 4 odnosiły się do oceny jakości reguły selektywnego szyfrowania relacji R_1 (E2). Jak już wiemy z poprzedniego punktu, reguła E1 dotyczy szyfrowania wszystkich mogących wchodzić w grę relacji. Zatem skuteczność takiej ochrony będzie na ogół lepsza, niż skuteczność ochrony przy zastosowaniu reguły E2. Podamy teraz dwa kolejne twierdzenia, które dotyczą oceny jakości ochrony przy wykorzystaniu reguły E1.

Twierdzenie 5. Jeżeli BD składa się z relacji $R_1, R_2, R_3, \ldots, R_n$ (rys. 14), gdzie n jest liczbą nieparzystą i do ochrony BD została zastosowana reguła kryptograficzna E1, a NU zna wszystkie relacje w postaci zaszyfrowanej $R'_1, R'_2, R'_3, \ldots, R'_n$, oraz spełnione są założenia Z1—Z5, to prawdopodobieństwo wyznaczenia klucza zastosowanego do relacji R_1 jest równe $\frac{1}{K_s}$, gdzie $K_s = |M_{1s}|$, a zbiór M_{1s} ma postać

$$M_{1s} = \left\{m_{1j_1} \in M_1; \quad \bigwedge_{g=2,4,\ldots,n-1} \bigvee_{\substack{1 \leqslant k_g \leqslant K_g \\ 1 \leqslant k_{g+1} \leqslant K_{g+1}}} f_1(R_{g(g+1)}(k_g, k_{g+1}), m_{1j_1}) \cap R'_1 = R'_1\right\}. \tag{3}$$

Dowód. Do zaszyfrowania relacji R_g i R_{g+1} jest użyta para kluczy należąca do zbioru $M_g \times M_{g+1}$ ($g = 2, 4, \ldots, n-1$). NU znając funkcje kryptograficzne (zał. Z2) może obliczyć dla każdej wartości $g = 2, 4, \ldots, n-1$ wszystkie relacje, które można utworzyć dla pary kluczy $m_{gk_g} \in M_g$, $m_{(g+1)k_{g+1}} \in M_{g+1}$, a więc relacje

$$R_g(k_g) = f_g^{-1}(R'_g, m_{gk_g}),\ R_{(g+1)}(k_{g+1}) = f_{g+1}^{-1}(R'_{g+1}, m_{(g+1)k_{g+1}}).$$

Z relacji tych tworzymy złożenie i oznaczamy je przez $R_{g(g+1)}(k_g, k_{g+1})$. Zatem otrzymujemy zbiór relacji

$$\{R_{g(g+1)}(k_g, k_{g+1});\quad 1 \leqslant k_g \leqslant K_g,\ 1 \leqslant k_{g+1} \leqslant K_{g+1}\}.$$

Zauważmy, że w zbiorze tym jest złożenie relacji $R_{g(g+1)}$. Interesujący nas zbiór kluczy M_1 można podzielić na dwa podzbiory: podzbiór kluczy, wśród których znajduje się klucz użyty do szyfrowania relacji R_1 (zbiór M_{1s}) oraz podzbiór, do którego klucz ten nie należy. Postępując podobnie jak przy dowodzie poprzedniego twierdzenia możemy stwierdzić, że z faktu przynależności klucza m_{1j_1} do zbioru M_{1s} wynika, iż dla każdego $g = 2, 4, \ldots, n-1$ musi istnieć para kluczy $(m_{gk_g}, m_{(g+1)k_{g+1}})$; $1 \leqslant k_g \leqslant K_g$, $1 \leqslant k_{g+1} \leqslant K_{g+1}$, takich że zachodzi inkluzja

$$f_1(R_{g(g+1)}(k_g, k_{g+1}), m_{1j_1}) \supset R'_1. \tag{33}$$

Ponieważ inkluzję (33) można zapisać w innej równoważnej postaci, mamy więc

$$f_1(R_{g(g+1)}(k_g, k_{g+1}), m_{1j_1}) \cap R'_1 = R'_1. \tag{34}$$

Uwzględniając założenie Z5 otrzymujemy ostateczną postać tezy tw. 5.

Na zakończenie tego punktu udowodnimy jeszcze twierdzenie, które dotyczyć będzie powiązań między liczbami K_p, K_r, K_s występującymi w twierdzeniach 2, 4 i 5.

Twierdzenie 6. Liczby K_p, K_r, K_s określone odpowiednio w twierdzeniach 2, 4 i 5 spełniają nierówności

$$K_1 \geqslant K_s \geqslant K_r \geqslant K_p \geqslant 1. \tag{35}$$

Dowód. Wykazanie słuszności tezy tego twierdzenia można sprowadzić do wykazania słuszności ciągu inkluzji $M_{1s} \supset M_{1r} \supset M_{1p}$. Pierwsza i ostatnia inkluzja zostały pominięte, przy czym poszczególne zbiory są określone w następujący sposób

$$M_{1s} = \{m_{1j_1} \in M_1;\quad \bigwedge_{g=2,4,\ldots,n-1} \bigvee_{\substack{1 \leqslant k_g \leqslant K_g \\ 1 \leqslant k_{g+1} \leqslant K_{g+1}}} f_1(R_{g(g+1)}(k_g, k_{g+1}), m_{1j_1}) \cap R'_1 = R'_1\}, \tag{36}$$

$$M_{1r} = \{m_{1j_1} \in M_1;\quad \bigwedge_{g=2,4,\ldots,n-1} \bigvee_{1 \leqslant k_g \leqslant K_g} f_1(R_{g(g+1)}(k_g), m_{1j_1}) \cap R'_1 = R'_1\}, \tag{37}$$

$$M_{1p} = \{m_{1j_1} \in M_1;\quad \bigwedge_{g=2,4,\ldots,n-1} f_1(R_{g(g+1)}, m_{1j_1}) \cap R'_1 = R'_1\}. \tag{38}$$

Tezę twierdzenia udowodnimy, jeżeli wykażemy, że zachodzi następujący ciąg implikacji $m \in M_{1p} \Rightarrow m \in M_{1r} \Rightarrow m \in M_{1s}$. Załóżymy, że $m \in M_{1p}$. Zatem z określenia (38) wynika, że

$$m \in M_{1p} \Leftrightarrow \bigwedge_{g=2,4,\ldots,n-1} f_1(R_{g(g+1)}, m) \cap R'_1 = R'_1 \Rightarrow$$

$$\Rightarrow \bigwedge_{g=2,4,\ldots,n-1} \bigvee_{1 \leqslant k_g \leqslant K_g} f_1(R_{g(g+1)}(k_g), m) \cap R'_1 = R'_1 \Leftrightarrow m \in M_{1r}.$$

Omawiana implikacja zachodzi, ponieważ istnieje klucz k_g ($1 \leqslant k_g \leqslant K_g$), dla którego spełniony jest warunek (37). Jest nim klucz, który został zastosowany do szyfrowania relacji R_g. Ostatnia interesująca nas implikacja ma postać

$$m \in M_{1r} \Rightarrow \bigwedge_{g=2,4,\ldots,n-1} \bigvee_{\substack{1 \leqslant k_g \leqslant K_g \\ 1 \leqslant k_{g+1} \leqslant K_{g+1}}} f_1(R_{g(g+1)}(k_g, k_{g+1}), m) \cap R_1' = R_1' \Leftrightarrow m \in M_{1s}.$$

Takie postawienie implikacji jest poprawne, gdyż istnieje co najmniej jeden klucz k_{g+1} ($1 \leqslant k_{g+1} \leqslant K_{g+1}$), dla którego słuszna jest ta implikacja. Jest nim klucz, który został zastosowany do szyfrowania relacji R_{g+1}.

W ten sposób wykazaliśmy słuszność postawionej tezy. Z udowodnionego twierdzenia wynika, że reguła kryptograficzna E1 całkowitego zaszyfrowania relacji semantycznej jest lepsza (odpowiada jej mniejsze prawdopodobieństwo określenia przez NU zastosowanego klucza) od reguły E2. Obie te reguły są jednak lepsze od szyfrowania wyłącznie relacji R_1, bez uwzględnienia powiązań semantycznych (tw. 1 i 2). Ocena jakości reguł E3 i E4 jest podobna jak reguł E1 oraz E2 i dlatego została pominięta.

5. PRZYKŁADOWA BAZA DANYCH Z KRYPTOGRAFICZNĄ OCHRONĄ JEJ ELEMENTÓW

W celu zilustrowania poprzednio przedstawionych problemów zajmiemy się przykładową bazą danych służącą do przechowywania informacji o pracownikach pewnego instytutu badawczego oraz o prowadzonych przez nich pracach badawczych. Rozważana baza danych będzie relacyjną bazą danych, tzn. zgromadzone w niej dane będą przechowywane w postaci relacji [1, 14]. Przyjmiemy, że logiczna baza danych składa się z trzech relacji *R*1, *R*2 i *R*3 podanych za pomocą tablicy 9.

Do pełnego określenia logicznej bazy danych niezbędne jest podanie definicji poszczególnych pól oraz relacji. Załóżmy dalej, że słuszne są następujące definicje.

Definicje pól:

[NRPR] — liczba dziesiętna o długości 2,
[NAZ] — sekwencja 10 liter,
[IMIĘ] — sekwencja 10 liter,
[STN] — sekwencja 4 liter,
[NRPB] — liczba dziesiętna o długości 3,
[SNZ] — sekwencja 3 liter,
[ZAROBEK] — liczba dziesiętna o długości 5.

Definicje relacji:

R1 (NRPR, NAZ, IMIE, STN), pole podstawowe[13] [NRPR],
R2 (NRPB, NAZ, SNZ), pole podstawowe [NRPB],
R3 (NRPB, NRPR, ZAROBEK), pole podstawowe [NRPB] × [NRPR].

Relacja R1 podaje nazwisko, imię i stopień naukowy poszczególnych pracowników instytutu badawczego. Natomiast relacja R2 określa dla poszczególnych prac badawczych nazwisko jej kierownika oraz skrót nazwy zleceniodawcy. W końcu relacja R3 dostarcza

[13] Pole X jest dla relacji $R(X, Y)$ polem podstawowym, jeżeli dla każdej jednostki danych należącej do pola X istnieje dokładnie jedna jednostka należąca do pola Y.

Tablica 9

Postać przykładowej bazy danych o trzech relacjach

R1

Nr kolejny pracownika [NRPR]	Nazwisko pracownika [NAZ]	Imię pracownika [IMIĘ]	Stopień naukowy [STN]
1	Adamski	Jerzy	dr
2	Kowalski	Adam	mgr
3	Kruk	Adam	mgr
4	Nowak	Tadeusz	dr
5	Wiśniewski	Jerzy	dr
6	Zalewski	Wojciech	mgr

R2

Nr problemu badawczego [NRPB]	Nazwisko kierownika pracy [NAZ]	Skrót nazwy zleceniodawcy [SNZ]
110	Wiśniewski	ATR
132	Adamski	ZZR
160	Wiśniewski	PAN

R3

Nr problemu badawczego [NRPB]	Nr kolejny pracownika [NRPR]	Zarobek miesięczny pracownika [ZAROBEK]
110	1	2000
110	2	3500
110	5	4500
132	1	4000
132	3	2000
132	4	6500
132	6	5500
160	3	3000
160	5	3500

nformacji o miesięcznym zarobku, który pracownik otrzymuje przy rozwiązywaniu okre-lonego problemu badawczego. Warto zwrócić uwagę, że powyższe relacje są zbudowane v ten sposób, że po pierwsze wszystkie pola w tych relacjach oprócz pola podstawowego ą od siebie funkcjonalnie niezależne[14], po drugie — są one w pełni funkcjonalnie za-eżne[15] od pola podstawowego. Dla każdej omawianej relacji spełnione są warunki:

- każde dwa pola nie będące polami podstawowymi są wazjemnie funkcjonalnie niezależne,
- każde pojedyncze pole nie będące polem podstawowym jest w pełni funkcjonalnie zależne od pola podstawowego określonej relacji.

[14] Pole Y w relacji $R(X, Y)$ jest funkcjonalnie zależne od pola X, jeżeli dla każdej jednostki danych)loku fizycznego) $x \in X$ istnieje tylko jedna jednostka danych (blok fizyczny) $y \in Y$, przy czym $(x, y) \in R(X, Y)$. W przypadku, gdy ten warunek nie jest spełniony, to pola są funkcjonalnie niezależne [15].

[15] Pole Y w relacji $R(X, Y)$ jest w pełni funkcjonalnie zależne od pola X, jeżeli pole Y nie jest funkcjo-alnie zależne od żadnego podzbioru jednostek danych (bloku fizycznego) pola X [14].

Powołując się na definicje wprowadzone przez Codda [14], można powiedzieć, że powyżej omawiane relacje są w trzeciej znormalizowanej postaci (3ZP).

Aby zbudować przestrzenie pracy użytkowników, muszą być określone pewne działania na relacjach (patrz Dodatek). Przy użyciu tych działań zdefiniujemy poszczególne PPU (LBD klas użytkowników dla omawianej BD jest identyczna z podaną powyżej LBD systemu operacyjnego). W dalszym ciągu zakładamy, że z bazy danych korzystać będzie trzech użytkowników: U1, U2, U3. I tak przyjmujemy, że przestrzenie pracy poszczególnych użytkowników są określone następująco:

PPU1 — relacja określająca miesięczny zarobek każdego z pracowników (dalej nazywana relacją R4),

PPU2 — cała BD oprócz jakiejkolwiek informacji o zarobkach pracowniczych i zleceniodawcach prac badawczych,

PPU3 — cała baza danych.

Opierając się na podanych w Dodatku działaniach na relacjach otrzymujemy następującą postać interesującej nas relacji R4:

R4(NAZ, ZAROBEK);

pole podstawowe [NAZ]

R1 : X

$$R4 = \bigcup_{X\langle NAZ\rangle} \{X\langle NAZ\rangle, [ZAROBEK] = RAZEM((R1\langle NAZ, NRPR\rangle * R3\langle NRPR, NRPB, ZAROBEK\rangle)\langle NAZ = X\langle NAZ\rangle, NRPB, ZAROBEK\rangle)\langle ZAROBEK\rangle\}.$$

Powyższy zapis określa nazwę relacji, zbiór pól, na których jest ona zdefiniowana oraz jej pole podstawowe (jest to pole [NAZ]). Zmienna X ma zakres zmienności taki sam jak relacja R1. Relacja R4 jest więc zbiorem bloków logicznych o postaci

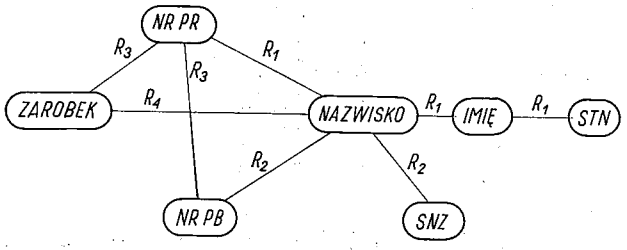
$$R4 = \bigcup_{X\langle NAZ\rangle} \{X\langle NAZ\rangle, ZAROBEK\},$$

przy czym pojedynczy blok fizyczny tej relacji określa zarobek miesięczny pracownika o danym nazwisku. Pole [ZAROBEK] w relacji R4 jest określone jako suma wszystkich zarobków miesięcznych otrzymywanych przy rozwiązywaniu poszczególnych problemów badawczych[16]. Zauważmy, że relacja R4 nie jest przechowywana w pamięci zewnętrznej (tak jak relacje R1, R2, R3), a jest tworzona na żądanie autoryzowanego użytkownika i po jej wykorzystaniu elementy relacji są wymazywane z pamięci operacyjnej.

Przyjmijmy dalej, że do ochrony elementów BD stosujemy przekształcenia kryptograficzne. Rozważmy teraz, jakie elementy BD i w jaki sposób powinny być zaszyfrowane, aby rozdział zasobów BD pomiędzy użytkownikami był zgodny z powyżej określonym ich przestrzeniami pracy. Dostęp pierwszego użytkownika U1 do BD jest ograniczony do relacji R4. Przypuśćmy, że kontrola tego dostępu jest przeprowadzana poprzez odpowiednio zaprojektowany system operacyjny. Z tego wynika, że użytkownik U1 nie posiada żadnego klucza kryptograficznego, a jego dostęp do BD odbywa się pośrednio poprzez

[16] Relacja R = (R1⟨NAZ, NRPR⟩*R3⟨NRPR, NRPB, ZAROBEK⟩)⟨NAZ = X⟨NAZ⟩, NRPB ZAROBEK⟩ określa dla pracownika o nazwisku należącym do X⟨NAZ⟩ numery prac badawczych, w jakich on bierze udział i jego zarobki miesięczne otrzymywane przy rozwiązywaniu tych problemów, Jeżeli relację R zrzutujemy na pole [ZAROBEK], to otrzymamy częściowe zarobki pracowników. Zarobki t musimy zsumować. Do tego celu służy procedura „RAZEM". Procedura ta jest zdefiniowana na relac o pojedynczym polu i sumuje elementy tego pola (pole to musi być zdefiniowane jako liczba [15]).

wywołanie procedury obliczającej relację R4. Stąd wnioskujemy, że dla tej procedury powinny być przechowywane odpowiednie klucze. Klucze te służą do rozszyfrowania relacji R1 i R3 (patrz definicja relacji R4) z poziomu systemu operacyjnego (dla przypomnienia LBD z poziomu systemu operacyjnego jest identyczna z LBD klas użytkowników), czyli do tworzenia odpowiedniej PPU. W celu wyznaczenia relacji dostępnych użytkownikowi U2 zastosujemy regułę kryptograficzną selektywnego szyfrowania relacji. Schemat powiązań semantycznych w omawianej BD jest pokazany na rys. 15.



Rys. 15. Ilustracja graficzna powiązań semantycznych w przykładowej bazie danych

Z określenia reguły selektywnego szyfrowania wynika, że niedostępnymi relacjami dla użytkownika U2 będą relacje R2, R3, R4 (rys. 15). Ponieważ relacja R4 jest chroniona przez odpowiednio zaprojektowany system operacyjny, więc szyfrowanymi elementami będą relacje R2 i R3. Zakładamy, że szyfrowanie relacji R2 będzie polegało na szyfrowaniu pola [SNZ], a szyfrowanie relacji R3 na szyfrowaniu pola [ZAROBEK]. Z definicji LBD wynika, że jednostki danych należące do pola [ZAROBEK] są 5-cyfrowymi liczbami dziesiętnymi. Zastosujemy do tego pola przekształcenie kryptograficzne w sposób poniżej podany.

Niech zarobki pracowników nie przekraczają liczby $2^{15} = 32\,768$ zł i są podawane z dokładnością do złotówki. Zatem liczbę określającą zarobek można przedstawić jako 15-bitowy ciąg binarny. Przyjmijmy, że kluczem będzie 15-bitowy ciąg binarny. Niech kryptogramem będzie suma modulo dwa tych dwóch ciągów binarnych przedstawiona jako liczba dziesiętna. Wyniki szyfrowania są pokazane poniżej dla wartości klucza $m = 13\,787$.

Tablica 10

Wyniki szyfrowania pola ZAROBEK

[ZAROBEK] pole w postaci jawnej	[ZAROBEK] pole w postaci zaszyfrowanej
2000	12811
3500	14455
4500	9295
4000	14971
2000	12811
6500	11455
5500	8359
3000	15971
3500	14455

Dla pola [SNZ] z relacji R2 proponujemy podobne przekształcenie kryptograficzne. Z określenia tego pola wynika, że jest to ciąg 3-literowy. Zakładamy, że obowiązuje następujące przedstawienie liter w postaci binarnej:

A — 00001	L — 01100	X — 10111
B — 00010	M — 01101	Y — 11000
C — 00011	N — 01110	Z — 11001
D — 00100	O — 01111	1 — 11010
E — 00101	P — 10000	2 — 11011
F — 00110	R — 10001	3 — 11100
G — 00111	S — 10010	4 — 11101
H — 01000	T — 10011	5 — 11110
I — 01001	U — 10100	6 — 11111
J — 01010	V — 10101	7 — 00000
K — 01011	W — 10110	

Po zastosowaniu klucza m = SX3 i po dodaniu do niego modulo dwa zakodowanej binarnie oryginalnej wiadomości, otrzymujemy kryptogramy podane w tablicy 11.

Tablica 11

Wyniki szyfrowania pola SNZ

[SNZ] pole w jawnej postaci	[SNZ] pole w postaci zaszyfrowanej
ATR	TDM
ZZR	DAM
PAN	BZS

Tablica 12

Postać zaszyfrowanych relacji R2′ oraz R3′

[NRPB]	[NAZ]	[SNZ]
110	Wiśniewski	TDM
132	Adamski	DAM
160	Wiśniewski	BZS

[NRPB]	[NRPR]	[ZAROBEK]
110	1	12811
110	2	14455
110	5	9295
132	1	14971
132	3	12811
132	4	11455
132	6	8359
160	3	15971
160	5	14455

Otrzymaliśmy więc relacje R2 i R3 w postaci zaszyfrowanej (tablica 12).

Powyższe relacje tworzą LBD istniejącą w pamięci operacyjnej systemu komputerowego. Dla dodatkowej ochrony informacji o zarobkach poszczególnych pracowników oraz nazw zleceniodawców przed NU zastosujemy szyfrowanie adresu dostępu. Zatem BD istniejąca w pamięci zewnętrznej ma postać określoną w tablicy 13.

Tablica 13

Baza danych po wprowadzeniu przekształceń kryptograficznych

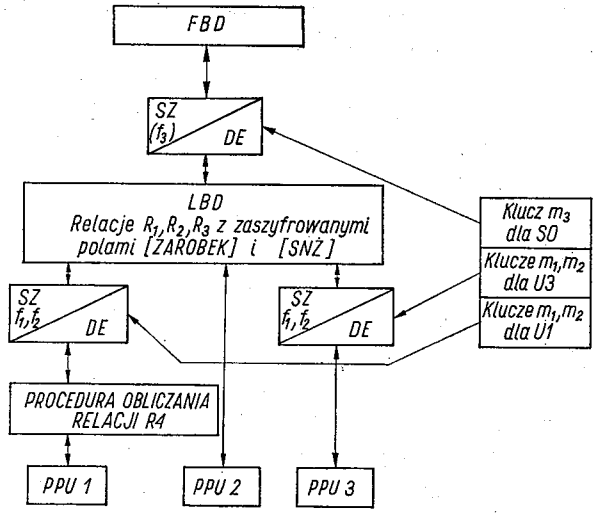
R1 w jawnej postaci

R2'

[NRPB]	[NAZ]	[ADRES]	[SNZ]
110	Wiśniewski	f_3(A21, m)	TDM
132	Adamski	f_3(A22, m)	DAM
160	Wiśniewski	f_3(A23, m)	BZS

[NRPB]	[NRPR]	[ADRES]	[ZAROBEK]
110	1	f_3(A31, m)	12811
110	2	f_3(A32, m)	14455
110	5	f_3(A33, m)	9295
132	1	f_3(A34, m)	14971
132	3	f_3(A35, m)	12811
132	4	f_3(A36, m)	11455
132	6	f_3(A37, m)	8359
160	3	f_3(A38, m)	15971
160	5	f_3(A39, m)	14455

Bazę danych z zastosowaną ochroną kryptograficzną przedstawia rys. 16. W tej prostej BD ocena prawdopodobieństwa poprawnego określenia relacji sematycznej może być przeprowadzona w przypadku użytkownika U2 (PPU1 = {R4}, a PPU3 = BD). Opierając się na założeniach Z1—Z5 podanych w poprzednim punkcie możemy stwierdzić, że prawdopodobieństwo odgadnięcia klucza zastosowanego do ochrony relacji R3 wynosi 2^{-15}, podobnie jak prawdopodobieństwo odgadnięcia klucza zastosowanego do ochrony relacji R2.



Rys. 16. Schemat bazy danych z ochroną kryptograficzną

6. WNIOSKI KOŃCOWE

Koncepcja zastosowania przekształceń kryptograficznych do ochrony zbiorów informacji przechowywanych w BD nie jest nowa. Już Feistel w swojej pracy [20] zaproponował zastosowanie metod kryptograficznych w tych celach. W systematyczny sposób problem kryptograficznej ochrony danych potraktował Gudes [23] wprowadzając wielopoziomowy model organizacji BD i w oparciu o ten model sprecyzował rodzaje przekształceń kryptograficznych, które mogą być zastosowane do ochrony elementów BD. W niniejszej pracy, oprócz przedstawienia uzyskanych dotychczas w literaturze wyników nad zastosowaniem przekształceń kryptograficznych do ochrony informacji w BD, udało się uzyskać kilka nowych wyników. Są one głównie związane z uwzględnieniem powiązań semantycznych przy badaniu jakości kryptograficznej ochrony informacji w BD. Takie kompleksowe ujęcie przekształceń kryptograficznych z uwzględnieniem powiązań semantycznych stworzyło podstawy do określenia reguł kryptograficznej ochrony relacji semantycznych, bloków fizycznych i pól. Końcowymi rezultatami rozważań była ocena jakości tych reguł z punktu widzenia prawdopodobieństwa złamania klucza zastosowano do ochrony określonych elementów BD. Okazało się, że najlepszą regułą jest reguła całkowitej ochrony relacji (jednostki danych). Niemniej jednak, jeżeli do szyfrowania poszczególnych relacji stosowany jest szyfr o bardzo dużej liczbie kluczy[17], to reguła selektywnego szyfrowania elementów BD wydaje się bardziej użyteczna od reguły całkowitego zaszyfrowania. Wynika to z dwóch powodów. Po pierwsze złamanie kluczy dla obu reguł sprowadza się do rozwiązywania bardzo skomplikowanego problemu algebraicznego, przy czym problem ten jest praktycznie nierozwiązywalny, a reguła selektywnego szyfrowania jest prostsza od reguły całkowitego szyfrowania elementów BD (mniej elementów jest szyfrowanych). Po drugie selektywne szyfrowanie wymaga mniejszej liczby elementarnych operacji wynikających z szyfracji i deszyfracji chronionych elementów BD, które są odczytywane, zapisywane bądź modyfikowane.

W niniejszej pracy było również przedstawione zastosowanie reguł kryptograficznych do ochrony bloków fizycznych i jednostek danych. Ponieważ reguły te wymagają dużej liczby elementarnych operacji przetwarzania, a zawartość BD ciągle się zmienia, więc wydaje się, że w praktyce reguły te nie znajdą szerszego zastosowania. Dlatego też jako podstawowy element BD, który będzie podlegał ochronie, zaproponowany został blok logiczny (relacja semantyczna). Tę samą koncepcję w ogólnikowej postaci sformułował Codd [14] w odniesieniu do relacyjnej BD. W niniejszej pracy przedstawiono ogólny model, który może być stosowany zarówno do BD o strukturze relacyjnej, hierarchicznej jak również sieciowej.

W pracy pominięty został w sposób zamierzony problem oceny jakości przekształceń kryptograficznych stosowanych do ochrony pamięci zewnętrznej i niszczących strukturę organizacyjną BD. Przekształceniami tymi są w ogólnym przypadku dowolne przekształcenia kryptograficzne, a ocena ich jakości została dokonana w pracach [20, 25, 32]. W niniejszej pracy rozważane były wyłącznie przekształcenia kryptograficzne zachowujące strukturę BD. Interesujące wydaje się uwzględnienie w tych rozważaniach takich przekształceń kryptograficznych, które nie tylko zachowują strukturę BD, ale również za-

[17] W praktyce liczba kluczy jest często równa 2^{64}.

chowują pewne operacje na nich wykonywane. Korzyścią z zastosowania takich przekształceń byłoby wyeliminowanie potrzeby deszyfracji podczas wykonywania na nich operacji.

Dodatek

SPIS DZIAŁAŃ NA RELACJACH DOPUSZCZALNYCH W RELACYJNEJ BD

W relacyjnej BD możemy zdefiniować następujące działania na relacjach [15]:

D1. Rzut relacji $R(X, Y)$ na pole $X(R(X, Y)\langle X\rangle)$.

D2. Iloczyn dwóch relacji. Jeżeli mamy dwie relacje $R1(X, Z)$, $R2(Z, Y)$, to iloczynem nazywamy relację $R3(X, Z, Y) = R1(X, Z)*R2(Z, Y)$, przy czym blok fizyczny $(d_1, d_2, d_3) \in R3$ wtedy i tylko wtedy, gdy bloki fizyczne $(d_1, d_2) \in R1$ oraz $(d_2, d_3) \in \in R2$ i pole Z jest polem podstawowym relacji R2.

D3. Dzielenie relacji $R1(X, Y)$ przez relację $R2(Y)$. Wynikiem tego działania jest relacja $R3(X)$, taka że dla każdego bloku fizycznego relacji R3 i dla każdego bloku fizycznego relacji R2 istnieje blok fizyczny należący do relacji R1 [15]. Działanie to zapisujemy w postaci $R1(X, Y): R2(Y) = R3(X)$.

LITERATURA

1. F. M. Bancilhon, N. Spyratos, *Protection of information in relational data bases*, Rapport de Recherche IRIA-LABORIA, No. 240, 1977.
2. R. Bayer, J. K. Metzger, *On the encipherment of search trees and random access files*, ACM Transactions on Database Systems, vol. 1, No. 1, March 1976, pp. 37—52.
3. Ch. W. Beardsley, *Is your computer insecure?*, IEEE Spectrum, January 1972, pp. 67—75.
4. J. L. Berg, *Data base directions — The next steps*, Proceedings of the Workshop of the NBS and the ACM held at Fort Landerdale, Florida, October 1975.
5. H. Biller, W. Glatthaar, *On the semantics of data bases: the semantics of data definition languages*, Lecture Notes in Computer Science, vol. 34, 1975, pp. 221—235.
6. H. Biller, W. Glatthaar, E. J. Neuhold, *On the semantics of data bases: The semantics of data manipulation languages, Modelling in Data Base Management Systems*, North Holland Publishing Company, 1976, pp. 239—267.
7. H. Biller, E. J. Neuhold, *The semantic interrelationship of data models*, Angewandte Informatik, No. 2, 1977, pp. 78—83.
9. H. Biller, E. J. Neuhold, *Concepts for the conceptual schema*, Proceedings of the IFIP Working Conference on Modelling in Data Base Management Systems, 1977.
10. R. G. Canning, *Data security in the CDB*, EDP Analyzer, vol. 8, No. 5, May 1970.
11. R. G. Canning, *Protecting valuable data — part I*, EDP Analyzer, vol. 11, No. 12, December 1973.
12. H. A. Chadwick, *Burning down the data center*, Datamation, October 1975, pp. 60—64.
13. D. Clements, L. J. Hoffman, *Computer assisted security system design*, Electronics Research Laboratory, Report ERL-M468, University of California, Berkeley, November 1974.
14. E. F. Codd, *Normalized data base structure: a brief tutorial*, Proceedings of ACM SIGFIDET Workshop on Data Description, Access and Control, 1971.

15. C. J. Date, *An introduction to database systems*, Addison-Wesley Publishing Company, Reading Massachusetts, 1976.
16. M. S. Davis, *Standards, management and security of astronomical data sets*, Reidel Publishing Company, Dordrecht, Holland, 1977, pp. 3—17.
17. P. B. Demitriades, *Administrative secrecy and privacy legislation*, Journal of Systems Management, vol. 27, No. 10, October 1976, pp. 24—29.
18. W. Diffie, M. E. Hellman, *Multiuser cryptographic techniques*, Proceedings of National Computer Conference, AFIPS, vol. 45, June 1976, pp. 109—112.
19. W. Diffie, M. E. Hellman, *New directions in cryptography*, IEEE Transactions on Information Theory, vol. IT-22, No. 6, November 1976, pp. 644—654.
20. H. Feistel, *Cryptography and computer privacy*, Scientific American, vol. 228, No. 5, May 1973.
21. V. D. Gligor. D. Maierd, *Representing data bases in segmented name spaces*, Proceedings of the International Conference on Data Bases, Improving, Usability and Responsiveness, Haifa, Israel, August 1978.
22. E. J. Grenier, R. L. Winkler, *Data security: technology and the regulatory process*, Sutherland, Asbill and Brennon, Washington 1975.
23. E. Gudes, H. S. Koch, F. A. Stahl, *The application of cryptography for data base security*, National Computer Conference, AFIPS, 1976, pp. 97—107.
24. M. A. Harrison, W. L. Ruzzo, J. D. Ullman, *Protection in operating system*, Communications of the ACM, vol. 19, No. 8, August 1976, pp. 461—471.
25. I. Ingemarsson, *Negotiable documents in automatic data processing INTERNSKRIFT LiTH-ISY-I-0153*, Linkoping University, 1977.
26. N. R. Nielsen, B. Ruder, D. H. Brandin, *Efective safequards for computer system integrity*, Proceedings National Computer Conference, AFIPS Press, 1976, pp. 75—84.
27. N. R. Nielsen, B. Ruder, *Computer system integrity safequards*, Information Processing, IFIP, North-Holland Publishing Company, 1977, pp. 337—341.
28. N. R. Nielsen, B. Ruder, *Efective computer system integrity safequards*, Stanford Research Institute Report, 1977.
29. D. B. Parker, *Computer abuse assessment*, Stanford Research Institute Project 5068, December 1975.
30. D. B. Parker, *Computer abuse perpetrators and vulnerabilities of computer systems*, Proceedings of the National Computer Conference, AFIPS Press, 1976, pp. 65—73.
31. R. L. Patrick, *Privacy, people, and credit service*, Datamation, January 1974, pp. 48—50.
32. J. Pieprzyk, *Zastosowanie metod kryptograficznych do zabezpieczenia danych w systemach komputerowych*, Praca doktorska, Instytut Podstaw Informatyki PAN, Warszawa, 1979.
33. J. B. Rule, *Electronic funds transfer and federal privacy policy*, Jurimetrics Journal, Fall, vol. 18, No. 1, 1977, pp. 56—79.
34. T. A. Schabeck, *Computer crime investigation- part 1*, Assets Protection Journal, vol. 2, No. 2 1977, pp. 11—16.
35. T. A. Schabeck, *Computer crime investigation — part 2*, Assets Protection Journal, vol. 2, No. 3 1977, pp. 6—11.
36. T. A. Schabeck, *The social responsibility of business for the accuracy, privacy and security o computerized information*, Assets Protection Journal, vol. 3, No. 2, 1978.
37. M. E. Senko, *The DDL in the context of a multilevel srructured description: DIAM II with FORAL* North Holland Publishing Company, Amsterdam 1975.
38. M. E. Senko, *FORAL LP for DIAM II: FORAL with light pen — a language primer*, IBM Researcl Laboratory, Technical Report RC 6328, Yorktown Heights, November 1976.
39. M. E. Senko, *DIAM as a detailed example of the ANSI SPARC architecture*, Proceedings of th IFIP, January 1976.
40. M. E. Senko, *Data structures and data accessing in data base systems, past, present, future*, IBM Systems Journal, vol. 16, No. 3, 1977, pp. 208—257.
41. M. E. Senko, *FORAL LP: making pointed queries with a light pen*, IFIP Conference, 1977,

42. B. Shneiderman, *Design, development and utilization perspectives on database management systems*, Information Processing and Management, vol. 13, 1977, pp. 23—33.
43. L. Stern, *Contingency planning: why? how? and how much?*, Datamation, vol. 20, No. 9, September 1974, pp. 83—95.
44. R. Turn, *Privacy and security in personal information databank systems*, Rand Corporation Report, No. R-1044-NSF, 1974.
45. R. Turn, W. H. Ware, *Privacy and security in computer systems*, American Scientist, vol. 63, No. 2, March-April 1975, pp. 196—203.
46. R. W. Watson, *Timesharing system design concepts*, McGraw-Hill Book Company, New York 1970.
47. H. Weiss, *Computer security an overview*, Datamation, vol. 20, No. 1, January 1974, pp. 42—47.
48. H. Wienier, *Outwitting the forger*, Computer Weekly International, July 1976, pp. 18—19.
49. H. M. Wood, *On-line password techniques*, Proceedings of the Trends and Applications Computer Security and Integrity, IEEE Computer Society, May 1977.

