

POLSKA AKADEMIA NAUK
KOMITET ELEKTRONIKI I TELEKOMUNIKACJI

KWARTALNIK
ELEKTRONIKI I TELEKOMUNIKACJI

ELECTRONICS AND
TELECOMMUNICATIONS
QUARTERLY

TOM 51 — ZESZYT 3

WARSZAWA 2005

RADA REDAKCYJNA

Przewodniczący

Prof. dr hab. inż. STEFAN HAHN
czł. rzecz. PAN

Członkowie

prof. dr hab. inż. DANIEL JÓZEF BEM — czł. koresp. PAN, prof. dr hab. inż. MICHAŁ BIAŁKO — czł. rzecz. PAN, prof. dr hab. inż. MAREK DOMAŃSKI, prof. dr hab. inż. ANDRZEJ HAŁAŚ, prof. dr hab. inż. JÓZEF MODELSKI, prof. dr inż. JERZY OSIOWSKI, prof. dr hab. inż. EDWARD SĘDEK, prof. dr hab. inż. MICHAŁ TADEUSIEWICZ, prof. dr hab. inż. WIESŁAW WOLIŃSKI — czł. koresp. PAN, prof. dr inż. MARIAN ZIENTALSKI

REDAKCJA

Redaktor Naczelny

prof. dr hab. inż. WIESŁAW WOLIŃSKI

Zastępca Redaktora Naczelnego

doc. dr inż. KRYSTYN PLEWKO

Sekretarz Odpowiedzialny

mgr ELŻBIETA SZCZEPANIAK

ADRES REDAKCJI

00-665 Warszawa, ul. Nowowiejska 15/19 Politechnika, pok. 470
Instytut Telekomunikacji, Gmach im. prof. JANUSZA GROSZKOWSKIEGO

Dyżury Redakcji: środa i piątki, godz. 14–16
tel/fax (022) 660 77 37

Telefony domowe: Redaktora Naczelnego: 812 17 65
Zast. Red. Naczelnego: 826 83 41
Sekretarza Odpowiedzialnego: 633 92 52

Ark. wyd. 11,56	Ark. druk. 9,25	Podpisano do druku w czerwcu 2005 r.
Papier offset. kl. III 80 g. B-1		Druk ukończono w czerwcu 2005 r.

Skład, druk i oprawa: Warszawska Drukarnia Naukowa PAN
00-656 Warszawa, ul. Śniadeckich 8
Tel./fax: 628-87-77

Szanowni Autorzy

„Kwartalnik Elektroniki i Telekomunikacji” — Electronics and Telecommunications Quarterly jest kontynuatorem tradycji powstałego 51 lat temu kwartalnika p.t. „Rozprawy Elektrotechniczne”.

Kwartalnik jest czasopismem Komitetu Elektroniki i Telekomunikacji Polskiej Akademii Nauk. Wydawany jest przez Warszawską Drukarnię Naukową PAN. Kwartalnik jest czasopismem naukowym, na którego łamach są publikowane artykuły i komunikaty prezentujące wyniki oryginalnych prac teoretycznych i doświadczalnych, a także przeglądowych. Związane są one z szeroko rozumianymi dziedzinami współczesnej elektroniki, telekomunikacji, mikroelektroniki, oproelektroniki, radiotechniki i elektroniki medycznej.

Autorami publikacji są wybitni naukowcy, znani specjaliści o wieloletnim doświadczeniu, a także młodzi badacze — głównie doktoranci.

Artykuły charakteryzują się oryginalnym ujęciem zagadnienia, interesującymi wynikami badań, krytyczną oceną teorii lub metod, omówieniem aktualnego stanu, lub postępu danej gałęzi techniki oraz omówieniem perspektyw rozwojowych. Sposób pisania matematycznej części artykułów zgodny jest z wytycznymi IEC (International Electronics Commision) oraz ISO (International Organization of Standarization).

Wszystkie publikowane w Kwartalniku artykuły są recenzowane przez znanych krajowych specjalistów, co zapewnia że publikacje te są uznawane jako autorski dorobek naukowy. Opublikowane w kwartalniku wyniki prac naukowych zrealizowanych w ramach „GRANTów” Komitetu Badań Naukowych spełnia więc jeden z wymogów stawianych tym pracom.

Czasopismo dociera do wszystkich zajmujących się elektroniką i telekomunikacją krajowych ośrodków naukowych oraz technicznych, a także szeregu instytucji zagranicznych. Jest ponadto prenumerowane przez liczne grono specjalistów i biblioteki.

Każdy Autor otrzymuje bezpłatnie 20 egzemplarzy nadbitek swojego artykułu, co ułatwia przesłanie go do indywidualnych wybranych przez Autora osób i instytucji w kraju lub za granicą. Ułatwia to dodatkowo fakt, że w Kwartalniku są publikowane artykuły w języku angielskim.

Nadesłane do redakcji artykuły są publikowane w terminie około pół roku, w przypadku sprawnej współpracy Autora z Redakcją. Wytyczne dla Autorów dotyczące formy publikacji są zamieszczone w zeszytach Kwartalnika, można je także otrzymać w siedzibie Redakcji.

Artykuły można dostarczać osobiście, lub pocztą pod adresem zamieszczanym na stronie redakcyjnej w każdym zeszycie.

Redakcja

M

D

K

Z.

A.

S.

S.

P.

In

M

D.

K.

Z.

A.

S.

S.

P.

Inf

SPIS TREŚCI

M. Gajer: Adaptacja algorytmu Rate Monotonic Scheduling dla potrzeb szeregowania zadań wieloprocesorowych	373
D. Krzemieniecki: Badanie i optymalizacja szerokopasmowego transformatora dwukierunkowego ...	387
K. Górecki, J. Zarębski: Elektrotermiczny model transformatora impulsowego dla programu SPICE ..	421
Z. Szczepniak: Elektroniczny moduł dziesięciokrotnego zwiększania dokładności przetwarzania sygnałów optoelektronicznego przetwornika położenia	439
A. Bartoszewicz: Analiza odporności nieliniowego regulatora prędkości przepływu danych w sieciach ATM o wielu źródłach	449
S. Szczepański, S. Kozieł: Linearyzacja wzmacniacza transkonduktancyjnego CMOS metodą aktywnego sprzężenia typu „feedforward”	465
S. Kozieł: Analiza i optymalizacja szumów filtrów OTA-C czasu ciągłego	479
P. Gawkowski, J. Sosnowski: Programowe mechanizmy detekcji i tolerowania błędów sprzętu — cz. II Eksperymentalna analiza pokrycia błędów	495
Informacje dla Autorów (w jęz. polskim)	509

CONTENTS

M. Gajer: The adaptation of Rate Monotonic Scheduling algorithm for scheduling of multiprocessor tasks	373
D. Krzemieniecki: Analysis and optimization of a broadband two-directional transformer	387
K. Górecki, J. Zarębski: The electrothermal model of the pulse transformer for SPICE	421
Z. Szczepniak: Electronic module for 10-times increment of processing accuracy of photoelectric transducer signals	439
A. Bartoszewicz: On the robustness of a nonlinear ABR flow controller for multi-source ATM networks	449
S. Szczepański, S. Kozieł: Active-Error feedforward technique for linearization of CMOS transconductance amplifier	465
S. Kozieł: Noise analysis and optimization of continuous-time OTA-C filters	479
P. Gawkowski, J. Sosnowski: Software implemented fault detection and fault tolerance mechanisms — Part II: Experimental evaluation of error coverage	495
Information for the Authors (w jęz. ang.)	512

Adaptacja algorytmu rate monotonic scheduling dla potrzeb szeregowania zadań wieloprocessorowych

MIROSŁAW GAJER

*Katedra Automatyki AGH
al. Mickiewicza 30, 30-059 Kraków
email: mgajer@ia.agh.edu.pl*

*Otrzymano 2004.06.09
Autoryzowano 2004.10.01*

Artykuł niniejszy poświęcony został zagadnieniom związanym z szeregowaniem jednoprocessorowych i wieloprocessorowych zadań dla systemu komputerowego zbudowanego z pewnej liczby procesorów. W systemach czasu rzeczywistego o ostrych ograniczeniach czasowych (ang. hard real-time systems) bardzo często występuje konieczność znalezienia planu szeregowania zbioru niezależnych i wyłączalnych zadań periodycznych. Ponadto zwykle już na etapie projektowania systemu należy wykazać w sposób formalny, że ograniczenia czasowe będą zawsze zachowane dla wszystkich zadań. W tym celu w przypadku zadań jednoprocessorowych powszechnie wykorzystywany jest algorytm RMS (ang. Rate Monotonic Scheduling) oraz związane z nim formalne metody dowodzenia szeregowalności zbioru zadań. Autor niniejszego artykułu zaproponował rozszerzenie zakresu stosowalności algorytmu RMS również dla przypadku zadań wieloprocessorowych. Podejście zaproponowane przez autora bazuje na konkatenaacji zadań jedno- i wieloprocessorowych oraz na transformacji wartości okresów wybranych zadań. W artykule rozważono zarówno przypadek zbioru zadań przeznaczonych dla procesorów arbitralnych, jak i dedykowanych.

Słowa kluczowe: szeregowanie zadań, Rate Monotonic Scheduling, systemy wieloprocessorowe, zadania wieloprocessorowe

1. WPROWADZENIE

Systemy czasu rzeczywistego stanowią obecnie dobrze zdefiniowaną klasę systemów komputerowych, charakteryzujących się pewnymi szczególnymi właściwościami, które odróżniają je w sposób zdecydowany od systemów komputerowych ogólnego przeznaczenia [1], [2]. W przypadku systemów czasu rzeczywistego o tzw. ostrych ograniczeniach czasowych (ang. hard real-time) program realizowany przez system

komputerowy jest podzielony na zadania, z których każde musi zostać zakończone przed upływem przewidzianego dla niego ograniczenia czasowego (ang. deadline). Wyniki poprawne pod względem ich analizy logicznej, lecz spóźnione (uzyskane po upływie ich ograniczenia czasowego) są uznawane za całkowicie bezużyteczne. Co więcej, naruszenie ograniczenia czasowego w systemie o ostrych ograniczeniach czasowych może pociągnąć za sobą bardzo poważne konsekwencje, związane z utratą kontroli nad sterowanym obiektem. Z oczywistych powodów projektant systemu czasu rzeczywistego nie ma możliwości eksperymentalnego sprawdzenia, czy dla każdego z zadań będzie dochowane jego ograniczenie czasowe, gdyż w praktyce tego typu eksperymenty byłyby zbyt kosztowne lub zbyt niebezpieczne (np. sterowanie lotem pocisku raketowego, nadzorowanie pracy reaktora jądrowego itp.). Zatem właściwość szeregowanego zbioru zadań polegająca na tym, że każde z zadań dochowa swoje ograniczenie czasowe musi zostać wykazana w sposób formalny już na etapie projektowania systemu czasu rzeczywistego o ostrych ograniczeniach czasowych.

W przypadku systemów czasu rzeczywistego najczęściej mamy do czynienia ze zbiorem zadań okresowych, z których każde jest aktywowane cyklicznie po upływie charakterystycznego dla niego odcinka czasu zwanego okresem zadania. Zwykle zadania takie są niezależne, w tym sensie, że żadne z nich nie pobiera danych wejściowych, które byłyby wynikami któregoś z pozostałych zadań. Zatem kolejność realizacji rozważanego zbioru zadań jest dowolna. Ponadto zadania występujące w systemach czasu rzeczywistego są zwykle zadaniami wywłaszczalnymi, tzn. ich realizacja może zostać przerwana w dowolnym momencie, przy czym kontekst wywłaszczonego zadania jest zapisywany na stosie procesora. Wywłaszczone zadanie może czekać dowolnie długo na wznowienie swojej realizacji, czyli do momentu, gdy procesor nie ma już w kolejce żadnych innych, ważniejszych zadań do wykonania.

W 1973 Liu i Layland zaproponowali algorytm zwany Rate Monotonic Scheduling (RMS), który pozwala na skuteczne szeregowanie rozważanych w artykule zbiorów niezależnych i wywłaszczalnych zadań okresowych [3]. W kolejnych latach algorytm RMS okazał się być niezwykle przydatnym narzędziem, upraszczającym w znacznym stopniu proces projektowania i analizy systemów czasu rzeczywistego o ostrych ograniczeniach czasowych. Algorytm ten został zaimplementowany w zdecydowanej większości systemów operacyjnych czasu rzeczywistego i z powodzeniem jest wykorzystywany w wielu implementacjach po dzień dzisiejszy [4], [5], [6], [7].

2. PODSTAWOWE WŁAŚCIWOŚCI ALGORYTMU RMS

Algorytm RMS przewidziany jest do szeregowania zbioru niezależnych, wywłaszczalnych i periodycznych zadań jednoprosesorowych. Niech dany będzie zbiór N zadań, z których każde posiada swój okres aktywacji T_i oraz czas wykonania C_i . Dla każdego z zadań wartość ilorazu C_i/T_i określa stopień wykorzystania czasu pracy procesora przez rozważane zadanie.

Działanie algorytmu RMS opiera się na następujących zasadach [8]. Zadaniom przydzielane są priorytety według zasady im krótszy okres zadania, tym wyższy jego priorytet. W danej chwili, spośród zadań będących w stanie gotowości (ang. ready) wykonywane jest to zadanie, które posiada najwyższy priorytet. Wykonywane zadanie zawsze zostaje wywłaszczone, jeżeli w stan gotowości wejdzie jakieś inne zadanie o wyższym priorytecie. Wywłaszczone zadanie może wznowić swoje wykonywanie tylko w sytuacji, gdy już żadne inne zadanie o wyższym priorytecie nie znajduje się aktualnie w stanie gotowości.

Warunek konieczny na szeregowalność rozpatrywanego zbioru zadań ma postać

$$C_1/T_1 + C_2/T_2 + \dots + C_N/T_N \leq 1 \quad (1)$$

Warunek ten jest bezpośrednią konsekwencją oczywistego faktu, polegającego na tym, że procesor nie może być wykorzystywany w stopniu większym niż 100%.

Z kolei warunek wystarczający na szeregowalność rozważanego zbioru zadań został podany przez Liu i Laylanda [3] i ma następującą postać

$$C_1/T_1 + C_2/T_2 + \dots + C_N/T_N \leq N(2^{1/N} - 1) \quad (2)$$

Dla dużych wartości N prawa strona nierówności (2) dąży do $\ln 2 = 0,693$. Oznacza to, że jeżeli procesor jest wykorzystywany przez zbiór zadań w mniej niż 69,3% to taki zbiór zadań jest zawsze szeregowalny algorytmem RMS.

W praktyce często występuje sytuacja, w której stopień wykorzystania procesora przez szeregowany zbiór zadań mieści się w przedziale zawartym pomiędzy 69,3% i 100%. W takiej sytuacji indywidualnie dla każdego z zadań należy policzyć czas zakończenia jego realizacji dla najgorszego przypadku, czyli w sytuacji, gdy równocześnie z nim zostały aktywowane wszystkie zadania o wyższych priorytetach. Jeżeli otrzymany czas zakończenia realizacji zadania jest mniejszy niż jego ograniczenie czasowe wówczas zadanie to również nigdy w przyszłości nie przekroczy swojego ograniczenia czasowego.

3. SZEREGOWANIE ZADAŃ WIELOPROCESOROWYCH

Zgodnie z posiadaną przez autora wiedzą, algorytm RMS nie był nigdy stosowany do szeregowania zadań wieloprocessorowych. Artykuł niniejszy stanowi próbę opracowania ogólnej metodyki szeregowania zbioru niezależnych, wywłaszczalnych i okresowych zadań wieloprocessorowych przy wykorzystaniu algorytmu RMS. Opracowana przez autora metodyka zostanie zaprezentowana na przykładzie systemu wieloprocessorowego zbudowanego z trzech procesorów. Jednakże uzyskane rezultaty z łatwością mogą zostać przeniesione na systemy o większej liczbie procesorów.

Zatem niech dany będzie zbiór trzech procesorów P_1 , P_2 i P_3 . Niech dany będzie również zbiór zadań wieloprocessorowych Z , który można podzielić na siedem wzajemnie rozłącznych podzbiorów, grupujących zadania tego samego typu: Z^{123} , Z^{12} , Z^{13} ,

Z^{23} , Z^1 , Z^2 i Z^3 . Górny indeks wskazuje numery procesorów potrzebnych do wykonania zadań należących do danego podzbioru. Zatem do realizacji zadań należących do podzbioru Z^{123} wymagana jest jednoczesna dostępność wszystkich trzech procesorów P1, P2 i P3. Z kolei do realizacji zadań należących do podzbioru Z^{13} wymagana jest jednoczesna dostępność procesorów P1 i P3. Zadania takie nazywa się zadaniami dla procesorów dedykowanych, ponieważ jest z góry określone, które z zadań wykonywane będzie przez które z procesorów [9].

Dwa różne zadania nazywa się zadaniami kompatybilnymi, jeżeli iloczyn zbiorów procesorów koniecznych do wykonywania każdego z nich jest zbiorem pustym [8]. Jeżeli warunek ten nie jest spełniony, wówczas zadania takie nazywa się niekompatybilnymi. W rozważanym przypadku zadania należące do podzbioru Z^{123} nie są kompatybilne z żadnym innym zadaniem. Również z oczywistych powodów nie są wzajemnie kompatybilne żadne dwa zadania należące do tego samego podzbioru.

W rozważanym przypadku zadań Z, przewidzianych dla trzech dedykowanych procesorów, istnieją trzy pary podzbiorów zadań kompatybilnych: $\{Z^{12}, Z^3\}$, $\{Z^{13}, Z^2\}$ i $\{Z^{23}, Z^1\}$. Na przykład dowolne z zadań należących do podzbioru Z^{12} może być jednocześnie wykonywane z dowolnym zadaniem należącym do podzbioru Z^3 , ponieważ rozważane zadania nie wymagają do swej realizacji ani jednego wspólnego procesora.

Zaproponowana przez autora metodyka szeregowania zadań wieloprocessorowych według algorytmu RMS zostanie zilustrowana na przykładzie zbioru zadań o charakterystykach zamieszczonych w tab. 1.

Spośród zadań zamieszczonych w tab. 1 wyłonił się następujący podzbiór zadań $\{z_1^{123}, z_2^{123}, z_3^{12}, z_{15}^3, z_4^{12}, z_{16}^3, z_5^{13}, z_{12}^2, z_6^{13}, z_{13}^2, z_7^{23}, z_9^1, z_8^{23}, z_{10}^1\}$ i $\{z_{11}^1, z_{14}^2, z_{17}^3\}$. Należy zwrócić szczególną uwagę na fakt polegający na tym, że w ramach każdego podzbioru zostały zgrupowane zadania wzajemnie kompatybilne.

W ramach wymienionych podzbiorów okresy niektórych zadań ulegają transformacji (zostają zmniejszone), w taki sposób, aby wszystkie zadania wchodzące w skład danego podzbioru posiadały identyczny okres. W wyniku transformacji wartości okresów niektórych zadań otrzymuje się nowy zbiór zadań o charakterystykach zamieszczonych w tab. 2.

Analizując parametry zadań zamieszczonych w tab. 2 widać, że nie uległy zmianie wartości okresów zadań: z_1^{123} , z_2^{123} , z_3^{12} , z_{16}^3 , z_5^{13} , z_6^{13} , z_7^{23} , z_8^{23} , z_{14}^2 . Natomiast wartości okresów pozostałych zadań, w ramach każdego z rozważanych podzbiorów, zostały odpowiednio zmniejszone (zatem zadania te będą aktywowane nieco częściej), tak aby wszystkie zadania kompatybilne, wchodzące w skład danego podzbioru miały równe wartości swoich okresów. Takie zmniejszenie wartości okresu zadania, czyli de facto zwiększenie częstotliwości jego aktywacji jest całkowicie dopuszczalne w każdym systemie czasu rzeczywistego ostrych ograniczeniach czasowych. Oznacza to bowiem, że dane pochodzące z sensorów, za pośrednictwem których jest monitorowany obiekt będący przedmiotem sterowania, są zbierane częściej. Również częściej są wysyłane do obiektu sygnały sterujące, będące wynikiem obliczeń przeprowadzanych w komputerowym systemie czasu rzeczywistego. To wszystko sprawia, że margines stabilności

Tabela 1

Przykładowy zbiór zadań podlegających szeregowaniu algorytmem RMS

The example set of tasks that are scheduled by the means of RMS algorithm

zadanie	procesory, na których zadanie jest wykonywane	czas wykonania zadania C_i [ms]	okres zadania T_i [ms]
z_1^{123}	P1, P2, P3	12	254
z_2^{123}	P1, P2, P3	25	1458
z_3^{12}	P1, P2	15	346
z_4^{12}	P1, P2	17	890
z_5^{13}	P1, P3	18	542
z_6^{13}	P1, P3	16	321
z_7^{23}	P2, P3	14	705
z_8^{23}	P2, P3	13	287
z_9^1	P1	18	765
z_{10}^1	P1	16	311
z_{11}^1	P1	19	178
z_{12}^2	P2	15	600
z_{13}^2	P2	16	336
z_{14}^2	P2	12	168
z_{15}^3	P3	17	367
z_{16}^3	P3	16	876
z_{17}^3	P3	14	189

sterowanego obiektu powiększa się, co jest zjawiskiem pod każdym względem korzystnym. Jedyną ujemną stroną transformacji okresu zadania, polegającej na zmniejszeniu jego wartości, stanowi zwiększenie wartości stosunku C/T dla tego zadania, czyli zadanie to będzie w większym stopniu wykorzystywało czas pracy procesora, niż miało to miejsce przed transformacją jego okresu.

W kolejnym etapie kompatybilne zadania, posiadające identyczne wartości okresów są łączone w większe jednostki nazwane przez autora superzadaniami. Każde superzadanie do swej realizacji wymaga jednoczesnej dostępności wszystkich trzech procesorów P1, P2 i P3. Superzadania, które powstały poprzez konkatencję zadań wchodzących w skład podzbiorów $\{y_1^{123}\}$, $\{y_2^{123}\}$, $\{y_3^{12}, y_{15}^3\}$, $\{y_4^{12}, y_{16}^3\}$, $\{y_5^{13}, y_{12}^2\}$, $\{y_6^{13}, y_{13}^2\}$, $\{y_7^{23}, y_9^1\}$, $\{y_8^{23}, y_{10}^1\}$ i $\{y_{11}^1, y_{14}^2, y_{17}^3\}$, zostały zebrane w tab. 3.