J. PIEPRZYK, D. RUTKOWSKI

APPLICATION OF CRYPTOGRAPHY FOR INFORMATION PROTECTION IN DATA BASE SYSTEM

Summary

In this work we present the application of cryptographic transformations for the purpose of enhancing security in data base system. This problem is treated with the aid of a multi-level model of data base organisation. Particular attention is paid to the influence of semantic connections between data base elements on the cryptographic transformation application. The paper also gives rules of cryptographic transformation applications as to conditions which define the kind of protection elements before unautorized user. Then we estimate the break probability of cipher for a particular rule of cryptographic transformations on condition that semantic connections between data base element are known. At the end we present the application of cryptographic transformation to a particular data base.

J. PIEPRZYK, D. RUTKOWSKI

APPLICATION DES TRANSFORMATIONS CRYPTOGRAPHIQUES POUR LA PROTECTION DES INFORMATIONS DANS LES BASES DE DONNÉES

Résumé

Dans l'étude on a présenté l'application des transformations cryptographiques pour la protection des éléments de la base de données. Le problème a été résolu à l'aide de la construction du modèle à plusieurs étages de la base de données. On a porté surtout attention sur l'influence des relations sémantiques entre les éléments particuliers de la base de données sur la transformation cryptographique appliquée. On a donné aussi les régles de l'application des transformations cryptograpghiques pour les conditions posées d'avance, déterminant le genre des éléments de la base de données protégés contre l'action de l'usager non autorisé. Ensuite on a évalué la probabilité du déchiffrement des régles particulières de réalisation des transformations cryptographiques sous condition que les relations sémantiques, se produisant entre les éléments de la base de données, soient connues. Vers la fin de l'étude, on a donné l'exemple de l'application des transformations cryptographiques au chiffrement des éléments de la base de données.

J. PIEPRZYK, D. RUTKOWSKI

ANWENDUNG KRYPTOGRAPHISCHER UMWANDLUNGEN ZUM SCHUTZ VON INFORMATIONEN IN DER DATENBASIS

Zusammenfassung

Im vorliegenden Artikel wurde die Anwendung von kryptographischen Umwandlungen in Datenbasen dargestellt. Dieses Problem wurde anhand eines mehrebenen Modells für die Datenbasisorganisation gelöst. Besondere Aufmerksamkeit wurde dem Einfluß der semantischen Bindungen zwischen den einzelnen Elementen der Datenbasis auf die angewandten kryptographischen Umwandlungen verliehen. Es wurden auch die Regeln für die Anwendung von kryptographischen Umwandlungen unter vorausgegebenen Bedingungen angeführt, die die Art der vor Mißbrauch geschützten Elements der Datenbasis bestimmen. Im weiteren wurde die Wahrscheinlichkeit eines Chiffrebruchs für die einzelnen Realisationsregeln der kryptographischen Umwandlungen unterder Bedingung abgeschätzt, daß die semantischen Bindungen zwischen den Elementen der Datenbasis bekannt sind. Abschließend wurde die Anwendung der kryptographischen Umwandlung zur Chieffrierung von Elementen einer Datenbasis dargetsellt.

Ю. ПЕПШИК, Д. РУТКОВСКИ

ПРИМЕНЕНИЕ КРИПТОГРАФИЧЕСКИХ ПРЕОБРАЗОВАНИЙ К ОХРАНЕ ИНФОРМАЦИИ В БАЗАХ ДАННЫХ

Резюме

Представлено применение криптографических преобразований к охране элементов информации базы данных. Проблемма решена на основании модели организации базы данных со многими уровнями. Особое внимание обращено на влияние семантических связей отдельных элементов базы данных на применяемое криптографическое преобразование. Приведены также правила применения криптографических преобразований при предопределенных условиях определяющих класс элементов базы данных охраняемых перед действием неуполномоченного потребителя.

Проведена оценка вероятности рвсшифровки для отдельных правил реализации криптографических преобразований при условии известности семантических связей элементов базы данных. В заключение представлено применение криптографических преобразований к зашифровке элементов примерной базы данных.

ROZPRAWY ELEKTROTECHNICZNE 1981, 27, z. 2, ss. 923-930

621.327.8: 621.394.9

Zasady optymalizacji testowania jakości kanałów transmisji danych

ANDRZEJ CIBORSKI (WROCŁAW), WITOLD KAZIMIERCZAK (WROCŁAW)
JAN PORĘBA (WROCŁAW)

Instytut Telekomunikacji i Akustyki Politechniki Wrocławskiej

Otrzymano 28.12.1979

W pracy omówiono zasady wyboru optymalnych metod oceny jakości kanałów transmisji danych. W tym celu zaproponowano czterostanowy proces semi-Markowa jako model kanału. Na podstawie tego modelu sformułowano i przeanalizowano kryteria wyboru metod oceny jakości kanałów.

1. WPROWADZENIE

Jednym ze sposobów zwiększenia efektywności działania systemów transmisji informacji cyfrowych jest optymalizacja testowania jakości kanałów. W ostatnich latach pojawiło się wiele prac dotyczących konkretnych metod oceny jakości kanałów. Metody te opierają się na pomiarach parametrów pierwotnych: tłumienności [1], zakłóceń impulsowych i krótkich przerw transmisji [2], szumów, itp. oraz parametrów wtórnych: stopy błędów [3, 4], zniekształceń telegraficznych [5, 6], zniekształceń amplitudy sygnału w różnych punktach kanału [7, 8], itp. Opracowano także metody, w których ocena jakości kanałów dokonywana jest na podstawie wyników pomiarów kilku parametrów, np. metody statystyczne zaproponowane w [9].

Duża liczba metod oceny jakości kanałów transmisji danych utrudnia wybór najlepszej metody. Dlatego też ważnym, z punktu widzenia optymalizacji systemów transmisji danych, jest wyznaczenie kryteriów, które ułatwiłyby ten wybór [10, 11]. Ma to istotne znaczenie, zarówno podczas projektowania, jak i udoskonalania systemów transmisji informacji cyfrowych.

W pracy zaproponowano i przeanalizowano kryteria wyboru metod oceny jakości kanałów transmisji danych dla przyjętego modelu stanów kanału. Określono także parametry, których znajomość jest potrzebna do efektywnego stosowania tych kryteriów.

2. PARAMETRY JAKOŚCI KANAŁU TRANSMISJI DANYCH

Podstawowym parametrem określającym jakość kanału transmisji danych jest prawdopodobieństwo przekłamania elementarnej informacji P_b. Inne parametry, takie jak ilość informacji o sygnale wyjściowym zawarta w sygnale wejściowym, czy też przepustowość

kanału, mimo że dokładniej opisują jakość kanału, są — w tym przypadku — niewygodne ze względu na trudności ich wyznaczania [12].

Dodatkową zaletą parametru P_b jest to, że zbadano jego związek z niektórymi parametrami pierwotnymi [13÷15] oraz wtórnymi [6, 7] kanału transmisji danych. Znana jest zależność parametru P_b od stosunku mocy sygnału do mocy szumu gaussowskiego [13] oraz od parametrów zakłóceń impulsowych [14, 15], występujących w kanale. W pracach [16, 17, 7] podano sposób wyznaczenia parametru P_b na podstawie pomiaru zniekształceń amplitudy sygnału na wyjściu dyskryminatora, a w pracy [6] określono także jego zależność od zniekształceń telegraficznych sygnału danych.

Badania częstości występowania błędów przeprowadzone dla różnych typów kanałów transmisji danych, wykazały, że parametr P_b zmienia się w czasie w sposób losowy [18÷20]; można go zatem traktować jako proces stochastyczny $P_b(t)$. W celu uniknięcia trudności identyfikacji tego procesu, przyporządkujmy mu dwustanowy proces stochastyczny $Z(t)$, który można określić jako odpowiedź na pobudzenie $P_b(t)$ układu o charakterystyce

$$F(x) = \begin{cases} s_o, & \text{jeżeli} \quad x < P_b^*, \\ s_1, & \text{jeżeli} \quad x \geqslant P_b^*, \end{cases} \tag{1}$$

gdzie P_b^* jest progową wartością prawdopodobieństwa błędu.

Stany s_o i s_1, w których znajduje się proces $Z(t)$, można utożsamiać odpowiednio ze stanami poprawnego działania i niesprawności kanału, pamiętając jednak o ich zależności od P_b^*.

Wygodnym modelem procesu $Z(t)$ jest proces odnowy, scharakteryzowany rozkładami prawdopodobieństw czasu przebywania w stanach s_o i s_1. Postać funkcyjną tych rozkładów analizowano w [21, 22] oraz w niektórych pracach dotyczących modeli błędów [23, 24].

3. MODEL STANÓW KANAŁU TRANSMISJI DANYCH

Przyjmując, że kryterium niesprawności jest określone zależnością (1), wszystkie metody oceny jakości kanałów transmisji danych można traktować jako testowanie hipotezy statystycznej o wartości parametru P_b. Podejmowanie decyzji na podstawie wnioskowania statystycznego jest obarczone tak zwanymi błędami pierwszego i drugiego rodzaju. W naszym przypadku błąd pierwszego rodzaju oznacza sygnalizację, że kanał jest w stanie s_1, mimo że w rzeczywistości znajduje się on w stanie s_o — oznaczmy taki stan przez s_{o1}. Natomiast stan odpowiadający błędowi drugiego rodzaju — oznaczmy go przez s_{1o} — polega na braku sygnalizacji stanu s_1. Pozostałe dwa stany, odpowiadające prawidłowej sygnalizacji, oznaczmy przez s_{oo} i s_{11}.

Niech $Z_1(t)$ będzie procesem stochastycznym o stanach s_{oo}, s_{o1}, s_{1o} i s_{11}. Wygodnym modelem tego procesu jest proces semi-Markowa [23, 25], opisany rozkładami prawdopodobieństw czasu przebywania w poszczególnych stanach i macierzą prawdopodobieństw przejść

$$M_p = \begin{bmatrix} p(s_{oo}/s_{oo}), & p(s_{o1}/s_{oo}), & p(s_{1o}/s_{oo}), & p(s_{11}/s_{oo}) \\ p(s_{oo}/s_{o1}), & p(s_{o1}/s_{o1}), & p(s_{1o}/s_{o1}), & p(s_{11}/s_{o1}) \\ p(s_{oo}/s_{1o}), & p(s_{o1}/s_{1o}), & p(s_{1o}/s_{1o}), & p(s_{11}/s_{1o}) \\ p(s_{oo}/s_{11}), & p(s_{o1}/s_{11}), & p(s_{1o}/s_{11}), & p(s_{11}/s_{11}) \end{bmatrix}, \tag{2}$$

gdzie $p(s_{kl}/s_{mn})$, $k, l, m, n = 0, 1$ są prawdopodobieństwami przejścia procesu ze stanów s_{mn} do stanów s_{kl}.

Przejścia między niektórymi stanami procesu $Z_1(t)$ nie występują w ogóle, bądź występują bardzo rzadko w porównaniu z innymi przejściami. Dlatego też

$$p(s_{oo}/s_{oo}) = p(s_{o1}/s_{o1}) = p(s_{1o}/s_{1o}) = p(s_{11}/s_{11}) = 0,$$

a także można przyjąć, że:

$$p(s_{1o}/s_{o1}) = p(s_{11}/s_{o1}) = p(s_{o1}/s_{1o}) = p(s_{o1}/s_{11}) = p(s_{1o}/s_{11}) = 0,$$

więc:

$$p(s_{oo}/s_{o1}) = p(s_{oo}/s_{11}) = 1,$$

ponieważ

$$\sum_{k=0}^{1}\sum_{l=0}^{1} p(s_{kl}/s_{mn}) = 1. \tag{3}$$

Oznaczmy przez P_{kl}, $k, l = 0, 1$, prawdopodobieństwa stacjonarne łańcucha Markowa włożonego w proces $Z_1(t)$. Prawdopodobieństwa te można łatwo wyznaczyć, znając macierz prawdopodobieństw przejść, z układu równań:

$$\left.\begin{aligned} P_{kl} &= \sum_{m=0}^{1}\sum_{n=0}^{1} p(s_{kl}/s_{mn})P_{mn}, \\ \sum_{m=0}^{1}\sum_{n=0}^{1} P_{mn} &= 1. \end{aligned}\right\} \tag{4}$$

Rozwiązaniem tego układu równań są przy powyższych uproszczeniach następujące wyrażenia na prawdopodobieństwa P_{kl}, dla $k, l = 0, 1$:

$$\left.\begin{aligned} P_{oo} &= \frac{1}{2+p(s_{11}/s_{1o})p(s_{1o}/s_{oo})}, \\ P_{o1} &= \frac{p(s_{o1}/s_{oo})}{2+p(s_{11}/s_{1o})p(s_{1o}/s_{oo})}, \\ P_{1o} &= \frac{p(s_{1o}/s_{oo})}{2+p(s_{11}/s_{1o})p(s_{1o}/s_{oo})}, \\ P_{11} &= \frac{p(s_{11}/s_{oo})+p(s_{11}/s_{1o})p(s_{1o}/s_{oo})}{2+p(s_{11}/s_{1o})p(_{1o}/s_{oo})}. \end{aligned}\right\} \tag{5}$$

Oznaczmy przez $f_{kl}(t)$, $k, l = 0, 1$ funkcje gęstości rozkładów prawdopodobieństw zasów przebywania procesu $Z_1(t)$ w stanach s_{kl}, $k, l = 0, 1$, oraz przez t_{kl}, $k, l = 0, 1$, dpowiadające im wartości średnie. Prawdopodobieństwa Π_{kl}, $k, l = 0, 1$, przebywania ego procesu w poszczególnych stanach można wyrazić zależnościami:

$$\Pi_{kl} = \frac{P_{kl}t_{kl}}{\sum_{m=0}^{1}\sum_{n=0}^{1} P_{mn}t_{mn}}, \quad k, l = 0, 1. \tag{6}$$

Podstawiając zależności (5) do (6), otrzymuje się prawdopodobieństwa Π_{kl}, $k, l = 0, 1$, jako funkcje prawdopodobieństw przejść.

4. KRYTERIA WYBORU METOD OCENY JAKOŚCI POJEDYNCZEGO KANAŁU TRANSMISJI DANYCH

Duża liczba metod oceny jakości kanału transmisji danych wymaga dokonania wyboru metody najlepszej. Wybór ten jest możliwy tylko w sensie konkretnego kryterium.

Najprostsze kryteria opierają się na analizie jednego parametru określającego jakość kanału. Jeżeli oznaczymy metody testowania przez x_i, $i = 1, 2, \ldots, n$, to jedno z tych kryteriów ma postać

$$K_1(x_i) = \Pi_{oo}^i, \tag{7}$$

gdzie Π_{oo}^i jest wartością parametru Π_{oo} dla metody x_i (w dalszym ciągu indeksem i będą oznaczane wielkości związane z metodą x_i).

W pracy [25] uściślono powyższe kryterium, wprowadzając parametr Q_i^i, który jest prawdopodobieństwem przebywania kanału w stanie s_{oo} i testowania jego jakości, proponując jako kryterium

$$K_2(x_i) = \Pi_{oo}^i - Q^i. \tag{8}$$

Istotną wadą przedstawionych kryteriów jest brak uwzględnienia prawdopodobieństw przebywania kanału w pozostałych stanach: s_{01}, s_{10} i s_{11} oraz strat z tym związanych. Straty te można uwzględnić przyjmując jako kryterium funkcję ryzyka [26].

Jeżeli oznaczymy przez $C_{kl}(t)$, $k, l = 0, 1$, straty[1] związane z przebywaniem kanału odpowiednio w stanach s_{kl}, $k, l = 0, 1$, w czasie t, to średnie straty, przypadające na jednokrotne przebywanie kanału w poszczególnych stanach, wyrażają się odpowiednio zależnościami

$$\bar{C}_{kl}^i = \int_0^\infty C_{kl}(t) f_{kl}^i(t)\,dt, \quad k, l = 0, 1, \tag{9}$$

a kryterium oparte na funkcji ryzyka przyjmuje postać

$$K_3(x_i) = \sum_{k=0}^{1}\sum_{l=0}^{1} P_{kl}^i \bar{C}_{kl}^i. \tag{10}$$

W przypadku gdy $C_{kl}(t) = C_{kl}t$, $k, l = 0, 1$, kryterium to upraszcza się do wyrażenia

$$K_3'(x_i) = \sum_{k=0}^{1}\sum_{l=0}^{1} \Pi_{kl}^i C_{kl}. \tag{11}$$

[1] Straty $C_{oo}(t)$ wynikają z zajętości kanału podczas testowania jego jakości w stanie s_{oo}.

Kryterium (10) nie uwzględnia kosztów opracowania, wdrożenia i utrzymania aparatury testującej jakość kanału daną metodą. Jeżeli koszty te, przeliczone na jednostkę czasu, oznaczymy przez C^i, to można następująco zmodyfikować kryterium (10)

$$K_4(x_i) = C^i + K_3(x_i). \tag{12}$$

Tak sformułowane kryterium uwzględnia wszystkie istotne parametry wpływające na jakość kanału, dlatego też na jego podstawie powinna być wybierana najlepsza metoda oceny tej jakości.

5. KRYTERIUM WYBORU METOD OCENY JAKOŚCI KANAŁÓW W SIECI TRANSMISJI DANYCH

Kryterium (12) pozwala na wybór optymalnej metody oceny jakości (tzn. takiej, dla której funkcja $K_4(x_i)$ osiąga minimum) dla pojedynczego kanału. Często zachodzi konieczność wyboru optymalnego zbioru metod oceny jakości dla wszystkich kanałów w sieci transmisji danych. Niżej przedstawiono kryterium umożliwiające taki wybór.

Rozbijmy zbiór wszystkich kanałów na klasy o zbliżonych wartościach parametrów użytkowych (szybkość transmisji, rodzaj modulacji, typ i przeznaczenie systemu informatycznego, w którym kanał pracuje, itp.). Oznaczmy indeksem j parametry kanałów j-tej klasy opisane zależnościami (5) i (9). Ryzyko dla kanału z tej klasy, po zastosowaniu do oceny jakości metody x_i, określa zależność

$$K_4^j(x_i) = C^{ij} + \sum_{k=0}^{1} \sum_{l=0}^{1} P_{kl}^{ij} \bar{C}_{kl}^{ij}, \tag{13}$$

gdzie C^{ij} jest wartością parametru C^i dla kanału j-tej klasy.

Oznaczmy przez x_{k_j}, $k_j \in \{1, 2, \ldots, n\}$, metodę, którą testujemy j-tą klasę. Średnie ryzyko dla kanału z sieci (uśrednione po wszystkich klasach) wyraża się zależnością

$$K(x_{k_1}, x_{k_2}, \ldots, x_{k_M}) = \sum_{j=1}^{M} \frac{l_j}{N} K_4^j(x_{k_j}), \tag{14}$$

gdzie l_j jest liczbą kanałów w j-tej klasie, M — liczbą klas kanałów, a N — liczbą kanałów w sieci.

Wyznaczając minimum funkcji $K(x_{k_1}, x_{k_2}, \ldots, x_{k_M})$, określa się optymalny zbiór metod oceny jakości wszystkich kanałów w sieci transmisji danych.

6. PODSUMOWANIE

W pracy zaproponowano sposób zobiektywizowania wyboru metod oceny jakości kanałów transmisji danych. W tym celu przyjęto czterostanowy proces semi-Markowa jako model kanału. Proces ten jest opisany rozkładami czasu przebywania kanału w poszczególnych stanach i macierzą prawdopodobieństw przejść.

Na podstawie zaproponowanego modelu sformułowano i przeanalizowano kryteria wyboru metod oceny jakości kanałów transmisji danych. Kryteria te pozwalają na wybór

metody optymalnej, pod warunkiem wyznaczenia parametrów modelu stanów kanału oraz strat związanych z przebywaniem kanału w poszczególnych stanach.

Parametry modelu stanów są zdetermimowane rodzajem kanału i metodą jego testowania. Mogą one być określone poprzez bezpośrednie pomiary lub symulację kanału. Zachodzi jednak potrzeba opracowania szczegółowej metodyki ich wyznaczania.

Straty związane z przebywaniem kanału w poszczególnych stanach są uzależnione od przeznaczenia systemu informatycznego, w którym kanał pracuje i mogą być określone przez użytkownika systemu.

Znając powyższe parametry można, na podstawie kryteriów zaproponowanych w pracy, określić najlepsze metody testowania jakości kanałów, co stanowi jeden z aspektów optymalizacji systemów transmisji danych.

LITERATURA

1. J. L. Holsinger i inni, *Line Level Signal Measuring Device*, United States Patent No. 4 055 733, 1977.
2. V. A. Bovtruk, V. A. Ponomarenko, *Ustrojstvo kontrolja kanalov svjazi*, Opisanie izobretenija k avtorskomu svidetel'stvu, No. 499 682, 1976.
3. L. C. Schooley, T. W. Mc Curnin, *Measurement of Bit Errors in Data Transmission System*, IEEE Transactions on Communications, vol. Com-24, No. 1, 1976, s. 144—146.
4. E. N. Dulaney i inni, *Apparatus for Continuous Assessment of Data Transmission Accuracy in a Communication System*, United States Patent No. 3 934 224, 1976.
5. Ju. I. Piskun, Ju. I. Popše, *Ustrojstvo izmerenija iskaženij startstopnych signalov*, Opisanie izobretenija k avtorskomu svidetel'stvu, No. 577 688, 1977.
6. V. E. Buchviner, *Ocenka kačestva radiosvjazi*, Svjaz', Moskva 1974.
7. B. J. Leon i inni, *A Bit Error Rate Monitor for Digital PSK Links*, IEEE Transactions on Communications, vol. Com-23, No. 5, 1975, s. 518—525.
8. V. A. Kulagin, S. A. Vorob'ev, *Ustrojstvo ocenki kačestva diskretnych signalov*, Opisanie izobretenija k avtorskomu svidetel'stvu, No. 564 725, 1977.
9. S. B. Guthery, A. C. Salazar, *Statistical Methods for Testing Quality of Communication Channels*, IEEE Transactions on Communications, vol. Com-23, No. 6, 1975, s. 670—675.
10. A. N. Zelinger, *Kriterii ocenki kačestva sistem svjazi*, Svjaz', Moskva 1974.
11. Ju. B. Okunev, V. G. Plotnikov, *Principy sistemnogo podchoda k projektirovaniju v technike svjazi*, Svjaz', Moskva 1976.
12. J. Seidler, *Systemy przesyłania informacji cyfrowych*, WNT, Warszawa 1976.
13. L. Knoch, T. Ekiert, *Modulacja i detekcja*, WKiŁ, Warszawa 1979.
14. S. P. Vol'fbejn, N. G. Veksler, *Pomechi pri peredače diskretnoj informacii*, Technika, Kiev 1973.
15. J. Lubacz, *Pewne związki analityczne parametrów szumu impulsowego z prawdopodobieństwem błędu w transmisji cyfrowej*, Rozprawy Elektrotechniczne, t. XXI. z. 1, 1975, s. 153—166.
16. D. J. Gooding, *Performance Monitor Techniques for Digital Receivers Based on Extrapolation of Error Rate*, IEEE Transactions on Communication Technology, vol. Com-16, No. 3, 1968, s. 380—387.
17. S. B. Weinstein, *Estimation of Small Probabilities by Linearization of the Tail a Probability Distribution Function*, IEEE Transactions on Communication Technology, vol. Com-19, No. 6, 1971, s. 1149—1155.
18. S. I. Samojlenko (red.), *Statistika ošibok pri peredače cifrovoj informacii*, Sbornik perevodov, Mir, Moskva 1966.

19. T. Nowicki, M. Bereziński, *Analiza rozkładu błędów w systemach transmisji danych*, Prace Instytutu Automatyki PAN, zeszyt 88, Warszawa 1970.
20. L. P. Purtov i inni, *Èlementy teorii peredači diskretnoj informacii*, Svjaz', Moskva 1972.
21. V. O. Švarcman, G. D. Michalev, *Rasčet nadežnostnych charakteristik traktov peredači dannych*, Svjaz', Moskva 1975.
22. V. P. Altarev, G. I. Šakun, P. I. Trofimov, *Processy otkazov i vosstanovlenij v sistemach PD*, Svjaz', Moskva 1977.
23. V. Ja. Turin, *Peredača informacii po kanalam s pamjat'ju*, Svjaz', vyp. 8, Moskva 1977.
24. L. N. Kanal, A. R. K. Sastry, *Models for Channels with Memory and Their Applications to Error Control*, Proceedings of the IEEE, vol. 66, No. 7, 1978, s. 724—744.
25. Ju. A. Lejbman, A. F. Levčenko, *Strategii prinjatija rešenij pri organizacii kontrolja kanalov PD*, Élektrosvjaz', No. 12, 1978, s. 52—54.
26. A. P. Sage, I. L. Melse, *Estimation Theory with Application to Communication and Control*, Mc Graw-Hill, New York 1972.

A. CIBORSKI, W. KAZIMIERCZAK, J. PORĘBA

PRINCIPLES OF OPTIMIZATION OF DATA TRANSMISSION CHANNELS QUALITY TESTING

Summary

In the paper, the selection principles of the most favorable methods of data transmission channels quality assessment are presented. For this purpose, the four-state Markov renewal process as a model of channel is proposed. The criteria of the methods selection of channels quality assessment for this model are formulated and analysed.

A. CIBORSKI, W. KAZIMIERCZAK, J. PORĘBA

PRINCIPES D'OPTIMISATION DU CONTRÔLE DE LA QUALITÉ DES VOIES DE TRANSMISSION DES DONNÉES

Résumé

Dans cet article on décrit les principes décidant du choix des modes optimaux de l'évaluation de la qualité des voies de transmission des données. Conformément à cela, on propose comme modèle de voie un processus renové de Markov à quatre états. En se basant sur ce modèle, on formule et on analyse les critères de l'évaluation de la qualité des voies.

A. CIBORSKI, W. KAZIMIERCZAK, J. PORĘBA

OPTIMALISATIONSPRINZIPIEN DER QUALITÄTSPRÜFUNG VON DATENÜBERTRAGUNGSKANÄLEN

Zusammenfassung

Es wurden Wählprinzipien für optimale Methoden zur Qualitätsbeurteilung der Datenübertragungskanäle angegeben. Zu diesem Zweck wurde als Kanalmodell ein 4-ständiger Markov-Erneuerungsprozeß vorgeschlagen. Auf Grund dieses Modells sind Wählkriterien für Beurteilungsmethoden der Kanalqualität bestimmt worden, und einer Analyse unterzogen.

А. ЦИБОРСКИ, В. КАЗИМЕРЧАК, Я. ПОРЕМБА

ПРИНЦИПЫ ОПТИМАЛИЗАЦИИ КОНТРОЛЯ КАЧЕСТВА КАНАЛОВ ПЕРЕДАЧИ ДАННЫХ

Р е з ю м е

Представлены принципы выбора оптимальных методов оценки качества каналов передачи данных. В качестве модели канала использован четырёхсостоянный полумарковский процесс. На основе этой модели представлены и анализированы критерии выбора методов оценки качества каналов.

TREŚĆ

A. L. Dobrucki, M. S. Piekarski: Synteza macierzy rozproszenia układów bezstratnych metodami przestrzeni stanów 619
M. Krakowski, G. Szymański, A. Ziętkowiak: Potential Distribution Along Underground Conductors Connected to Station Earth Electrode 633
M. Siwczyński, L. Topór-Kamiński: Analiza i realizacja aktywnych obwodów z okresowo zmiennymi parametrami 643
K. Adamiak: Odwrotne zagadnienia brzegowe typu Dirichleta dla równania Laplace'a . . 657
K. Miączyński: Algorytm sterowania dyskretnego układem Leonarda 671
J. Sykulski: Analysis of Transformer Tank Shielding Effectivenees 687
K. Kluszczyński: Składowe aktywne i zerowe prądów w obwodach elektromagnetycznych 697
C. Królikowski, A. Kamińska-Pranke, R. Niewiedział: Próba opisu charakterystyk napięciowo-prądowych wyładowania łukowego stabilizowanego strugą gazu przy wykorzystaniu równań magnetohydrodynamiki 713
A. Boubakeur: Influences des barrières sur l'amorçage de moyens intervalles d'air pointe-plan 729
B. Miedziński, H. Kieloch, T. Żołnierczyk: Zachowanie się kontaktronów gazowanych w środowisku ciekłego azotu 745
P. Dymarski: Metody poprawiania stabilności algorytmów numerycznego rozwiązywania sztywnych układów równań różniczkowo-algebraicznych 759
M. Kazimierczuk: Analiza, projektowanie i badania eksperymentalne wzmacniacza klasy E w optymalnych warunkach pracy 795
A. Kączkowski, A. Milewski: Metody wyznaczania zespolonej pulsacji rezonansowej współosiowego rezonatora ze skupioną pojemnością niejednorodnie wypełnionego dielektrykiem 823
J. Lubacz, W. Burakowski: Multidimensional Birth-Death Processes in the Analysis of Multi-User-Class Systems 837
J. Lubacz, W. Burakowski, M. Pióro: Traffic Analysis of a Data Transmission Channel Concentrator 847
A. Badach, R. Sobczak, M. Zientalski: Zasady tworzenia sieci programów sterujących przesyłaniem informacji cyfrowych w systemach telekomunikacyjnych 863
J. Pieprzyk, D. Rutkowski: Zastosowanie przekształceń kryptograficznych do ochrony informacji w bazach danych 885
A. Ciborski, W. Kazimierczak, J. Poręba: Zasady optymalizacji testowania jakości kanałów transmisji danych 923

CONTENTS — TABLE DES MATIÈRES — INHALT

A. L. Dobrucki, M. S. Piekarski: A Scattering Matrix Synthesis of Lossless Networks Via State-Space Techniques 631
Synthèse de scattering matrice des circuits sans dissipation dans l'espace d'états 632
Synthese der Streumatrix verlustfreier Schaltungen mit Hilfe von Zustandsraummethoden 632
M. Krakowski, G. Szymański, A. Ziętkowiak: Rozkład potencjału wzdłuż przewodów podziemnych dołączonych do uziomu stacyjnego 641
Répartition du potential le long des conducteurs enterrés raccordés à la prise de terre de la station 641

Potentialverteilung in erdverlegten, an Stationserder angeschalten Leitungen 641
M. Siwczyński, L. Topór-Kamiński: Analysis and Realization of Linear Periodically Time-Varying Networks 654
Analyse et simulation des circuits actifs à paramètres périodiquement variables 655
Analyse und Realisation linearer Kreise mit zeitveränderlichen Parametern 655
K. Adamiak: Reverse Edge Problems of Dirichlet Type for Laplace Equation 668
Problèmes des limites inverses du type de Dirichlet pour l'équation de Laplace 668
Reziprokes Randproblem Dirichlet-Typs für Laplace-Gleichungen 668
K. Miączyński: Sample Data Algorithm for Leonhard Drive System 686
Algorithme de la commande directe du système de Leonard 686
Algorithmus der Abtaststeuerung eines Leonard-Antriebes 686
J. Sykulski: Analiza skuteczności ekranowania kadzi transformatorów 695
Analyse de l'efficacité du blindage de la cuve des transformateurs 695
Zur Analyse der Effektivität der Abschirmung eines Transformator-Kessels 695
K. Kluszczyński: Active and Zero Components in Electro-Magnetic Circuits 711
Composantes actives et neutres des courants dans les circuits électromagnétiques 711
Aktiv- und Nullkomponenten der elektromagnetischen Stromkreise 712
C. Królikowski, A. Kamińska-Pranke, R. Niewiedział: An Attempt to Describe Voltage-Current Characteristics of an Arc with Gas Jet Stabilisation When Utilizing Equations of Magnetohydrodynamics 726
Essai pour déterminer les caractéristiques tension-courant de l'arc électrique stabilisé par le gaz de travail en utilisant les équations magnétohydrodynamiques 726
Analytischer Bestimmungsversuch von Strom-Spannungs-Kennlinien des mit Arbeitsgas stabilisierten Lichtbogens unter Ausnutzung magnetohydrodynamischer Gleichungen . . . 726
A. Boubakeur: The Influence of Barriers on Dielectric Strength of Mean Air Gaps in Point-Plane Arrangements 743
Wpływ przegród na wytrzymałość elektryczną powietrza dla średnich odstępów układu „ostrze-płyta" 743
Einfluss von Trennwänden auf die elektrische Luftfestigkeit bei mittleren Abständen der Anordnung „Spitze-Platte" 744
B. Miedziński, H. Kieloch, T. Żołnierczyk: Behavior of Gas Filled Reeds in Liquid Nitrogen Environment 756
Comportement des ILS à gaz dans le milieu d'azote liquide 757
Eigenschaften der Schutzrohrkontakte im Medium flüssigen Stickstoffes 757
P. Dymarski: Methods for Improving the Stability of an Algorithm for Solution of Stiff Differential-Algebraic Systems 792
Méthodes d'amélioration de la stabilité d'un algorithme destiné à résoudre les systèmes d'équations algébriques différentielles 792
Methoden für die Stabilitätsverbesserung des Verfahrens zur Lösung steifer Systeme von Algebraisch-Differentialgleichungen 792
M. Kazimierczuk: Analysis, Design and Experimental Results of the *E* Class Amplifier for Optimum Operation 820
Analyse, projection et essais experimentaux de l'amplificateur de la classe *E* dans les conditions de travail optimales 820
Analyse, Entwerfen und experimentelle Untersuchungen eines *E*-Klasse-Verstärkers unter optimalen Arbeitsbedingungen 820
A. Kączkowski, A. Milewski: Methods of a Complex Resonant Frequency Determination of Re-Entrent Resonators with Dielectric Sample of Optional Diameter 835
Méthodes de détermination de la pulsation de résonance complexe du résonateur coaxial à capacité recueillie du diélectrique rempli non hétérogéniquement 835
Bestimmungsmethoden für die Resonanzkomplexpulsation eines koaxialen, mit dielektrischem Körper inhomogen gefüllten Resonators mit aggregierter Kapazität 836

J. Lubacz, W. Burakowski: Wielowymiarowe procesy urodzin i śmierci w analizie systemów obsługi z wieloma klasami użytkowników 845
Processus polydimensionnels naissance/mort à l'analyse des systèmes de service des usagers partagés en classes 845
Mehrdimensionale Geburts- und Todesprozesse bei der Analyse der Bediennugssysteme mit mehreren Teilnehmerklassen 846
J. Lubacz, W. Burakowski, M. Pióro: Analiza ruchowa koncentratora kanałów teleinformatycznych 861
Analyse du traffic dans la voie de transmission des données par concentrateur 861
Verkehrsanalyse eines Konzentrators fernmeldeinformativer Kanäle 861
A. Badach, R. Sobczak, M. Zientalski: Principles of Setting up Networks of Computer Programs for Control of Digital Information Transmission in Telecommunications Systems 882
Principes de la formation du réseau des programmes commandant la transmission des informations numériques dans les systèmes de télécommunications 882
Prinzipien der Netzbildung von Programmen zur Steuerung der Übertragung von digitalen Informationen in Nachrichtensystemen 883
J. Pieprzyk, D. Rutkowski: Application of Cryptography for Information Protection in Data Base System 921
Application des transformations cryptographiques pour la protection des informations dans les bases de données 921
Anwendung kryptographischer Umwandlungen zum Schutz von Informationen in der Datenbasis 922
A. Ciborski, W. Kazimierczak, J. Poręba: Principles of Optimization of Data Transmission Channels Quality Testing 929
Principes d'optimisation du contrôle de la qualite des voies de transmission des données . . 929
Optimalisationsprinzipien der Qualitätsprüfung von Datenübertragungskanälen 929

СОДЕРЖАНИЕ

А. Л. Добруцки, М. С. Пекарски: Синтез матрицы цепей без потерь методами пространства состояний 632
М. Краковски, Г. Шиманьски, А. Зенткович: Распределение потенциала в подземных проводах присоединенных к заземлению станции 641
М. Сивчыньски, Л. Топур-Каминьски: Анализ и реализация линейных цепей с периодически изменяющимися параметрами 655
К. Адамяк: Обратная краевая задача типа Дирихле для уравнения Лапласса . . . 669
Я. Сыкульски: Анализ эффективности экранирования баков трансформаторов . 686
К. Клющиньски: Активные и нулевые составляющие токов в электромагнитных цепях 695
Ч. Круликовски, А. Каминьска-Пранке, Р. Неведзял: Попытка описания вольтамперных характеристик дуги стабилизированной потоком газа на основании уравнений магнитогидродинамики 712
А. Боубакэур: Влияние барьеров на напряжение перекрытия воздушного промежутка для средних расстояний острие-плоскость 727
Б. Медзински, Х. Кельох, Т. Жолнерчик: Поведение газовых герконов в среде жидкого азота 744
П. Дымарски: Методы улучшения устойчивости алгоритма решения жёстких систем алгебро-дифференциальных уравнений 758
М. Казимерчук: Анализ, проектирование и экспериментальные исследования усилителя класса Е в оптимальном режиме 793
821

А. Кончкрвски, А. Милевски: Методы определения комплексной резонансной частоты соосного резонатора с сосредоточенной емкостью с неоднородным диэлектриком 836
Ю. Любач, В. Бураковски: Многомерные процессы рождения и гибели в анализе систем массового обслуживания со многими классами пользователей. 846
Ю. Любач, В. Бураковски, М. Пюро: Анализ трафика в концентраторе каналов передачи данных 862
А. Бадах, Р. Собчак, М. Зентальски: Принципы построения сетей программ управляющих передачей дискретных сообщений 883
Ю. Пепшик, Д. Рутковски: Применение криптографических преобразований к охране информации в базах данных 922
А. Циборски, В. Казимерчак, Я. Поремба: Принципы оптимализации контроля качества каналов передачи данных 930

Wytyczne dla Autorów

Komitet Redakcyjny prosi o przestrzeganie następujących wytycznych przy opracowywaniu maszynopisów artykułów nadsyłanych do opublikowania.