Ponieważ do realizacji wszystkich superzadań wymagana jest jednoczesna dostępność wszystkich trzech procesorów, każde superzadanie może de facto zostać potraktowane, jako zwyczajne zadanie jednoprocessorowe, z tym że do jego realizacji potrzebny byłby specjalny procesor, który w swoim wnętrzu zawierałby trzy jednostki oblicze-

Tabela 2

Przykładowy zbiór zadań otrzymany po transformacji okresów wybranych zadań

The example set of tasks obtained after transformation of periods of chosen tasks

zadanie	procesory, na których zadanie jest wykonywane	czas wykonania zadania C_i [ms]	okres zadania T_i [ms]
y_1^{123}	P1, P2, P3	12	254
y_2^{123}	P1, P2, P3	25	1458
y_3^{12}	P1, P2	15	346
y_4^{12}	P1, P2	17	876
y_5^{13}	P1, P3	18	542
y_6^{13}	P1, P3	16	321
y_7^{23}	P2, P3	14	705
y_8^{23}	P2, P3	13	287
y_9^1	P1	18	705
y_{10}^1	P1	16	287
y_{11}^1	P1	19	168
y_{12}^2	P2	15	542
y_{13}^2	P2	16	321
y_{14}^2	P2	12	168
y_{15}^3	P3	17	346
y_{16}^3	P3	16	876
y_{17}^3	P3	14	168

niowe (procesory P1, P2 i P3). Potraktowanie superzadań jako zadań jednoprocessorowych, realizowanych wszakże przez jednostkę obliczeniową specjalnego typu (klastery wieloprocessorowy) otwiera drogę do bezpośredniego zastosowania algorytmu RMS do szeregowania takich zadań. Z punktu widzenia systemu operacyjnego szeregującego superzadania według algorytmu RMS fakt wewnętrznej budowy jednostki obliczeniowej, dla której szeregowane są zadania, jest całkowicie nieistotny. Również nieistotna dla programu szeregującego superzadania jest wewnętrzna struktura superzadań, które w istocie mogą składać się z wielu mniejszych zadań (zarówno jednoprocessorowych, jak i wieloprocessorowych). Jednak podczas tworzenia superzadania wszystkie zadania wchodzące w jego skład zostają połączone w jedną nierozzerwalną całość. Dlatego każde superzadanie jest zadaniem trójprocessorowym, a posiadając zbiór zadań wyłącznie trójprocessorowych można je potraktować wszystkie, jako zadania jednoprocessorowe, z tym że wymagające do swej realizacji specjalnej jednostki obliczeniowej, która w swym wnętrzu kryje trzy procesory. Jednakże fakt ten jest niewidoczny dla zewnętrznego obserwatora (programu szeregującego zadania). Następnie mając dany zbiór za-

Tabela 3

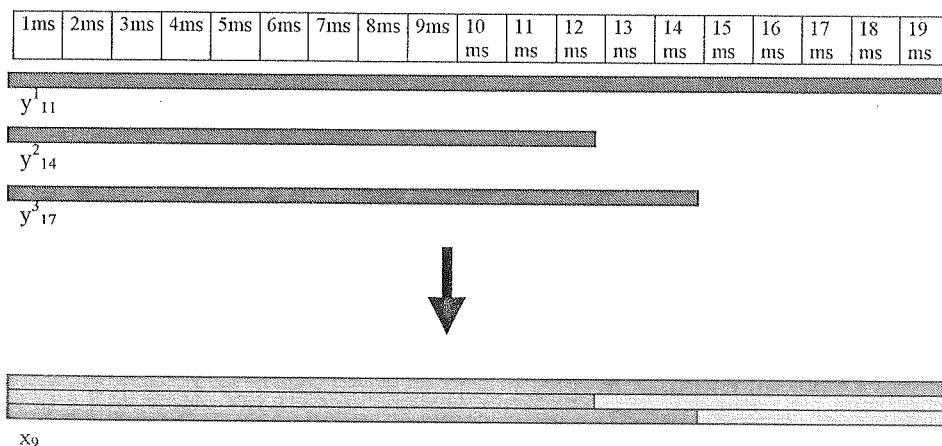
Superzadania powstałe na drodze konkatencji zadań wchodzących
w skład tego samego podzbioru zadań

Supertasks created by the concatenation of tasks that belong to the same subset of tasks

superzadanie	czas wykonania zadania C_i [ms]	okres zadania T_i [ms]	stopień wykorzystania czasu pracy jednostki obliczeniowej C_i/T_i
x_1	12	254	0,047
x_2	25	1458	0,017
x_3	17	346	0,049
x_4	17	876	0,019
x_5	18	542	0,033
x_6	16	321	0,050
x_7	18	705	0,026
x_8	16	287	0,056
x_9	19	168	0,113

dań wyłącznie jednoprocessorowych (wymagających co prawda jednostki obliczeniowej specjalnego typu), można do ich szeregowania zastosować już bezpośrednio algorytm RMS.

Na rys. 1 przedstawiono, w jaki sposób jednoprocessorowe zadania y_{11}^1 , y_{14}^2 , y_{17}^3 , posiadające czasy realizacji odpowiednio 19, 12 i 14 ms oraz okres aktywacji 168 ms zostają połączone w jedno superzadanie o czasie realizacji 19 ms i okresie aktywacji 168 ms.



Rys. 1. Ilustracja sposobu, w jaki trzy zadania jednoprocessorowe tworzą jedno superzadanie

Fig. 1. The illustration of the way in which three uniprocessor tasks create one supertask

Na rys. 1 kolorem ciemniejszym zaznaczono te odcinki czasu, w których dane jednostki obliczeniowe trójprocesorowego klastra wykonują obliczenia związane z realizacją zadań y_{11}^1 , y_{14}^2 , y_{17}^3 , natomiast kolorem jaśniejszym zaznaczono odcinki czasu, w których poszczególne jednostki obliczeniowe znajdują się w stanie pracy jałowej (nie wykonują żadnych użytecznych obliczeń).

Sumaryczny stopień wykorzystania czasu pracy jednostki obliczeniowej przez szeregowany zbiór superzadań $\{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9\}$ wynosi 0,410. Jest to znacznie mniej niż prawa strona nierówności (2), czyli $N(2^{1/N}-1)$, która dla $N=9$ przyjmuje wartość 0,721. W związku z tym rozważany zbiór superzadań jest szeregowalny według algorytmu RMS.

4. BINARYZACJA OKRESÓW ZADAŃ

Binaryzacja okresów zadań jest techniką często stosowaną w połączeniu z algorytmem RMS. Celem binaryzacji okresów zadań jest zmniejszenie rozmiarów tzw. horyzontu czasowego systemu [11]. Horyzont czasowy jest zdefiniowany jako najmniejsza wspólna wielokrotność okresów szeregowanych zadań i jest to przedział czasu, po upływie którego cały przebieg pracy systemu czasu rzeczywistego powtarza się cyklicznie. Ponieważ w wyniku przeprowadzenia binaryzacji okresów zadań otrzymuje się zadania o okresach binarnych, stanowiących wielokrotność (będącą naturalną potęgą liczby dwa) pewnego okresu podstawowego r (tzn. występują wyłącznie zadania o okresach $2r$, $4r$, $8r$, $16r$, $32r$, $64r$ itd.), najmniejsza wspólna wielokrotność okresów zadań jest równa wartości najdłuższego z okresów. Jest to wartość zwykle o kilka rzędów wielkości mniejsza niż przed binaryzacją, co bardzo ułatwia przeprowadzenie dalszej czasowej analizy pracy systemu i przebiegu jego interakcji z otoczeniem [10].

Binaryzacja okresów zadań jest przeprowadzana poprzez transformację okresów zadań. W wyniku binaryzacji okresy zadań zostają zmniejszone do wartości równych najbliższej im wielokrotności okresu podstawowego r , zgodnie ze wzorem

$$T_i^b = r \cdot 2^{\lceil \log_2(\frac{T_i}{r}) \rceil} \quad (3)$$

We wzorze (3) T_i oznacza wartość okresu zadania i przed binaryzacją, a T_i^b wartość okresu tego zadania po binaryzacji.

Ujemnym skutkiem zastosowania binaryzacji okresów zadań jest zwiększenie wartości stopnia wykorzystania procesora C/T przez zadania, dla których, w wyniku przeprowadzenia binaryzacji okresów, wartości tych okresów uległy zmniejszeniu. Ponieważ wartość okresu bazowego r może zostać wybrana jako dowolna liczba rzeczywista, istnieje zatem nieskończenie wiele binaryzacji, które w różnym stopniu powodują zwiększenie wykorzystania czasu pracy jednostki obliczeniowej przez zbiór szeregowanych zadań. Spośród wszystkich możliwych binaryzacji należy wybrać tę, która powoduje jak najmniejszy wzrost stopnia wykorzystania czasu pracy jednostki obliczeniowej. Taką binaryzację optymalną poszukuje się wśród tzw. binaryzacji istotnych.

Dla każdej wartości okresu zadania, wchodzącego w skład szeregowanego zbioru zadań istnieje binaryzacja istotna, która nie powoduje zmiany wartości okresu tego zadania. Dla zadania o okresie T_i okres bazowy związanej z nim binaryzacji istotnej wynosi r_i i wylicza się go zgodnie z następującym wzorem

$$r_i = \frac{T_i}{2^{\lceil \log_2(\frac{T_i}{T_{MIN}}) \rceil}} \quad (4)$$

We wzorze (4) T_{MIN} oznacza najmniejszą wartość okresu zadania, spośród wszystkich zadań podlegających szeregowaniu.

W tab. 4 zamieszczono wartości okresów superzadań, uzyskane dla każdej z dziewięciu binaryzacji istotnych.

Tabela 4

Wartości okresów superzadań po binaryzacjach

The values of periods of supertask after binarizations

	$r_1=$ 127	$r_2=$ 91,125	$r_3=$ 86,5	$r_4=$ 109,5	$r_5=$ 135,5	$r_6=$ 160,5	$r_7=$ 88,125	$r_8=$ 143,5	$r_9=$ 168
T_1^b	254	182,25	173	219	135,5	160,5	176,25	143,5	168
T_2^b	1016	1458	1384	876	1084	1284	1410	1148	1344
T_3^b	254	182,25	346	219	271	321	176,25	287	336
T_4^b	508	729	692	876	542	642	705	574	672
T_5^b	508	364,5	346	438	542	321	352,5	287	336
T_6^b	254	182,25	173	219	271	321	176,25	287	168
T_7^b	508	364,5	692	438	542	642	705	574	672
T_8^b	254	182,25	173	219	271	160,5	176,25	287	168
T_9^b	127	91,125	86,5	109,5	135,5	160,5	88,125	143,5	168

Z kolei w tab. 5 zamieszczono wartości stopnia wykorzystania czasu pracy jednostki obliczeniowej przez każde z superzadań o zbinaryzowanych okresach, dla każdej z rozważanych binaryzacji istotnych.

Spośród binaryzacji istotnych najlepszą binaryzacją okazała się binaryzacja z okresem bazowym r_1 wynoszącym 127 ms. Binaryzacja ta powoduje wzrost stopnia wykorzystania czasu pracy jednostki obliczeniowej do wartości 0,518. Jest to w dalszym ciągu znacznie mniej niż wynosi prawa strona nierówności (2) dla $N=9$, a zatem zbiór superzadań o zbinaryzowanych wartościach okresów $\{b_1, b_2, b_3, b_4, b_5, b_6, b_7, b_8, b_9\}$ jest w dalszym ciągu szeregowalny. Charakterystyki superzadań o binarnych okresach zostały zebrane w tab. 6.

Tabela 5

Stopień wykorzystania czasu pracy jednostki obliczeniowej dla każdego superzadania

The value of processor time usage coefficient for each supertask

	$r_1=$ 127	$r_2=$ 91,125	$r_3=$ 86,5	$r_4=$ 109,5	$r_5=$ 135,5	$r_6=$ 160,5	$r_7=$ 88,125	$r_8=$ 143,5	$r_9=$ 168
T_1^b	0,047	0,066	0,069	0,055	0,089	0,075	0,068	0,084	0,071
T_2^b	0,025	0,017	0,018	0,029	0,023	0,019	0,018	0,022	0,019
T_3^b	0,067	0,093	0,049	0,078	0,063	0,053	0,096	0,059	0,051
T_4^b	0,033	0,023	0,025	0,019	0,031	0,026	0,024	0,030	0,026
T_5^b	0,035	0,049	0,052	0,041	0,033	0,056	0,051	0,063	0,054
T_6^b	0,063	0,088	0,092	0,073	0,059	0,050	0,091	0,056	0,095
T_7^b	0,035	0,049	0,026	0,041	0,033	0,028	0,026	0,031	0,027
T_8^b	0,063	0,088	0,092	0,073	0,059	0,100	0,091	0,056	0,095
T_9^b	0,150	0,209	0,220	0,174	0,140	0,118	0,216	0,132	0,113
suma	0,518	0,682	0,643	0,583	0,530	0,525	0,681	0,533	0,551

Tabela 6

Charakterystyki superzadań o zbinaryzowanych okresach

The characteristics of supertasks with binarized periods

superzadanie o binarnej wartości okresu	czas wykonania zadania C_i [ms]	okres zadania T_i [ms]	stopień wykorzystania czasu pracy jednostki obliczeniowej C_i/T_i
b_1	12	254	0,047
b_2	25	1016	0,025
b_3	17	254	0,067
b_4	17	508	0,033
b_5	18	508	0,035
b_6	16	254	0,063
b_7	18	508	0,035
b_8	16	254	0,063
b_9	19	127	0,150

5. UWAGI KOŃCOWE

W artykule zaproponowano nowatorską metodę szeregowania zbioru niezależnych, wywłaszczalnych i periodycznych zadań jedno- i wieloprocesorowych, z wykorzystaniem do tego celu popularnego algorytmu RMS. Zgodnie z wiedzą autora algorytm RMS nie był jak dotąd nigdy stosowany do szeregowania zadań wieloprocesorowych, a w związku z tym jedynym obszarem jego zastosowań było szeregowanie zadań jednoprocessorowych. Zaproponowana przez autora metoda stanowi zatem istotne rozszerzenie obszarów zastosowań algorytmu RMS. Jest to ważne głównie z tego powodu, iż rozwiązania wieloprocesorowe są coraz częściej stosowane w nowoczesnych systemach czasu rzeczywistego o ostrych ograniczeniach czasowych, gdzie rozdział zadań do poszczególnych jednostek obliczeniowych często stanowi jedyną drogę, prowadzącą do zwiększenia mocy obliczeniowej systemu i tym samym dochowania ograniczeń czasowych występujących w nim zadań [12].