1. *Tematyka, charakter i oryginalność artykułów.* „Rozprawy Elektrotechniczne" są przeznaczone do ogłaszania prac naukowych wchodzących w zakres szeroko pojętej elektrotechniki, zwłaszcza jeżeli traktują one omawiane zagadnienia w sposób możliwie kompletny i zmierzają do osiągnięcia celu informacyjno-dydaktycznego. W związku z tym publikowane prace mogą być w mniejszym lub większym stopniu pracami nieorygnalnymi, jeśli chodzi o treść w nich zawartą, co nie zwalnia Autorów od obowiązku ujęcia tematu w sposób oryginalny. Twórczy wkład Autora może polegać wówczas np. na: porównaniu opisywanych teorii, metod lub systemów, krytycznej ich ocenie, wyciągnięciu własnych wniosków co do celowości takiego lub innego działania, zastosowaniu własnej klasyfikacji zagadnień, prognostyce itp. W szczególności „Rozprawy" ogłaszają takie prace naukowe, które zawierają: analizę porównawczą teorii, metod lub systemów, syntetyczne ujęcia określonego zagadnienia naukowego, omówienie obecnego stanu wybranego działu elektrotechniki, omówienie jego postępów lub tendencji rozwojowych itp.

2. *Wymagania ogólne.* Artykuły należy nadsyłać w maszynopisie, w dwóch egzemplarzach. Maszynopis powinien być napisany jednostronnie przez czarną taśmę, na maszynie do pisania z niezabrudzonymi i nieuszkodzonymi znakami. Dopuszcza się odręczne czytelne uzupełnienie lub pisanie atramentem albo długopisem kolorem czarnym lub ciemnoniebieskim znaków specjalnych oraz znaków w językach, których alfabetów nie ma na maszynach do pisania, np. znaków matematycznych, chemicznych, liter greckich. Maszynopis powinien być napisany na papierze do maszyn do pisania koloru białego, formatu A-4; numeracja ciągła na wszystkich kartkach. W artykułach należy stosować Międzynarodowy Układ Jednostek SI.

3. *Sposób pisania tekstu.* Tekst w maszynopisie powinien być pisany bez używania wyróżnień, a w szczególności nie dopuszcza się spacjowania (rozstrzelania), podkreślania i pisania tekstów dużymi literami. Proponowane wyróżnienia Autor może wprowadzić do maszynopisu (zwykłym ołówkiem) za pomocą przyjętych znaków adiustacyjnych (podkreślenie linią przerywaną — spacjowanie, podkreślenie linią ciągłą — pogrubienie, podkreślenie wężykiem — kursywa itp.). Na jednej kartce maszynopisu powinno być 30 wierszy po około 60 znaków łącznie z odstępami. Marginesy każdej kartki powinny mieć następujące wymiary: górny — ok. 25 mm, lewy — ok. 35 mm. Tekst maszynopisu powinien być napisany z podwójnym odstępem między wierszami; tytuły i podtytuły małymi literami. Akapity należy rozpoczynać wcięciem równym trzem uderzeniom maszyny do pisania.

4. *Sposób pisania tablic.* Tablice powinny być napisane na oddzielnych arkuszach, w układzie zbliżonym do układu zecerskiego. Tytuły rubryk pionowych i poziomych powinny być napisane małymi literami z podwójnym odstępem między wierszami. Przypisy (notki) dotyczące tablic należy pisać bezpośrednio pod tablicą. Tablice należy numerować kolejno liczbami arabskimi; u góry każdej tablicy podać tytuł.

5. *Sposób pisania wzorów matematycznych.* Rozmieszczenie znaków, cyfr, liter i odstępów powinno być zbliżone do rozmieszczenia elementów druku. Wskaźniki i wykładniki powinny być napisane wyraźnie i być prawidłowo obniżone lub podwyższone w stosunku do linii wiersza podstawowego. Znaki nad literami i cyframi: strzałki, linie, kropki, daszki itp. powinny być napisane dokładnie nad tymi elementami, do których się odnoszą. Numery wzorów należy umieszczać z prawej strony.

6. *Przygotowanie materiału ilustracyjnego.* Rysunki, wykresy i fotografie należy wykonywać zgodnie z obowiązującymi normami, na oddzielnych arkuszach, z podaniem kolejnych numerów rysunków. W maszynopisie artykułu na marginesie, obok właściwego tekstu, należy podać jedynie odnośny numer rysunku, a na oddzielnym arkuszu wykaz podpisów pod rysunki. Wszystkie rysunki, wykresy i fotografie należy nazywać w tekście rysunkami (skrót: rys.). U samego dołu rysunku (a przy fotografiach na odwrocie) należy

wpisać czytelnie numer rysunku, tytuł pracy i nazwisko autora. Ostateczne wykonanie rysunków obowiązuje Redakcję.

7. *Streszczenia.* Do każdej nadsyłanej pracy należy dołączyć krótkie streszczenie (analizę) w języku polskim (w 5 egz.) oraz streszczenia w językach obcych: angielskim, francuskim, niemieckim i rosyjskim (w 2 egz.). W razie niemożności przygotowania streszczeń w językach obcych Autor powinien przynajmniej podać niezbędną do wykonania tłumaczenia obcojęzyczną terminologię.

8. *Literatura.* Po zakończeniu artykułu należy podać wykaz literatury wymieniając w następującej kolejności: pierwsze litery imion, nazwisko autora, po czym po przecinku pełny tytuł dzieła lub artykułu; dalej, w przypadku książki — wydawcę, miejsce wydania i rok, a w przypadku artykułu — tytuł czasopisma, numer zeszytu, rok wydania i ewent. numer strony. Pozycje wykazu powinny być ponumerowane.

Uwagi

1. Niestosowanie się Autora do podanych wyżej wytycznych pociągnie za sobą konieczność potrącenia z honorarium autorskiego kosztów związanych z doprowadzeniem dostarczonych materiałów do postaci wymaganej przez Redakcję.

2. Autorowi przysługuje bezpłatnie 25 egz. odbitek pracy. Dodatkowe egzemplarze Autor może zamówić w Redakcji na własny koszt.

3. **Autora obowiązuje korekta autorska, którą należy zwracać w ciągu 3 dni** pod adresem Redakcji.

Cena zł 40,—

Rozprawy Elektrotechniczne

Kwartalnik

CZASOPISMO
ROZPRAWY ELEKTROTECHNICZNE

Warunki prenumeraty
Cena prenumeraty krajowej
półrocznie zł 80,—
rocznie zł 160,—

Prenumeratę **na kraj** przyjmują Oddziały RSW „Prasa—Książka—Ruch" oraz urzędy pocztowe i doręczyciele w terminach:
— do 25 listopada na I półrocze roku bieżącego i na cały rok następny,
— do 10 czerwca na II półrocze roku bieżącego.
Jednostki gospodarki uspołecznionej, instytucje, organizacje i wszelkiego rodzaju zakłady pracy zamawiają prenumeratę w miejscowych Oddziałach RSW „Prasa—Książka—Ruch", w miejscowościach zaś, w których nie ma Oddziałów RSW — w urzędach pocztowych.
Czytelnicy indywidualni opłacają prenumeratę wyłącznie w urzędach pocztowych i u doręczycieli.
Prenumeratę ze zleceniem wysyłki **za granicę** przyjmują RSW „Prasa—Książka—Ruch", Centrala Kolportażu Prasy i Wydawnictw, ul. Towarowa 28, 00-958 Warszawa, konto NBP XV Oddział w Warszawie Nr 1153-201045-139-11 — w terminach podanych dla prenumeraty krajowej.
Prenumerata ze zleceniem wysyłki za granicę jest droższa od prenumeraty krajowej o 50% dla zleceniodawców indywidualnych i o 100% dla zleceniodawców instytucji i zakładów pracy.
Bieżące i archiwalne numery można nabyć lub zamówić we Wzorcowni Wydawnictw Naukowych PAN-Ossolineum-PWN, Pałac Kultury i Nauki (wysoki parter) 00-901 Warszawa oraz w księgarniach naukowych „Domu Książki".

A subscription order stating the period of time, along with the subscriber's name and address can be sent to your subscription agent or directly to Foreign Trade Enterprise Ars Polona — Ruch, 00-068 Warszawa, 7 Krakowskie Przedmieście, P. O. Box 1001, Poland. Please send payments to the account of Ars Polona — Ruch in Bank Handlowy S. A., 7 Traugutt Street, 00-067 Warszawa, Poland.

Rozpr. Elektrot. T. 27, z. 3, 617—936, Warszawa 1981

Index 37483

POLSKA AKADEMIA NAUK
KOMITET ELEKTROTECHNIKI

PL ISSN 0035-9386

ROZPRAWY ELEKTROTECHNICZNE

TOM XXVIII — ZESZYT 4

KWARTALNIK

PAŃSTWOWE WYDAWNICTWO NAUKOWE
WARSZAWA 1981

POLSKA AKADEMIA NAUK
KOMITET ELEKTROTECHNIKI

ROZPRAWY ELEKTROTECHNICZNE

TOM XXVII—ZESZYT 4

KWARTALNIK

PAŃSTWOWE WYDAWNICTWO NAUKOWE
WARSZAWA 1981

RADA REDAKCYJNA

Przewodniczący

prof. dr hab. JANUSZ GROSZKOWSKI — członek rzeczywisty PAN

Członkowie

prof. dr hab. LUDWIK BADIAN — członek korespondent PAN, prof. TADEUSZ CHOLEWICKI — członek rzeczywisty PAN, doc. dr ZDZISŁAW KACHLICKI, prof. dr hab. TADEUSZ KACZOREK, prof. dr LEONARD KNOCH, doc. dr hab. ANDRZEJ KUSY, prof. JAN MANITIUS, prof. dr WITOLD NOWICKI, prof. dr WITOLD ROSIŃSKI — członek korespondent PAN, prof. dr JERZY SEIDLER — członek korespondent PAN, prof. dr RYSZARD SIKORA, prof. dr hab. JANUSZ TUROWSKI, prof. dr STEFAN WĘGRZYN — członek rzeczywisty PAN

REDAKCJA

Redaktor naczelny

prof. dr WITOLD NOWICKI

Zastępca redaktora naczelnego

prof. dr hab. TADEUSZ KACZOREK

Sekretarz odpowiedzialny

red. JANINA MATHEISEL

ADRES REDAKCJI

00-661 Warszawa, Plac Jedności Robotniczej 1
Politechnika, Instytut Telekomunikacji, Gmach Elektroniki, pok. 483

Dyżury Redakcji: środy i piątki, godz. 11-13
tel. 28 89 81 lub 210 07-737

Telefony domowe: redaktor naczelny – 17 37 03
zastępca redaktora naczelnego – 44 88 34
sekretarz odpowiedzialny – 33 84 97

PAŃSTWOWE WYDAWNICTWO NAUKOWE
Warszawa, Miodowa 10

Nakład 650	Oddano do składania 19.III.1982 r.
Ark. druk. 17,75 + 1,0 luź. wkł. Ark. wyd. 22,75	Podpisano do druku w lutym 1983 r.
Papier druk. sat. kl. IV 71 g 70×100	Druk ukończono w marcu 1983 r.
Zamówienie nr 2245/12/82	Cena zł 40.— Z-82

Drukarnia im. Rewolucji Październikowej, Warszawa

ROZPRAWY ELEKTROTECHNICZNE 1981, 27, z. 4, ss. 939—948

621.319.7

Pole elektrostatyczne dwóch płaskich ekwipotencjalnych taśm

KAZIMIERZ ADAMIAK (KIELCE)

Instytut Elektrotechniki Politechniki Świętokrzyskiej

Otrzymano 10.3.1980

W pracy rozpatrzono zagadnienie brzegowe dla równania Laplace'a w przypadku gdy powierzchnie, na których zadane zostały warunki typu Dirichleta, są powierzchniami wewnętrznymi obszaru, w którym rozwiązanie jest poszukiwane. Rozwiązanie zostało sprowadzone do układu całkowych równań Fredholma I rodzaju. Wykonano obliczenia pola pochodzącego od dwu cienkich ekwipotencjalnych taśm. Do obliczeń wykorzystano metodę różnic skończonych.

1. WSTĘP

Obliczanie pól elektromagnetycznych, opisywanych równaniami Maxwella, polega na rozwiązywaniu zagadnień brzegowych dla równań różniczkowych cząstkowych II rzędu. Dla problemów stacjonarnych lub quasi-stacjonarnych są to równania eliptyczne typu Laplace'a, Poissona lub Helmholtza z warunkami brzegowymi Dirichleta lub Neumanna. Szczególnie często występuje równanie Laplace'a, a mianowicie wszędzie tam, gdzie zachodzi potrzeba wyznaczania pól bezźródłowych i bezwirowych. Równanie Laplace'a opisuje między innymi potencjały: pola elektrostatycznego w obszarze bez ładunków, pola przepływowego prądów i pola magnetycznego w obszarach bezprądowych. Konieczność wyznaczania wymienionych pól zachodzi w wielu dziedzinach elektrotechniki, np. w technice wysokich napięć [1, 2], w maszynach elektrycznych [3], metrologii [4, 5] oraz w fizyce stosowanej.

Literatura na temat rozwiązywania równania Laplace'a jest bardzo bogata [6, 7]. Dotyczy ona jednak głównie klasycznie sformułowanych zagadnień brzegowych, tzn. poszukiwania rozwiązania równania Laplace'a w obszarze $D \subset R^n$, $n = 2, 3$, gdy warunki brzegowe zadane są na powierzchni S stanowiącej brzeg obszaru D. Jeśli D znajduje się wewnątrz powierzchni S, mówi się o wewnętrznym, jeśli nie — o zewnętrznym zagadnieniu brzegowym.

Zagadnienie znacznie się komplikuje, jeśli S stanowi zbiór powierzchni, z których część leży wewnątrz obszaru D. Rozróżnienie zagadnienia wewnętrznego i zewnętrznego nie jest wtedy możliwe, wobec tego mówi się o mieszanym problemie brzegowym [8]. Najbardziej znaną metodą rozwiązywania mieszanych zagadnień brzegowych jest metoda wielokrotnych równań całkowych [8]. Prowadzi ona jednak do dość zawiłych obliczeń, a ponadto daje się zastosować jedynie w najprostszych układach współrzędnych: karte-

zjańskim, walcowym i sferycznym. Niekiedy stosuje się metody iteracyjne [9]. Bardziej uniwersalne są metody oparte na całkowych równaniach Fredholma: II rodzaju [10, 11] i I rodzaju [2, 12—19]. Równanie I rodzaju dla pól elektrostatycznych łatwo formułuje się wychodząc z interpretacji fizycznej. Jest to znany związek między ładunkiem i potencjałem [7]. Rozwiązanie otrzymanego równania napotyka jednak na pewne trudności. Analitycznie udało się je rozwiązać jedynie w przypadku naładowanego, przewodzącego dysku znajdującego się w polu elektrycznym [12]. Nieco większą skuteczność ma metoda sprowadzenia równania I rodzaju do równania II rodzaju [13, 14]. W pracy [19] opisano metodę szacowania pojemności na podstawie równania Fredholma I rodzaju.

Najpopularniejszym sposobem rozwiązywania równań całkowych jest metoda różnic skończonych, dzięki czemu zamiast równania całkowego otrzymuje się układ równań algebraicznych. Dla równań I rodzaju metoda taka może zawieść z powodu złego uwarunkowania [20], mimo to jest często proponowana [15—18].

W pracy sformułowano układ całkowych równań Fredholma I rodzaju dla mieszanego zagadnienia Dirichleta dla równania Laplace'a. Otrzymane równanie zostało następnie rozwiązane dla problemu dwóch ekwipotencjalnych równoległych taśm. Osobliwość występująca w jądrze równania została usunięta metodą podaną w pracy [21]. Do rozwiązania układu równań algebraicznych zastosowany został regularyzujący algorytm Krianiewa [22]. W rozważanym przypadku niestabilność obliczeń nie wystąpiła, ale podana metoda może być zastosowana tam, gdzie złe uwarunkowanie równania jest silniejsze i obserwuje się niestabilność obliczeń [16].

2. ZAGADNIENIE DIRICHLETA DLA RÓWNANIA LAPLACE'A

Niech dane będzie równanie

$$\nabla^2 V = 0 \tag{1}$$

z warunkami brzegowymi

$$V|_S = f(P), \quad P \in S. \tag{2}$$

Rozwiązaniem równania (1) jest [23]

$$V(P) = \int_S \mu(Q) E(P, Q) dS, \tag{3}$$

$$P \in D, \quad Q \in S,$$

gdzie

$$E(P, Q) = \begin{cases} \ln \dfrac{1}{r_{PQ}} & n = 2, \\ \dfrac{1}{r_{PQ}} & n = 3; \end{cases}$$

$\mu(Q)$ jest gęstością powierzchniową potencjału.

Wzór (3) obowiązuje w całym obszarze D, a więc w szczególności na powierzchni S. Ponieważ dla $P \in S$, $V(P)$ jest funkcją znaną z warunku (2), równanie (3) można potraktować jako równanie całkowe z niewiadomą $\mu(Q)$. Po obliczeniu $\mu(Q)$, wyrażenie (3) służy

do wyznaczenia $V(P)$ w dowolnym punkcie obszaru D. Zadanie sprowadza się do zwykłego całkowania.

Podana metoda jest słuszna zarówno dla zagadnienia klasycznego jak też i mieszanego. W tym ostatnim przypadku, przy założeniu, że warunki brzegowe zostały zadane na n powierzchniach

$$S = \bigcup_{i=1}^{n} S_i,$$

otrzymuje się układ równań

$$\sum_{i=1}^{n} \int_{S_i} \mu_i(Q) E(P, Q) dS = V_j(P), \tag{4}$$

$$P \in S_j, \quad Q \in S_i.$$

Jądra równań (3) i (4) zawierają osobliwość dla $r_{PQ} = 0$. Stanowi to dużą trudność, jeśli do rozwiązania tych równań zastosować numeryczną metodę różnic skończonych. Aby osobliwość usunąć, zamiast posługiwać się przy obliczaniu całek wzorami Simpsona

$$\int_a^b \mu(x) \ln \frac{1}{x-y} dx = \sum_{i=1}^{n} \mu(x_i) \ln \frac{1}{x_i - y} h,$$

$$\int_a^b \int_a^d \mu(x, y) \frac{1}{\sqrt{(x-s)^2 + (y-t)^2}} dxdy = \sum_{i=1}^{n} \sum_{j=1}^{m} \frac{\mu(x_i, y_j)}{\sqrt{(x_i - s)^2 + (y_j - t)^2}} hh_1,$$

$$h = \frac{b-a}{n}, \quad h_1 = \frac{d-c}{m},$$

należy posłużyć się wzorami następującymi, będącymi uogólnieniami zależności podanych w pracy [21]

$$\int_a^b \mu(x) \ln \frac{1}{x-y} dx = \sum_{i=1}^{n} \mu(x_i) K(x_i, y), \tag{5a}$$

$$\int_a^b \int_c^d \frac{\mu(x, y)}{\sqrt{(x-s)^2 + (y-t)^2}} dxdy = \sum_{i=1}^{n} \sum_{j=1}^{m} \mu(x_i, y_j) L(x_i, y_j, s, t), \tag{5b}$$

gdzie:

$$K(x_i, y) = \int_{x_i - h/2}^{x_i + h/2} \ln \frac{1}{x_i - y} dx,$$

$$L(x_i, y_j, s, t) = \int_{x_i - h/2}^{x_i + h/2} \int_{y_j - h_1/2}^{y_j + h_1/2} \frac{dxdy}{\sqrt{(x_i - s)^2 + (y_j - t)^2}}.$$

Dla $x_i = y$ oraz $x_i = s$ i $y_j = t$, otrzymuje się

$$K(x_i, y) = h(1 - \ln h/2), \tag{6a}$$

$$L(x_i, y_j, s, t) = 2h \operatorname{Arsh}(h_1/h) - 2h_1 \ln[\operatorname{th}(0{,}5 \operatorname{Arsh} h_1/h)]. \tag{6b}$$

3. REGULARYZACJA CAŁKOWEGO RÓWNANIA FREDHOLMA I RODZAJU

Znaną własnością całkowych równań Fredholma I rodzaju jest ich złe uwarunkowanie [20]. Aby uchronić się od niestabilności obliczeń, przy rozwiązywaniu stosuje się różne metody regularyzacji. Jedną z nich jest iteracyjny algorytm Krianiewa [22].

Niech dane będzie operatorowe równanie I rodzaju

$$Ax = y, \tag{7}$$

gdzie A jest operatorem liniowym, ograniczonym, dodatnim, samosprzężonym.

Przyjmuje się dowolne przybliżenie zerowe x_0. Zakładając n-te przybliżenie rozwiązania, przybliżenie $n+1$ znajduje się z zależności

$$A + \varepsilon B x_{n+1} = y + \varepsilon B x_n, \tag{8}$$

gdzie:

ε — dowolna mała stała dodatnia,

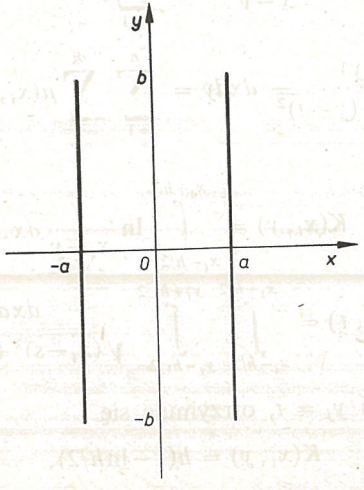
B — operator liniowy, ograniczony, samosprzężony, dodatni.

Można wykazać, że przy $n \to \infty$, x_{n+1} dąży do rozwiązania równania (7) [22].

4. ZAGADNIENIE DWU PŁASKICH EKWIPOTENCJALNYCH TAŚM

Bardzo często w zastosowaniach występuje problem znalezienia pola pochodzącego od dwóch cienkich, nieskończenie długich, ekwipotencjalnych taśm. Może to być np. model przetwornika pojemnościowego lub rezystancyjnego [24], lub linii paskowej do prowadzenia fal elektromagnetycznych [25].

Do obliczeń przyjęto model układu przedstawiony na rys. 1. Zakłada się, że taśmy mają potencjały odpowiednio $-V_0/2$ i $V_0/2$ oraz że cała otaczająca przestrzeń jest elektrycznie jednorodna.



Rys. 1. Model układu dwóch płaskich taśm przyjęty do analizy

Zadanie wyznaczenia potencjału pola elektrostatycznego polega na rozwiązaniu równania Laplace'a z warunkami

$$V\Big|_{\substack{x=a\\|y|\leqslant b}} = \frac{V_0}{2}, \tag{9a}$$

$$V\Big|_{\substack{x=-a\\|y|\leqslant b}} = -\frac{V_0}{2}. \tag{9b}$$

Ze względu na nieparzystość zagadnienia względem osi x i parzystość względem y

$$\mu(y)|_{x=-a} = -\mu(y)|_{x=a},$$

$$\mu(y) = \mu(-y)$$

i układ równań (4) sprowadzi się do równania

$$\int_0^b K(y, s)\mu(s)ds = V_0/2, \tag{10}$$

gdzie

$$K(y, s) = \ln \sqrt{\frac{[(y-s)^2+4a^2][(y+s)^2+4a^2]}{(y-s)^2(y+s)^2}}.$$

Równanie (10) jest równaniem całkowym z jądrem słabo osobliwym. Operator całkowy z takim jądrem spełnia warunki narzucone na operator w równaniu (7) [26]. Można więc zastosować metodę podaną w p.3.

W celu uniwersalizacji rozważań wprowadzić można parametry bezwymiarowe

$$a' = \frac{a}{b}, \quad y' = \frac{y}{b}, \quad s' = \frac{s}{b}.$$

Otrzymuje się wówczas

$$\int_0^1 K(y', s')\mu(s')ds' = V_0/2. \tag{10a}$$

Znak' będzie dalej pomijany.

Równanie (10) zastąpiono, wykorzystując metodę różnic skończonych oraz wzory (5a) i (6a), układem równań algebraicznych i rozwiązano za pomocą emc ODRA 1305. W charakterze operatora B wzięty został operator macierzowy

$$B = \{b_{ij}\},$$

$$b_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}.$$

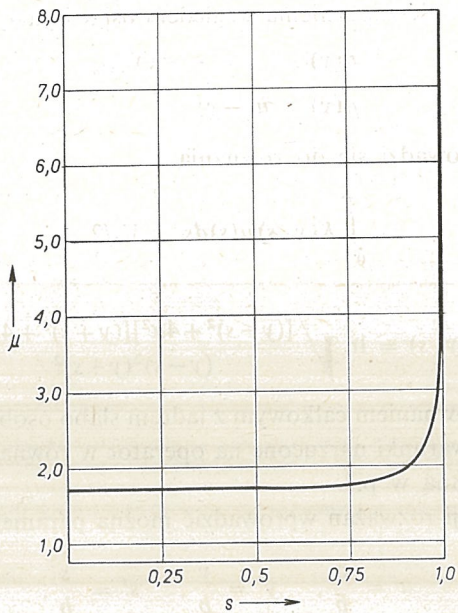
Okazało się, że rozwiązywanie rozparywanego równania jest stabilne nawet dla $\varepsilon \doteq 0$. Dla uzyskania wyników wystarczy wtedy 1 iteracja. Stosowanie $\varepsilon \neq 0$ nie powoduje błędów, przedłuża jednak czas obliczeń.

Wykres $\mu(s)$ dla $n = 100$, $a = 0{,}1$ oraz $V_0 = 1{,}0$ [V] przedstawia rys. 2. Rozkład gęstości potencjału $\mu(s)$ jest dość równomierny wewnątrz taśmy. Natomiast w pobliżu krawędzi staje się silnie nierównomierny.

Po rozwiązaniu równania wyznaczono parametr

$$C_n = \frac{4\pi}{V_0 h} \sum_{i=1}^{n} \mu(s_i) \tag{11}$$

mający ważną interpretację fizyczną. Dla pola elektrostatycznego jest to pojemność elektryczna, dla pól przepływowych — rezystancja między dwiema elektrodami taśmowymi.



Rys. 2. Rozkład gęstości powierzchniowej potencjału $\mu(s)$

Tablica 1

Wartości pojemności taśm dla różnych n

n	C_n
10	11,8011
20	11,7401
30	11,7213
40	11,7118
50	11,7062
60	11,7024
70	11,6999
80	11,6974
90	11,6961
100	11,6943
110	11,6930
120	11,6922
130	11,6917

Tablica 2

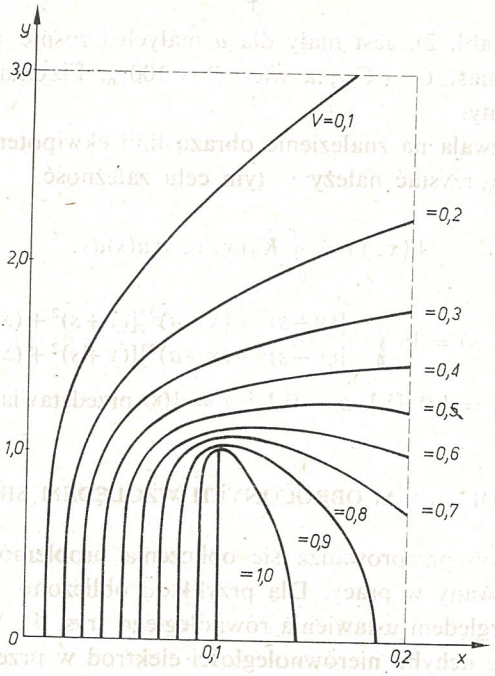
Porównanie dokładnej i przybliżonej wartości pojemności taśm

a	C	C_0	B
0,01	104,6552	100,0	4,45
0,02	52,6989	50,0	5,12
0,05	21,9566	20,0	8,91
0,1	11,6943	10,0	14,49
0,2	6,4912	5,0	22,97
0,5	3,2622	2,0	38,69
1,0	2,1149	1,0	52,72
2,0	1,4891	0,5	66,42
5,0	1,0468	0,2	80,89
10,0	0,8514	0,1	88,25

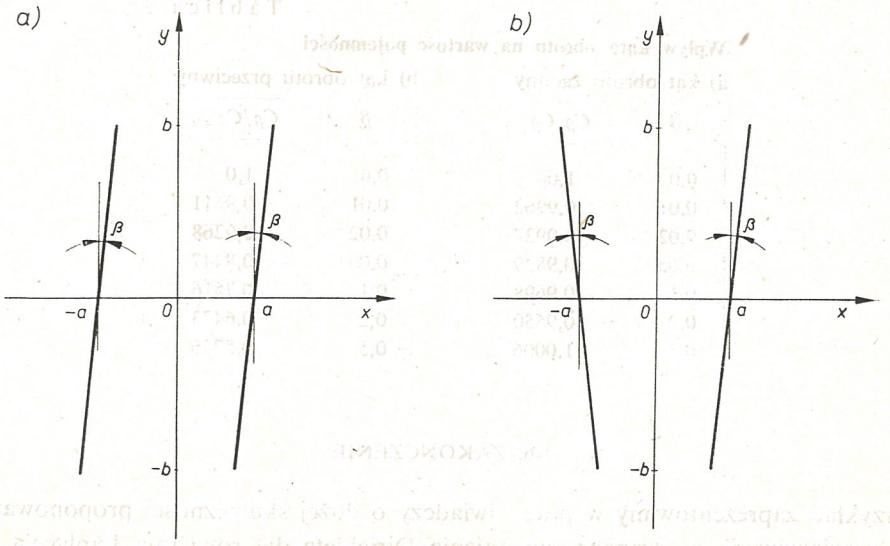
Otrzymane wyniki porównano ze wzorem przybliżonym [7]

$$C_0 = \frac{1}{a}. \tag{12}$$

Tabl. 1 zawiera wartości pojemności C_n dla różnej liczby punktów dyskretyzacji — n.