Istota zaproponowanej przez autora metody polega na transformacji okresów wybranych zadań, polegającej na zmniejszeniu ich wartości, w taki sposób, aby uzyskać pewną liczbę podzbiorów zadań o identycznych wartościach okresów. Następnie w ramach każdego z rozważanych podzbiorów, szeregowane zadania jedno- i wieloprocesorowe podlegają konkatenacji, w związku z czym tworzą jedno tzw. superzadanie, które do swego wykonania potrzebuje jednoczesnej dostępności wszystkich jednostek obliczeniowych występujących w systemie. Bezpośrednie zastosowanie algorytmu RMS do szeregowania zbioru superzadań jest możliwe, dzięki potraktowaniu wszystkich superzadań, jako zadań jednoprocessorowych, do realizacji których potrzebna jest specjalna jednostka obliczeniowa — klastr składający się z pewnej liczby podjednostek. Jednakże wewnętrzna struktura klastra jest całkowicie nieistotna dla programu szeregującego superzadania, który traktuje je jako niezależne, wywłaszczalne i periodyczne zadania przeznaczone dla jednej jednostki obliczeniowej, przy czym wewnętrzna struktura rozważanej jednostki obliczeniowej nie ma w tym miejscu żadnego znaczenia.

W miarę możliwości w kolejnym etapie można podjąć próbę dokonania binaryzacji okresów szeregowanych superzadań w celu skrócenia horyzontu czasowego systemu.

W artykule zilustrowano sposób adaptacji algorytmu RMS na przykładzie szeregowania zbioru niezależnych, wywłaszczalnych i periodycznych zadań dla trzech dedykowanych procesorów. Zaproponowana przez autora metoda może zostać z łatwością uogólniona, zarówno dla przypadku większej liczby procesorów, jak i dla systemów, w których zamiast procesorów dedykowanych występują procesory arbitralne. Ponieważ w przypadku systemów z procesorami arbitralnym, żadne z zadań nie jest z góry przydzielone do danego procesora, należy dokonać na wstępie alokacji zadań, polegającej na ich przypisaniu do odpowiednich procesorów. Postępując w ten sposób problem szeregowania zadań dla procesorów arbitralnych zostaje przetransformowany do problemu szeregowania zadań dla procesorów dedykowanych, gdzie metodyka dalszego postępowania została powyżej w artykule już szczegółowo omówiona.

6. BIBLIOGRAFIA

1. T. S z m u c: *Modele i metody inżynierii oprogramowania systemów czasu rzeczywistego*. Uczelniane Wydawnictwa Naukowo-Dydaktyczne AGH, Kraków, 2001
2. K. G. S h i n, P. R a m a n a t h a n: *Real-time computing: A new discipline of computer science and engineering*. Proceedings of the IEEE, vol. 82, no. 1, January 1994, pp. 6-24
3. L. L i u, J. W. L a y l a n d: *Scheduling algorithms for multiprogramming in a hard real time environment*. Journal on Assoc. Comput. Mach., vol. 20, no. 1, 1973, pp. 46-61
4. T. S t o y e n k o, P. B a k e r: *Real-time schedulability- analyzable mechanisms in Ada9X..* Proceedings of the IEEE, vol. 82, no. 1, January 1994, pp. 95-107
5. K. R a m a n a t h a n, J. A. S t a n k o v i c: *Scheduling algorithms and operating systems support for real-time systems*. Proceedings of the IEEE, vol. 82, no. 1, January 1994, pp. 55-67
6. A. Z a l e w s k i: *What every engineer needs to know about rate monotonic scheduling: A tutorial*. Odessa, TX, 1995, pp. 321-335
7. A. S. T a n e n b a u m: *Rozproszone systemy operacyjne*. PWN, Warszawa, 1997
8. J. V a r r i e t: *Scheduling inverted-ordered task with non- uniform deadlines subject tonon-zero-communication delay*. Parallel Computing 25 (1999), pp. 3-21
9. O. K w o n, K. Y. C h w a: *Approximation algorithms for general parallel task scheduling*. Information Processing Letters 81 (2002), pp. 143-150
10. A. C z a j k a, J. N a w r o c k i: *Szeregowanie zadań o okresach binarnych w systemach silnie uwarunkowanych czasowo*. I Krajowa Konferencja Metody i systemy komputerowe w badaniach naukowych i projektowaniu inżynierskim, Kraków, 1997, ss. 669-676
11. J. N a w r o c k i, A. C z a j k a: *Binaryzacja okresów zadań cyklicznych*. VII Konferencja Systemy Czasu Rzeczywistego, Kraków, 2000, ss. 41-51
12. K. W a l e: *How VME facilitates the implementation of real-time DSP systems*. Real Time Magazine, 1/1997, pp. 21-28

M. GAJER

THE ADAPTATION OF RATE MONOTONIC SCHEDULING ALGORITHM FOR
SCHEDULING OF MULTIPROCESSOR TASKS

S u m m a r y

The popularity and ubiquity of real-time systems with hard real-time constraints forced the extensive development of task scheduling theory. In the case of real-time systems with hard real-time constraints it does not suffice that the task produces logically correct results but these results must be delivered within their time constraints.

In such systems even logically correct results but delivered with the violation of their time constraints are totally useless. Moreover, the consequences of violation of time constraints can very often be quite severe and can cause the great economic losses and even losses of human lives, e.g. in the case of control systems of nuclear reactors, space ships etc.

The main goal of the task scheduling theory is to prove at the stage of the system project that the time constraints for all tasks will always be met under any possible circumstances.

In the case of the real-time systems with hard real-time constraints there is very often a necessity of scheduling a set of independent, pre-emptive and periodic tasks. The most popular algorithm for scheduling such set of independent, pre-emptive and periodic tasks is the Rate Monotonic Scheduling algorithm.

In the case of Rate Monotonic Scheduling each task is assigned a priority. There are several rules basing on which the priorities are assigned to the tasks and then the tasks are being scheduled. First of all, the shorter the period of task is the higher priority it is assigned. Then, in a given moment, among all the tasks actually in a ready state the one is being executed that has the highest priority. If some task with higher priority enters into the ready state the task being executed is automatically pre-empted and the task with higher priority begins its execution. The pre-empted task can restart its execution only in the case if there is actually no other task with higher priority in the ready state.

The Rate Monotonic Scheduling is adequate for scheduling uniprocessor tasks. This author has not known so far any method of adaptation of Rate Monotonic Scheduling theory for the purpose of scheduling multiprocessor tasks. According with this author's knowledge the proposed by himself method of scheduling of the set of uniprocessor and multiprocessor tasks is the first method of the kind and thus has totally pioneer character.

The clue of the method proposed by this author is concatenation of uniprocessor and multiprocessor tasks. In order to achieve this the periods of some tasks must be transformed, i.e. they must be shortened in such a way that several subsets of tasks are made. Then each subset of tasks is treated as a uniprocessor task and for the set of such tasks (called by this author supertasks) the Rate Monotonic Scheduling algorithm can be used directly.

The method developed by this author was illustrated on the example of scheduling set of tasks for three dedicated processors. The method can be easily extended both for the case of greater number of processors and for the systems with arbitrary processors.

Keywords: Task Scheduling Theory, Rate Monotonic Scheduling, Multiprocessor Systems, Multiprocessor Tasks

Badanie i optymalizacja szerokopasmowego transformatora dwukierunkowego

DARIUSZ KRZEMIENIECKI

*Gdańskie Zakłady Teleelektroniczne TELKOM-TELMOR
80-425 Gdańsk ul. Mickiewicza 5/7
e-mail: dariusz.krzemieniecki@telmor.pl*

Otrzymano 2004. 01.28

Autoryzowano 2004.10.11

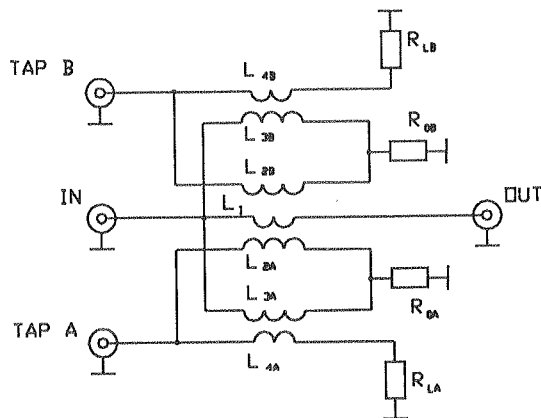
W pracy zamieszczono analizę teoretyczną, wyniki pomiarów i metody optymalizacji szerokopasmowego transformatora dwukierunkowego. Przedstawiono dwie możliwe konfiguracje transformatora, oraz zbadano wpływy takich czynników jak: przenikalność początkowa, rezystancja balastowa, pojemności międzyzwojowe i montażowe. Do badań posłużono się transformatorem symetrycznym 15-decybelowym.

Słowa kluczowe: transformator kierunkowy, konfiguracja podstawowa, konfiguracja odwrócona

1. WPROWADZENIE

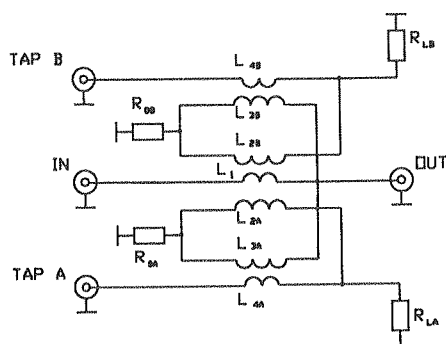
Transformatory kierunkowe w różnych konfiguracjach stosowane są powszechnie w urządzeniach telewizji kablowej (CATV), najczęściej w odgałęźnikach. Wykorzystywane w urządzeniach CATV transformatory, to ściślej mówiąc transformatory jednokierunkowe — z jedną linią odgałęźną. Zostały one dokładnie opisane w [1], [5]. W niniejszej pracy jest zaprezentowany transformator dwukierunkowy — z dwiema liniami odgałęźnymi. Jest to konstrukcja nowa, nigdzie dotąd nie stosowana. Dotychczas, aby skonstruować odgałęźnik dwukrotny posługiwano się kaskadą dwóch transformatorów jednokierunkowych. Transformator dwukierunkowy jest propozycją i alternatywą dla istniejących tradycyjnych rozwiązań. Transformator taki umożliwia wykonanie odgałęźnika dwukrotnego o jednakowych tłumieniach w liniach odgałęźnych, co w przypadku zastosowania kaskady transformatorów jednokierunkowych stanowi problem w tym sensie, że kaskada jest możliwa do wykonania, jednak jest to o wiele bardziej skomplikowane a w przypadku transformatora dwukierunkowego jest natomiast rzeczą

naturalną. Ponadto transformator taki (choć bardziej złożony niż transformator jednokierunkowy) wymaga o 30% mniej materiałów i o tyleż mniejszej pracochłonności niż tradycyjna konstrukcja, składająca się z kaskady dwóch transformatorów jednokierunkowych.



Rys. 1. Transformator dwukierunkowy w konfiguracji podstawowej

Fig. 1. Two-directional transformer in basic configuration



Rys. 2. Transformator dwukierunkowy w konfiguracji odwróconej

Fig. 2. Two-directional transformer in inverse configuration

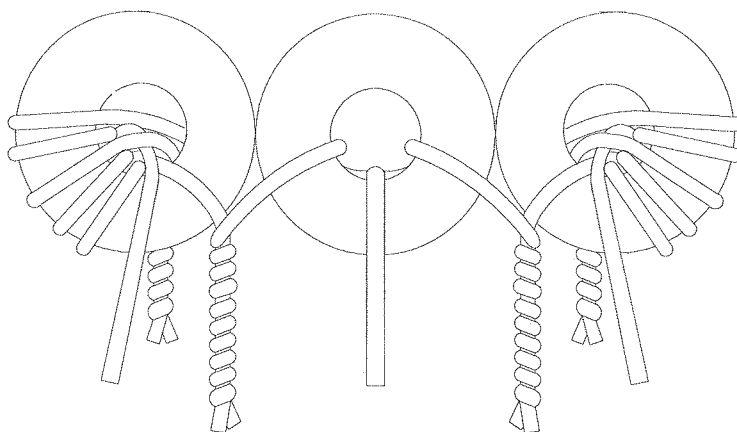
Należy tu jeszcze dodać, że wykorzystując ideę transformatora dwukierunkowego można konstruować analogiczne transformatory o większej liczbie odgałęzień. Wówczas oszczędności na materiałach i pracochłonności są bliskie 50%.

Rys. 1 ilustruje transformator dwukierunkowy w konfiguracji zwanej tu podstawową ze względu na analogię do klasyfikacji przedstawionej w [5]. Rys. 2 przedstawia transformator dwukierunkowy w konfiguracji zwanej tu odwróconą. Dodajmy, że przy opisie transformatora jednokierunkowego [5] wyróżnione zostały cztery konfiguracje

włączając tzw. konfiguracje z rezystorami zerowymi. W przypadku transformatora dwukierunkowego taki podział też jest możliwy z tym, że dla konfiguracji bez rezystorów zerowych (R_{0A} , R_{0B}), transformatory takie charakteryzują się brakiem dopasowania w liniach sprzężonych, a zatem są bezużyteczne. Dlatego rozpatrywać będziemy tylko dwie konfiguracje z rys. 1, 2 zawierające rezystory zerowe.

W pracy dokonano szczegółowej, kompleksowej analizy transformatora dwukierunkowego, badając możliwe czynniki, jakie mają wpływ na jego funkcjonowanie. Praca adresowana jest głównie do konstruktorów, którzy mogliby wykorzystać opisywany transformator zarówno jako odrębne urządzenie, jak również współpracujące w innych urządzeniach.

Omawiany transformator wykonano na bazie małych (rzędu 3×3 mm) rdzeni ferrytowych jednootworowych połączonych w trójkę, na które nawija się uzwojenia zgodnie z rys. 1. Widok takiego transformatora przedstawia rys. 3. a sposób montażu ilustruje fotografia (rys. 4).



Rys. 3. Widok transformatora dwukierunkowego

Fig. 3. View of the two-directional transformer

Analizowany transformator jest transformatorem symetrycznym, a jego uzwojenia L_{1A} , L_{2A} , L_{3A} mają po 0.5 zwoja, uzwojenia L_{3A} , L_{3B} po 4.5 zwoja, natomiast uzwojenia L_{4A} , L_{4B} po 1.5 zwoja. Transformator ten opisano za pomocą następujących parametrów: tłumienie w linii głównej IN-OUT, tłumienie w linii sprzężonej IN-TAP, tłumienie oddzielenia (między wyjściem a odgałęzieniem) OUT-TAP, tłumienie przenikowe (między odgałęzieniami) TAP A-B, oraz tłumienności niedopasowania na wejściu IN, wyjściu OUT i odgałęzieniu TAP. W niniejszej pracy wielkości te wyrażone są w decybelach (dB). Wielkości te zdefiniowane są w punkcie 2.

Do badań użyto rdzenia o przenikalności początkowej $\mu_i = 1400$, oraz dla porównania rdzenia o przenikalności $\mu_i = 2500$. Oba rdzenie są produkcji tajwańskiej.

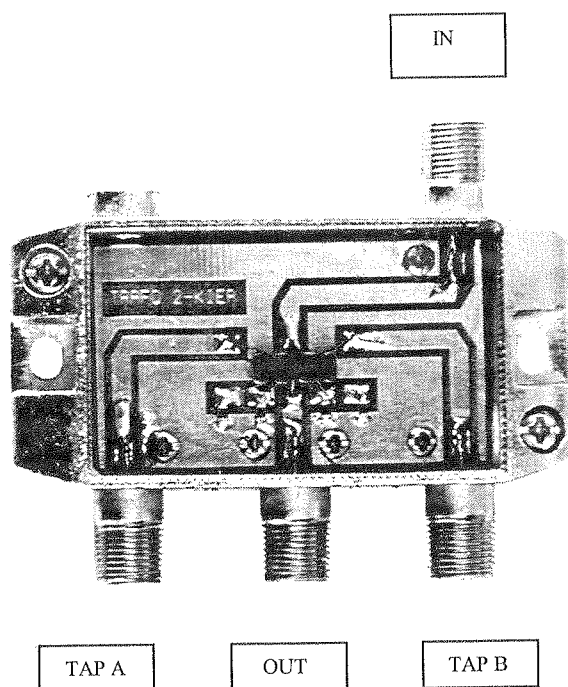
Jako narzędzia programowego do obliczenia teoretycznych charakterystyk transformatora, oraz w celu zbadania wpływów niektórych czynników na jego parametry użyto pakietu MATLAB.

Ze względu na fakt, że na odgałęźniki pracujące w CATV narzucone są pewne wymagania, wykresy pomiarowe zawierają orientacyjne krzywe:

18 dB dla 5–40 MHz, 18–1.5 dB/oct dla 40–862 MHz,

14 dB dla 5–40 MHz, 14–1.5 dB/oct dla 40–862 MHz.

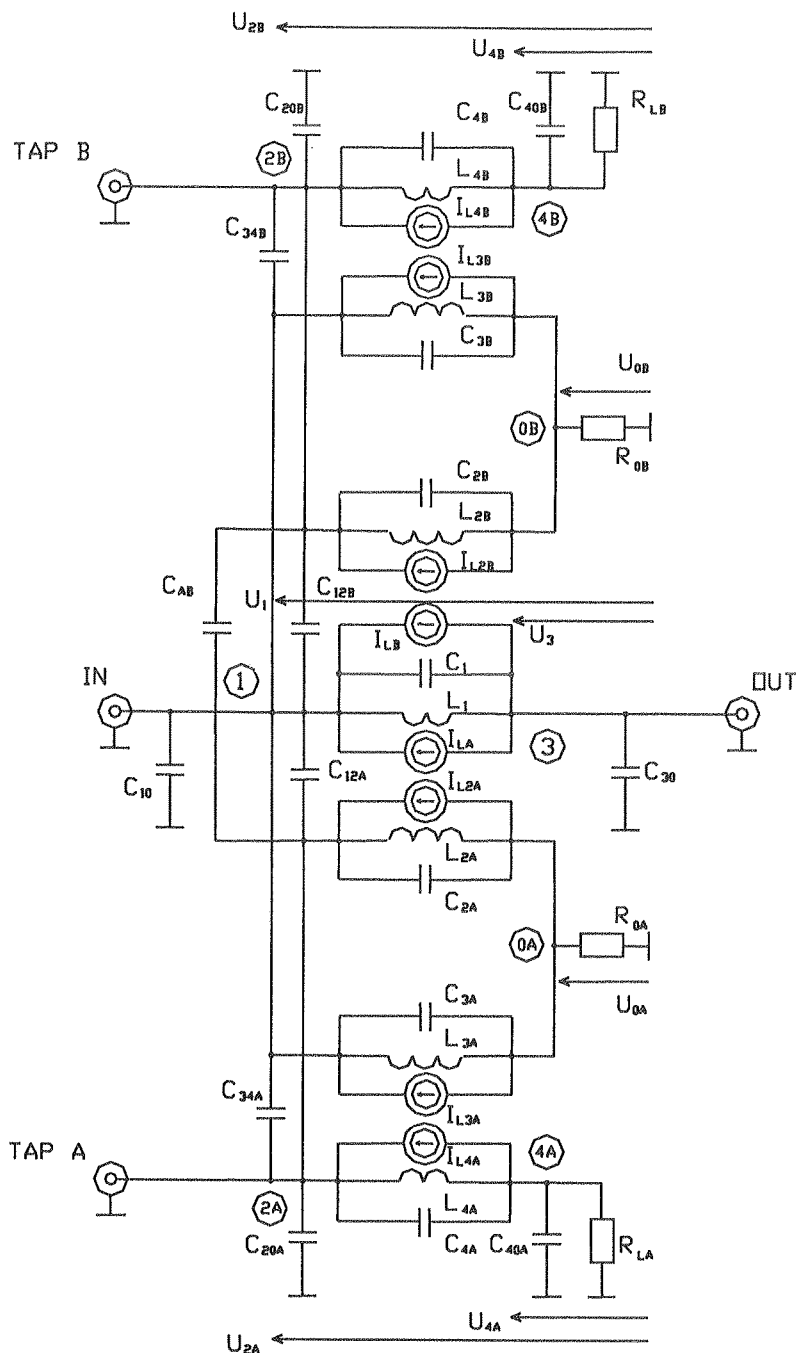
Krzywe te stanowią szablon, który musi spełnić charakterystyka tłumienności niedopasowania dla odgałęźników 2 i 3 kategorii.



Rys. 4. Widok zmontowanego układu transformatora

Fig. 4. View of the assembled transformer

W dalszej części pracy użyto sformułowań typu „poprawa”, bądź „pogorszenie” parametrów. Poprawa parametrów takich, jak tłumienie w linii głównej (IN-OUT) i w linii sprzężonej (IN-TAP) oznacza zmniejszenie ich wartości ujemnej wyrażonej w decybelach. Natomiast dla takich parametrów, jak tłumienność niedopasowania na wejściu (IN), wyjściu (OUT), odgałęzieniu (TAP) i tłumienie oddzielenia (OUT-TAP), przez poprawę parametrów rozumiemy wzrost ich wartości ujemnej wyrażonej w decybelach.