Rys. 3. Obraz linii ekwipotencjalnych w otoczeniu taśm



Rys. 4. Taśmy obrócone o kąt β

a) kąt obrotu zgodny, b) kąt obrotu przeciwny

Przy n rosnącym $C_n \to C$. Zbieżność jest dość wolna i dlatego dla uzyskania dostatecznie dokładnych wyników konieczne jest założenie dużej wartości n ($n \geqslant 100$).

Błąd przybliżonego wzoru (12)

$$B = \frac{C - C_0}{C} 100\% \tag{13}$$

zależy od wartości a (tabl. 2). Jest mały dla a małych i rośnie przy a rosnącym. Dla $a \to \infty$, $C_0 \to 0$, natomiast $C \to C_\infty$, a więc $B \to 100\%$. Fizycznie C_∞ oznacza pojemność odosobnionej taśmy.

Znajomość $\mu(s)$ pozwala na znalezienie obrazu linii ekwipotencjalnych w otaczającej taśmy przestrzeni. Wykorzystać należy w tym celu zależność

$$V(x, y) = \int_0^1 K_1(x, y, s)\mu(s)ds, \tag{14}$$

$$K_1(x, y, s) = \ln \sqrt{\frac{[(y-s)^2+(x+a)^2][(y+s)^2+(x+a)^2]}{[(y-s)^2+(x-a)^2][(y+s)^2+(x-a)^2]}}.$$

Wyniki obliczeń dla $V_0 = 1{,}0$ [V], $a = 0{,}1$ i $n = 100$ przedstawia rys. 3.

5. POLE TAŚM OBRÓCONYCH WZGLĘDEM SIEBIE

W identyczny sposób przeprowadza się obliczenia problemów bardziej skomplikowanych niż zaprezentowany w pracy. Dla przykładu obliczono pojemność dwóch taśm obróconych o kąt β względem ustawienia równoległego (rys. 4). W ten sposób rozpatrywać można zagadnienie uchybu nierównoległości elektrod w przetworniku pojemnościowym lub rezystancyjnym. Wyniki obliczeń zamieszczono w tabl. 3.

Tablica 3

Wpływ kąta obrotu na wartość pojemności

a) kąt obrotu zgodny

β	$C_\beta/C_{\beta=0}$
0,0	1,0
0,01	0,9962
0,02	0,9927
0,05	0,9829
0,1	0,9698
0,2	0,9550
0,5	1,0006

b) kąt obrotu przeciwny

β	$C_\beta/C_{\beta=0}$
0,0	1,0
0,01	0,9611
0,02	0,9268
0,05	0,8447
0,1	0,7516
0,2	0,6473
0,5	0,5736

6. ZAKOŃCZENIE

Przykład zaprezentowany w pracy świadczy o dużej skuteczności proponowanej metody rozwiązywania mieszanego zagadnienia Dirichleta dla równania Laplace'a. Mimo, że problem sprowadza się do rozwiązania całkowego równania Fredholma I rodzaju,

zazwyczaj wymagającego regularyzacji, niestabilność obliczeń nie wystąpiła. Sformułowanie równań jest bardzo proste, jeśli posłużyć się interpretacją fizyczną. Czas obliczeń jest krótki (na emc ODRA 1305 dla $n = 100$ wynosi ok. 3 min). W podobny sposób rozwiązuje się zagadnienia o dowolnej geometrii. Do rozwiązania otrzymanych równań całkowych zastosować też można inne metody numeryczne, np. metodę momentów [27].

LITERATURA

1. S. Szpor, H. Dzierżek, W. Winiarski, *Technika wysokich napięć*, Warszawa 1967.
2. E.S. Koleczickij, N.A. Melikow, *Rasczot elektrostaticzeskogo pola ekranow słożnoj formy*, Elektriczestwo, no. 2, 43—47 (1974).
3. M. Dąbrowski, *Pola i obwody magnetyczne maszyn elektrycznych*, Warszawa 1971.
4. B.A. Łopatin, *Konduktometria*, Nowosibirsk 1964.
5. W.N. Bałaklejew, N.A. Zołotariew, *Rasczot krajewoj jomkosti datczika postrojennogo po tipu płoskogo kondensatora*, IVUZ Elektromechanika, no. 5, 453 (1975).
6. P. Moon, D.E. Spencer, *Teoria pola*, Warszawa 1966.
7. W.R. Smythe, *Static and dynamic electricity*, New York, Toronto, London 1950.
8. I.N. Sneddon, *Mixed boundary value problems in potential theory*, New York 1966.
9. M. Ehrich, *Beitrag zur Kapazitätsberechnung von Zylinderkondensatoren*, Arch. f. Elektr., 58, 129—133 (1976).
10. L.A. Drużkin, *Integralnyje urawnienia elektrostatiki i krajewyje zadaczi*, Moskwa 1967.
11. A. Majewski, *Obliczanie rozkładu ładunków na ciałach przewodzących*, Arch. Elektr., 27, 63—70 (1978).
12. E.T. Copson, *On the problem of the electified disc*, Proc. Edinbg. Math. Soc., 8, 14—19 (1946).
13. W.E. Wiliams, *The reduction of boundary value problems to Fredholm integral equations of the second kind*, ZAMP 13, 133—152 (1962).
14. D.L. Jain, R.P. Kanwal, *An integral equation method for solving mixed boundary value problems*, SIAM J. Appl. Math., 20, 642—658 (1971).
15. A. Majewski, *Efektywna metoda obliczania nabiegunników*, Arch. Elektr., 23, 445—453 (1974).
16. E.S. Koleczickij, *Rasczoty elektrostaticzeskich polej s ispolzowaniem integralnych urawnienij pierwogo roda*, Elektriczestwo, no. 8, 21—25 (1975).
17. G. De Mey, *Integral equation solution of potential problems in geometries with a slotted boundary*, Int. J. Electronics, 41, 221—226 (1976).
18. A. Majewski, *O obliczaniu pojemności ciał przewodzących*, Arch. Elektr., 27, 467—470 (1978).
19. N.N. Lebiediew, J.P. Skalskaja, *Tocznyje wierchniaja i niżniaja granicy dla jomkosti prowodnika imiejuszczego formu lenty swiornutoj w kolco*, Żurn. Tiech. Fiz., 48, 1147—1153 (1978).
20. A.N. Tichonow, W.J. Arsenin, *Metody reszenia niekorektnych zadacz*, Moskwa 1974.
21. S.S. Sastry, *Numerical solution of Fredholm integral equations with a logarithmic singularity*, Int. J. Num. Math. Engng., 10, 1202—1207 (1976).
22. A.W. Krianiew, *Iteracjonnyj metod reszenia niekorektnych zadacz*, Żurn. Wyczisl. Mat. i Mat. Fiz., 14, 25—35 (1974).
23. H. Marcinkowska, *Wstęp do teorii równań różniczkowych cząstkowych*, Warszawa 1972.
24. Z. Kędryna, *Problemy metrologiczne w procesach wyznaczania konduktywności elektrolitu*, Prace Nauk. Inst. Metr. Elektr. Pol. Wrocławskiej, nr 15, Studia i Mat., nr 7, 5—19 (1977).
25. J.F. Geer, *The electric field due to stripline conductors*, J. Appl. Phys., 47, 5313—5319 (1976).
26. W. Kołodziej, *Wybrane rozdziały analizy matematycznej*, Warszawa 1970.
27. R.F. Harrington, *Field computation by moment methods*, New York 1968.

K. ADAMIAK

ELECTROSTATIC FIELD OF TWO FLAT EQUIPOTENTIAL TAPES

Summary

The paper considers the edge problem for Laplace equation when the surfaces, on which Dirichlet type conditions are assumed, are the internal surfaces of the area for which a solution is sought. The solution has been reduced to a system of Fredholm integral equations of 1st kind. Calculations have been made of a field coming from 2 thin equipotential tapes. The method of finite differences was used in the calculations.

K. ADAMIAK

CHAMP ÉLECTROSTATIQUE DE DEUX BANDES ÉQUIPOTENTIELLES PLATES

Résumé

Dans le travail on a étudié le problème des limites de l'équation de Laplace pour le cas où les surfaces soumises aux conditions du type Dirichlet sont des surfaces intérieures de l'etendue, pour laquelle la solution est recherchée. La résolution a été réduite à un système d'équations de Fredholm du I-er genre. On a calculé le champ déterminé par deux minces bandes équipotentielles. Dans les calculs on a utilisé la méthode des différences finies.

K. ADAMIAK

ELEKTROSTATISCHES FELD ZWEIER FLACHEN ÄQUIPOTENTIELLEN BÄNDER

Zusammenfassung

Es wird das Randproblem für die Laplace-Gleichung für den Fall erörtert, wenn die mit Dirichlet--Bedingungen behafteten Flächen die Innenflächen des jenigen Bereiches sind, worin die Lösung gesucht wird. Die Lösung erhielt die Form des Fredholmschen Systems I. Art der Integralgleichungen. Ein Feld wurde berechnet, das von zwei dünnen äquipotentiellen Bändern stammt. Hierzu wurde die Methode endlicher Differenzen angewandt.

К. АДАМЯК

ЗЛЕКТРОСТАТИЧЕСКОЕ ПОЛЕ ДВУХ ПЛОСКИХ ЭКВИПОТЕНЦИАЛЬНЫХ ЛЕНТ

Резюме

Рассмотрены краевые вопросы для уравнения Лапласса в случае когда поверхностью на которой заданы условия типа Дирихле являются внутренние поверхности области поиска решения. Решение сведено к системе интегральных уравнений Фредгольма I порядка. Расчитано поле генерированное двумя тонкими эквипотенциальными лентами. При расчетах использован метод конечных разностей.

ROZPRAWY ELEKTROTECHNICZNE 1981, 27, z. 4, ss. 949—958

621.372.821

Analiza wypierania prądu i wyznaczanie impedancji w układzie dwóch płaskich przewodów

WOJCIECH LIPIŃSKI (SZCZECIN), MODEST GRAMZ (SZCZECIN)

Instytut Automatyki Przemysłowej Politechniki Szczecińskiej

PIOTR KRASOŃ (SZCZECIN)

Ośrodek Elektronicznej Techniki Obliczeniowej Politechniki Szczecińskiej

Otrzymano 10.4.1980

W pracy przeprowadzono analizę rozkładu gęstości prądu w układzie dwóch płaskich przewodów. Podana została metoda wyznaczania impedancji przewodów dla układów opisanych równaniem całkowym Fredholma drugiego rodzaju o symetrycznym jądrze. Wykonane przykłady obliczeń numerycznych przedstawione zostały w postaci wykresów.

1. CEL PRACY

Celem pracy jest analiza rozkładu gęstości prądu w układzie dwóch płaskich przewodów, których wymiary geometryczne podane zostały na rys. 3. Długość przewodów, w których płynie prąd sinusoidalnie zmienny $i(t) = I\sqrt{2}\sin\omega t$, jest dostatecznie duża, aby możliwa była dwuwymiarowa analiza pola w danym przekroju $z = \text{const}$. Wysokość przewodów $2b$ przyjęta została jako znacznie mniejsza od głębokości wnikania fali elektromagnetycznej. Rozkład gęstości prądu w przewodzie jest przy uwzględnieniu ostatniego uproszczenia funkcją tylko jednej współrzędnej x, co umożliwia dokładniejszą analizę numeryczną. Podana zostanie metoda wyznaczenia impedancji przewodów dla układów opisanych równaniem całkowym Fredholma drugiego rodzaju o symetrycznym jądrze.

2. WYBÓR METODY OBLICZEŃ

Wybór metody obliczania rozkładu prądów w przewodach wiodących prąd zmienny wiąże się z dwoma zagadnieniami:

— sformułowania równań i warunków brzegowych,

— wyboru metody rozwiązywania zagadnienia brzegowego.

W celu sformułowania równań i warunków brzegowych wprowadza się zwykle pomocnicze funkcje potencjalne.

Istnieją zasadniczo dwa odrębne sformułowania [1, 3]:

— przy pomocy potencjału wektorowego pola magnetycznego **A** i potencjału skalarnego pola elektrycznego Φ,
— przy pomocy potencjału wektorowego pola elektrycznego **T** i potencjalu skalarnego pola magnetycznego U.

Kryterium wyboru funkcji potencjalnych jest prostota uzyskanych równań, głównie zaś ich liczba, co wiąże się z liczbą niezerowych składowych wektora **A** lub **T** (por. [1]). W analizowanym zagadnieniu dwuwymiarowym do opisu pola stosowany będzie potencjał wektorowy pola magnetycznego **A**, który w wyniku przyjętych założeń posiada jedną składową z-ową. Rozpatrywany układ jest nieograniczony przestrzennie, zatem warunki brzegowe powinny być sformułowane na podstawie własności pola magnetycznego w nieskończoności.

Metody rozwiązywania zagadnień brzegowych, występujących przy analizie wypierania prądów w przewodnikach, można podzielić na trzy grupy.

1. *Bezpośrednie rozwiązywanie równania różniczkowego cząstkowego przy zadanych warunkach brzegowych i początkowych.* W przypadku analizy zagadnień wiroprądowych mamy do czynienia z równaniem przewodnictwa, gdy rozpatrywany jest stan nieustalony, lub z równaniem Helmholtza, gdy rozważany jest stan ustalony przy zastosowaniu metody symbolicznej. Podstawowymi metodami rozwiązywania tych równań są:
— metoda rozdzielenia zmiennych,
— metoda funkcji Greena,
— metoda przekształceń całkowych,
należące do grupy metod analitycznych oraz numeryczna metoda siatek. W przypadku silnego wypierania prądu stosuje się do analizy pola na zewnątrz obszarów przewodzących metodę odwzorowań konforemnych.

2. *Sprowadzenie zagadnienia do równoważnego równania całkowego, które z kolei rozwiązuje się jedną z przybliżonych metod w realizacji numerycznej.* Najczęściej stosowanymi metodami są:
— metoda Fredholma,
— metoda iteracyjna wyznaczania szeregu von Neumanna,
— metoda przybliżeń Nyströma,
— metoda Bubnowa-Galerkina i metoda najmniejszych kwadratów.

3. *Sformułowanie odpowiedniego zagadnienia wariacyjnego lub zastosowanie procedury Galerkina.* Metoda ta zyskała ostatnio na popularności dzięki dogodnej realizacji numerycznej zwanej metodą elementów skończonych [2, 4]. W niektórych układach, obliczenia można przeprowadzić bez użycia EMC, uzyskując proste, szacunkowe zależności, łatwe do opracowania przy pomocy kalkulatora [9, 10].

W przypadku układów nieograniczonych przestrzennie korzystne jest sprowadzenie zagadnienia do równania całkowego. Obszar całkowania pokrywa się bowiem z obszarem, w którym płynie prąd, a zatem jest obszarem ograniczonym. Zmniejsza to znacznie czas obliczeń numerycznych. Wpływ nieograniczonego przestrzennie obszaru powietrznego uwzględniony jest w kształcie jądra, które byłoby inne, gdyby np. w pobliżu znajdowały się materiały ferromagnetyczne. Zastosowanie, podobnie jak w pracy [12], metody rozdzielenia zmiennych byłoby znacznie bardziej pracochłonne.

3. RÓWNANIE CAŁKOWE I OKREŚLENIE IMPEDANCJI PRZEWODÓW

Pomijając prąd przesunięcia z równań Maxwella

$$\mathrm{rot}\underline{\mathbf{E}} = -j\omega\underline{\mathbf{B}} \quad (\mathrm{div}\underline{\mathbf{B}} = 0) \quad \text{i} \quad \mathrm{rot}\underline{\mathbf{H}} = \underline{\mathbf{J}} \quad (\mathrm{div}\underline{\mathbf{J}} = 0),$$

przy przyjęciu klasycznie zdefiniowanego potencjału wektorowego $\underline{\mathbf{B}} = \mathrm{rot}\underline{\mathbf{A}}$ $(\mathrm{div}\underline{\mathbf{A}} = 0)$ oraz zależności $\underline{\mathbf{B}} = \mu\underline{\mathbf{H}}$, $\underline{\mathbf{J}} = \gamma\underline{\mathbf{E}}$, otrzymuje się równanie różniczkowe $\Delta\underline{\mathbf{A}} = -\mu\underline{\mathbf{J}}$ oraz zależności:

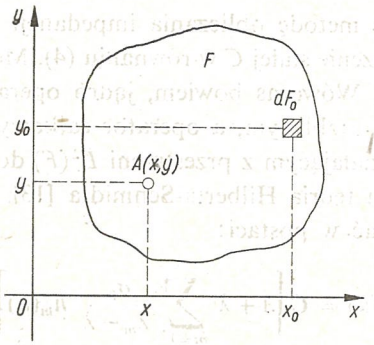
$$\left.\begin{aligned}\mathrm{rot}(\underline{\mathbf{E}}+j\omega\underline{\mathbf{A}}) &= 0\\ \mathrm{div}(\underline{\mathbf{E}}+j\omega\underline{\mathbf{A}}) &= 0\end{aligned}\right\} \Rightarrow \begin{aligned}\underline{\mathbf{J}} &= -j\omega\gamma\underline{\mathbf{A}}+\mathrm{grad}\underline{\Phi}\\ \Delta\underline{\Phi} &= 0\end{aligned} \qquad (1)$$

W zagadnieniach płaskorównoległych wektory $\underline{\mathbf{J}} = \underline{J}\mathbf{1}_z$, $\underline{\mathbf{A}} = \underline{A}\mathbf{1}_z$ mają tylko jedną składową nie będącą funkcją równoległej do osi przewodów zmiennej z. Potencjał $\underline{\Phi}$ ma więc postać

$$\underline{\Phi} = \underline{C}z, \quad \text{gdzie } \underline{C} \text{ — stała.} \qquad (2)$$

Uwzględniając, że spełniające zadane dla danego zagadnienia warunki brzegowe rozwiązanie równania $\Delta\underline{A} = -\mu\underline{J}$ ma postać (rys. 1)

$$\underline{A}(x, y) = -\mu \int_F \underline{J}(x_o, y_o)\, K(x_o, y_o, x, y)\, dF_o, \qquad (3)$$



Rys. 1. Układ współrzędnych i przekrój przewodu w płaskorównoległym zagadnieniu wiroprądowym

otrzymuje się po podstawieniu (2) i (3) do pierwszego z wyrażeń (1) równanie całkowe Fredholma drugiego rodzaju

$$\underline{J}(x, y) = \underline{C}+j\omega\mu\gamma \int_F \underline{J}(x_o, y_o)\, K(x_o, y_o, x, y)\, dF_o, \qquad (4)$$

$$\underline{J}(u) = \underline{C}+\lambda \int_F \underline{J}(v) K(v, u)\, dF_o. \qquad (4a)$$

Przykładowo w podanym na rys. 2 układzie dwóch symetrycznych względem osi y przewodów umieszczonych w nieograniczonej przestrzeni otrzymuje się

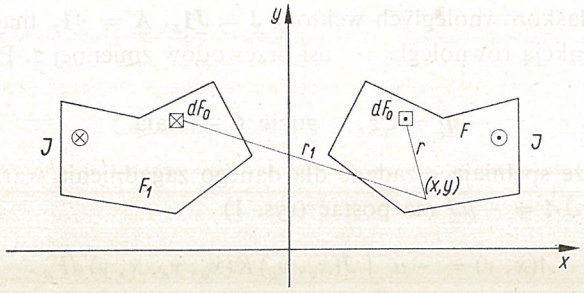
$$K(x_o, y_o, x, y) = \frac{1}{2\pi} \ln \frac{r}{r_1}. \qquad (5)$$

Nieznana zespolona stała $\underline{C}$ występuje jako czynnik odpowiadający niejednorodności równania i obliczenie jej wartości związane jest z warunkiem dodatkowym określającym całkowity prąd w danym przewodzie

$$\underline{I} = \int_F \underline{J}(x_o, y_o) dF_o. \tag{6}$$

W przypadku prądu stałego stała $\underline{C}$, która jest funkcją parametru $\omega\mu\gamma$, przyjmuje wartość równą gęstości prądu

$$\underline{C}(\omega = 0) = \underline{C}_o = \underline{J}(\omega = 0) = \underline{J}_o. \tag{7}$$



Rys. 2. Symetryczne względem osi y zagadnienie wiroprądowe

W pracy [5] podano prostą metodę obliczania impedancji przewodów wiodących prąd zmienny w oparciu o wyznaczenie stałej $\underline{C}$ w równaniu (4). Metoda ta może być stosowana do układów symetrycznych. Wówczas bowiem, jądro operatora całkowego $K(u, v)$ jest symetrycznym jądrem słabo osobliwym, a operator całkowy staje się operatorem samosprzężonym pełnociągłym działającym z przestrzeni $L^2(F)$ do przestrzeni $L^2(F)$. Dla tego typu operatorów słuszna jest teoria Hilberta-Schmidta [13], na mocy której rozwiązanie równania (4a) można zapisać w postaci:

$$\underline{J}(u) = C\left[1 + \lambda \sum_{m=1}^{\infty} \frac{\sigma_m}{\lambda_m - \lambda} h_m(u)\right], \tag{8}$$

gdzie: $\sigma_m = \int_F h_m(u)\, du$, a λ_m i $h_m(u)$ są odpowiednio wartościami i funkcjami własnymi operatora całkowego.

Spełniając warunek (6) i wykonując całkowanie otrzymujemy zależność określającą stałą $\underline{C}$

$$\frac{\underline{I}}{\underline{C}} = F + \lambda^2 \sum_{m=1}^{\infty} \frac{\sigma_m^2}{\lambda_m^2 - \lambda^2} + \lambda \sum_{m=1}^{\infty} \frac{\lambda_m \sigma_m^2}{\lambda_m^2 - \lambda^2}. \tag{9}$$

Po przekształceniu otrzymujemy

$$\frac{\underline{C}}{\underline{I}} = \left|\frac{\underline{C}}{\underline{I}}\right|^2 \left\{F + \lambda^2 \sum_{m=1}^{\infty} \frac{\sigma_m^2}{\lambda_m^2 - \lambda^2} - \lambda \sum_{m=1}^{\infty} \frac{\lambda_m \sigma_m^2}{\lambda_m^2 - \lambda^2}\right\}. \tag{9a}$$

Powyższy wzór zostanie wykorzystany do obliczenia impedancji.

Indukcyjność przewodów na jednostkę długości obliczamy ze wzoru

$$L = \frac{2W_m}{I^2},$$

gdzie W_m jest uśrednioną w czasie wartością energii pola magnetycznego rozpatrywanego układu.

Energię tę można wyznaczyć przy pomocy zależności

$$W_m = \frac{1}{2}\,\mathrm{Re}\left\{\int_F \underline{A}^*(u)\,\underline{J}(u)\,du\right\}. \tag{10}$$

Podstawiając

$$\underline{A}^* = j\frac{\underline{C}^* - \underline{J}^*}{\omega\gamma}$$

otrzymujemy

$$W_m = \frac{1}{2}\,\frac{\underline{I}}{\omega\gamma}\,\mathrm{Im}\{\underline{C}\}. \tag{10a}$$

Zatem reaktancja indukcyjna wyraża się wzorem

$$\omega L = \frac{1}{\gamma \underline{I}}\,\mathrm{Im}\{\underline{C}\}.$$

W analogiczny sposób oblicza się rezystancję przewodów na jednostkę długości

$$R = \frac{P}{I^2}, \tag{11}$$

gdzie

$$P = \frac{1}{\gamma}\int_F \underline{J}(u)J^*(u)\,du. \tag{11a}$$

Po wykonaniu całkowania otrzymujemy

$$P = \frac{|\underline{C}|^2}{\gamma}\left\{F + \lambda^2 \sum_{m=1}^{\infty} \frac{\sigma_m^2}{\lambda_m^2 - \lambda^2}\right\}, \tag{11b}$$

co można zapisać w postaci

$$P = \frac{\underline{I}}{\gamma}\,\mathrm{Re}\{\underline{C}\}. \tag{11c}$$

Ostatecznie, impedancja przewodów na jednostkę długości $\underline{Z} = R + j\omega L$ jest równa

$$\underline{Z} = \frac{1}{\gamma}\,\frac{\underline{C}}{\underline{I}}. \tag{12}$$

Jeśli przyczyną płynięcia prądów jest nie tylko przyłożone napięcie $\underline{U}$, ale również pole magnetyczne pochodzące od innych przewodów i określone potencjałem $A_0(x, y)$ w równaniach (3) (4) dochodzą dodatkowe człony odpowiadające A_0. W zagadnieniach płaskorównoległych znajomość stałej $\underline{C}$ pozwala określić impedancję przewodu bez konieczności przeprowadzenia dodatkowych całkowań, np. według twierdzenia Poyntinga.

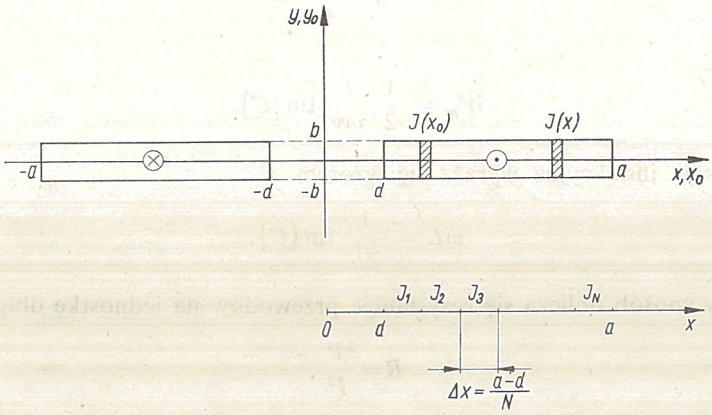
4. UKŁAD DWÓCH PŁASKICH PRZEWODÓW

Wymiary geometryczne układu podano na rys. 3. Rozkład gęstości prądu w przewodzie opisany jest odpowiednio równaniem całkowym Fredholma drugiego rodzaju o jądrze symetrycznym

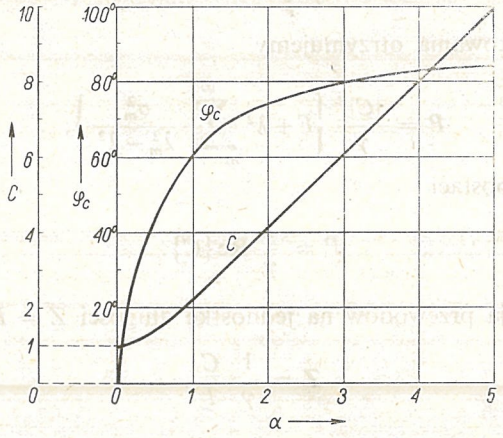
$$\underline{J}(x) = \underline{C} + j\frac{\omega\mu_o\gamma ab}{\pi}\int_{d/a}^{1} \underline{J}(x_o)\ln\left|\frac{x-x_o}{x+x_o}\right| d\left(\frac{x_o}{a}\right), \tag{13}$$

które po przyjęciu

$$x = d + k\Delta x, \quad d = N_1\Delta x,$$
$$x_o = d + n\Delta x, \quad a - d = N\Delta x,$$



Rys. 3. Dwa przewody taśmowe z prądem sinusoidalnie zmiennym



Rys. 4. Zmiany: a) amplitudy gęstości prądu dla $a = 2d$, b) fazy gęstości prądu dla $a = 2d$

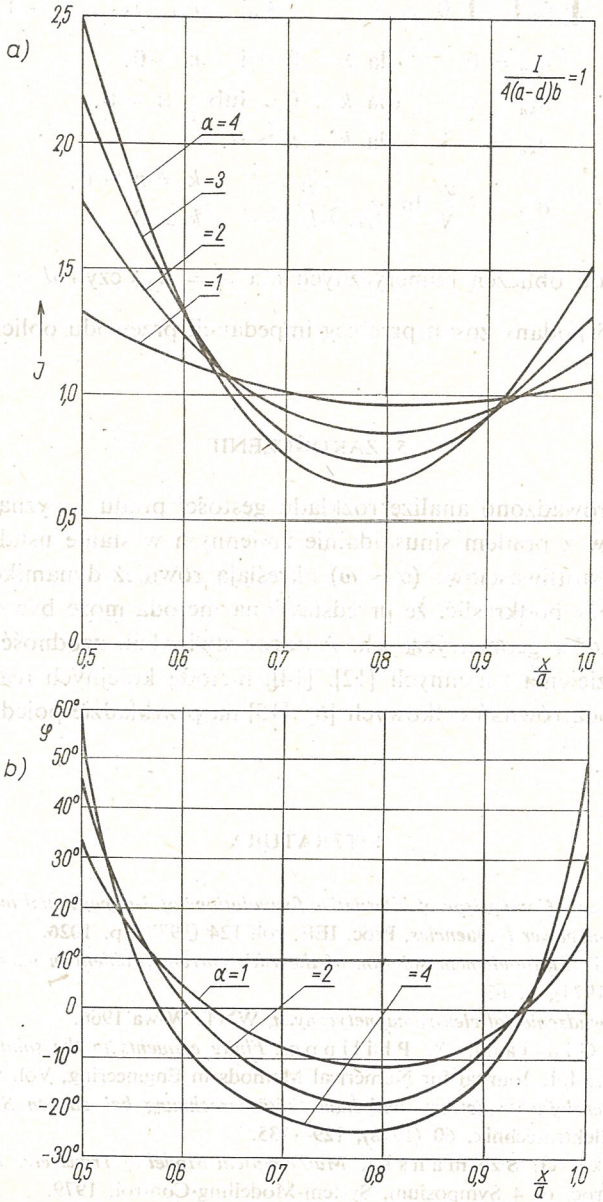
można przy oznaczeniu

$$\alpha = \frac{\omega\mu_o\gamma(a-d)b}{\pi}$$

zapisać jako

$$\frac{1}{N}\sum_{n=1}^{N}\underline{J}_n = \frac{\underline{I}}{2(a-d)b} \qquad k = 1, 2, 3 \ldots N, \tag{14}$$

$$C + j\frac{\alpha}{N}\sum_{n=1}^{N}\underline{J}_n \ln\left|\frac{k-n}{2N_1+k+n}\right| - \underline{J}_k = 0; \qquad n \neq k.$$



Rys. 5. Zmiana impedancji przewodów dla $a = 2d$

Powyższy układ równań można zapisać w postaci:

$$[a_{kn}]\cdot\begin{bmatrix}\underline{C}\\ \underline{J}_1\\ \underline{J}_2\\ \dots\\ \underline{J}_n\\ \dots\\ \underline{J}_N\end{bmatrix}=\begin{bmatrix}N\\0\\0\\\dots\\0\\\dots\\0\end{bmatrix},\qquad [a_{kn}]=\begin{bmatrix}0 & 1 & 1 & \dots & 1\\ 1 & -1 & a_{12} & \dots & a_{1N}\\ 1 & a_{21} & -1 & \dots & a_{2N}\\ \dots & & & & \\ 1 & a_{n1} & a_{n2} & \dots & a_{nN}\\ \dots & & & & \\ 1 & a_{N1} & a_{N2} & \dots & -1\end{bmatrix}; \tag{14a}$$

$$\begin{aligned} a_{kn} &= 0 && \text{dla } k=0 \quad \text{i} \quad n=0,\\ a_{kn} &= 1 && \text{dla } k=0 \quad \text{lub} \quad n=0,\\ a_{kn} &= -1 && \text{dla } k=n>0,\\ a_{kn} &= j\frac{\alpha}{N}\ln\left|\frac{k-n}{2N_1+k+n}\right|, && k\neq n>0,\quad k>0. \end{aligned}$$

Wyniki przykładu obliczeń numerycznych dla $N = N_1$, czyli $d = \frac{a}{2}$, przedstawiono na rys. 4. Na rys. 5 podany został przebieg impedancji przewodu obliczony według wzoru (12).

5. ZAKOŃCZENIE

W pracy przeprowadzono analizę rozkładu gęstości prądu i wyznaczono impedancję płaskich przewodów z prądem sinusoidalnie zmiennym w stanie ustalonym. Otrzymane charakterystyki częstotliwościowe ($\alpha \sim \omega$) określają również dynamikę układu w stanie nieustalonym. Należy podkreślić, że przedstawiona metoda może być stosowana w przypadku innych układów geometrycznych. Autorzy stwierdzili zgodność wyników uzyskanych metodą rozdzielenia zmiennych [12], [14], metodą kolejnych reakcji prądów wirowych [8] oraz metodą równań całkowych [6], [15] na przykładzie pojedynczej taśmy przewodzącej.

LITERATURA

1. C. J. Carpenter, *Comparison of alternative formulations of 3-dimensional magnetic field and eddy-current problems at power frequencies*, Proc. IEE, vol. 124 (1977), p. 1026.
2. M. V. K. Chari, *Finite-element solution of the eddy-current problem in magnetic structures*, IEEE Trans., PAS-93 (1974), p. 62.
3. R. Colin, *Prowadzenie fal elektromagnetycznych*, WNT, W-wa 1966.
4. J. Donea, S. Giuliani, A. Philippe, *Finite elements in the solution of electromagnetic induction problems*, Int. Journal for Numerical Methods in Engineering, Vol. 8 (1974), p. 359.
5. M. Ehrich, *Vereinfachte Verlust- und Induktivitätsberechnung bei ebenen Stromverdrängungsproblemen*, Arch. f. Elektrotechnik, 60 (1978), 129—135.
6. M. Krakowski, G. Szymański, *Mathematical Model of Transverse Flux Induction Heating of Thin Plates*, Proc. of 4 Symposium, System-Modelling-Control, 1979.
7. M. Krakowski, *Elektrotechnika teoretyczna*, część 2, PWN, Warszawa-Poznań 1979.

8. W. Lipiński, *Analiza dynamiki pola elektromagnetycznego w środowiskach przewodzących*, P.N.P.S nr 63, Szczecin, 1976, rozprawa habilitacyjna w Politechnice Gdańskiej.
9. J. Purczyński, *Wykorzystanie metody rozdzielenia zmiennych i metod wariacyjnych do analizy stacjonarnych pól elektromagnetycznych*, P.N.P.S nr 74, Szczecin, 1977, rozprawa habilitacyjna w Politechnice Gdańskiej.
10. P. Rolicz, *Zastosowanie bezpośrednich metod analizy funkcjonalnej do analizy nieustalonych procesów elektromagnetycznych*, rozprawa habilitacyjna, Poznań 1977.
11. R. Sikora, *Teoria pola elektromagnetycznego*, WNT, Warszawa 1977.
12. J. Sonntag, *Abschirmung einer Doppelleitung durch einen parallelen Bandleiter*, Arch. F. Elektrotechnik, 61 (1979), 117—121.
13. W. Kołodziej, *Wybrane rozdziały analizy matematycznej*, WNT, Warszawa 1970.
14. H. Morawska-Koraszewska, *Prądy wirowe w szynie metalowej umieszczonej w równomiernym polu magnetycznym*, Rozprawa doktorska, Łódź, 1979.
15. W. Lipiński, M. Gramz, P. Krasoń, *Ein Beitrag zur Berechnung der Stromverdrängung in einem Bandleiter*, Archiv f. Elektrot., 62 (1980), 51—55.

W. LIPIŃSKI, M. GRAMZ, P. KRASOŃ

CURRENT DISPLACEMENT ANALYSIS AND CALCULATION OF IMPEDANCE IN A SYSTEM OF TWO FLAT CONDUCTORS

Summary

In the paper a distribution of eddy-current density has been analysed in a system consisting of two flat conductors. A method of impedance calculation has been given for systems which are described by the integral Fredholm equation of the second kind with symmetrical kernel. Examples of the numerical computation have been shown in the form of graphs.

W. LIPIŃSKI, M. GRAMZ, P. KRASOŃ

ANALYSE DU DÉPLACEMENT DU COURANT ET LA DÉTERMINATION DE L'IMPÉDANCE DANS LE SYSTÈME DE DEUX CONDUCTEURS PLATS

Résumé

Dans l'article on a analysé la répartition de la densité du courant dans le système de deux conducteurs plats. On a présenté la méthode de la détermination de l'impédance des conducteurs pour les systèmes décrits par l'équation intégrale de Fredholm du second type à noyau symétrique. Les exemples des calculs numériques sont illustrés dans les graphiques.

W. LIPIŃSKI, M. GRAMZ, P. KRASOŃ

ANALYSE DER STROMVERDRÄNGUNG UND IMPEDANZBESTIMMUNG IN ANORDNUNG VON ZWEIFLACHEN LEITERN

Zusammenfassung

In der Arbeit wurde eine Analyse der Stromdichteverteilung in einer Anordnung von zwei flachen Leitern durchgeführt. Es wurde eine Berechnungsmethode für die Impedanz der Leiter für die mittels Fredholmscher Integralgleichung zweiter Art mit symmetrischen Kern beschriebenen Anordnungen angeführt. Die gewonnenen Ergebnisse der numerischen Auswertung wurden graphisch dargestellt.

М. ЛИПИНЬСКИ, М. ГРАМЗ, П. КРАСОНЬ

АНАЛИЗ ВЫТЕСНЕНИЯ ТОКА И ОПРЕДЕЛЕНИЕ ИМПЕДАНСА В СИСТЕМЕ ДВУХ ПЛОСКИХ ПРОВОДНИКОВ

Р е з ю м е

Проведен анализ распределения плотности тока в системе двух плоских проводников. Приведен метод определения импеданса проводников для системов описанных интегральным уравнением Фредгольма второго порядка с симметрическим ядром. Примеры числовых расчетов представлены в форме графиков.

ROZPRAWY ELEKTROTECHNICZNE 1981, 27, z. 4. ss. 959—966

621.3.014.4

Dyfuzja nieustalonego pola elektromagnetycznego w przewodzącym elemencie sferycznym

JERZY GOŁĘBIOWSKI (BIAŁYSTOK)

Instytut Elektrotechniki Politechniki Białostockiej

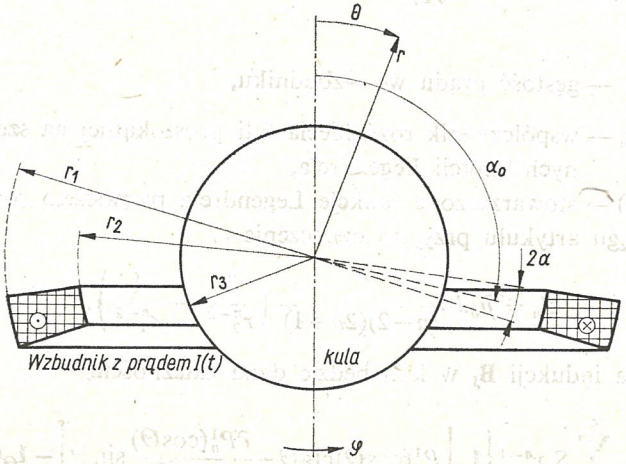
Otrzymano 17.1.1980

W pracy będącej kontynuacją artykułów [1,5] wychodząc z wektorowego równania dyfuzji wyznaczono nieustalony przebieg indukcji $B(t)$ w środku kuli przewodzącej przy skokowym załączeniu prądu wzbudzenia. Uzyskano rozwiązanie bezpośrednie, nie bazujące na określeniu transmitancji operatorowej.

1. PODSTAWIENIE PROBLEMU

Podane w artykule zależności są kontynuacją opracowań [1, 4], nie są jednak z nimi związane bezpośrednio. W pracy [1] wyznaczono w stanie ustalonym i przy wymuszeniu sinusoidalnie zmiennym transmitancję widmową i operatorową układu oraz charakterystyki amplitudowo — fazowe pola jako funkcje danego punktu przestrzeni. Częstotliwościowy opis pośrednio określił też dynamikę stanu nieustalonego układu.

Celem niniejszego artykułu jest bezpośrednia analiza dyfuzji pola poprzez zastosowanie metody rozdzielenia zmiennych przestrzennych i czasu. Badany układ przedstawiono na rys. 1. Wewnątrz cewki sferycznej znajduje się pełna kula przewodząca o przewodności



Rys. 1. Kula przewodząca i wzbudnik pola o symetrii kulistej

właściwej γ i przenikalności magnetycznej μ_o. Postawionym zadaniem będzie wyznaczenie nieustalonego przebiegu indukcji $B(t)$ w środku kuli przy skokowym załączeniu prądu wzbudzenia. Należy podkreślić, że punktem wyjścia jest wektorowe równanie dyfuzji

$$\Delta \mathbf{A} = \mu\gamma \frac{\partial \mathbf{A}}{\partial t} \tag{1}$$

a nie transmitancja częstotliwościowa, jak to miało miejsce w [1].

2. SKŁADOWA USTALONA PRZEBIEGU

Jeżeli jako wymuszenie przyjmiemy jednostkowy skok prądu wzbudzenia $I(t) = I_o \mathbf{1}(t)$, to w stanie ustalonym, w uzwojeniach cewki popłynie prąd stały i prądy wirowe w kuli zanikną. Ponieważ dodatkowo przyjęto przenikalność magnetyczną kuli $\mu_r = 1$, w stanie ustalonym układ zachowa się tak, jakby kuli nie było. Potencjał cewki sferycznej bez kuli w stanie ustalonym został wyznaczony w pracy [6] i wynosi on odpowiednio:

A. obszar powietrza bez wymuszeń, $r \leqslant r_2$

$$A_{\mathrm{I}} = \mu_o J \sum_{n=1}^{\infty} \frac{C_n P_n^1(\cos\Theta)}{(n-2)(2n+1)} \left(\frac{1}{r_2^{n-2}} - \frac{1}{r_1^{n-2}} \right) r^n, \tag{2}$$

B. obszar wzbudzenia, $r_2 \leqslant r \leqslant r_1$

$$A_{\mathrm{II}} = \mu_o J \sum_{n=1}^{\infty} \frac{C_n P_n^1(\cos\Theta)}{(n-2)(2n+1)} \left(\frac{2n+1}{n+3} r^2 - \frac{r^n}{r_1^{n-2}} - \frac{n-2}{n+3} \frac{r_2^{n+3}}{r^{n+1}} \right), \tag{3}$$

C. powietrze bez wymuszeń, $r \geqslant r_1$

$$A_{\mathrm{III}} = \mu_o J \sum_{n=1}^{\infty} \frac{C_n P_n^1(\cos\Theta)}{(n+3)(2n+1)} \frac{r_1^{n+3} - r_2^{n+3}}{r^{n+1}}, \tag{4}$$

gdzie:

$J = \dfrac{J_o}{\alpha(r_1^2 - r_2^2)}$ — gęstość prądu we wzbudniku,

C_n — współczynnik rozwinięcia fali prostokątnej na szereg stowarzyszonych funkcji Legendre'a,

$P_n^1(\cos\Theta)$ — stowarzyszone funkcje Legendre'a pierwszego rodzaju.

W dalszym ciągu artykułu przyjęto oznaczenie

$$S_n = \mu_o J \frac{C_n}{(n-2)(2n+1)} \left(\frac{1}{r_2^{n-2}} - \frac{1}{r_1^{n-2}} \right). \tag{5}$$

Składowa ustalona indukcji $\mathbf{B}_{\mathrm{I}}$ w kuli będzie dana zależnością

$$\mathbf{B}_{\mathrm{I}} = \mathrm{rot}(A_{\mathrm{I}} \mathbf{1}_\varphi) = \sum_{n=1}^{\infty} S_n r^{n-1} \left\{ \mathbf{1}_r \left[P_n^1(\cos\Theta)\operatorname{ctg}\Theta - \frac{\partial P_n^1(\cos\Theta)}{\partial(\cos\Theta)} \sin\Theta \right] - \mathbf{1}_\Theta (n+1) P_n^1(\cos\Theta) \right\}.$$

Jej wartość w środku układu sferycznego wynosi

$$B_1(r=0,\Theta=0^0)=2S_1=-\frac{2}{3}\mu_o J C_1(r_2-r_1)=\frac{2\mu_o C_1 I_o}{3\alpha(r_1+r_2)}. \tag{6}$$

3. SKŁADOWA PRZEJŚCIOWA PRZEBIEGU

Podczas trwania stanu nieustalonego wewnątrz przewodzącej kuli płyną prądy wirowe, wywołujące własne pole elektromagnetyczne. Tym razem nie należy więc pomijać kuli znajdującej się w cewce. Obowiązuje układ jak na rys. 1.

A. Obszar kuli przewodzącej, $r \leqslant r_3$, $\mu_r = 1$

Składowa zaburzeniowa potencjału A' w kuli będzie opisana wektorowym równaniem dyfuzji

$$\Delta \mathbf{A}' = \mu_o \gamma \frac{\partial \mathbf{A}'}{\partial t},$$

gdzie $\mathbf{A}' = A'(r,\Theta,t)\mathbf{1}_\varphi$.

Przyjmując zgodnie z metodą separacji zmiennych rozwiązanie w postaci

$$A'(r,\Theta,t)=\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}U_n(r,\Theta)\,T_m(t),$$

otrzymano dwa równania:

$$\frac{dT_m}{dt}+\frac{M_m^2}{\mu_o\gamma}T_m=0 \quad \text{(liniowe, jednorodne, rzędu pierwszego)},$$

$$\Delta U_n + M_m^2 U_n = 0 \quad \text{(Helmholtza)}.$$

Ich rozwiązania odpowiednio wynoszą

$$T_m(t)=C_m e^{-\frac{M_m^2}{\mu_o\gamma}t}$$

oraz zgodnie z pracą [1]

$$U_n(r,\Theta)=\sum_{n=1}^{\infty}\left[\frac{A_n}{\sqrt{r}}J_{n+0,5}(M_m r)+\frac{B_n}{\sqrt{r}}N_{n+0,5}(M_m r)\right]\left[C_n P_n^1(\cos\Theta)+D_n Q_n^1(\cos\Theta)\right].$$

Potencjał na osi z (czyli dla $\Theta = 0^0$) i w początku układu jest skończony, zatem

$$Q_n^1(\cos 0^0)=\infty \Rightarrow D_n = 0,$$

$$\frac{1}{\sqrt{r}}N_{n+0,5}(M_m r)\big|_{r=0}=\infty \Rightarrow B_n = 0.$$

Ostatecznie składowa przejściowa potencjału A' wewnątrz przewodzącej kuli dana jest zależnością

$$A'=\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\frac{D_1(m,n)}{\sqrt{r}}J_{n+0,5}(M_m r)P_n^1(\cos\Theta)e^{-\frac{M_m^2}{\mu_o\gamma}t}. \tag{7}$$

B. Obszar poza kulą, $r \geqslant r_3$

Zgodnie z prawami dynamiki układów fizycznych składowa przejściowa jest składową swobodną, niezależną od wymuszenia. Nie wyróżnia więc ona obszaru wzbudzenia. Dodatkowo własności fizyczne przestrzeni są wszędzie takie same ($\mu = \mu_o$). Wynika z tego stwierdzenie, że poza kulą dla składowej zaburzeniowej jest jednolity obszar opisany równaniem Laplace'a. Nie jest bowiem wyróżniany obszar wzbudzenia

$$\Delta A'' = 0.$$

Rozwiązanie tego równania ma postać

$$A'' = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} [A_n r^n + B_n r^{-(n+1)}][C_n P_n^1(\cos\Theta) + D_n Q_n^1(\cos\Theta)] C_m e^{-\frac{M_m^2}{\mu_o \gamma} t}.$$

Na osi z ($\Theta = 0^0$) i w nieskończoności potencjał zeruje się:

$$Q_n^1(\cos 0^0) = \infty \Rightarrow D_n = 0,$$
$$r^n|_{r=\infty} = \infty \Rightarrow A_n = 0.$$

Ostatecznie składowa zaburzeniowa potencjału wektorowego A'' dla $r \geqslant r_3$ wyraża się zależnością

$$A'' = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} D_2(m, n) r^{-(n+1)} P_n^1(\cos\Theta) e^{-\frac{M_m^2}{\mu_o \gamma} t}. \tag{8}$$

Składową przejściową indukcji $\mathbf{B}'$ w kuli wyznaczono ze wzoru

$$\mathbf{B}' = \operatorname{rot}(A' \mathbf{1}_\varphi). \tag{9}$$

Po podstawieniu (7) w (9) i wykorzystaniu związków:

$$\frac{\partial}{\partial\Theta}[P_n^1(\cos\Theta)\sin\Theta] = P_n^1(\cos\Theta)\cos\Theta + \cos\Theta\sin\Theta,$$

$$\frac{\partial}{\partial r}[\sqrt{r} J_{n+0,5}(M_m r)] = M_m \sqrt{r} J_{n-0,5}(M_m r) - \frac{n}{\sqrt{r}} J_{n+0,5}(M_m r),$$

ostatecznie otrzymano

$$\mathbf{B}'(r, \Theta, t) =$$

$$= \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} D_1(m, n) r^{-1,5} J_{n+0,5}(M_m r) \left[P_n^1(\cos\Theta)\operatorname{ctg}\Theta - \frac{\partial P_n^1(\cos\Theta)}{\partial(\cos\Theta)} \sin\Theta \right] e^{-\frac{M_m^2}{\mu_o \gamma} t} \mathbf{1}_r +$$

$$+ \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} D_1(m, n)[n r^{-1,5} J_{n+0,5}(M_m r) - M_m r^{-0,5} J_{n-0,5}(M_m r)] P_n^1(\cos\Theta) e^{-\frac{M_m^2}{\mu_o \gamma} t} \mathbf{1}_\Theta.$$

By podać wartość składowej zaburzeniowej w początku układu sferycznego wyznaczono granice:

$$\lim_{r \to 0} r^{-1,5} J_{1,5}(M_m r) = \frac{M_m^{1,5}}{2^{1,5} \Gamma(2,5)} = \sqrt{\frac{2}{\pi}} \frac{M_m^{1,5}}{3},$$

$$\lim_{r \to 0} \sum_{n=2}^{\infty} r^{-1,5} J_{n+0,5}(M_m r) = 0,$$

$$\lim_{r=0} r^{-0,5} J_{0,5}(M_m r) = \sqrt{\frac{2}{\pi}}\, M_m^{0,5},$$

$$\lim_{r\to 0} \sum_{n=2}^{\infty} r^{-0,5} J_{n-0,5}(M_m r) = 0.$$

Z powyższych granic wynika redukcja sumy określającej składową zaburzeniową do pierwszego składnika względem n w początku układu współrzędnych

$$\mathbf{B}'(r=0) = \frac{2}{3}\sqrt{\frac{2}{\pi}} \sum_{m=1}^{\infty} D_1(m, n=1) M_m^{1,5} \{\cos\Theta \mathbf{1}_r - \sin\Theta \mathbf{1}_\Theta\} e^{-\frac{M_m^2}{\mu_o \gamma} t}.$$

Uwzględniając, że wektor $\mathbf{B}'$ skierowany jest wzdłuż osi $\Theta = 0^0$, można wyznaczyć jego długość w początku układu współrzędnych. Jest to zarazem wartość składowej przejściowej w tym punkcie

$$B'(r=0, \Theta=0^0, t) = \frac{2}{3}\sqrt{\frac{2}{\pi}} \sum_{m=1}^{\infty} D_1(m, n=1) M_m^{1,5} e^{-\frac{M_m^2}{\mu_o \gamma} t}. \tag{10}$$

4. WYZNACZENIE STAŁYCH NA PODSTAWIE WARUNKÓW GRANICZNYCH

Całkowity przebieg potencjału będzie odpowiednio sumą wzorów (2) oraz (7) i (8) dla $r \leqslant r_3$

$$A_1 = A_I + A^1 = A_I + \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} D_1(m, n) r^{-0,5} J_{n+0,5}(M_m r) P_n^1(\cos\Theta) e^{-\frac{M_m^2}{\mu_o \gamma} t}, \tag{11}$$

dla $r_3 \leqslant r \leqslant r_2$

$$A_2 = A_I + A'' = A_I + \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} D_2(m, n) r^{-(n+1)} P_n^1(\cos\Theta) e^{-\frac{M_m^2}{\mu_o \gamma} t}. \tag{12}$$

A. Warunki brzegowe

Na powierzchniach granicznych potencjały są ciągłe wraz ze składowymi stycznymi wektora natężenia pola magnetycznego (brak prądów powierzchniowych)

$$A'(r_3) = A''(r_3) \quad \text{oraz} \quad \frac{\partial}{\partial r}(rA')\bigg|_{r=r_3} = \frac{\partial}{\partial r}(rA'')\bigg|_{r=r_3} \quad \text{dla } t > 0,$$

co daje równania:

$$D_1(m, n) J_{n+0,5}(M_m r_3) - D_2(m, n) r_3^{-(n+0,5)} = 0,$$

$$D_1(m, n)[M_m r_3 J_{n-0,5}(M_m r_3) - n J_{n+0,5}(M_m r_3)] + D_2(m, n) n r_3^{-(n+0,5)} = 0.$$

Aby powyższy jednorodny układ równań posiadał rozwiązania nietrywialne $D_1(m, n) \neq 0$ i $D_2(m, n) \neq 0$, potrzeba i wystarcza by jego wyznacznik zerował się

$$W = 0. \tag{13}$$

Związek (13) jest równaniem wartości własnych. Po obliczeniu wyznacznika W przybiera on postać

$$J_{n-0,5}(M_m r_3) = 0. \tag{14}$$

Jeżeli przez γ_m oznaczymy pierwiastki równania

$$J_{n-0,5}(x) = 0, \tag{15}$$

to wartości własne można przedstawić jako

$$M_m = \frac{\gamma_m}{r_3}. \tag{16}$$

B. Warunek początkowy

W momencie załączenia skoku jednostkowego potencjał w kuli był zerowy

$$A_1(r, \Theta, t = 0) = 0 \quad \text{dla } 0 \leqslant r \leqslant r_3. \tag{17}$$

Z zależności (11) i (17) wynika

$$A'(r, \Theta, t = 0) = -A_1(r, \Theta) \quad \text{dla } 0 \leqslant r \leqslant r_3. \tag{18}$$

Kładąc (7) (dla $t = 0$) i (2) w (18) oraz stosując oznaczenie (5) otrzymano

$$r^{-0,5} \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} D_1(m, n) J_{n+0,5}(M_m r) P_n^1(\cos\Theta) = -\sum_{n=1}^{\infty} S_n r^n P_n^1(\cos\Theta),$$

co daje po uproszczeniu i zastosowaniu (16) zależność

$$\sum_{m=1}^{\infty} D_1(m, n) J_{n+0,5}\left(\frac{\gamma_m}{r_3} r\right) = -S_n r^{n+0,5} \quad \text{dla } 0 \leqslant r \leqslant r_3.$$

Stanowi ona szereg Fouriera-Bessela ortogonalny z wagą $\varphi(r) = r$ w przedziale $\langle 0, r_3 \rangle$. $\sqrt{\gamma_m}$ jest pierwiastkiem równania $J_{n-0,5}(x) = 0$, spełnia więc dla $k = -(n+0,5)$ równanie [2]

$$\gamma_m J'_{n+0,5}(\gamma_m) - k J_{n+0,5}(\gamma_m) = 0.$$

$D_1(m, n)$ wyznaczono więc z zależności

$$D_1(m, n) = \frac{-2S_n \int_0^{r_3} r^{n+1,5} J_{n+0,5}\left(\frac{\gamma_m}{r_3} r\right) dr}{r_3^2 J_{n+0,5}^2(\gamma_m)}$$

i stąd

$$D_1(m, n) = -\frac{2S_n r_3^{n+0,5} J_{n+1,5}(\gamma_m)}{\gamma_m J_{n+0,5}^2(\gamma_m)}. \tag{19}$$

Ze względu na redukcję sumy określającej składową swobodną indukcji B' do pierwszego składnika względem n ($n = 1$) w początku układu współrzędnych równanie wartości własnych (14) przybierze postać

$$J_{0,5}(M_m r_3) = 0 = \sqrt{\frac{2}{\pi M_m r_3}} \sin(M_m r_3) \Rightarrow M_m = \frac{m\pi}{r_3}; \quad m \in C.$$

Z powyższego oraz z (16) wynika

$$\gamma_m = M_m r_3 = m\pi. \tag{20}$$

Z (19) wyznaczono $D_1(m, n = 1)$

$$D_1(m, n = 1) = -\frac{2S_1 r_3^{1,5} J_{2,5}(\gamma_m)}{\gamma_m J_{1,5}^2(\gamma_m)}. \tag{21}$$

gdzie z (5) mamy:

$$\left.\begin{aligned} S_1 &= \frac{1}{3}\mu_o J C_1(r_1 - r_2) = \frac{\mu_o C_1 J_o}{3\alpha(r_1 + r_2)}, \\ J_{2,5}(\gamma_m) &= J_{2,5}(m\pi) = -\frac{3}{m\pi^2}\sqrt{\frac{2}{m}}(-1)^m, \\ J_{1,5}^2(\gamma_m) &= J_{1,5}^2(m\pi) = \frac{2}{m\pi^2}. \end{aligned}\right\} \tag{22}$$

Wstawiając (22) do (21) otrzymano

$$D_1(m, n = 1) = \frac{\sqrt{2}}{\pi}\frac{1}{m^{1,5}}\frac{\mu_o C_1 I_o r_3^{1,5}}{\alpha(r_1 + r_2)}(-1)^m. \tag{23}$$

Umieszczając (23) w (10) oraz wykorzystując (20) otrzymamy nieustaloną składową indukcji B' w środku kuli

$$B' = \frac{4\mu_o C_1 I_o}{3\alpha(r_1 + r_2)} \sum_{m=1}^{\infty} (-1)^m e^{-\frac{m^2\pi^2}{\mu_o \gamma r_3^2} t}. \tag{24}$$

Dodając wzory (6) i (24) otrzymano całkowity przebieg indukcji w środku kuli przewodzącej

$$B_1 = B_I + B' = \frac{2\mu_o C_1 I_o}{3\alpha(r_1 + r_2)}\left[1 + 2\sum_{m=1}^{\infty}(-1)^m e^{-\frac{m^2\pi^2}{\mu_o \gamma r_3^2} t}\right]. \tag{25}$$

Jest to wynik zgodny ze wzorem podanym w pracy [1], co związane jest z jednoznacznością metody częstotliwościowej i czasowej. Zgodnie z ogólną teorią dyfuzji pola elektromagnetycznego do obszarów przewodzących indukcja $B(t)$ zmienia się jak funkcja theta Jacobiego

$$B(t) = \frac{2\mu_o C_1}{3\alpha(r_1 + r_2)} I_o \vartheta_o\left(u = 0, \tau = j\pi\frac{t}{\mu_o \gamma r_3^2}\right).$$

LITERATURA

1. J. Gołębiowski, W. Lipiński, *Analiza dynamiki ekranu elektromagnetycznego o symetrii kulistej*, Rozprawy Elektrotechniczne, z. 3, 1980.
2. E. Kącki, *Równania różniczkowe cząstkowe w elektrotechnice*, WNT, Warszawa 1971.
3. E. Kącki, *Termokinetyka*, WNT, Warszawa 1967.
4. W. Lipiński, J. Gołębiowski, J. Wasik, *Dynamik des Wirbelstrommagnetfelds in abgeschirmten induktiren Messwergeberen*, IET Z. elektr. Inform-u Energietechnik, 10 (1980), 3.
5. W. Lipiński, J. Gołębiowski, *Modelling of electromagnetic shields dynamics*. Materiały — Proceedings of the Fourth International Symposium — System, Modelling, Control, Zakopane, Poland, October 8—13, 1979.

6. W. Lipiński, J. Purczyński, R. Sikora, *Zastosowanie całkowego przekształcenia Mellina do obliczeń pola magnetycznego cewki kulistej*, Arch. Elektrot., t. XXII, z. 2, 1973.
7. P. Moon, D.E. Spencer, *Teoria pola*, PWN, Warszawa 1966.
8. W.R. Smythe, *Static and dynamic electricity*, New York 1950.

J. GOŁĘBIOWSKI

DIFFUSION OF A TRANSIENT ELECTROMAGNETIC FIELD IN A SPHERICAL CONDUCTING ELEMENT

Summary

In this paper, a continuation of articles [1, 5], the induction transient $B(t)$ in the centre of a sphere on exciting current step feeding was determined from the vector diffusion equation. The equation was obtained directly without the need to determine the operating transfer function.

J. GOŁĘBIOWSKI

ANALYSE DE LA DYNAMIQUE DE DIFFUSION INSTABLE DU CHAMP MAGNÉTIQUE DANS LE SYSTÈME BIDIMENSIONNEL DE LA SPHÈRE CONDUCTRICE

Résumé

Dans l'étude, qui constitue la suite des articles [1, 5], on a déterminé, à partir de l'équation vectorielle de la diffusion, le régime transitoire de l'induction $B(t)$ au centre d'une sphère, pendant l'enclenchement du courant d'excitation. On a obtenu une solution directe qui n'est pas basée sur la transmittance opératrice.

J. GOŁĘBIOWSKI

DIFFUSION DES NICHTSTATIONÄREN ELEKTROMAGNETISCHEN FELDES IN EINER LEITENDEN SPHERISCHEN ABSCHIRMUNG

Zusammenfassung

In vorliegender Arbeit, die eine Ergänzung der Artikel [1, 5] bildet, wird als Lösung der verktorialen Diffusionsgleichung der Verlauf der nichtstationären magnetischen Induktion im Mittelpunkt der leitenden Kugel bestimmt. Während der Berechnung wird die zeitliche Änderung des erregenden Stromes als Einheitsfunktion vorausgesetzt. Es wird eine direkte Lösung im Zeitbereich ohne Anwendung der Übertragungsfunktion für die analysierte Anordnung angegeben.

Е. ГОЛЭМБИОВСКИ

АНАЛИЗ ДИНАМИКИ НЕУСТАНОВИВШЕГОСЯ ПРОЦЕССА ДИФФУЗИИ ЭЛЕКТРОМАГНИТНОГО ПОЛЯ В ЭЛЕКТРОПРОВОДЯЩЕМ СФЕРИЧЕСКОМ ЭЛЕМЕНТЕ

Резюме

В труде являющемся продолжением [1, 5], исходя из векторного уравнения диффузии определена характеристика переходного процесса магнитной индукции в середине шара при скачкообразном изменении тока возбуждения. Задача решена без применения частотных методов анализа

ROZPRAWY ELEKTROTECHNICZNE 1981, 27, z. 4, ss. 967—978

621.315.1

Effects of Short-Circuit Currents in an Overhead a—c Transmission Line on Underground Conductors

WOJCIECH MACHCZYŃSKI (POZNAŃ), GRZEGORZ SZYMAŃSKI (SZAMOTUŁY)

Instytut Elektrotechniki Przemysłowej Politechniki Poznańskiej

Otrzymano 10.1.1980

A method of calculation of currents and potentials excited along an underground conductor by the short-circuit current in an overhead a—c transmission line with earth wires is presented. The method permits the calculation of currents and potentials along an underground conductor if the fault-to-ground occurs in the proximity of the underground conductor. The resultant electrical effects of a short-circuit current on an underground conductor are obtained by the superposition of conductive and inductive effects resulting from the conductive and inductive couplings between the underground conductor and the power line.

1. INTRODUCTION

The calculation of currents and potentials along underground conductors occurs in many practical cases which result from a joint usage of the same right-of-way by electric power transmission lines and other ground return circuits, such as cables, pipelines, etc.

A line-to-ground fault in a—c transmission lines with an effectively grounded neutral may give rise to high potentials and voltages in underground conductors which can cause damage to equipment or coating and even hazard to personnel. The problem of the interference of overhead transmission lines on underground conductors becomes significant in the case of extended proximity and high transmission voltages.

Therefore, the problem of calculation of currents and potentials produced in underground conductors by fault currents in nearby transmission lines is of great importance. A number of publications have been devoted to this subject [1—6]. The results derived in [3] permit the calculation of the current and potential excited by the short-circuit current in a parallel overhead transmission line on the assumption that the line includes one conductor only, which is treated as an equivalent conductor carrying the short-circuit current.

The purpose of this paper is to determine the current and potential along an extensive underground conductor due to ground fault in an a—c transmission line with earth wires. The considerations are based on the assumptions that the earth is uniformly conducting and that all conductors are parallel to the earth's surface. The underground conductor

is considered to be infinite and buried parallel to the overhead transmission line. It is assumed that the system underground conductor-overhead transmission line is linear.

It is also supposed that the currents and voltages vary with time as $\exp(j\omega t)$. Therefore the alternating component of the short-circuit current is taken into account.

List of principal symbols

a — length of span,
b — horizontal separation between underground conductor and power line,
d — depth at which conductor is buried,
G_i — unit-length leakage conductance of conductor insulation or coating,
h — height of overhead conductor above earth's surface,
$I(x)$ — current along underground conductor,
I_p — current in overhead conductor, short-circuit current in power line,
I_c — earth-wire current,
I_t — tower current,
$k^2 = j\omega\mu_o\gamma$,
l — length of power line, distance from substation to place of fault,
r — external radius of underground conductor,
s — distance from earth electrode to underground conductor,
$V(x)$ — potential along underground conductor,
V_p — voltage between phase conductor and ground,
V_c — voltage between earth wire and ground,
Z_o — characteristic impedance of underground conductor,
Z_{11} — unit-length self impedance of underground conductor,
Z_{12} — unit-length mutual impedance between underground conductor and phase conductor,
Z_{12c} — unit-length mutual impedance between underground conductor and earth wire,
Z_{op} — self impedance of phase conductor,
Z_{oc} — self impedance of earth wire,
Z_{om} — mutual impedance between phase conductor and earth wire,
Γ — propagation coefficient,
γ — earth conductivity,
μ_o — permeability of vacuum,
ω — angular frequency.

2. GENERAL EQUATION

The propagation coefficient of an underground conductor of infinite length is the solution of the transcendental equation [2]

$$\Gamma^2\left[G_i^{-1}+\frac{1}{\pi\gamma}\ln\frac{1{,}12}{\sqrt{2\mathrm{rd}(\Gamma^2+k^2)}}\right]=Z_i+\frac{j\omega\mu_o}{\pi}\ln\frac{1{,}12}{\sqrt{2\mathrm{rd}(\Gamma^2+k^2)}}, \tag{1}$$

with the requirement $\mathrm{Re}(\Gamma) > 0$.

Denoting

$$Z = Z_i + \frac{j\omega\mu_o}{\pi} \ln \frac{1{,}12}{\sqrt{2rd(\Gamma^2 + k^2)}}, \tag{2}$$

$$Y^{-1} = G_i^{-1} + \frac{1}{\pi\gamma} \ln \frac{1{,}12}{\sqrt{2rd(\Gamma^2 + k^2)}}, \tag{3}$$

the propagation coefficient becomes

$$\Gamma = \sqrt{ZY}; \quad \mathrm{Re}(\Gamma) > 0. \tag{4}$$

More detailed data concerning the propagation coefficient are given in [2].

The characteristic impedance of an underground conductor is given by

$$Z_o = \sqrt{\frac{Z}{Y}}; \quad \mathrm{Re}(Z_o) > 0, \tag{5}$$

where Z and Y are taken from eqns. (2) and (3) respectively.

As shown in reference [2], the circuit in Fig. 1 may be regarded as an approximate equivalent circuit to an underground conductor with the parameters given by eqns. (2) and (1). The underground conductor shown in Fig. 1 is modeled as a ladder network,

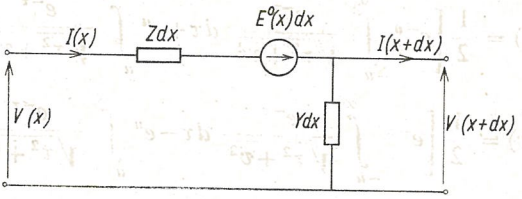


Fig. 1. Equivalent circuit for an underground conductor

with series impedances per unit length and shunt admittances per unit length. The quantity $E^o(x)$ is the electric intensity of the primary electric field, which is excited by the factors causing the field in general, such as energy sources, currents which leave earth electrode, currents in overhead conductors, etc.

From Fig. 1 follows the differential equation

$$\frac{d^2I(x)}{dx^2} - \Gamma^2 I(x) = -YE^o(x). \tag{6}$$

It is shown in [2] that the expression for the current $I(x)$ along an underground conductor takes the form

$$I(x) = \frac{1}{2Z_o}\left[e^{-\Gamma x} \int_{-\infty}^{x} E^o(v)\, e^{\Gamma v} dv + e^{\Gamma x} \int_{x}^{\infty} E^o(v)\, e^{-\Gamma v} dv\right]. \tag{7}$$

Equation (7) permits the determination of the current and potential excited along underground conductors by conductive or inductive influence of the transmission line [2], [3].