Rys. 5. Model pełny transformatora dwukierunkowego w konfiguracji podstawowej

Fig. 5. Complete model of the two-directional transformer in basic configuration

2. TRANSFORMATOR W KONFIGURACJI PODSTAWOWEJ

2.1. ANALIZA TEORETYCZNA

Poniżej przedstawiona jest skrócona analiza teoretyczna układu transformatora przedstawionego na rys. 5. Do analizy transformatora wykorzystano metodę potencjałów węzłowych. Badany transformator składa się z trzech transformatorów prostych, które tworzone są przez trójkę cewek L_1, L_{2A}, L_{2B} , oraz pary cewek L_{3A}, L_{4A} i L_{3B}, L_{4B} . Dla trójki cewek L_1, L_{2A}, L_{2B} , wyznaczamy najpierw równania napięciowe:

$$U_{2A} - U_{0A} = j\omega L_{2A} I_{L2A} + j\omega M_{2B} I_{L2B} + j\omega M_{1B} I_{L1} \quad (1)$$

$$U_{2B} - U_{0B} = j\omega M_{2A} I_{L2A} + j\omega L_{2B} I_{L2B} + j\omega M_{1A} I_{L1} \quad (2)$$

$$U_1 - U_3 = j\omega M_A I_{L2A} + j\omega M_B I_{L2B} + j\omega L_1 I_{L1} \quad (3)$$

gdzie elementy M to indukcyjności wzajemne jakie występują w analizowanym układzie cewek. Na podstawie równań (1)...(3) otrzymujemy równania prądowe:

$$I_{L1A} = (U_{2A} - U_{0A})\xi_{110} + (U_{2B} - U_{0B})\xi_{120} + (U_1 - U_3)\xi_{130} \quad (4)$$

$$I_{L2B} = (U_{2A} - U_{0A})\xi_{210} + (U_{2B} - U_{0B})\xi_{220} + (U_1 - U_3)\xi_{230} \quad (5)$$

$$I_{L1} = (U_{2A} - U_{0A})\xi_{310} + (U_{2B} - U_{0B})\xi_{320} + (U_1 - U_3)\xi_{330} \quad (6)$$

gdzie elementy ξ stanowią transkonduktancje układu cewek. Następnie piszemy analogiczne równania napięciowe dla pary cewek L_{3A}, L_{4A} :

$$U_{2A} - U_{4A} = j\omega L_{4A} I_{L4A} + j\omega M_{43A} I_{L3A} \quad (7)$$

$$U_1 - U_{0A} = j\omega L_{34A} M_{34A} + j\omega L_{3A} I_{L3A} \quad (8)$$

a stąd otrzymujemy równania prądowe:

$$I_{L4A} = (U_{2A} - U_{4A})\xi_{11A} + (U_1 - U_{0A})\xi_{12A} \quad (9)$$

$$I_{L3A} = (U_{2A} - U_{4A})\xi_{21A} + (U_1 - U_{0A})\xi_{22A} \quad (10)$$

Podobnie dla pary cewek L_{3B}, L_{4B} :

$$U_{2B} - U_{4B} = j\omega L_{4B} I_{L4B} + j\omega M_{43B} I_{L3B} \quad (11)$$

$$U_1 - U_{0B} = j\omega L_{34B} M_{34B} + j\omega L_{3B} I_{L3B} \quad (12)$$

i stąd równania prądowe:

$$I_{L4B} = (U_{2B} - U_{4B})\xi_{11B} + (U_1 - U_{0B})\xi_{12B} \quad (13)$$

$$I_{L3B} = (U_{2B} - U_{4B})\xi_{21B} + (U_1 - U_{0B})\xi_{22B} \quad (14)$$

W równaniach cewek (1),(2),(3),(7),(8),(11) i (12) występują jako współczynniki prądów, elementy macierzy impedancyjnych pojedynczych transformatorów prostych.

W celu rozwiązania równań węzłowych macierze te zostały odwrócone, czego rezultatem są macierze admitancyjne, których elementy występują jako współczynniki napięć w równaniach (4),(5),(6),(9),(10),(13),(14).

Równania węzłowe transformatora z rys. 5 mają postać:

— dla węzła 1:

$$\begin{aligned} U_{1Yf} = & U_1 (\xi_{22A} + \xi_{22B} + \xi_{330} + j\omega C_1 + j\omega C_{3A} + j\omega C_{3B} + j\omega C_{30A} + j\omega C_{30B} + \\ & + j\omega C_{12A} + j\omega C_{12B} + j\omega C_{34A} + j\omega C_{34B}) + \\ & + U_{2A}(\xi_{21A} + \xi_{310} - j\omega C_{12A} - j\omega C_{34A}) + \\ & + U_{2B}(\xi_{21B} + \xi_{320} - j\omega C_{12B} - j\omega C_{34B}) + \\ & - U_{4A}\xi_{21A} - U_{4B}\xi_{21B} - U_3(\xi_{330} + j\omega C_1) + \\ & - U_{0A}(\xi_{22A} + \xi_{310} + j\omega C_{3A}) - U_{0B}(\xi_{22B} + \xi_{320} + j\omega C_{3B}) \end{aligned} \quad (15)$$

— dla węzła 2A:

$$\begin{aligned} U_{2AYf} = & U_1(\xi_{12A} + \xi_{130} - j\omega C_{12A} - j\omega C_{34A}) + \\ & + U_{2A}(\xi_{11A} + \xi_{110} + j\omega C_{2A} + j\omega C_{4A} + j\omega C_{12A} + j\omega C_{34A} + j\omega C_{20A} + j\omega C_{AB}) + \\ & + U_{2B}(\xi_{120} - j\omega C_{AB}) - U_{4A}(\xi_{11A} + j\omega C_{4A}) + \\ & - U_3\xi_{130} - U_{0A}(\xi_{12A} + \xi_{110} + j\omega C_{2A}) - U_{0B}\xi_{120} \end{aligned} \quad (16)$$

— dla węzła 2B:

$$\begin{aligned} U_{2AYf} = & U_1(\xi_{12B} + \xi_{230} - j\omega C_{12B} - j\omega C_{34B}) + U_{2A}(\xi_{210} - j\omega C_{AB}) + \\ & + U_{2B}(\xi_{11B} + \xi_{220} + j\omega C_{2B} + j\omega C_{4B} + j\omega C_{12B} + j\omega C_{34B} + j\omega C_{20B} + j\omega C_{AB}) + \\ & - U_{4B}(\xi_{11B} + j\omega C_{4B}) - U_3\xi_{230} - U_{0A}\xi_{210} + \\ & - U_{0B}(\xi_{12B} + \xi_{220} + j\omega C_{2B}) \end{aligned} \quad (17)$$

— dla węzła 3:

$$\begin{aligned} U_{3Yf} = & -U_1(\xi_{110} + j\omega C_1) - U_{2A}\xi_{310} - U_{2B}\xi_{320} + U_3(\xi_{330} + j\omega C_1 + j\omega C_{30}) + \\ & + U_{0A}\xi_{310} + U_{0B}\xi_{320} \end{aligned} \quad (18)$$

— dla węzła 4A:

$$\begin{aligned} U_{4AYL} = & -U_1\xi_{12A} - U_{2A}(\xi_{11A} + j\omega C_{4A}) + U_{4A}(\xi_{11A} + j\omega C_{4A} + j\omega C_{40A}) + \\ & + U_{0A}\xi_{12A} \end{aligned} \quad (19)$$

— dla węzła 4B:

$$\begin{aligned} U_{4BYL} = & -U_1\xi_{12B} - U_{2B}(\xi_{11B} + j\omega C_{4B}) + U_{4B}(\xi_{11B} + j\omega C_{4B} + j\omega C_{40B}) + \\ & + U_{0B}\xi_{12B} \end{aligned} \quad (20)$$

— dla węzła 0A:

$$U_{0A}y_0 = -U_1(\xi_{22A} + \xi_{130} + j\omega C_{3A}) - U_{2A}(\xi_{21A} + \xi_{110} + j\omega C_{2A}) - U_{2B}\xi_{120} + U_{4A}\xi_{21A} - U_{3}\xi_{130} + U_{0A}(\xi_{22A} + \xi_{110} + j\omega C_{2A} + j\omega C_{3A}) + U_{0B}\xi_{120} \quad (21)$$

— dla węzła 0B:

$$U_{0B}y_0 = -U_1(\xi_{22B} + \xi_{230} + j\omega C_{3B}) - U_{2A}\xi_{210} - U_{2B}(\xi_{21B} + \xi_{220} + j\omega C_{2B}) + U_{4B}\xi_{21B} - U_{3}\xi_{230} + U_{0A}\xi_{210} + U_{0B}(\xi_{22B} + \xi_{220} + j\omega C_{2B} + j\omega C_{3B}) \quad (22)$$

przy czym y_f jest admitancją linii transmisyjnej, natomiast y_L stanowi tzw. admitancję balastową (w omawianym przypadku $y_L = 1/R_L$).

Równania węzłowe (15)...(22) wyznaczają macierz admitancyjną transformatora z rys. 5 na podstawie której po eliminacji węzłów 4A, 4B, 0A, 0B otrzymujemy macierz rozproszenia korzystając z równania macierzowego:

$$[S] = \{z_f[Y] + [I]\}^{-1}\{[I] - z_f[Y]\} \quad (23)$$

gdzie $[Y]$ jest macierzą admitancyjną układu po eliminacji w. wym. węzłów, $[I]$ jest macierzą jednostkową, z_f impedancją linii transmisyjnej (w omawianym przypadku $z_f = 75 \Omega$). Rozważane macierze mają wymiar 4×4 .

Przenikalność jest w rozważanym przypadku wielkością zmienną w funkcji częstotliwości. Jest ona potrzebna do wyznaczenia indukcyjności transformatora. Do wyznaczenia przenikalności posłużono się wzorem zapożyczonym z [1], [2]:

$$\mu = 1 + \frac{\mu_i}{1 + j \frac{f}{f_m}} \quad (24)$$

gdzie: μ_i jest przenikalnością początkową, f_m częstotliwością relaksacji. Poszczególne indukcyjności obliczono ze wzoru:

$$L_i = \mu_i \cdot L_{0i}$$

gdzie L_{0i} jest indukcyjnością cewki nawiniętej na rdzeń bez uwzględnienia jego przenikalności. Indukcyjności wzajemne transformatora obliczono za pomocą wzoru:

$$M_{ij} = k\mu \sqrt{L_{0i} \cdot L_{0j}}$$

gdzie k jest współczynnikiem sprzężenia (w obliczeniach przyjęto $k = 0.99$).

Analizowane parametry transformatora, wyrażone w decybelach, definiujemy następująco:

— tłumienie w linii głównej IN-OUT:

$$S_{31[\text{dB}]} = 20 \log S_{31}$$

— tłumienie w linii sprzężonej IN-TAP:

$$S_{1A[\text{dB}]} = 20 \log S_{1A}$$

$$S_{1B[\text{dB}]} = 20 \log S_{1B}$$

$$\begin{aligned}
 (21) \quad & \text{— tłumienie przenikowe TAP A-B:} & S_{AB[\text{dB}]} &= 20 \log S_{AB} \\
 & \text{— tłumienie oddzielenia AT-TAP:} & S_{3A[\text{dB}]} &= 20 \log S_{3A} \\
 & & S_{3B[\text{dB}]} &= 20 \log S_{3B} \\
 (22) \quad & \text{— tłumienność niedopasowania na wejściu IN:} & S_{11[\text{dB}]} &= 20 \log S_{11} \\
 & \text{— tłumienność niedopasowania na wyjściu OUT:} & S_{33[\text{dB}]} &= 20 \log S_{33} \\
 & \text{— tłumienność niedopasowania na odgałęzieniu TAP:} & S_{AA[\text{dB}]} &= 20 \log S_{AA} \\
 & & S_{BB[\text{dB}]} &= 20 \log S_{BB}
 \end{aligned}$$

gdzie parametry S_{ij} są to elementy macierzy rozproszenia obliczone na podstawie wzoru (23).