2.1. Conductive energisation

The current and potential along an underground conductor due to the conductive energisation as shown in Fig. 2. are derived in [2]

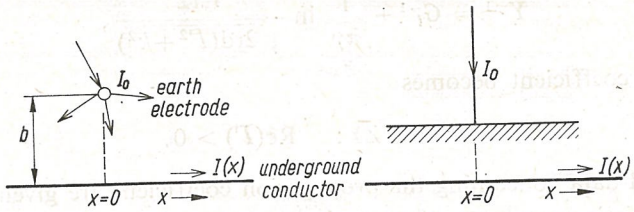


Fig. 2. Conductive energisation of underground conductor by current leaving a point electrode and equivalent elementary system

$$I(x) = \frac{\gamma I_o}{4\pi\gamma}\,[\Psi(\Gamma x, \Gamma s) + \Psi(\Gamma x, \Gamma s')], \tag{8}$$

$$V(x) = \frac{\Gamma I_o}{4\pi\gamma}\,[\Omega(\Gamma x, \Gamma s) + \Omega(\Gamma x, \Gamma s')], \tag{9}$$

where

$$\Omega(u, v) = \frac{1}{2}\left[e^{-u}\int_{-u}^{\infty}\frac{e^{-\tau}}{\sqrt{\tau^2+v^2}}\,d\tau + e^{u}\int_{u}^{\infty}\frac{e^{-\tau}}{\sqrt{\tau^2+v^2}}\,d\tau\right], \tag{10}$$

$$\Psi(u, v) = \frac{1}{2}\left[e^{-u}\int_{-u}^{\infty}\frac{e^{-\tau}}{\sqrt{\tau^2+v^2}}\,d\tau - e^{u}\int_{u}^{\infty}\frac{e^{-\tau}}{\sqrt{\tau^2+v^2}}\,d\tau\right], \tag{11}$$

and s and s' are the distances from the underground conductor to the earth electrode and its image, respectively

$$s = \sqrt{(h-d)^2+b^2}, \tag{12}$$

$$s' = \sqrt{(h+d)^2+b^2}. \tag{13}$$

Properties of the functions Ψ and Ω are discussed in [2] and [6], while tables of these functions are given in [6] for complex u and v.

2.2. Inductive energisation

The current and potential along an underground conductor buried parallel to an overhead conductor which is energised by current I_o, as shown in Fig. 3, are determined in [3].

As shown in [3], the current $I(x)$ can be written in the form

$$I(x) = \begin{cases} \dfrac{Z_{12}I_o}{2Z_{11}}\,(1-e^{\Gamma x}), & x < 0, \\ -\dfrac{Z_{12}I_o}{2Z_{11}}\,(1-e^{-\Gamma x}), & x > 0. \end{cases} \tag{14}$$

The potential excited along the underground conductor is given by

$$V(x) = \frac{Z_{12} Z_o I_o}{2Z_{11}} e^{-\Gamma|x|}. \tag{15}$$

The circuit from Fig. 3 permits the examination of the inductive energisation of an underground conductor by a line as in Fig. 4.

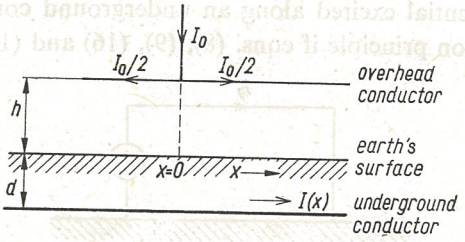


Fig. 3. Infinite underground conductor with inductive energisation

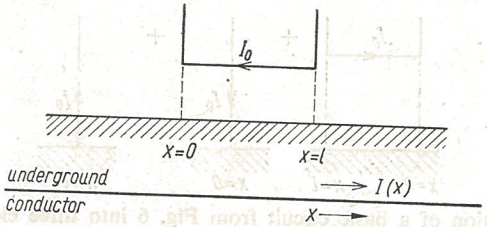


Fig. 4. Underground conductor energised by a line

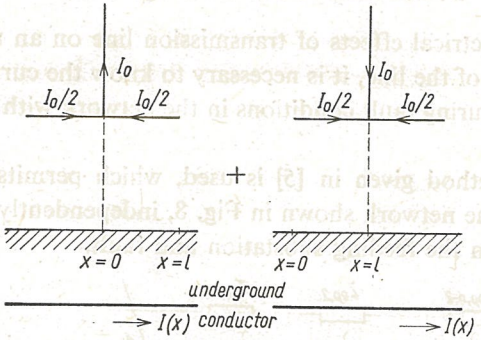


Fig. 5. Decomposition of circuit from Fig. 4 into two elementary systems

The circuit from Fig. 4 can be decomposed in a manner shown in Fig. 5, into two elementary systems. Thus the current and potential along an underground conductor according to eqns. (14) and (15) can be written in the form

$$I(x) = \frac{Z_{12} I_o}{2Z_{11}} [\mathrm{sign}(x)(1-e^{-\Gamma|x|}) - \mathrm{sign}(x-l)(1-e^{-\Gamma|x-l|})], \tag{16}$$

$$V(x) = -\frac{Z_{12} Z_o I_o}{2Z_{11}} [e^{-\Gamma|x|} - e^{-\Gamma|x-l|}]. \tag{17}$$

2.3. Effect of short-circuit current

The eqns. (8), (9), (16) and (17) permit the examination of a basic ground-return equivalent circuit with line-to-ground fault as in Fig. 6. The circuit shown in Fig. 6 can be decomposed into three elementary systems, each with inductive or conductive energisations, as shown in Fig. 7.

The current and potential excited along an underground conductor can be obtained by use of the superposition principle if eqns. (8), (9), (16) and (17) are applied.

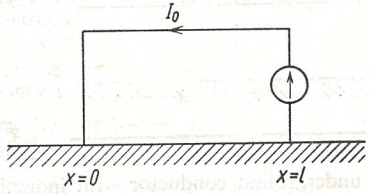


Fig. 6. Basic ground-return equivalent circuit with line-to-ground fault

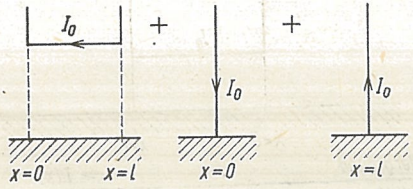


Fig. 7. Decoposition of a basic circuit from Fig. 6 into three elementary systems

2.4. Current distribution along transmission line

To determine the electrical effects of transmission line on an underground conductor located in the proximity of the line, it is necessary to know the current distribution among earth wires and towers during fault conditions in the network with an effectively grounded neutral.

In the paper the method given in [5] is used, which permits the calculation of the current distribution in the network shown in Fig. 8, independently of the location of fault and the distance between the feeding substation and fault.

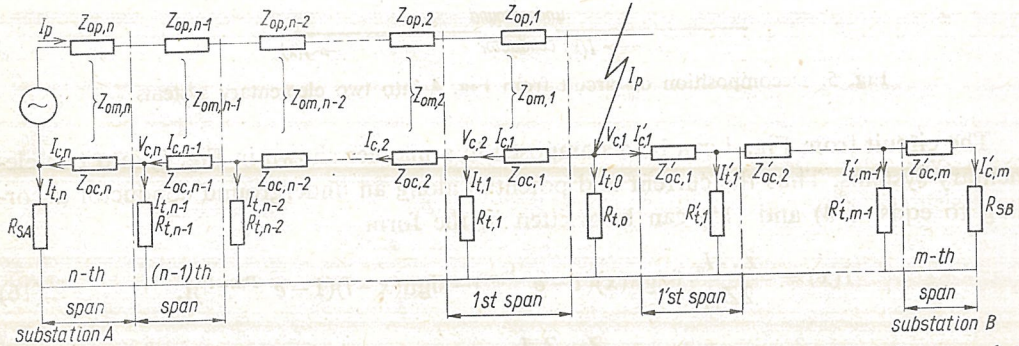


Fig. 8. Equivalent circuit of a faulty transmission line supplied from one side, consisting of n spans between the feeding substation A and the fault, and of m spans between the fault and the substation B

According to the method used, all currents and voltages in the transmission line can be obtained as per-unit complex values of the chosen reference quantity I_p. In this method the equivalent circuit shown in Fig. 9 represents one span of the transmission line.

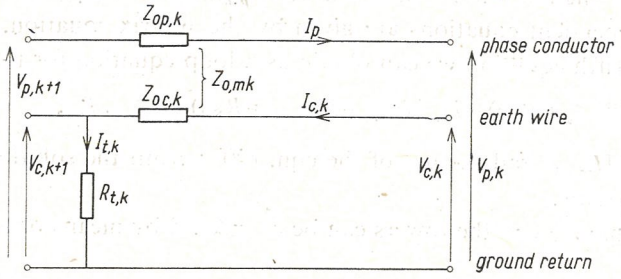


Fig. 9. Equivalent circuit of the k-th span of the transmission line [5]

To solve for the unknown voltages and currents of this network, loop and node equations can be used together with the assumed reference

$$I_p = 1+j0. \tag{18}$$

Relations describing the basic ground-return equivalent circuit of one span can be written in matrix form as follows [5]

$$\begin{bmatrix} V_{p,k+1} \\ V_{c,k+1} \\ I_p \\ I_{c,k+1} \end{bmatrix} = \begin{bmatrix} 1 & 0 & Z_{op,k} & -Z_{om,k} \\ 0 & 1 & Z_{om,k} & -Z_{oc,k} \\ 0 & 0 & 1 & 0 \\ 0 & \dfrac{1}{R_{t,k}} & \dfrac{Z_{om,k}}{R_{t,k}} & -\left(1+\dfrac{Z_{oc,k}}{R_{t,k}}\right) \end{bmatrix} \begin{bmatrix} V_{p,k} \\ V_{c,k} \\ I_p \\ I_{c,k} \end{bmatrix} \tag{19}$$

or

$$[L_k] = [S_k][R_k], \tag{20}$$

where:

$[L_k]$ — the column vector of the left-hand side quantities,

$[S_k]$ — the matrix of the constants of the equivalent circuit representing the k-th span,

$[R_k]$ — the column vector of the right-hand side quantities of eqn. (19).

The boundary conditions at the fault in the transmission line shown in Fig. 8 are

$$[R_1] = \begin{bmatrix} V_{p,1} \\ V_{c,1} \\ I_p \\ I_{c,1} \end{bmatrix} = \begin{bmatrix} Z_{Fl}(I_p - I_{c,1}) \\ Z_{Fl}(I_p - I_{c,1}) \\ I_p \\ I_{c,1} \end{bmatrix}, \tag{21}$$

where Z_{Fl} — resultant impedance of the ladder network extending beyond the fault and the resistance of the faulty tower footing $R_{t,o}$.

The boundary conditions before the n-th span are

$$[L_{n-1}] = [R_n] = \begin{bmatrix} V_{p,n} \\ V_{c,n} \\ I_p \\ I_{c,n} \end{bmatrix} = \begin{bmatrix} V_{p,n} \\ R_{t,n-1}(I_{c,n-1} - I_{c,n}) \\ I_p \\ I_{c,n} \end{bmatrix}. \tag{22}$$

According to eqn. (20) we can write

$$[L_{n-1}] = [S_{n-1}][S_{n-2}] \dots [S_2][S_1][R_1]. \tag{23}$$

In eqn. (22) occur the five unknown quantities $V_{p,n}$, I_p, $I_{c,1}$, $I_{c,n}$, $I_{c,n-1}$.

The three independent equations are given by the matrix equation, the fourth one is eqn. (18) and the fifth equation we can obtain as a loop equation for the n-th span

$$I_p(Z_{om,n}+R_{SA})-I_{c,n}(Z_{oc,n}+R_{t,n-1}+R_{SA})+I_{c,n-1}R_{t,n-1} = 0. \tag{24}$$

Knowledge of $[L_{n-1}]$ and the use of the eqn. (20) permit the solution of the values of all earth wire currents.

The currents that flow in the towers can be calculated by means of Kirchhoff's current law

$$I_{c,k}-I_{c,k-1}+I_{t,k-1} = 0. \tag{25}$$

3. EFFECTS ON AN UNDERGROUND CONDUCTOR DUE TO FAULT CURRENT IN A TRANSMISSION LINE

Consider an underground conductor which is buried at a depth d, parallel to an overhead transmission line, as in Fig. 10. It is assumed that fault current I_p flows in the faulted phase. In the segments of the earth wire flow currents $I_{c,i}$ and $I'_{c,i}$ respectively, and the

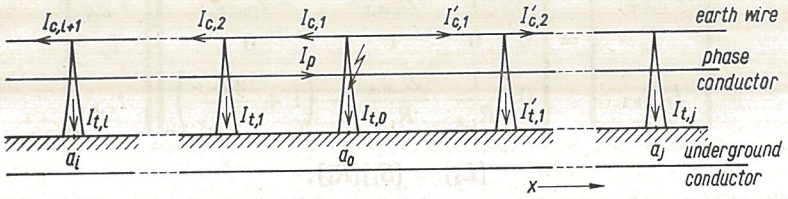


Fig. 10. Infinite underground conductor in the proximity of a transmission line with ground fault

currents $I_{t,i}$ and $I'_{t,i}$ leave the tower groundings, which are assumed to be in the form of the point electrode buried at the depth h below the earth's surface and separated by the distance s from the underground conductor.

Since the magnetic coupling exists between the underground conductor and the overhead phase conductor and earth wire, the current and potential along the underground conductor are excited by the currents flowing in the overhead conductors as shown in Fig. 4.

In addition to the inductive energisation, a conductive energisation of the underground conductor exists. The conductive energisation of the conductor is related to the currents leaving the tower groundings as shown in Fig. 2.

Thus we can get the current and potential along the underground conductor due to fault in a transmission line using the superposition principle.

To determine the current and potential excited along the underground conductor, first consider the circuit shown in Fig. 9. This circuit can be decomposed into elementary systems with conductive and inductive energisations as shown in Fig. 11.

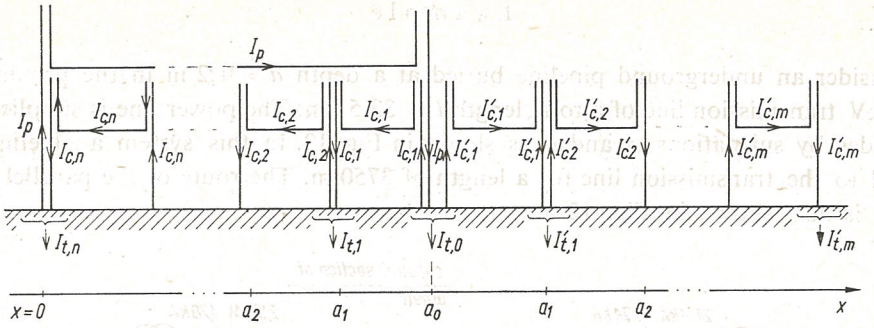


Fig. 11. Decomposition of the circuit of a faulty transmission line supplied from one side into elementary systems with inductive and conductive energisations

Thus the current and potential excited along the underground conductor by currents in the circuit in Fig. 10 decomposed as shown in Fig. 11 take the form

$$I(x) = \frac{Y}{4\pi\gamma}\left\{\sum_{i=0}^{n} I_{t,i}\{\Psi[\Gamma(x-a_i), \Gamma s_i] + \Psi[\Gamma(x-a_i), \Gamma s'_i]\} + \right.$$

$$\left. + \sum_{j=1}^{m} I'_{t,j}\{\Psi[\Gamma(x-a_j), \Gamma s_j] + \Psi[\Gamma(x-a_j), \Gamma s'_j]\}\right\} - \frac{Z_{12} I_p}{2Z_{11}}[\operatorname{sign}(x)(1-e^{-\Gamma|x|}) -$$

$$- \operatorname{sign}(x-a_o)(1-e^{-\Gamma|x-a_o|})] + \frac{Z_{12c}}{2Z_{11}} \sum_{i=0}^{n} I_{c,i+1}[\operatorname{sign}(x-a_{i+1})(1-e^{-\Gamma|x-a_{i+1}|}) -$$

$$- \operatorname{sign}(x-a_i)(1-e^{-\Gamma|x-a_i|})] - \frac{Z_{12c}}{2Z_{11}} \sum_{j=0}^{m} I'_{c,j+1}[\operatorname{sign}(x-a_j)(1-e^{-\Gamma|x-a_j|}) +$$

$$- \operatorname{sign}(x-a_{j+1})(1-e^{-\Gamma|x-a_{j+1}|})], \tag{26}$$

$$V(x) = \frac{\Gamma}{4\pi\gamma}\left\{\sum_{i=0}^{n} I_{t,i}\{\Omega[\Gamma(x-a_i), \Gamma s_i] + \Omega[\Gamma(x-a_i), \Gamma s'_i]\} + \right.$$

$$\left. + \sum_{j=1}^{m} I'_{t,j}\{\Omega[\Gamma(x-a_j), \Gamma s_j] + \Omega[\Gamma(x-a_j), \Gamma s'_j]\}\right\} + \frac{Z_{12} Z_o I_p}{2Z_{11}}[e^{-\Gamma|x|} - e^{-\Gamma|x-a_o|}] +$$

$$- \frac{Z_{12c} Z_o}{2Z_{11}} \sum_{i=0}^{n} I_{c,i+1}[e^{-\Gamma|x-a_{i+1}|} - e^{-\Gamma|x-a_i|}] + \frac{Z_{12c} Z_o}{2Z_{11}} \sum_{j=0}^{m} I'_{c,j+1}[e^{-\Gamma|x-a_j|} - e^{-\Gamma|x-a_{j+1}|}]. \tag{27}$$

The eqns. (26) and (27) are determined for the network shown in Fig. 10 e.q. in the case of ground fault supplied from one side. When the ground fault occurs in the network supplied from both sides, we can solve for the current and potential along the underground conductor using the superposition principle.

To examine the electrical effects of an a—c transmission line on underground conductor a computer program was developed. Some computations are presented below.

Example

Consider an underground pipeline buried at a depth $d = 1{,}2$ m in the proximity of a 110 kV transmission line of a total length $l = 33{,}5$ km. The power line is supplied from both sides by substations A and B as shown in Fig. 12. In this system a pipeline runs parallel to the transmission line for a length of 3750 m. The route of the parallel stretch studied is also shown in Fig. 12.

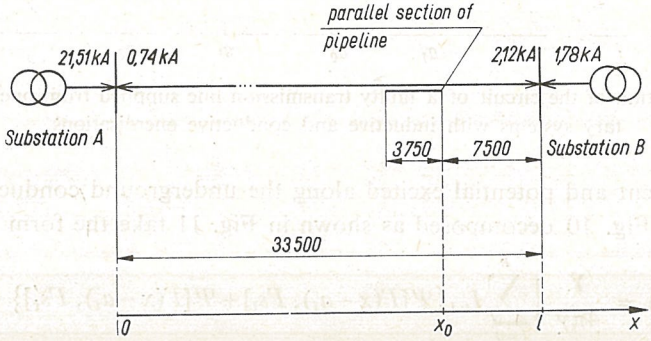


Fig. 12. The route of the parallel stretch studied

The remaining data on the transmission line-underground pipeline arrangement are as follows:

pipeline coating conductance $G_i = 10^{-3}$ S/m,
pipeline diameter $D = 600$ mm,
ground conductivity $\gamma = 10^{-3}$ S/m,
horizontal separation $b = 10$ m,
distance from the overhead conductor to the pipeline $d_{cp} = 16$ m,
length of a span $a = 300$ m,
type of earth wires $1 \times$AFl-70.

Case a. A fault to ground in a section of the transmission line located in the vicinity of the underground pipeline.

Suppose that the fault to ground occurs at point x_o, e.g. at the end of the approach section of the underground pipeline. In this case the fault currents from substations A and B are $I_{pA} = 2{,}72$ kA and $I_{pB} = 1{,}53$ kA, respectively. The potential of the underground pipeline at the point x_0, calculated using eqn. (27), is

$$|V(x_o)| = 1194 \text{ V}.$$

It should be noted that in this case the potential along the pipeline is due to both conductive and inductive effects of a—c transmission line.

Case b. A fault to ground far from the approach section of the underground pipeline.

Suppose that the ground fault occurs at the end of the transmission line in substation B. In this case the fault current from substation A $I_{pA} = 2{,}12$ kA. The potential of the underground pipeline at the point x_o, calculated using eqn. (27), is

$$|V(x_o)| = 678 \text{ V}.$$

It should be noted that in this case the potential along the pipeline is excited by the inductive effects of the phase and earth wire currents only.

5. CONCLUSIONS

A method of calculation of current and potential along an infinite underground conductor due to short-circuit effects in a nearby a—c transmission line with earth wires is presented in the paper.

The relations obtained in the paper permit the solution of many problems concerning the interference of transmission lines on underground conductors.

The method presented permits the determination of the harmful effects on an underground conductor due to both the conductive influence of currents leaving the tower groundings and the inductive influence of currents flowing in the phase conductor and earth wire, respectively. The computer program developed permits the determination of the current and potential along an underground conductor, independently of the location of ground fault and the distance between the feeding point and fault.

REFERENCES

1. H. Böcker, D. Oeding, *Induktionsspannungen an Pipelines in Trassen von Hochspannungsleitungen*, Elektrizitätswirtschaft, 1966, H. 5, pp. 157—170.
2. M. Krakowski, *Currents and potentials along extensive underground conductors*, Proc. IEE, Vol. 155, No. 9, Sept. 1968, pp. 1299—1304.
3. M. Krakowski, *Currents and potentials along underground conductors due to short-circuit current in a power line*, Archiwum Elektrot., T. XXIII, Z. 1, pp. 25—39.
4. M. Krakowski, *Obwody ziemnopowrotne*, WNT, Warszawa 1979.
5. S. A. Sebo, *Zero-sequence current distribution along transmission lines*, IEEE Trans. on Power, Apparatus and Systems, Vol. PAS-88, No. 6, June 1969, pp. 910—919.
6. E. D. Sunde, *Earth conduction effects in transmission systems*, New York, Dover Publications, 1968.

W. MACHCZYŃSKI, G. SZYMAŃSKI

ODDZIAŁYWANIE PRĄDÓW ZWARCIOWYCH W LINII ELEKTROENERGETYCZNEJ NA PRZEWODY PODZIEMNE

Streszczenie

W pracy przedstawiono metodę obliczania prądów i potencjałów wywołanych w przewodzie podziemnym przez prąd zwarcia doziemnego w pobliskiej linii elektroenergetycznej. Prądy i potencjały wzdłuż podziemnego przewodu otrzymano jako łączny wynik oddziaływania galwanicznego prądów wypływających z uziomów oraz oddziaływania indukcyjnego prądów płynących przewodami fazowymi i odgromowymi linii elektroenergetycznej.

W. MACHCZYŃSKI, G. SZYMAŃSKI

INFLUENCE ÉLECTRIQUE DU COURANT DE COURT-CIRCUIT DANS LA LIGNE DE HAUTE TENSION SUR LES CONDUCTEURS ENTERRÉS

Résumé

Dans l'article on a présenté la méthode du calcul de l'influence électrique du courant de court-circuit dans la ligne aérienne sur les conducteurs enterrés parallèles. Dans le calcul on a pris en considération l'influence conductive et inductive, ayant égard à la distribution réelle du courant dans le câble de garde. On a déterminé le courant et le potentiel dans le conducteur enterré par la superposition de deux influences.

W. MACHCZYŃSKI, G. SZYMAŃSKI

ELEKTRISCHE BEEINFLUSSUNG DER ERDVERLEGTEN LEITUNGEN DURCH ERDKURZSCHLUßSTROM IN HOCHSPANNUNGSLEITUNGEN

Zusammenfassung

In der Arbeit wurde die Methode für die Berechnung der elektrischen Beeinflussung einer erdverlegten Leitung durch eine parallellaufende Hochspannungsleitung bei Erdkurzschluss dargestellt. Es wurden induktive und ohmsche Beeinflussungen des Erdseils, der Maste und Erdrückleitung betrachtet unter Berücksichtigung der wirklichen Stromverteilung des Erdkurzschlußstromes. Die in der erdverlegten Leitung erregten Potentiale und Ströme wurden als Superposition der beiden Beeinflussungen erhalten.

В. МАХЧИНЬСКИ, Г. ШИМАНЬСКИ

ЭЛЕКТРОМАГНИТНОЕ ВЛИЯНИЕ ТОКА КОРОТКОГО ЗАМЫКАНИЯ В ЛИНИИ ВЫСОКОГО НАПРЯЖЕНИЯ НА ПОДЗЕМНЫЕ ПРОВОДНИКИ

Резюме

Представлен метод вычисления токов и потенциалов вызванных в подземном проводнике током короткого замыкания параллельной электроэнергетической линии. Рассмотрено проводимостное и индуктивное влияние, учитывая действительное распределение тока короткого замыкания в электроэнергетической линии.

ROZPRAWY ELEKTROTECHNICZNE 1981, 27, z. 4, ss. 979—1003

621.313

Pole magnetyczne wzbudzane przez stojan maszyny dwufazowej

JACEK GIERAS (BYDGOSZCZ), PIOTR PAUL (BYDGOSZCZ)

Instytut Telekomunikacji i Elektrotechniki Akademii Techniczno-Rolniczej w Bydgoszczy

Otrzymano 7.11.1977

Wyznaczono przestrzennoczasowe rozkłady przepływu, indukcji oraz obciążenia liniowego stojanów różnych maszyn dwufazowych, tzn. silników wykonawczych, prądnic tachometrycznych, silników jednofazowych z fazą pomocniczą pracy oraz fazą rozruchową. Rozważania dotyczą zarówno maszyn o biegunach utajonych jak i o biegunach jawnych. Uzwojenie znajdujące się żłobkach sprowadzono do równoważnej warstwy prądowej rozłożonej równomiernie na powierzchni przyszczelinowej rdzenia stojana. Z uwagi na wpływ zjawisk krańcowych na parametry maszyn dwufazowych, tzn. maszyn małych, rozpatrzono rozkład obciążenia liniowego oraz indukcji zarówno wzdłuż podziałki biegunowej jak i wzdłuż osi wału maszyny.

1. WSTĘP

Praca dotyczy wyznaczania rozkładu przepływu (krzywej wzbudzenia), obciążenia liniowego oraz indukcji magnetycznej (krzywej pola) stojana różnego rodzaju maszyn dwufazowych. Rozkład powyższych wielkości rozpatrzono wzdłuż podziałki biegunowej oraz wzdłuż osi maszyny. Maszyny dwufazowe są na ogół maszynami małymi, więc zjawiska brzegowe na krańcach rdzeni stojana i wirnika są tutaj zagadnieniem bardzo ważnym, którego — tak jak w maszynach dużej i średniej mocy — nie można pominąć. Rozpatrzono zarówno stojany o uzwojeniach rozłożonych (biegunach utajonych) jak i skupionych (biegunach jawnych). Obok maszyn o biegunach utajonych ostatnio bowiem coraz szersze rozpowszechnienie — ze względu na prostą technologię wytwarzania obwodu magnetycznego — uzyskują maszyny o biegunach jawnych, nie tylko jako silniki zwartozwojowe, ale również silniki jednofazowe z fazą pomocniczą *RC* oraz silniki wykonawcze i prądnice tachometryczne.

Wykaz ważniejszych oznaczeń

$A_{1\nu}$ — amplituda fali obciążenia liniowego stojana,
a_1 — wartość chwilowa obciążenia liniowego,
B — indukcja magnetyczna,
b_1 — wartość chwilowa indukcji magnetycznej,
$F_{1\nu}$ — amplitudy fal przepływu, indukcji lub obciążenia liniowego stojana (oznaczenie ogólne),

f_1 — wartość chwilowa przepływu, indukcji lub obciążenia liniowego stojana (oznaczenie ogólne),
I_1 — wartość skuteczna prądu uzwojenia stojana,
k_c — współczynnik Cartera,
k_n — współczynnik nasycenia obwodu magnetycznego,
L_i — długość idealna rdzenia stojana,
L_2 — długość idealna rdzenia wirnika,
m_1 — liczba faz stojana,
N — liczba zwojów w żłobku,
n — harmoniczne rozkładu pola wzdłuż osi maszyny,
p — liczba par biegunów,
q_1 — liczba żłobków na biegun i fazę,
s — rozpiętość cewki,
t_1 — podziałka żłobkowa,
z_1 — liczba zwojów szeregowych,
α — kąt przestrzenny między osiami magnetycznymi sąsiednich uzwojeń fazowych, współczynnik wypierania strumienia ku krańcom biegunów,
δ — szczelina powietrzna,
$\underline{\zeta}_{bv}$ — zespolony współczynnik rozłożenia uzwojenia w żłobkach,
ζ_{qv} — współczynnik grupowy,
ζ_{sv} — współczynnik skrótu,
Θ_1 — amplituda fali przepływu stojana,
ϑ_1 — wartość chwilowa przepływu stojana,
μ_o — przenikalność magnetyczna próżni,
v — harmoniczne rozkładu pola wzdłuż podziałki biegunowej,
Φ_v — argument zespolonego współczynnika $\underline{\zeta}_{bv}$,
τ — podziałka biegunowa,
ω_1 — pulsacja prądu w uzwojeniach stojana.

Indeksy

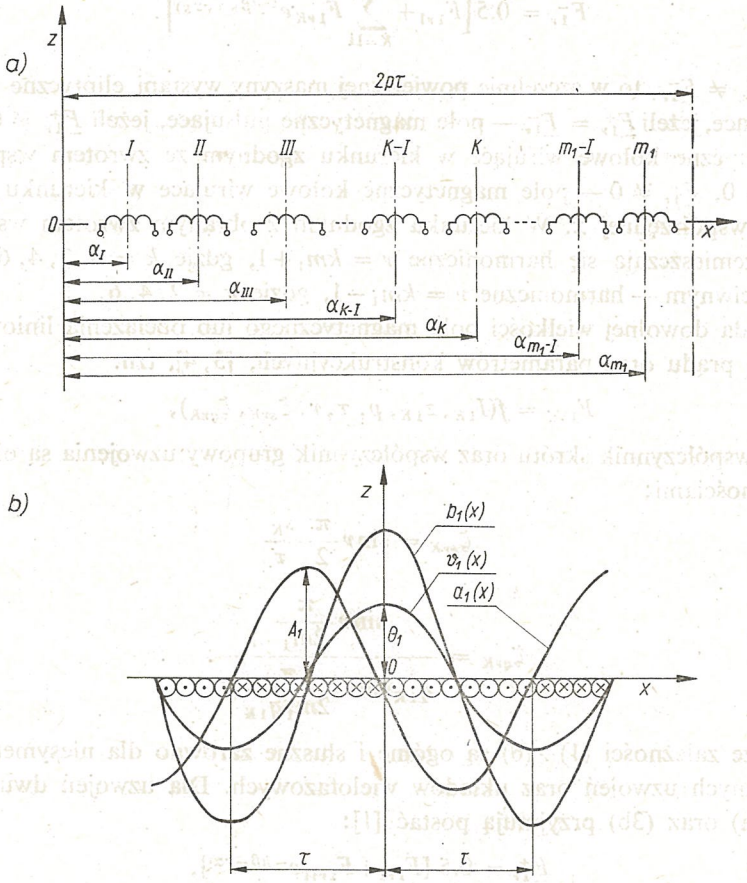
1 — stojan,
2 — wirnik,
K — numer fazy ($K = \mathrm{I}, \mathrm{II}, \ldots, m_1$),
n — harmoniczna n,
$\dot{z}$ — żłobek,
v — harmoniczna v,
„+" — kierunek zgodny ze zwrotem współrzędnej x,
„−" — kierunek przeciwny do zwrotu współrzędnej x.

2. ZALEŻNOŚCI PODSTAWOWE

Niech stojan maszyny indukcyjnej posiada uzwojenie m_1-fazowe. Pomiędzy dowolnie obranym początkiem układu współrzędnych a osią magnetyczną fazy K istnieje przesunięcie w przestrzeni o kąt α_K [9] — rys. 1a. Na ogół początek układu współrzędnych znaj-

duje się w osi symetrii magnetycznej uzwojenia fazy I (w połowie długości i na powierzchni przyszczelinowej rdzenia stojana), zatem dla fazy I kąt $\alpha_I = 0$. Pomiędzy wartościami maksymalnymi prądów płynących w uzwojeniach istnieje przesunięcie fazowe β_K, tzn.

$$\underline{I}_{1K} = \sqrt{2} I_{1K} e^{j(\omega_1 t - \beta_K)}, \tag{1}$$



Rys. 1. Uzwojenie rozwiniętego na płaszczyznę stojana maszyny indukcyjnej: a) rozmieszczenie poszczególnych uzwojeń fazowych wzdłuż obwodu; b) rozkłady indukcji magnetycznej $b_1(x)$, przepływu $\vartheta_1(x)$ oraz obciążenia liniowego $a_1(x)$

więc przyjmuje się również $\beta_I = 0$. Przy symetrycznym rozmieszczeniu poszczególnych uzwojeń fazowych kąt przestrzenny $\alpha_K = (K-I)\dfrac{2\pi}{m_1}$.