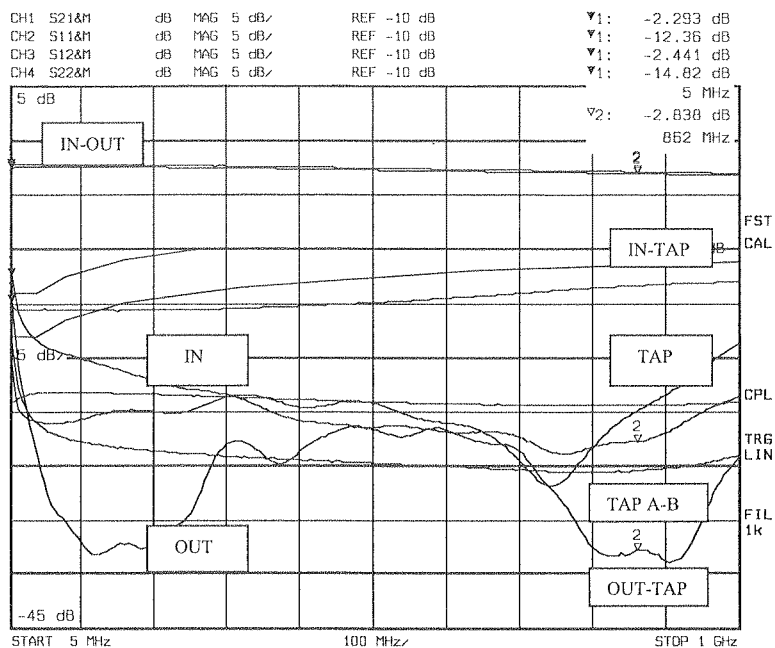
Rys. 6 i 7 przedstawiają charakterystyki transformatora pomierzone i teoretyczne. Zostały one pomierzone i obliczone dla rezystancji balastowej $R_L = 300 \, \Omega$, zerowej $R_0 = 200 \, \Omega$ oraz następujących pojemności transformatora: $C_1 = C_3 = C_{2A} = C_{2B} = 0 \, \text{pF}$, $C_{3A} = C_{3B} = 0.3 \, \text{pF}$, $C_{12A} = C_{12B} = C_{34A} = C_{34B} = 0.075 \, \text{pF}$, oraz $C_{10} = C_{20A} = C_{20B} = C_{30A} = C_{30B} = C_{40A} = C_{40B} = C_{4A} = C_{4B} = 0.05 \, \text{pF}$. Wielkości te zostały oszacowane na podstawie pomiarów rezonansowych i obowiązują dla pomiarów i obliczeń transformatora w obu konfiguracjach. Należy tu jednak podkreślić, że przyjęte pojemności są orientacyjne, gdyż ich pomiar na wysokich częstotliwościach nie jest w pełni dokładny zwłaszcza z uwagi na ich małą wartość. Wpływ tych pojemności jest jednak znaczący dla częstotliwości powyżej 500 MHz, dlatego zostały one uwzględnione w modelu teoretycznym. Do obliczeń przyjęto pojemności C_1 , C_3 , C_{2A} , C_{2B} jako zerowe, ponieważ w rozważanym układzie reprezentują one jedynie pojemności odcinka przewodu przewleczonego przez otwór w rdzeniu.

Różnice między przebiegami teoretycznymi a pomierzonymi są dość duże. Wynikają one z faktu, iż zastosowany wzór (24) zapewnia wg [2] dokładność 10%. Ponadto dla wysokich częstotliwości dochodzi niedoskonałość złączy F, wpływ układu pomiarowego i układu przeniesienia masy.

Charakterystyczną cechą omawianego transformatora (również w konfiguracji odwróconej) jest tendencja do malenia tłumienia w linii sprzężonej wraz ze wzrostem częstotliwości. Można to wyeliminować poprzez zamontowanie równoległych kondensatorów w liniach sprzężonych. Jednakże wyżej wymieniona tendencja niekoniecznie musi stanowić wadę. W sieciach CATV może być ona zaletą ponieważ rekompensuje tłumienie kabla koncentrycznego które w funkcji częstotliwości zachowuje się dokładnie odwrotnie.

2.2. WPLYW REZYSTORA BALASTOWEGO I ZEROWEGO

O parametrach układu decydują w znacznym stopniu rezystancja balastowa R_L i zerowa R_0 . Na podstawie obliczeń można dobrać je tak, aby parametry układu były jak najlepsze. Rys. 8 ilustruje przebiegi parametrów transformatora w funkcji zmiennej rezystancji balastowej wyrażonej w ohmach dla częstotliwości 5 MHz, natomiast rys. 9 dla częstotliwości 862 MHz. Jak widać, wpływ rezystancji balastowej dla obu częstotliwości 5 i 862 MHz jest podobny. Rezystor balastowy praktycznie nie ma wpływu na

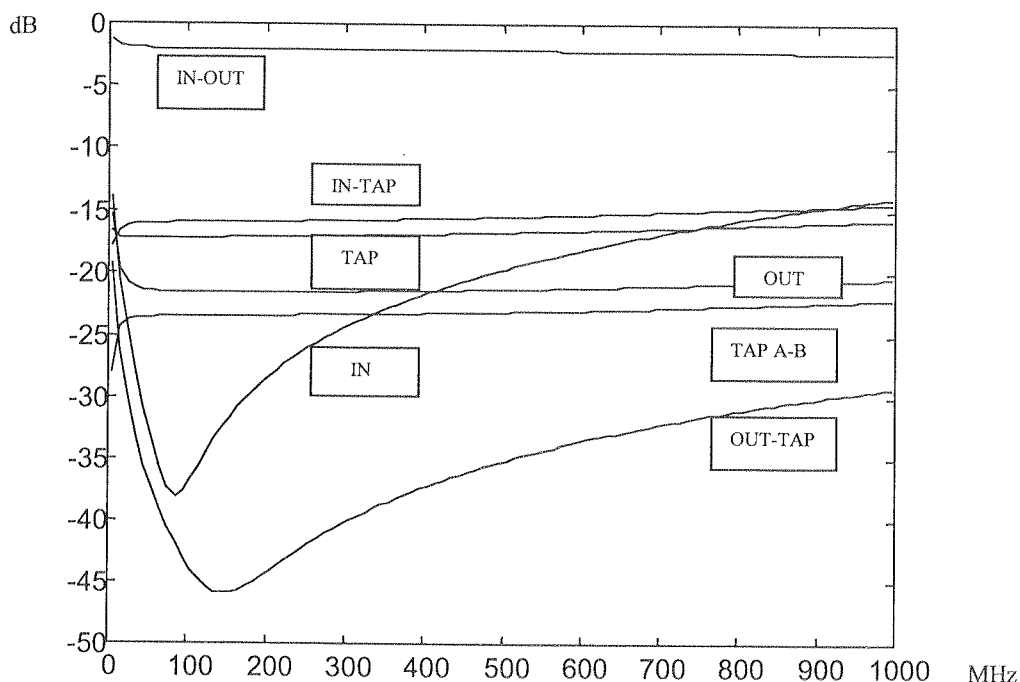


Rys. 6. Pomierzone charakterystyki transformatora dwukierunkowego w konfiguracji podstawowej

Fig. 6. Measured characteristics of the two-directional transformer in basic configuration

tłumienia w linii głównej i tłumienności niedopasowania na wejściu i wyjściu. Wyraźny jest za to wpływ na tłumienność niedopasowania na odgałęzieniu, gdzie dla rezystancji $R_L = 200 \Omega$ jej wielkość osiąga optimum. Ze względu na tłumienie przenikowe i tłumienie oddzielenia, najkorzystniej byłoby stosować jak największe wartości rezystancji R_L , jednak jej wzrost powoduje również wzrost tłumienia w linii sprzężonej. Dlatego do pomiarów z rys. 6 i 7 przyjęto $R_L = 300 \Omega$. Rys. 10. ilustruje wpływ rezystancji zerowej na parametry dla częstotliwości 5 MHz. Takie parametry jak tłumienie w linii głównej, tłumienność niedopasowania na wejściu, wyjściu oraz tłumienie w linii sprzężonej od ok. 200Ω są w przybliżeniu stałe. Tłumienność niedopasowania na odgałęzieniu osiąga minimum dla $R_0 = 150 \Omega$, a tłumienie przenikowe rośnie wraz ze wzrostem R_0 . Dla częstotliwości 862 MHz przebieg poszczególnych parametrów jest podobny z tym, że tłumienność niedopasowania na odgałęzieniu osiąga minimum dla $R_0 = 100 \Omega$, a tłumienie oddzielenia osiąga minimum dla $R_0 = 150 \Omega$. Jednak widać z rys. 10, że dla małych wartości R_0 tłumienność niedopasowania na odgałęzieniu i tłumienie przenikowe są wówczas za małe, dlatego wartością rezystancji zerowej, która „godzi” wszystkie parametry jest $R_0 = 200 \Omega$.

Chodzi o to, że powyższe wartości rezystancji balastowej R_L i zerowej R_0 dobrano mając na względzie wszystkie parametry transformatora. Możliwe są, rzecz jasna, również inne rozwiązania podyktowane szczególnymi potrzebami np. gdy zależy nam



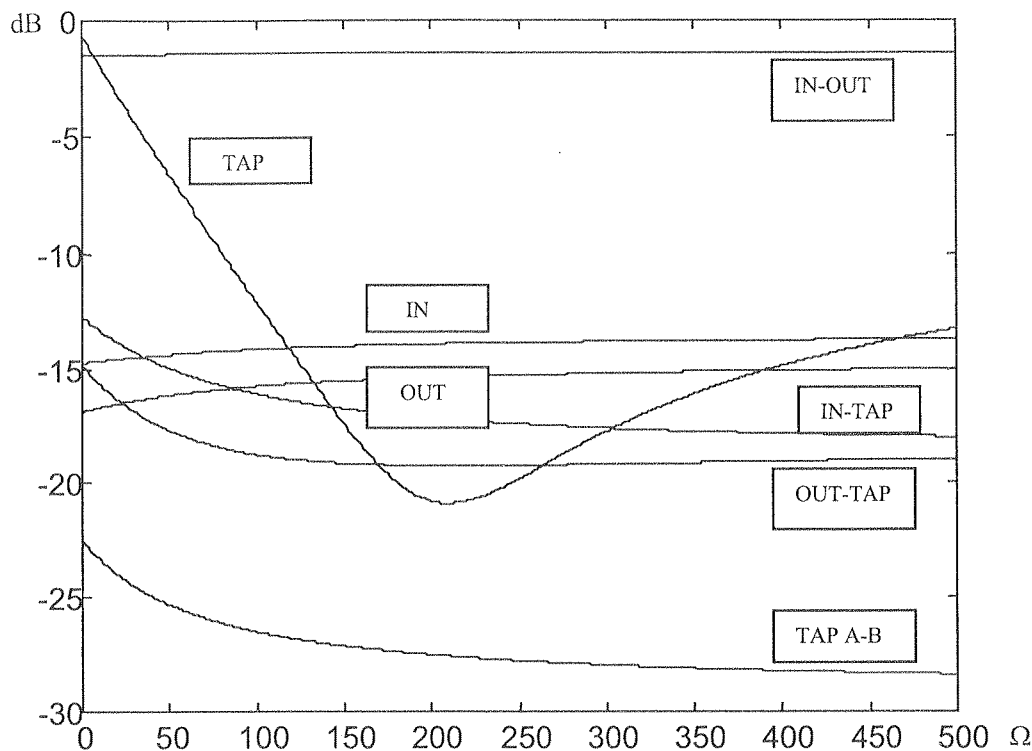
Rys. 7. Obliczone charakterystyki transformatora dwukierunkowego w konfiguracji podstawowej

Fig. 7. Calculated characteristics of the two-directional transformer in basic configuration

na konkretnym parametrze i dla konkretnych częstotliwości. I tak np. w celu uzyskania jak najlepszego tłumienia przenikowego (dla niskich częstotliwości), korzystniej jest zastosować większe wartości rezystancji balastowej i zerowej. Odbyłoby się to jednak kosztem pogorszenia tłumienności niedopasowania na odgałęzieniu.