Rozkład indukcji, natężenia pola magnetycznego, przepływu lub obciążenia liniowego stojana można wyrazić zależnością ogólną

$$\underline{f}_1(x,t) = \sum_{\nu=1}^{\infty} \left[\underline{F}^{+}_{1\nu} e^{j\left(\omega_1 t - \nu \frac{\pi}{\tau} x\right)} + \underline{F}^{-}_{1\nu} e^{j\left(\omega_1 t + \nu \frac{\pi}{\tau} x\right)} \right], \tag{2}$$

w której amplitudy zespolone wielkości przemieszczających się w kierunku zgodnym ze zwrotem współrzędnej x oraz przeciwnym wynoszą:

$$\underline{F}_{1\nu}^{+} = 0{,}5\left[F_{1\nu\mathrm{I}} + \sum_{K=\mathrm{II}}^{m_1} F_{1\nu K} e^{-j(\beta_K - \nu\alpha_K)}\right], \tag{3a}$$

$$\underline{F}_{1\nu}^{-} = 0{,}5\left[F_{1\nu\mathrm{I}} + \sum_{K=\mathrm{II}}^{m_1} F_{1\nu K} e^{-j(\beta_K + \nu\alpha_K)}\right]. \tag{3b}$$

Jeżeli $\underline{F}_{1\nu}^{+} \neq \underline{F}_{1\nu}^{-}$, to w szczelinie powietrznej maszyny wystąpi eliptyczne pole magnetyczne wirujące, jeżeli $\underline{F}_{1\nu}^{+} = \underline{F}_{1\nu}^{-}$ — pole magnetyczne pulsujące, jeżeli $\underline{F}_{1\nu}^{+} \neq 0$, $\underline{F}_{1\nu}^{-} = 0$ — pole magnetyczne kołowe wirujące w kierunku zgodnym ze zwrotem współrzędnej x, jeżeli $\underline{F}_{1\nu}^{+} = 0$, $\underline{F}_{1\nu}^{-} \neq 0$ — pole magnetyczne kołowe wirujące w kierunku przeciwnym do zwrotu współrzędnej x. W kierunku zgodnym z obranym zwrotem współrzędnej x (rys. 1) przemieszczają się harmoniczne $\nu = km_1 + 1$, gdzie $k = 0, 2, 4, 6, \ldots$, w kierunku przeciwnym — harmoniczne $\nu = km_1 - 1$, gdzie $k = 2, 4, 6, \ldots$

Amplituda dowolnej wielkości pola magnetycznego lub obciążenia liniowego stojana jest funkcją prądu oraz parametrów konstrukcyjnych, [3, 4], tzn.

$$F_{1\nu K} = f(I_{1K}, z_{1K}, p, \tau, \nu, \zeta_{s\nu K}, \zeta_{q\nu K}), \tag{4}$$

przy czym współczynnik skrótu oraz współczynnik grupowy uzwojenia są określone znanymi zależnościami:

$$\zeta_{s\nu K} = \sin\nu\frac{\pi}{2}\frac{s_K}{\tau} \tag{5}$$

$$\zeta_{q\nu K} = \frac{\sin\nu\dfrac{\pi}{2m_1}}{q_{1K}\sin\nu\dfrac{\pi}{2m_1 q_{1K}}}. \tag{6}$$

Powyższe zależności (1)÷(6) są ogólne i słuszne zarówno dla niesymetrycznych jak i symetrycznych uzwojeń oraz układów wielofazowych. Dla uzwojeń dwufazowych zależności (3a) oraz (3b) przyjmują postać [1]:

$$\underline{F}_{1\nu}^{+} = 0{,}5\,[F_{1\nu\mathrm{I}} + F_{1\nu\mathrm{II}} e^{-j(\beta - \nu\alpha)}], \tag{7a}$$

$$\underline{F}_{1\nu}^{-} = 0{,}5[F_{1\nu\mathrm{I}} + F_{1\nu\mathrm{II}} e^{-j(\beta + \nu\alpha)}]. \tag{7b}$$

Uzwojenia fazy I oraz II mogą zawierać różne liczby zwojów szeregowych $z_{1\mathrm{I}}$, $z_{1\mathrm{II}}$ oraz składać się z różnej liczby $q_{1\mathrm{I}}$, $q_{1\mathrm{II}}$ cewek o różnej rozpiętości $s_\mathrm{I} \leqslant \tau$, $s_\mathrm{II} \leqslant \tau$. Prądy płynące przez te uzwojenia mogą posiadać różne wartości skuteczne $I_{1\mathrm{I}}$, $I_{1\mathrm{II}}$ oraz przesunięcia fazowe β między sobą (np. w silnikach jednofazowych z fazą pomocniczą zwartą).

Jeżeli uzwojenia fazowe są rozmieszczone symetrycznie wzdłuż osi x, tzn. pomiędzy ich osiami magnetycznymi istnieje przesunięcie $\alpha = 2\pi/4 = \pi/2$, to amplitudy zespolone będą równe:

$$\underline{F}_{1\nu}^{+} = 0{,}5\left[F_{1\nu\mathrm{I}} + F_{1\nu\mathrm{II}} e^{-j\left(\beta - \nu\frac{\pi}{2}\right)}\right], \tag{8a}$$

$$\underline{F}_{1\nu}^{-} = 0{,}5\left[F_{1\nu\mathrm{I}} + F_{1\nu\mathrm{II}} e^{-j\left(\beta + \nu\frac{\pi}{2}\right)}\right]. \tag{8b}$$

Zależności (8a), (8b) są słuszne dla silników indukcyjnych z fazą pomocniczą R, L, C, silników dwufazowych wykonawczych sterowanych amplitudowo, fazowo lub amplitudowo-fazowo oraz prądnic tachometrycznych. Dla dwufazowego układu symetrycznego prądów zasilających, tzn. w którym $\beta = \frac{\pi}{2}$ oraz $I_{1\mathrm{I}} z_{1\mathrm{I}} = I_{1\mathrm{II}} z_{1\mathrm{II}} = I_1 z_1$ amplitudy zespolone przyjmują postać:

$$\underline{F}^{+}_{1\nu} = 0{,}5F_{1\nu}\left[1+e^{j\frac{\pi}{2}(\nu-1)}\right] = 0{,}5F_{1\nu}\frac{\sin\frac{\pi}{2}(\nu-1)}{\sin\frac{\pi}{4}(\nu-1)}\,e^{j(\nu-1)\frac{\pi}{4}}, \tag{9a}$$

$$\underline{F}^{-}_{1\nu} = 0{,}5F_{1\nu}\left[1+e^{-j\frac{\pi}{2}(\nu+1)}\right] = 0{,}5F_{1\nu}\frac{\sin\frac{\pi}{2}(\nu+1)}{\sin\frac{\pi}{4}(\nu+1)}\,e^{-j(\nu+1)\frac{\pi}{4}}. \tag{9b}$$

Zależności (9a), (9b) są słuszne dla silników, w których występuje kołowe wirujące pole magnetyczne.

3. WYŻSZE HARMONICZNE UZWOJEŃ ROZMIESZCZONYCH SYMETRYCZNIE

Na ogół wszystkie maszyny dwufazowe — z wyjątkiem silników jednofazowych z fazą pomocniczą zwartą — posiadają w stojanie uzwojenia rozmieszczone symetrycznie, tzn. o osiach magnetycznych przesuniętych względem siebie o kąt $\alpha = \frac{\pi}{2}$. W celu bliższej analizy pola magnetycznego stojana tych maszyn, zostaną rozpatrzone dwa przypadki skrajne, kiedy kąt $\beta = 0$ oraz kiedy $\beta = \frac{\pi}{2}$. Równania (8a), (8b) dla $\beta = 0$ przyjmują postać:

$$\underline{F}^{+}_{1\nu} = 0{,}5\left(F_{1\nu\mathrm{I}} + F_{1\nu\mathrm{II}} e^{j\nu\frac{\pi}{2}}\right), \tag{10a}$$

$$F^{-}_{1\nu} = 0{,}5\left(F_{1\nu\mathrm{I}} + F_{1\nu\mathrm{II}} e^{-j\nu\frac{\pi}{2}}\right). \tag{10b}$$

1° Jedno z uzwojeń jest rozwarte:

— jeżeli $F_{1\nu\mathrm{I}} \neq 0$, $F_{1\nu\mathrm{II}} = 0$, to dla wszystkich harmonicznych wystąpi pole pulsujące, które interpretuje się jako superpozycję dwóch pól o jednakowych amplitudach $0{,}5F_{1\nu\mathrm{I}}$ wirujących w przeciwnych kierunkach;

— jeżeli $F_{1\nu\mathrm{I}} = 0$, $F_{1\nu\mathrm{II}} \neq 0$, to wystąpi również dla wszystkich harmonicznych pole pulsujące, które interpretuje się jako superpozycję dwóch pól o jednakowych co do wartości bezwzględnej, lecz przeciwnych znaków amplitudach $0{,}5jF_{1\nu\mathrm{II}}$, $-0{,}5jF_{1\nu\mathrm{II}}$ wirujących w przeciwnych kierunkach; strzałka fali pulsującej, np. przepływu $F_{1\nu\mathrm{II}}$ przy $F_{1\nu\mathrm{I}} = 0$, będzie przesunięta o $\frac{\pi}{2}$ względem strzałki fali pulsującej $F_{1\nu\mathrm{I}}$ przy $F_{1\nu\mathrm{II}} = 0$.

2° W obydwu uzwojeniach płyną takie prądy, że $F_{1\nu\mathrm{I}} = F_{1\nu\mathrm{II}}$:

— harmoniczne nieparzyste $\nu = 1, 3, 5, 7, \ldots$ tworzą dwa pola pulsujące, które interpretuje się jako superpozycję czterech pól wirujących parami w przeciwnych kierunkach o amplitudach $0{,}5F_{1\nu I}$ oraz $0{,}5jF_{1\nu II}$, $-0{,}5jF_{1\nu II}$, przy czym strzałki o jednakowych długościach fal pulsujących $F_{1\nu I}$, $F_{1\nu II}$, np. przepływów, będą przesunięte o $\frac{\pi}{2}$ (przypadek ten można również sprowadzić do dwóch pól wirujących o amplitudach $\frac{1}{\sqrt{2}} F_{1\nu I}$);

— dla harmonicznych parzystych $\nu = 2, 6, 10, 14, \ldots$ fale, np. przepływów uzwojeń fazy I oraz fazy II wzajemnie znoszą się;

— harmoniczne parzyste $\nu = 4, 8, 12, 16, \ldots$ tworzą pole pulsujące, które interpretuje się jako superpozycją dwóch pól kołowych o jednakowych amplitudach $F_{1\nu I} = F_{1\nu II}$ wirujących w przeciwnych kierunkach.

Mimo, że analiza zależności (10a), (10b) wykazuje występowanie przestrzennych harmonicznych parzystych, nie mają jednak te harmoniczne znaczenia praktycznego.

Dla kąta $\beta = \frac{\pi}{2}$ równania (8a), (8b) są następujące:

$$\underline{F}^{+}_{1\nu} = 0{,}5\left[F_{1\nu I} + F_{1\nu II} e^{j\frac{\pi}{2}(\nu-1)}\right], \tag{11a}$$

$$\underline{F}^{-}_{1\nu} = 0{,}5\left[F_{1\nu I} + F_{1\nu II} e^{-j\frac{\pi}{2}(\nu+1)}\right]. \tag{11b}$$

1° W obydwu uzwojeniach płyną prądy takie, że amplitudy $F_{1\nu I} = F_{1\nu II}$:

— harmoniczne nieparzyste $\nu = 1, 5, 9, 13, \ldots$ tworzą pola kołowe współbieżne;

— harmoniczne nieparzyste $\nu = 3, 7, 11, 15, \ldots$ natomiast — pola kołowe przeciwbieżne.

Zatem w maszynie o polu kołowym

$$\underline{F}^{+}_{1\nu} = F_{1\nu} e^{j(\nu-1)\frac{\pi}{4}} \quad \text{dla} \quad \frac{\nu-1}{4} = 0, 1, 2, 3, \ldots, \tag{12a}$$

$$\underline{F}^{-}_{1} = F_{1\nu} e^{-j(\nu+1)\frac{\pi}{4}} \quad \text{dla} \quad \frac{\nu+1}{4} = 0, 1, 2, 3, \ldots, \tag{12b}$$

gdyż tylko dla $(\nu \mp 1)/4 = 0, 1, 2, 3, \ldots$ wyrażenie $\dfrac{\sin\frac{\pi}{2}(\nu \mp 1)}{\sin\frac{\pi}{4}(\nu \mp 1)}$ jest różne od zera i równe 2 — zależności (9a), (9b).

2° W obydwu uzwojeniach płyną prądy takie, że $F_{1\nu I} \neq F_{1\nu II}$; wtedy harmoniczne nieparzyste tworzą pola wirujące eliptyczne.

W kierunku zgodnym ze zwrotem współrzędnej x wirują harmoniczne $\nu = 2k+1$, gdzie $k = 0, 2, 4, 6, \ldots$, w kierunku przeciwnym — harmoniczne $\nu = 2k-1$, gdzie $k = 2, 4, 6, \ldots$ Spośród wyższych harmonicznych największy udział ma trzecia harmoniczna, objawiająca się występowaniem punktu przegięcia krzywej momentu przy 1/3

prędkości synchronicznej. Zakleśnięcie w krzywej momentu utrudnia lub uniemożliwia rozruch silnika pod obciążeniem. W celu ograniczenia trzeciej harmonicznej wykonuje się specjalne, tzw. uzwojenia z kompensacją trzeciej harmonicznej [6] lub uzwojenia sinusoidalne [6, 8].

4. KRZYWA WZBUDZENIA STOJANA O BIEGUNACH UTAJONYCH

Rozkład przepływu stojana wzdłuż podziałki biegunowej — zgodnie z zależnościami (2), (7a), (7b) — można zapisać następująco

$$\underline{\vartheta}_1(x,t) = \sum_{\nu=1}^{\infty}\left[\underline{\Theta}_{1\nu}^{+} e^{j\left(\omega_1 t - \nu\frac{\pi}{\tau}x\right)} + \underline{\Theta}_{1\nu}^{-} e^{j\left(\omega_1 t + \nu\frac{\pi}{\tau}x\right)}\right], \tag{13}$$

gdzie:

$$\underline{\Theta}_{1\nu}^{+} = 0{,}5\,[\Theta_{1\nu\mathrm{I}} + \Theta_{1\nu\mathrm{II}} e^{-j(\beta-\nu\alpha)}], \tag{14a}$$

$$\underline{\Theta}_{1\nu}^{-} = 0{,}5\,[\Theta_{1\nu\mathrm{I}} + \Theta_{1\nu\mathrm{II}} e^{-j(\beta+\nu\alpha)}]. \tag{14b}$$

Dla stojana o biegunach utajonych amplitudy przepływów uzwojeń fazy I lub II są równe

$$\Theta_{1\nu K} = \frac{2\sqrt{2}}{\pi}\,\frac{z_{1K}\zeta_{s\nu K}\zeta_{q\nu K}}{p}\,\frac{1}{\nu}\,I_{1K}. \tag{15}$$

Wzór (15) nie uwzględnia użłobkowania rdzenia stojana.

5. KRZYWA OBCIĄŻENIA LINIOWEGO STOJANA O BIEGUNACH UTAJONYCH

Krzywa obciążenia liniowego stojana jest krzywą różniczkową krzywej wzbudzenia

$$\underline{a}_1(x,t) = \frac{d[\vartheta_1(x,t)]}{dx} = \sum_{\nu=1}^{\infty}\left(-j\nu\frac{\pi}{\tau}\right)\left[\underline{\Theta}_{1\nu}^{+} e^{j\left(\omega_1 t - \nu\frac{\pi}{\tau}x\right)} - \underline{\Theta}_{1\nu}^{-} e^{j\left(\omega_1 t + \nu\frac{\pi}{\tau}x\right)}\right] =$$

$$= \sum_{\nu=1}^{\infty}\left[\underline{A}_{1\nu}^{+} e^{j\left(\omega_1 t - \nu\frac{\pi}{\tau}x\right)} + \underline{A}_{1\nu}^{-} e^{j\left(\omega_1 t + \nu\frac{\pi}{\tau}x\right)}\right], \tag{16}$$

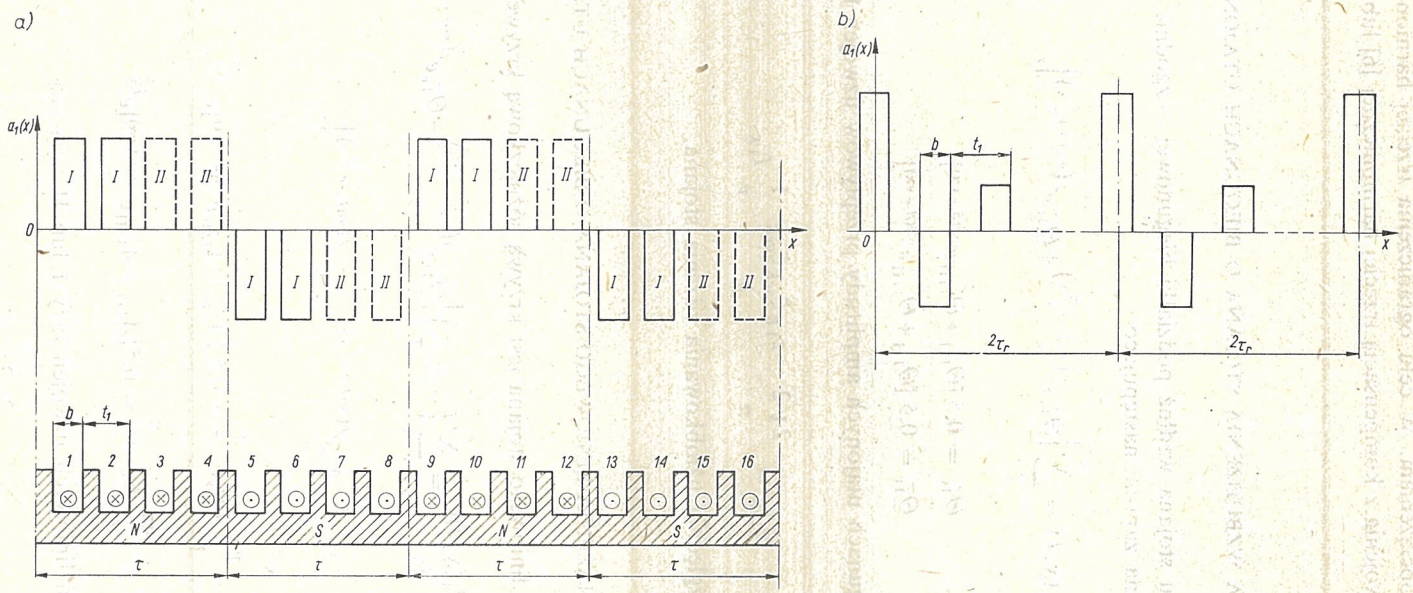
gdzie:

$$\underline{A}_{1\nu}^{+} = \frac{\nu\pi}{\tau}\,\underline{\Theta}_{1\nu}^{+} e^{-j\frac{\pi}{2}} = 0{,}5\,[A_{1\nu\mathrm{I}} + A_{1\nu\mathrm{II}} e^{-j(\beta-\nu\alpha)}] e^{-j\frac{\pi}{2}}, \tag{17a}$$

$$\underline{A}_{1\nu}^{-} = \frac{\nu\pi}{\tau}\,\underline{\Theta}_{1\nu}^{-} e^{j\frac{\pi}{2}} = 0{,}5[A_{1\nu\mathrm{I}} + A_{1\nu\mathrm{II}} e^{-j(\beta+\nu\alpha)}] e^{j\frac{\pi}{2}}. \tag{17b}$$

Amplitudy obciążeń liniowych uzwojeń fazy I lub II wynoszą

$$A_{1\nu K} = \frac{2\sqrt{2}z_{1K}\zeta_{s\nu K}\zeta_{q\nu K}}{\tau p}\,I_{1K}. \tag{18}$$



Rys. 2. Rozkład obciążenia liniowego żłobków stojana maszyny dwufazowej: a) o uzwojeniu jednowarstwowym ($2p = 4$, $q_1 = 2$); b) o uzwojeniu dwuwarstwowym [7]

Krzywa wzbudzenia (13) jest przesunięta względem krzywej obciążenia liniowego (16) o kąt $\frac{\pi}{2}$ — rys. 1b.

Krzywa obciążenia liniowego (16) uzyskana z krzywej wzbudzenia (13) opisuje rozkład gęstości liniowej prądu w nieskończenie cienkiej, równomiernej warstwie, która symuluje uzwojenie idealne, rozłożone równomiernie na bezżłobkowej powierzchni przyszczelinowej rdzenia stojana. Można wykazać, że istnieje możliwość sprowadzenia każdego uzwojenia rzeczywistego rozłożonego w żłobkach do równoważnej, nieskończenie cienkiej, równomiernie rozłożonej warstwy prądowej. Na rys. 2a przedstawiono uzwojenie jednowarstwowe, wykonane z cewek średnicowych. Obciążenie liniowe żłobka o szerokości b zawierającego N_K przewodów fazy I lub II wynosi

$$\underline{A}_{\dot{z}K} = \frac{N_K}{b}\sqrt{2}I_{1K}e^{j(\omega_1 t-\beta_K)}. \tag{19}$$

Rozpatrując pojedyncze uzwojenie fazowe widać, że krzywa obciążenia liniowego spełnia warunek $f(x) = -f(x+\tau)$. Przyjmuje ona następujące wartości w przedziałach:

$0 \leqslant x \leqslant 0{,}5(t_1-b)$	$\underline{a}_{1K}(x,t) = 0,$
$0{,}5(t_1-b) \leqslant x \leqslant 0{,}5(t_1+b)$	$\underline{a}_{1K}(x,t) = \underline{A}_{\dot{z}K},$
$0{,}5(t_1+b) \leqslant x \leqslant 0{,}5(3t_1-b)$	$\underline{a}_{1K}(x,t) = 0,$
$0{,}5(3t_1-b) \leqslant x \leqslant 0{,}5(3t_1+b)$	$\underline{a}_{1K}(x,t) = \underline{A}_{\dot{z}K},$
$0{,}5(3t_1+b) \leqslant x \leqslant 0{,}5(5t_1-b)$	$\underline{a}_{1K}(x,t) = 0,$
.....	.....
$0{,}5[(q_{1K}+1)t_1-b] \leqslant x \leqslant 0{,}5[(q_{1K}+1)t_1+b]$	$\underline{a}_{1K}(x,t) = \underline{A}_{\dot{z}K},$
$0{,}5[(q_{1K}+1)t_1+b] \leqslant x \leqslant \tau+0{,}5(t_1-b)$	$\underline{a}_{1K}(x,t) = 0.$

Tak opisaną funkcję rozkłada się następnie w przedziale $0 \leqslant x \leqslant \tau$ na szereg Fourierа

$$\underline{a}_{1K}(x,t) = \underline{A}_{\dot{z}K}\sum_{\nu=1}^{\infty}\left(a_{1\nu}\cos\nu\frac{\pi}{\tau}x + c_{1\nu}\sin\nu\frac{\pi}{\tau}x\right),$$

którego współczynniki wynoszą:

$$a_{1\nu} = \frac{4q_{1K}}{\nu\pi}\frac{\sin\nu\frac{\pi}{2}\frac{q_{1K}t_1}{\tau}}{q_{1K}\sin\nu\frac{\pi}{2}\frac{t_1}{\tau}}\sin\nu\frac{\pi}{2}\frac{b}{\tau}\cos\nu\frac{\pi}{2}\frac{t_1 q_{1K}}{\tau} =$$

$$= \frac{2bq_{1K}}{\tau}\zeta_{q\nu K}\frac{\sin\nu\frac{\pi}{2}\frac{b}{\tau}}{\nu\frac{\pi}{2}\frac{b}{\tau}}\cos\nu\frac{\pi}{2}\frac{t_1 q_{1K}}{\tau}$$

$$c_{1\nu} = \frac{4q_{1K}}{\nu\pi}\frac{\sin\nu\frac{\pi}{2}\frac{q_{1K}t_1}{\tau}}{q_{1K}\sin\nu\frac{\pi}{2}\frac{t_1}{\tau}}\sin\nu\frac{\pi}{2}\frac{b}{\tau}\sin\nu\frac{\pi}{2}\frac{t_1 q_{1K}}{\tau} =$$

$$= \frac{2bq_{1K}}{\tau}\zeta_{q\nu K}\frac{\sin\nu\frac{\pi}{2}\frac{b}{\tau}}{\nu\frac{\pi}{2}\frac{b}{\tau}}\sin\nu\frac{\pi}{2}\frac{t_1 q_{1K}}{\tau}.$$

Ponieważ $m_1 q_{1K} t_1 = \tau$, więc współczynnik grupowy (6) wynosi

$$\zeta_{q\nu K} = \frac{\sin\nu\frac{\pi}{2}\frac{q_{1K}t_1}{\tau}}{q_{1K}\sin\nu\frac{\pi}{2}\frac{t_1}{\tau}}.$$

Współczynnik skrótu natomiast $\zeta_{s\nu} = 1$, gdyż przedmiotem rozważań jest uzwojenie z cewek średnicowych.

Szereg Fouriera wygodniej jest doprowadzić do postaci

$$\underline{a}_{1K}(x,t) = \underline{A}_{\dot{z}K}\sum_{\nu=1}^{\infty} C_{1\nu}\cos\left(\nu\frac{\pi}{\tau}x + \Phi_\nu - \frac{\pi}{2}\right), \tag{20}$$

gdzie:

$$C_{1\nu} = \sqrt{a_{1\nu}^2 + c_{1\nu}^2},$$

$$\Phi_\nu = \operatorname{arctg}\frac{a_{1\nu}}{c_{1\nu}}.$$

Uwzględniając, że $N_K q_{1K} = z_{1K}/p$, otrzymuje się ostatecznie krzywą obciążenia liniowego stojana przez pojedyncze uzwojenie fazowe w postaci

$$\underline{a}_{1K}(x,t) = \sum_{\nu=1}^{\infty}\left[\underline{A}_{1\nu K}^{+}e^{j\left(\omega_1 t - \beta_K - \nu\frac{\pi}{\tau}x\right)} + \underline{A}_{1\nu K}^{-}e^{j\left(\omega_1 t - \beta_K + \nu\frac{\pi}{\tau}x\right)}\right], \tag{21}$$

gdzie:

$$\underline{A}_{1\nu K}^{+} = 0{,}5A_{1\nu K}e^{-j\Phi_\nu}e^{j\frac{\pi}{2}}, \tag{21a}$$

$$\underline{A}_{1\nu K}^{-} = 0{,}5A_{1\nu K}e^{j\Phi_\nu}e^{-j\frac{\pi}{2}}, \tag{21b}$$

przy czym

$$A_{1\nu K} = \frac{2\sqrt{2}z_{1K}\zeta_{q\nu K}\zeta_{b\nu K}}{\tau p}I_{1K}. \tag{22}$$

Współczynnik zespolony

$$\underline{\zeta}_{b\nu K} = \zeta_{b\nu K}e^{\mp j\Phi_\nu} = \frac{\sin\nu\frac{\pi}{2}\frac{b}{\tau}}{\nu\frac{\pi}{2}\frac{b}{\tau}}e^{\mp j\Phi_\nu} \tag{23}$$

uwzględnia rozkład uzwojenia w żłobkach. Jego moduł $\zeta_{b\nu K}$ wskazuje o ile amplitudy fal obciążenia liniowego stojana są mniejsze od tychże amplitud w przypadku krzywej idealnej obciążenia liniowego otrzymanej przez zróżniczkowanie krzywej wzbudzenia (13). Dla uzwojenia idealnego, tzn. w którym $b \to 0$ oraz $q_{1K}t_1 \to m_1 q_{1K} t_1 = \tau$, moduł $\zeta_{b\nu K} \to 1$

oraz argument $\Phi_\nu \to 0$. Zatem wyrażenie (22) dla uzwojenia rzeczywistego przechodzi w wyrażenie (18) dla uzwojenia idealnego.

Sumując geometrycznie krzywe obciążenia liniowego fazy I oraz fazy II otrzyma się wypadkową krzywą obciążenia liniowego stojana — wyrażenie (16), w którym:

$$\underline{A}_{1\nu}^{+} = 0{,}5[A_{1\nu I} + A_{1\nu II} e^{-j(\beta - \nu\alpha)}] e^{-j\Phi_\nu} e^{j\frac{\pi}{2}}, \tag{24a}$$

$$\underline{A}_{1\nu}^{-} = 0{,}5[A_{1\nu I} + A_{1\nu II} e^{-j(\beta + \nu\alpha)}] e^{j\Phi_\nu} e^{-j\frac{\pi}{2}}. \tag{24b}$$

Różnica w przesunięciach przestrzennych $-\pi$ dla pola współbieżnego oraz π dla pola przeciwbieżnego — wyrażenia (17a), (17b), (24a), (24b) — wynika z innego obrania początku układu współrzędnych na rys. 1 oraz na rys. 2.

Podobnie dla uzwojenia dwuwarstwowego ze skrótem, przebieg obciążenia liniowego poszczególnych żłobków w przedziale $2p\tau$ powtarza się co najmniej jeden raz, przy czym impulsy prostokątne reprezentujące obciążenie liniowe żłobków mają niejednakową wysokość — rys. 2b [7]. Przedział, w którym przebieg obciążenia liniowego odpowiada pełnemu okresowi wynosi $2\tau_r$, przy czym $\tau_r \neq \tau$; w przedziale $2\tau_r$ znajduje się r żłobków. Funkcję przedstawioną na rys. 2b rozkłada się następnie w przedziale $-\tau_r \leqslant x \leqslant \tau_r$ na szereg Fouriera

$$\underline{a}_1(x, t) = \sum_{\dot{z}=1}^{r} \underline{A}_{\dot{z}} \sum_{\nu=1}^{\infty} \left(a_{1\nu} \cos\nu \frac{\pi}{\tau_r} x + c_{1\nu} \sin\nu \frac{\pi}{\tau_r} \right). \tag{25}$$

w którym:

$$a_{1\nu} = \frac{\zeta'_{b\nu}}{\tau_r} \sum_{\dot{z}=1}^{r} N_{\dot{z}} \cos\left[\nu \frac{\pi}{\tau_r} (\dot{z} - 1) t_1\right],$$

$$c_{1\nu} = \frac{\zeta'_{b\nu}}{\tau_r} \sum_{\dot{z}=1}^{r} N_{\dot{z}} \sin\left[\nu \frac{\pi}{\tau_r} (\dot{z} - 1) t_1\right],$$

$$\zeta'_{b\nu} = \frac{\sin\nu \frac{\pi}{2} \frac{b}{\tau_r}}{\nu \frac{\pi}{2} \frac{b}{\tau_r}},$$

$$\underline{A}_{\dot{z}} = \frac{N_{\dot{z}}}{b} \sqrt{2} I_{1\dot{z}} e^{j\omega_1 t}, \tag{26}$$

gdzie: $N_{\dot{z}}$ — liczba przewodów w żłobku, $I_{1\dot{z}}$ — prąd w pojedynczym przewodzie.

Podstawowym ograniczeniem w takim postępowaniu [7] jest nieuwzględnienie przesunięć fazowych prądów w górnej i dolnej warstwie (w jednym żłobku mogą znajdować się przewody różnych faz), zastosowanie tylko do zasilania symetrycznego (równość wartości skutecznych prądów fazowych) oraz możliwość pojawienia się w krzywej $\underline{a}_1(x, t)$ — równanie (25) harmonicznych przestrzennych, które w rzeczywistości nie występują. Szereg (25) przekształca się następnie do postaci (20).

6. ROZKŁAD OBCIĄŻENIA LINIOWEGO WZDŁUŻ OSI MASZYNY

Obciążenie liniowe wzdłuż prętów uzwojeń ułożonych w żłobkach posiada wartość stałą. Poza rdzeniem stojana, rozkład obciążenia liniowego wzdłuż połączeń czołowych zależy od konstrukcji uzwojenia. Przyjmując do rozważań zamiast pojedynczej maszyny indukcyjnej ciąg maszyn identycznych, w których osie y pokrywają się ze sobą i prądy w prętach mają zwroty przeciwne, można rozkład osiowy obciążenia liniowego uważać za funkcję okresową spełniającą warunki Dirichleta. Taka interpretacja rozkładu osiowego obciążenia liniowego została zaproponowana przez Wooda i Concordię [10], którzy zakładają, że stojany idealne, bez połączeń czołowych, stykają się ze sobą powierzchniami czołowymi rdzeni. W związku z tym osiowy rozkład obciążenia liniowego ma kształt okresowej funkcji prostokątnej o okresie równym długości idealnej L_i rdzenia stojana. Zatem $\underline{a}_1(x, y, t) = \underline{a}_1(x, t)$ dla $0 \leqslant y \leqslant 0{,}5L_i$ oraz $\underline{a}_1(x, y, t) = -a_1(x, t)$ dla $0{,}5L_i \leqslant y \leqslant L_i$ (początek układu współrzędnych znajduje się w połowie długości rdzenia). Rozkładając tak opisaną funkcję na szereg Fouriera

$$\underline{a}_1(x, y, t) = \sum_{n=1,3,5}^{\infty} a_{1n}\underline{a}_1(x, t)\cos n\frac{\pi}{L_i}y, \tag{27}$$

otrzymuje się

$$a_{1n} = \frac{4}{n\pi}(-1)^{\frac{n-1}{2}}, \tag{28}$$

gdzie $n = 1, 3, 5, \ldots$

Przedstawienie osiowego rozkładu obciążenia liniowego za pomocą funkcji prostokątnej [10] ma ograniczone zastosowanie i nie może być stosowane, jeżeli zamierza się rozpatrywać pole elektromagnetyczne poza rdzeniem w obszarach $y \geqslant \mp 0{,}5L_i$.

7. PRZESTRZENNOCZASOWY ROZKŁAD OBCIĄŻENIA LINIOWEGO

Wstawiając równanie (16) do równania (27) otrzymuje się rozkład przestrzennoczasowy obciążenia liniowego stojana

$$\underline{a}_1(x, y, t) = \sum_{\nu=1}^{\infty}\sum_{n=1}^{\infty} a_{1n}\left[\underline{A}_{1\nu}^{+}e^{j\left(\omega_1 t - \nu\frac{\pi}{\tau}x\right)} + \underline{A}_{1\nu}^{-}e^{j\left(\omega_1 t + \nu\frac{\pi}{\tau}x\right)}\right]\cos n\frac{\pi}{L_i}y. \tag{29}$$

Uzwojenia stojana o biegunach utajonych są zatem zastąpione warstwą prądową o gęstości liniowej wyrażonej zależnością (29). Powyższa zależność (29) jest wykorzystywana do wyznaczania rozkładu pola elektromagnetycznego w maszynach indukcyjnych o biegunach utajonych w stojanie. Zakładając nieskończenie wielką przenikalność magnetyczną rdzenia stojana, składowa magnetyczna pola w kierunku osi x na powierzchni przyszczelinowej rdzenia stojana wynosi [2]

$$\underline{H}_{1x}(x, y, z = 0, t) = -\underline{a}_1(x, y, t). \tag{30}$$

8. KRZYWA POLA STOJANA O BIEGUNACH UTAJONYCH

W przypadku stojana o biegunach utajonych, rozkład indukcji magnetycznej wzdłuż podziałki biegunowej — według klasycznego ujęcia Richtera — można wyznaczyć na podstawie krzywej wzbudzenia (13)

$$\underline{b}_1(x,t) = \frac{\vartheta_1(x,t)}{\delta k_c k_n} = \sum_{\nu=1}^{\infty}\left[\underline{B}_{1\nu}^{+} e^{j\left(\omega_1 t - \nu\frac{\pi}{\tau}x\right)} + \underline{B}_{1\nu}^{-} e^{j\left(\omega_1 t + \nu\frac{\pi}{\tau}x\right)}\right], \tag{31}$$

gdzie:

$$\underline{B}_{1\nu}^{+} = \frac{\underline{\Theta}_{1\nu}^{+}}{\delta k_c k_n}, \tag{32a}$$

$$\underline{B}_{1\nu}^{-} = \frac{\underline{\Theta}_{1\nu}^{-}}{\delta k_c k_n}. \tag{32b}$$

Do wyznaczenia krzywej pola, podobnie jak i krzywych wzbudzenia oraz obciążenia liniowego, potrzebna jest znajomość prądów w uzwojeniach fazy I oraz II.

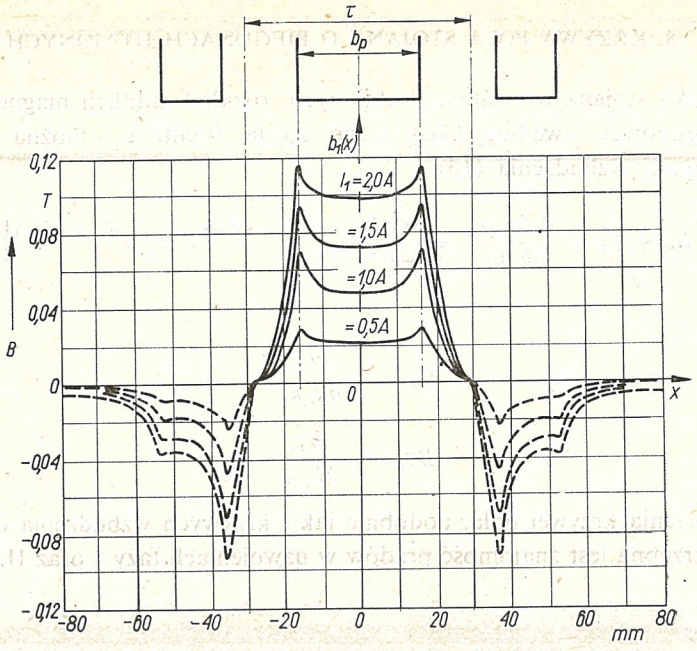
9. ROZKŁAD INDUKCJI MAGNETYCZNEJ UZWOJENIA STOJANA O BIEGUNACH JAWNYCH

Uzwojenia stojana o biegunach jawnych dają dyskretny rozkład gęstości liniowej prądu na powierzchni przyszczelinowej. Sprowadzenie uzwojeń stojana o biegunach jawnych do równomiernie rozłożonej warstwy prądowej — jak dla maszyn o biegunach utajonych — (rozdz. 5) nie ma uzasadnienia. W konsekwencji równanie (30) nie może być punktem wyjściowym procedury wyznaczania rozkładu pola elektromagnetycznego w maszynie. Zamiast składowej stycznej H_x natężenia pola magnetycznego na powierzchni przyszczelinowej stojana, lepiej jest rozpocząć od wyznaczania składowej normalnej H_z na powierzchni przyszczelinowej wirnika, w oparciu o krzywą pola w szczelinie powietrznej. Do tego celu najlepiej jest wykorzystać wyznaczony doświadczalnie rozkład indukcji magnetycznej pod biegunem. Badanie rozkładu pola magnetycznego w rzeczywistej maszynie indukcyjnej o biegunach jawnych jest bardzo trudne, m.in. z uwagi na małą szczelinę powietrzną, która na ogół jest zawarta w przedziale od 0,1 do 0,5 mm. Z tego powodu wykonano model fizyczny, w którym element aktywny symulujący stojan jest pakietowany z blach elektrotechnicznych gorącowalcowanych o grubości 0,5 mm w kształcie litery *E*. Szerokość kolumny środkowej z uzwojeniem o liczbie zwojów 615 jest dwa razy większa od kolumn skrajnych nieuzwojonych; dlatego też nie bierze się pod uwagę rozkładu pola magnetycznego zmierzonego na modelu dla $x > \mp\frac{\tau}{2}$. W celu wyeliminowania wpływu użłobkowania wirnika na rozkład pola magnetycznego stojana, przyjęto wirnik o parametrach rozłożonych, który symulują gładkie płyty prostokątne: miedziana — od strony szczeliny powietrznej oraz stalowa.

Dane modelu:

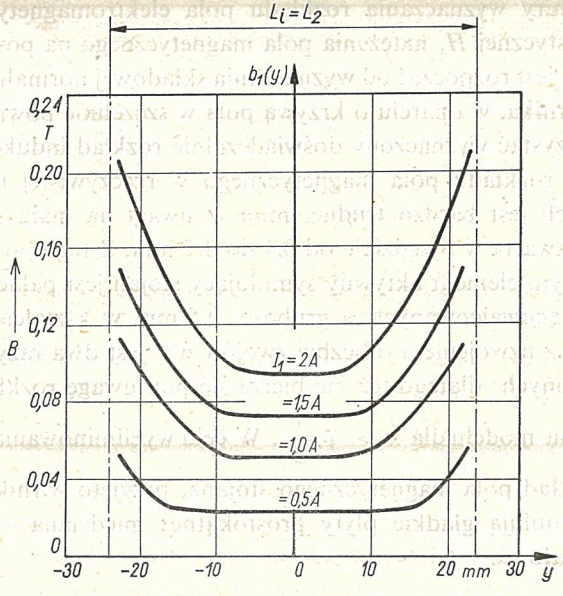
— podziałka biegunowa $\tau = 60$ mm,

— szerokość bieguna $b_p = 32$ mm,



Rys. 3. Rozkład składowej normalnej indukcji magnetycznej w szczelinie powietrznej wzdłuż podziałki biegunowej stojana o biegunach jawnych przy różnych prądach — pomiary sondą Halla na modelu, $\delta = 2$ mm, $d = 5$ mm, $y = 0$

Rozkład indukcii odpowiada warunkom w rzeczywistej maszynie indukcyjnej o biegunach jawnych tylko w przedziale $-0{,}5\,\tau \leqslant x \leqslant 0{,}5\,\tau$



Rys. 4. Rozkład składowej normalnej indukcji magnetycznej w szczelinie powietrznej wzdłuż współrzędnej y przy różnych prądach stojana o biegunach jawnych — pomiary sondą Halla na modelu, $\delta = 2$ mm, $d = 5$ mm, $x = 0$

— długość rdzenia bieguna $L_1 = 48$ mm,
— długość rdzenia wirnika $L_2 = 48$ mm,
— grubość litego rdzenia wirnika (płyty stalowej) 5 mm,
— grubość nakładki miedzianej $d = 5$ mm,
— szczelina powietrza $\delta = 2$ mm.

Składową normalną indukcji magnetycznej na powierzchni wirnika mierzono za pomocą teslomierza hallotronowego TH 23. Na rys. 3 przedstawiono rozkład składowej normalnej indukcji magnetycznej wzdłuż współrzędnej x ($y = 0$), na rys. 4 natomiast — wzdłuż współrzędnej y ($x = 0$) przy różnych prądach w uzwojeniu stojana.

9.1. Rozkład indukcji magnetycznej wzdłuż podziałki biegunowej

W pierwszym przybliżeniu można przyjąć, że rozkład składowej normalnej indukcji magnetycznej pod biegunem ma kształt prostokąta — rys. 5a. Jak wynika z badań eksperymentalnych — rys. 3, strumień magnetyczny wychodzi poza rdzeń bieguna przyjmując wartości zerowe dla $x = \mp\frac{\tau}{2}$, a nie dla $x = \mp\frac{b_p}{2}$. Wobec tego można uważać, że indukcja poza szerokością b_p nabiegunnika maleje prostoliniowo — rys. 5b lub wykładniczo — rys. 5c. Oprócz wychodzenia strumienia poza rdzeń bieguna, zachodzi również zjawisko wypierania strumienia pod biegunem, ku jego krawędziom. Przy dużym nasileniu tego zjawiska rozkład indukcji zmierzony pod biegunem będzie zbliżony do jednej z trzech funkcji okresowych przedstawionych na rys. 6. Po wyznaczeniu eksperymentalnym, należy rozkład indukcji pod biegunem przybliżyć do odpowiedniej, najbardziej adekwatnej funkcji okresowej przedstawionej na rys. 5 lub rys. 6. Funkcje według rys. 5 oraz rys. 6 opisujące rozkład indukcji K-tej fazy należy następnie rozłożyć w przedziale $0 \leqslant x \leqslant \tau$ na szereg Fouriera

$$\underline{b}_{1K}(x, t) = \sum_{\nu=1,3,5}^{\infty} b_{1\nu}\underline{B}_K(t)\cos\nu\frac{\pi}{\tau}x. \tag{33}$$

Pomijając zjawisko wypierania strumienia ku krawędziom biegunów, indukcja pod biegunem N przyjmuje wartość stałą równą $\underline{b}_{1K}(x, t) = \underline{B}_K(t)$, pod biegunem S natomiast $\underline{b}_{1K}(x, t) = -\underline{B}_K(t)$. Przy uwzględnieniu zjawiska wypierania strumienia ku krawędziom biegunów, indukcja pod biegunem N zmienia się wzdłuż współrzędnej x według zależności $\underline{b}_{1K}(x, t) = \underline{B}_K(t)\,\mathrm{ch}\,2\alpha\frac{x}{b_p}$, pod biegunem S zaś — $b_{1K}(x, t) = -\underline{B}_K(t)\,\mathrm{ch}\,2\alpha\frac{x}{b_p}$. Współczynnik $\alpha \geqslant 0$ wyraża stopień wypierania strumienia pod biegunem ku jego krawędziom. Dla $\alpha = 0$ nie występuje wypieranie strumienia ku krawędziom. Jedna z metod obliczania rozkładu indukcji pod biegunem jest podana w pracy [5]. Pomiędzy biegunami różnoimiennymi indukcja magnetyczna przyjmuje następujące wartości:

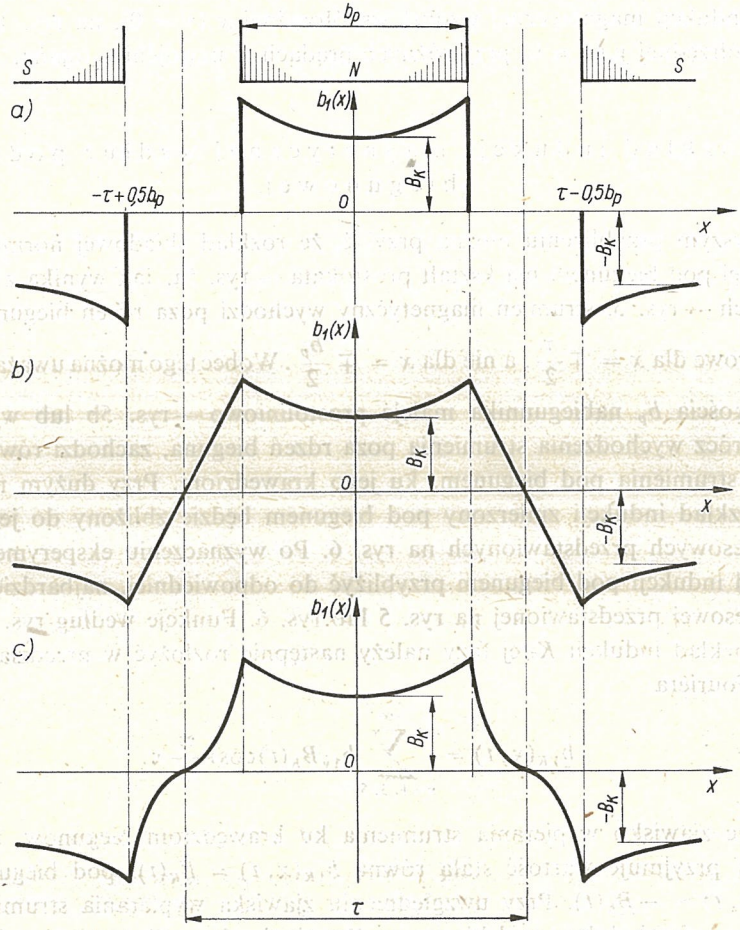
— wg rys. 5a oraz rys. 6a $\underline{b}_{1K}(x, t) = 0$,

— wg rys. 5b $\underline{b}_{1K}(x, t) = \frac{2\underline{B}_K(t)}{\tau - b_p}\left(\frac{\tau}{2} \mp x\right)$,

— wg rys. 5c $$\underline{b}_{1K}(x,t) = \underline{B}_K(t)\left(\frac{2}{\tau - b_p}\right)^3 \left(\frac{\tau}{2} \mp x\right)^3,$$

— wg rys. 6b $$\underline{b}_{1K}(x,t) = \frac{2\underline{B}_K(t)}{\tau - b_p}\left(\frac{\tau}{2} \mp x\right)\operatorname{ch}\alpha,$$

— wg rys. 6c $$\underline{b}_{1K}(x,t) = \underline{B}_K(t)\left(\frac{2}{\tau - b_p}\right)^3 \left(\frac{\tau}{2} \mp x\right)^3 \operatorname{ch}\alpha.$$



Rys. 5. Aproksymacje krzywej rozkładu indukcji magnetycznej stojana jawnobiegunowego wzdłuż podziałki biegunowej bez uwzględnienia zjawiska wypierania strumienia ku krawędziom biegunów

Przyjęto dla aproksymacji według rys. 5c oraz rys. 6c, że indukcja w przestrzeni międzybiegunowej maleje wykładniczo w trzeciej potędze wraz ze współrzędną x osiągając w osi poprzecznej q wartość zerową.

Współczynnik $b_{1\nu}$ w równaniu (33) dla funkcji przedstawionych na poszczególnych rysunkach jest wyrażony zależnościami:

— wg rys. 5a

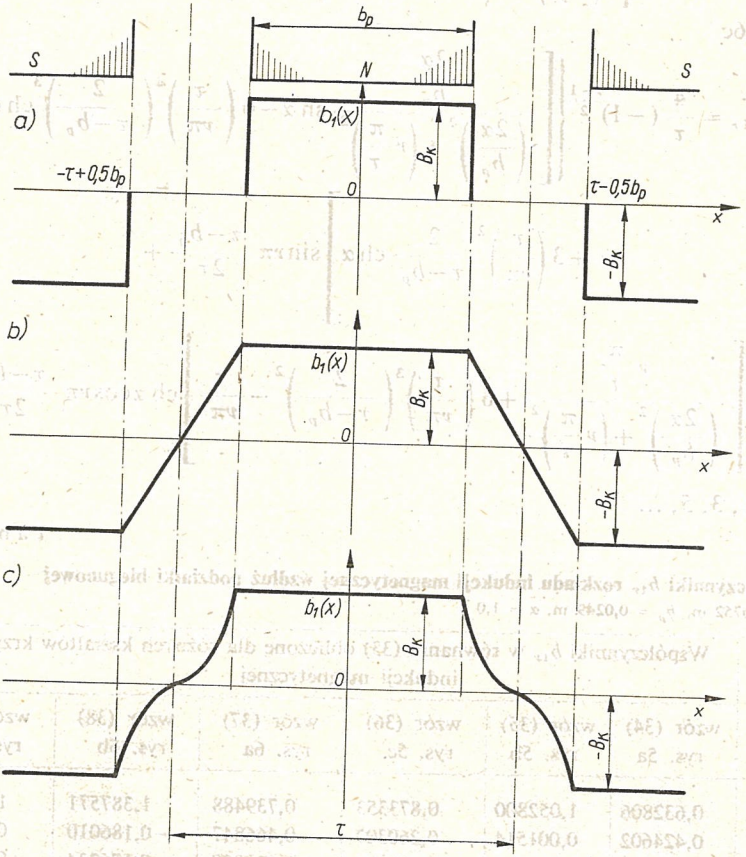
$$\underline{b}_{1\nu} = \frac{4}{\nu\pi}(-1)^{\frac{\nu-1}{2}} \cos\nu\pi \frac{\tau - b_p}{2\tau} = \frac{4}{\nu\pi}\sin\nu\frac{\pi}{\tau}\frac{b_p}{2}, \tag{34}$$

— wg rys. 5b

$$b_{1\nu} = \frac{4\tau}{(\nu\pi)^2} \frac{2}{\tau - b_p} (-1)^{\frac{\nu-1}{2}} \sin \nu\pi \frac{\tau - b_p}{2\tau}, \tag{35}$$

— wg rys. 5c

$$b_{1\nu} = \frac{24\tau^2}{(\nu\pi)^3} \left(\frac{2}{\tau - b_p}\right)^2 (-1)^{\frac{\nu-1}{2}} \left[\cos \nu\pi \frac{\tau - b_p}{2\tau} + \right.$$
$$\left. + \left(\frac{1}{2} \frac{\nu\pi}{\tau} \frac{\tau - b_p}{2} - \frac{\tau}{\nu\pi} \frac{2}{\tau - b_p}\right) \sin \nu\pi \frac{\tau - b_p}{2\tau}\right], \tag{36}$$



Rys. 6. Aproksymacje krzywej rozkładu indukcji magnetycznej stojana jawnobiegunowego wzdłuż podziałki biegunowej z uwzględnieniem zjawiska wypierania strumienia ku krawędziom biegunów

— wg rys. 6a

$$b_{1\nu} = \frac{4}{\tau} (-1)^{\frac{\nu-1}{2}} \frac{1}{\left(\frac{2\alpha}{b_p}\right)^2 + \left(\nu \frac{\pi}{\tau}\right)^2} \left(\frac{2\alpha}{b_p} \operatorname{sh} \alpha \sin \nu\pi \frac{\tau - b_p}{2\tau} + \right.$$
$$\left. + \nu \frac{\pi}{\tau} \operatorname{ch} \alpha \cos \nu\pi \frac{\tau - b_p}{2\tau}\right), \tag{37}$$

wg rys. 6b

$$b_{1\nu} = \frac{4}{\tau}(-1)^{\frac{\nu-1}{2}} \left\{ \left[\frac{\frac{2\alpha}{b_p}}{\left(\frac{2\alpha}{b_p}\right)^2 + \left(\nu\frac{\pi}{\tau}\right)^2} \operatorname{sh}\alpha + \frac{2}{\tau - b_p}\left(\frac{\tau}{\nu\pi}\right)^2 \operatorname{ch}\alpha \right] \sin\nu\pi\frac{\tau - b_p}{2\tau} + \right.$$

$$\left. + \left[\frac{\frac{\nu\pi}{\tau}}{\left(\frac{2\alpha}{b_p}\right)^2 + \left(\nu\frac{\pi}{\tau}\right)^2} - \frac{\tau}{\nu\pi} \right] \operatorname{ch}\alpha\cos\nu\pi\frac{\tau - b_p}{2\tau} \right\}, \tag{38}$$

— wg rys. 6c

$$b_{1\nu} = \frac{4}{\tau}(-1)^{\frac{\nu-1}{2}} \left\{ \left[\frac{\frac{2\alpha}{b_p}}{\left(\frac{2\alpha}{b_p}\right)^2 + \left(\nu\frac{\pi}{\tau}\right)^2} \operatorname{sh}\alpha - 6\left(\frac{\tau}{\nu\pi}\right)^4\left(\frac{2}{\tau - b_p}\right)^3 \operatorname{ch}\alpha + \right.\right.$$

$$\left. + 3\left(\frac{\tau}{\nu\pi}\right)^2 \frac{2}{\tau - b_p} \operatorname{ch}\alpha \right] \sin\nu\pi\frac{\tau - b_p}{2\tau} +$$

$$\left. + \left[\frac{\nu\frac{\pi}{\tau}}{\left(\frac{2\alpha}{b_p}\right)^2 + \left(\nu\frac{\pi}{\tau}\right)^2} + 6\left(\frac{\tau}{\nu\pi}\right)^3\left(\frac{2}{\tau - b_p}\right)^2 - \frac{\tau}{\nu\pi} \right] \operatorname{ch}\alpha\cos\nu\pi\frac{\tau - b_p}{2\tau} \right\}, \tag{39}$$

gdzie $\nu = 1, 3, 5, \ldots$

Tablica 1

Współczynniki $b_{1\nu}$ rozkładu indukcji magnetycznej wzdłuż podziałki biegunowej

$\tau = 0{,}0752$ m, $b_p = 0{,}0249$ m, $\alpha = 1{,}0$

ν	Współczynniki $b_{1\nu}$ w równaniu (33) obliczone dla różnych kształtów krzywych indukcji magnetycznej					
	wzór (34) rys. 5a	wzór (35) rys. 5b	wzór (36) rys. 5c	wzór (37) rys. 6a	wzór (38) rys. 6b	wzór (39) rys. 6c
1	0,632806	1,052800	0,873353	0,739488	1,387571	1,118617
3	0,424602	0,001514	0,260307	0,466847	−0,186010	0,212382
5	0,131498	−0,041560	−0,087017	0,090808	−0,176234	−0,246416
7	−0,086829	−0,021757	−0,072495	−0,172600	−0,072188	−0,150206
9	−0,141462	−0,000505	−0,010982	−0,209897	0,007612	−0,008303
11	−0,061986	0,008468	0,022237	−0,073303	0,035413	0,056717
13	0,044807	0,006383	0,020387	0,082490	0,023199	0,044690
15	0,084792	0,000303	0,002951	0,129430	−0,000943	0,002993
17	0,041520	−0,003494	−0,009635	0,055069	−0,014392	−0,023919
19	−0,029310	−0,003022	−0,009463	−0,051871	−0,011306	−0,021174
21	−0,060474	−0,000216	−0,001392	−0,093048	−0,000066	−0,001772
23	−0,031717	0,001880	0,005294	−0,044170	0,007673	0,012985
25	0,021240	0,001764	0,005462	0,036746	0,006693	0,012351

W tablicy 1 zestawiono współczynniki $b_{1\nu}$ w równaniu (33) dla harmonicznych od $\nu = 1$ do $\nu = 25$ obliczone według wzorów (34)÷(39) dla maszyny o wymiarach $\tau = 75{,}2$ mm, $b_p = 24{,}9$ mm oraz $\alpha = 1$.

Równanie (33) przedstawia krzywą pola stojana o biegunach wydatnych i uzwojeniu jednofazowym. Wypadkowa krzywa pola uzwojenia dwufazowego jest sumą geometryczną krzywych pola uzwojenia fazy I oraz fazy II. Przedstawia ją zależność (31), w której

$$\underline{B}^{+}_{1\nu} = 0{,}5 b_{1\nu}[B_{\mathrm{I}} + B_{\mathrm{II}} e^{-j(\beta - \nu\alpha)}], \tag{40a}$$

$$\underline{B}^{-}_{1\nu} = 0{,}5 b_{1\nu}[B_{\mathrm{I}} + B_{\mathrm{II}} e^{-j(\beta + \nu\alpha)}]. \tag{40b}$$

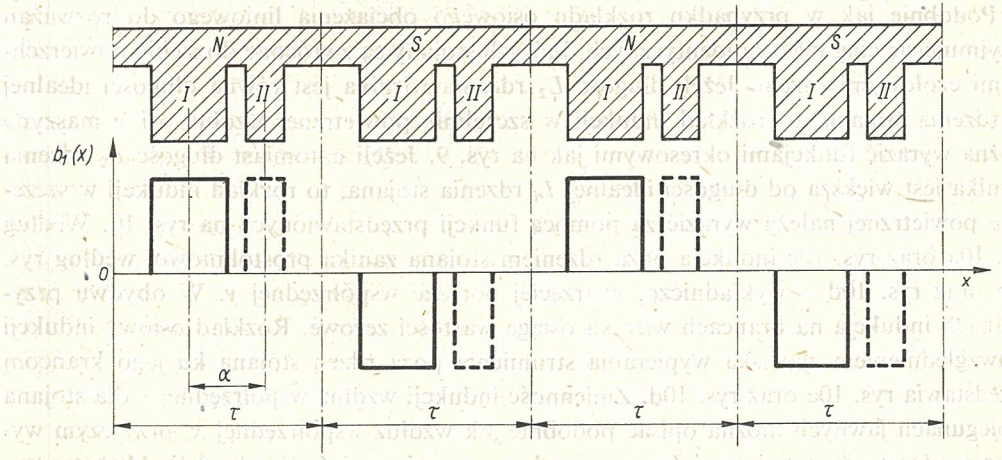
Amplitudy B_{I}, B_{II} zmiennej w czasie indukcji w szczelinie powietrznej najlepiej jest wyznaczyć eksperymentalnie ustawiając sondę teslomierza w osi symetrii poszczególnych biegunów. Na etapie projektowania maszyny można je wyznaczyć według zależności

$$B_{1K} = \mu_o \frac{\sqrt{2} I_{1K} z_{1K}}{2\delta k_c k_n}. \tag{41}$$

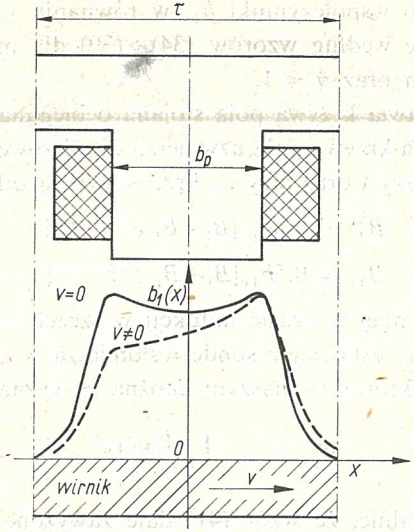
Stwierdzono doświadczalnie, że wzór (41) daje zawyżone wartości indukcji magnetycznej w osi symetrii bieguna, zbliżone do wartości występujących na krańcach (dotyczy rozkładu prostokątnego indukcji pod biegunem).

Szczególnym przypadkiem będzie silnik jednofazowy z fazą pomocniczą zwartą, w którym między osiami magnetycznymi nieekranowanej a ekranowanej zwojem zwartym części bieguna istnieje przesunięcie kątowe $\alpha < \frac{\pi}{2}$. Każdy biegun należy tutaj rozpatrywać jako dwa o nierównych szerokościach bieguny należące do dwu różnych faz, przedzielone żłobkiem, w którym umieszczony jest zwój zwarty. Rozkład indukcji w szczelinie powietrznej takiego silnika przedstawiono w przybliżeniu na rys. 7.

Należy zaznaczyć, że podczas ruchu wirnika rozkład składowej normalnej indukcji wzdłuż podziałki biegunowej ulega okdsztalceniu jak na rys. 8, szczególnie pod tą częścią bieguna, na którą wirnik nabiega. Odkształcenie to jest wyraźnie zauważalne dopiero



Rys. 7. Rozkład indukcji magnetycznej wzdłuż obwodu stojana silnika jednofazowego z fazą pomocniczą zwartą

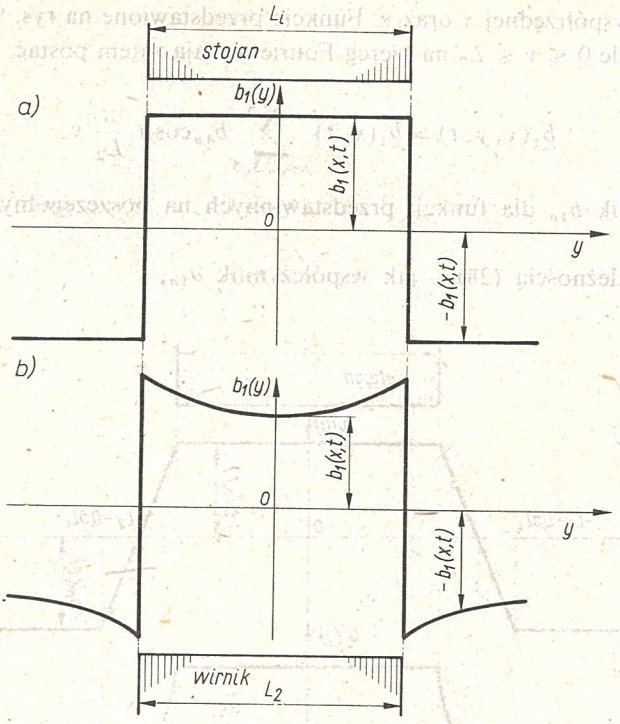


Rys. 8. Oddziaływanie prądów wirowych rotacji w wirniku na rozkład składowej normalnej indukcji magnetycznej pod biegunem jawnym

przy prędkości liniowej wirnika większej od $15\,\frac{\text{m}}{\text{s}}$, a więc praktycznie nie występuje w małych maszynach indukcyjnych, gdyż nawet dla $n = 50\,\frac{\text{obr}}{\text{s}}$, średnicy wirnika 80 mm, prędkość liniowa $v = 12{,}6\,\frac{\text{m}}{\text{s}}$.

9.2. Rozkład indukcji wzdłuż osi maszyny

Podobnie jak w przypadku rozkładu osiowego obciążenia liniowego do rozważań przyjmuje się ciąg maszyn identycznych, których stojany są zwrócone do siebie powierzchniami czołowymi rdzeni. Jeżeli długość L_2 rdzenia wirnika jest równa długości idealnej L_i rdzenia stojana, to rozkład indukcji w szczelinie powietrznej wzdłuż osi y maszyny można wyrazić funkcjami okresowymi jak na rys. 9. Jeżeli natomiast długość L_2 rdzenia wirnika jest większa od długości idealnej L_i rdzenia stojana, to rozkład indukcji w szczelinie powietrznej należy wyrazić za pomocą funkcji przedstawionych na rys. 10. Według rys. 10a oraz rys. 10c indukcja poza rdzeniem stojana zanika prostoliniowo, według rys. 10b oraz rys. 10d — wykładniczo, w trzeciej potędze współrzędnej y. W obydwu przypadkach indukcja na krańcach wirnika osiąga wartości zerowe. Rozkład osiowy indukcji z uwzględnieniem zjawiska wypierania strumienia poza rdzeń stojana ku jego krańcom przedstawia rys. 10c oraz rys. 10d. Zmienność indukcji wzdłuż współrzędnej y dla stojana o biegunach jawnych można opisać podobnie jak wzdłuż współrzędnej x, przy czym wymiar τ odpowiada wymiarowi L_2, wymiar b_p — wymiarowi L_i (rozdz. 9.1). Można przyjąć, że współczynnik α wypierania strumienia ku krańcom rdzenia stojana ma jednakowe



Rys. 9. Aproksymacje krzywej rozkładu indukcji stojana jawnobiegunowego wzdłuż osi maszyny dla długości rdzenia wirnika równej długości rdzenia stojana

Tablica 2

Współczynniki b_{1n} rozkładu indukcji magnetycznej wzdłuż osi maszyny

L_i = 0,033 m, L_2 = 0,039 m, α = 1,0

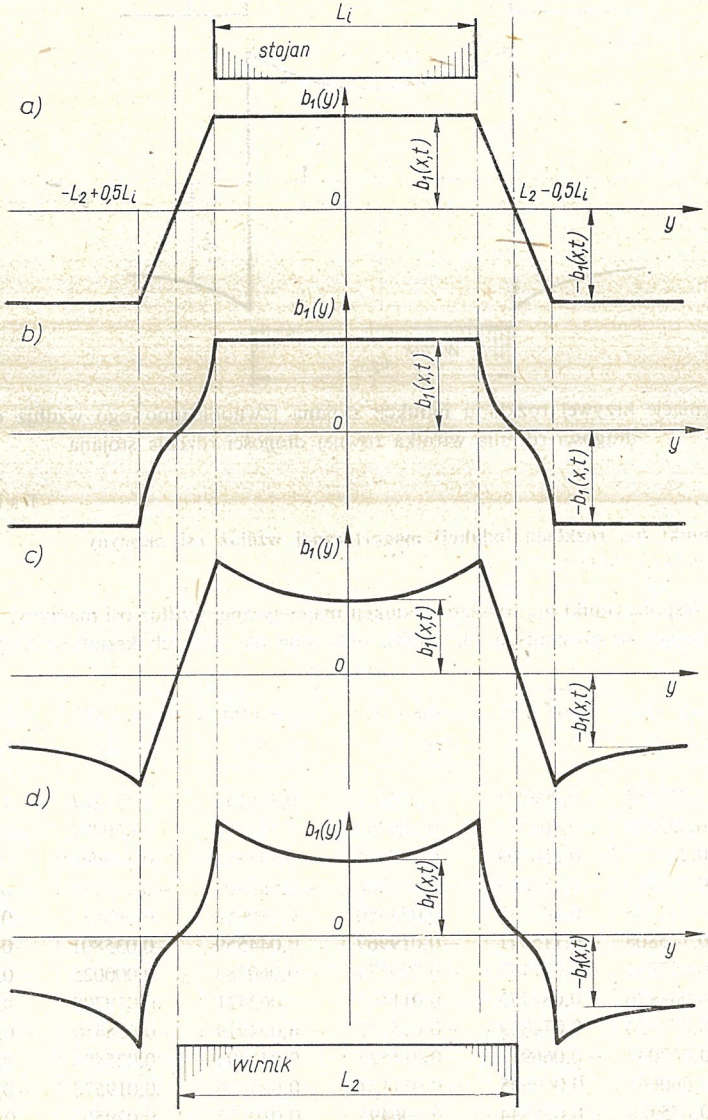
n	Współczynniki b_{1n} rozkładu indukcji magnetycznej wzdłuż osi maszyny, występujące w równaniach (42), (44), obliczone dla różnych kształtów krzywych indukcji					
	wzór (28) rys. 9a	wzór (43) rys. 9b	wzór (35) rys. 10a	wzór (36) rys. 10b	wzór (38) rys. 10c	wzór (39) rys. 10d
1	1,273885	0,906231	1,265616	0,819144	1,435364	1,412924
3	−0,424628	−0,406313	−0,389480	−0,343278	−0,650292	−0,603858
5	0,254777	0,250709	0,197504	0,148381	0,342968	0,273324
7	−0,181984	−0,180489	−0,106786	−0,048401	−0,187738	−0,100775
9	0,141543	0,140837	0,053360	−0,010655	0,094531	−0,002036
11	−0,115808	−0,115421	−0,019969	0,044559	−0,035801	0,062082
13	0,097991	0,097757	−0,000274	−0,060284	0,000025	−0,091373
15	−0,084926	−0,084773	0,011103	0,062421	0,019209	0,097695
17	0,074934	0,074829	−0,015137	−0,054919	−0,026459	−0,087636
19	−0,067047	−0,066971	0,014524	0,041494	0,025493	0,067326
21	0,060661	0,060605	−0,011121	−0,025566	−0,019578	−0,042394
23	−0,055386	−0,055344	0,006495	0,010052	0,029503	0,017655
25	0,050955	0,050922	−0,001881	0,002840	−0,003373	0,003216

wartości wzdłuż współrzędnej x oraz y. Funkcje przedstawione na rys. 9 oraz rys. 10 rozłożone w przedziale $0 \leqslant y \leqslant L_2$ na szereg Fouriera mają zatem postać

$$\underline{b}_1(x, y, t) = \underline{b}_1(x, t) \sum_{n=1,3,5}^{\infty} b_{1n} \cos n \frac{\pi}{L_2} y, \tag{42}$$

gdzie współczynnik b_{1n} dla funkcji przedstawionych na poszczególnych rysunkach jest określony:

— dla rys. 9a zależnością (28) — jak współczynnik a_{1n},



Rys. 10. Aproksymacje krzywej rozkładu indukcji stojana jawnobiegunowego wzdłuż osi maszyny dla długości rdzenia wirnika większej od długości rdzenia stojana