2.3. WPLYW POJEMNOŚCI MIĘDZYWOJOWYCH I MONTAŻOWYCH

W niniejszym punkcie przedstawiony jest wpływ pojemności transformatora na jego parametry. Generalnie można powiedzieć, że pojemności międzyzwojowe C_1 , C_{2A} , C_{2B} , C_{3A} , C_{3B} , C_{4A} , C_{4B} , C_{12A} , C_{12B} , C_{34A} , C_{34B} , oraz montażowe, C_{10} , C_{20A} , C_{20B} , C_{30A} , C_{30B} , C_{40A} , C_{40B} mają charakter pasożytniczy. Wynika to zarówno z badań jak i obliczeń. Pojemności C_{10} , C_{30} ujemnie wpływają na tłumienności niedopasowania na wejściu i wyjściu, oraz na tłumienie w linii głównej (rys. 11). Pojemności C_{20A} , C_{20B} , C_{40A} , C_{40B} pogarszają tłumienność niedopasowania na odgałęzieniu i tłumienie w linii sprzężonej (rys. 12). Wpływ tych pojemności jest jednak znikomy ze względu na ich małe wartości. Również wpływ pojemności C_{4A} , C_{4B} jest nieznaczny, ponieważ 1.5 zwoja stanowi małą pojemność. Niezależnie od tego ich wpływ na parametry transformatora jest minimalny zakładając nawet większe zmiany pojemności jak pokazano na rys. 13. Pojemności C_1 , C_{2A} , C_{2B} pominięto, gdyż są tak małe, że nie mają jakie-

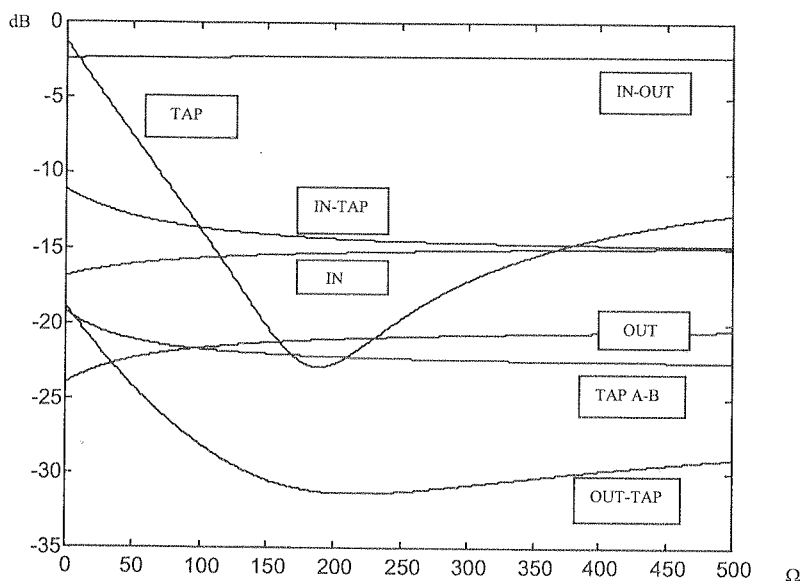


Rys. 8. Wpływ rezystancji balastowej R_L na charakterystyki transformatora w konfiguracji podstawowej dla częstotliwości 5 MHz

Fig. 8. Influence of ballast resistance R_L to characteristics of the transformer in basic configuration for frequency 5 MHz

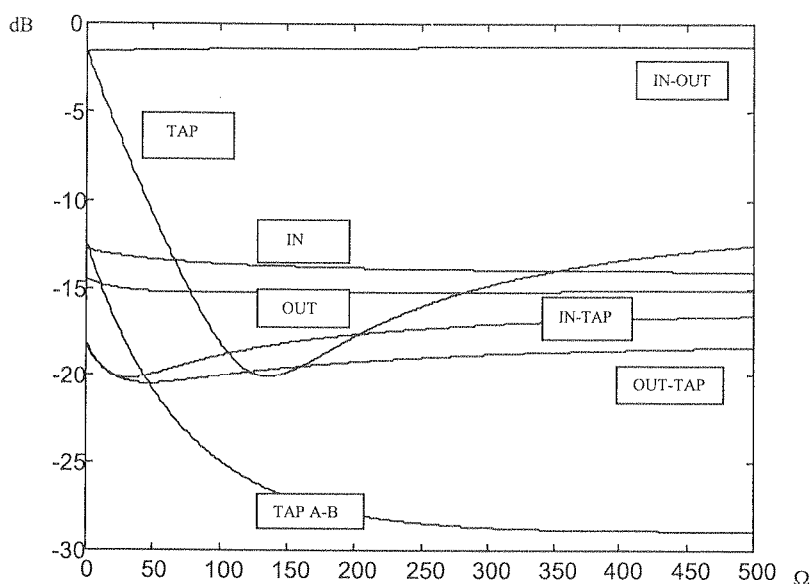
gokolwiek wpływu na parametry transformatora. Największy wpływ mają pojemności międzyzwojowe C_{3A} , C_{3B} , C_{34A} , C_{34B} . Rys. 14. ilustruje wpływ pojemności międzyzwojowych C_{3A} , C_{3B} , natomiast rys. 15 pokazuje wpływ C_{34A} , C_{34B} na parametry transformatora przy hipotetycznych zmianach wartości pojemności od 0 do 1 pF.

Pojemności C_{3A} , C_{3B} znacząco wpływają na tłumienie w linii sprzężonej, tłumienność niedopasowania na wejściu i tłumienie oddzielenia, jednak na pozostałe parametry praktycznie nie oddziałują. To te właśnie pojemności (obok C_{34A} , C_{34B}) powodują opisaną wcześniej tendencję układu do malenia tłumienia w linii sprzężonej wraz ze wzrostem częstotliwości. Pojemności C_{34A} , C_{34B} wpływają praktycznie na wszystkie parametry, choć w mniejszym stopniu na tłumienie w linii głównej i tłumienie przenikowe. Analogiczny wpływ na parametry układu mają pojemności C_{12A} , C_{12B} , chociaż ze względu na fakt iż jest to pojemność równoległych odcinków przewodów, pojemności te nie odgrywają ważącej roli. Wpływ tych pojemności jest negatywny. Kwestią dyskusji jednak pozostaje wpływ pojemności C_{3A} , C_{3B} , C_{34A} , C_{34B} na tłumienie w linii sprzężonej, gdyż zależy to od wykorzystania transformatora.



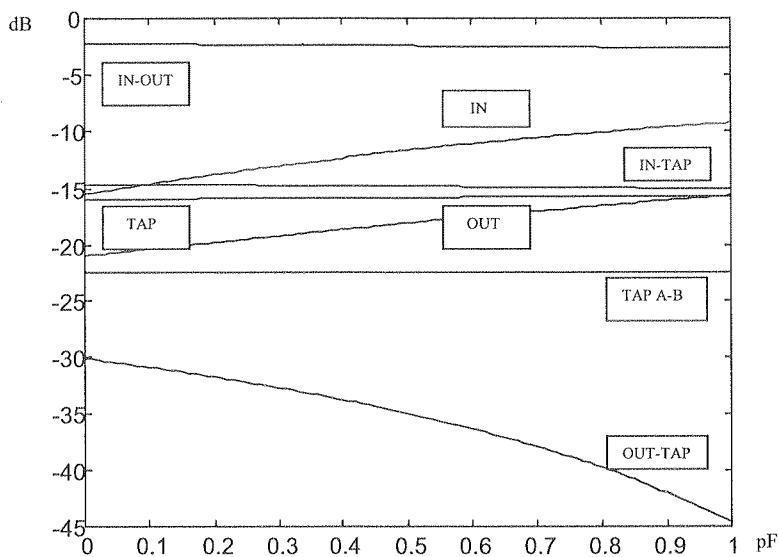
Rys. 9. Wpływ rezystancji balastowej R_L na charakterystyki transformatora w konfiguracji podstawowej dla częstotliwości 862 MHz

Fig. 9. Influence of ballast resistance R_L to characteristics of the transformer in basic configuration for frequency 862 MHz



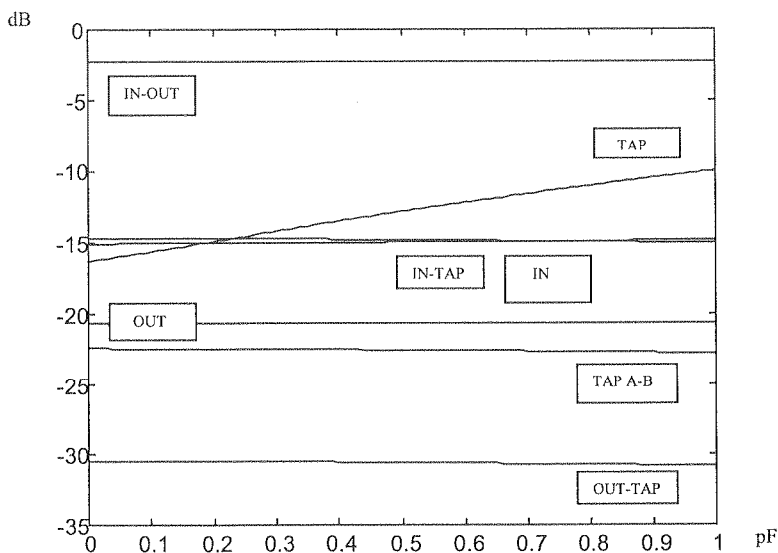
Rys. 10. Wpływ rezystancji zerowej R_0 na charakterystyki transformatora dla częstotliwości 5 MHz

Fig. 10. Influence of zero resistance R_0 to characteristics of the transformer in basic configuration for frequency 5 MHz



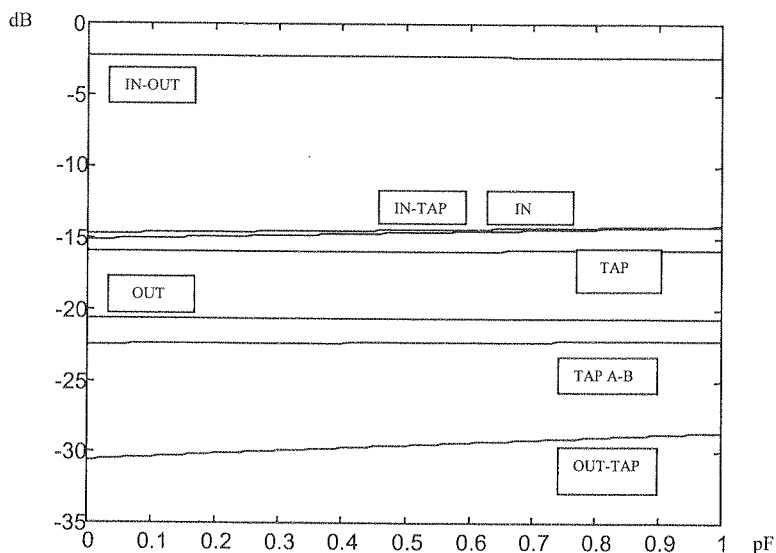
Rys. 11. Wpływ pojemności montażowej C_{30A} na charakterystyki transformatora w konfiguracji podstawowej dla częstotliwości 862 MHz

Fig. 11. Influence of assembly capability C_{30A} to characteristics of the transformer in basic configuration for frequency 862 MHz



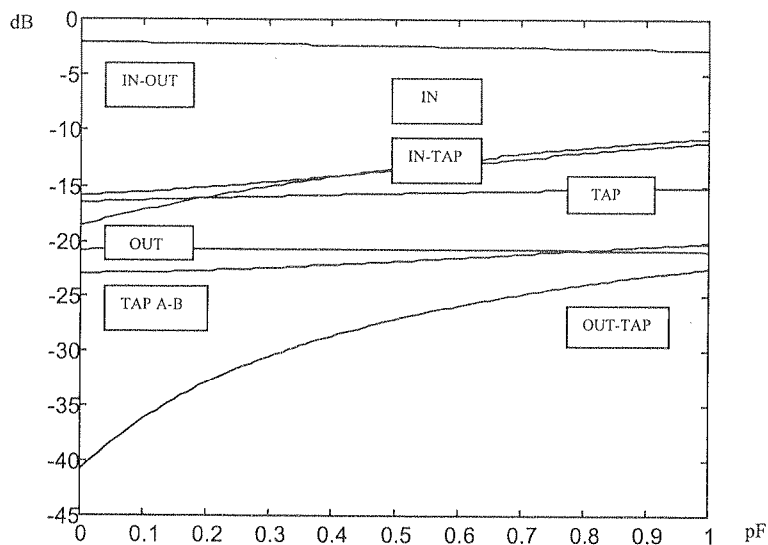
Rys. 12. Wpływ pojemności montażowej C_{40A} na charakterystyki transformatora w konfiguracji podstawowej dla częstotliwości 862 MHz

Fig. 12. Influence of assembly capability C_{40A} to characteristics of the transformer in basic configuration for frequency 862 MHz



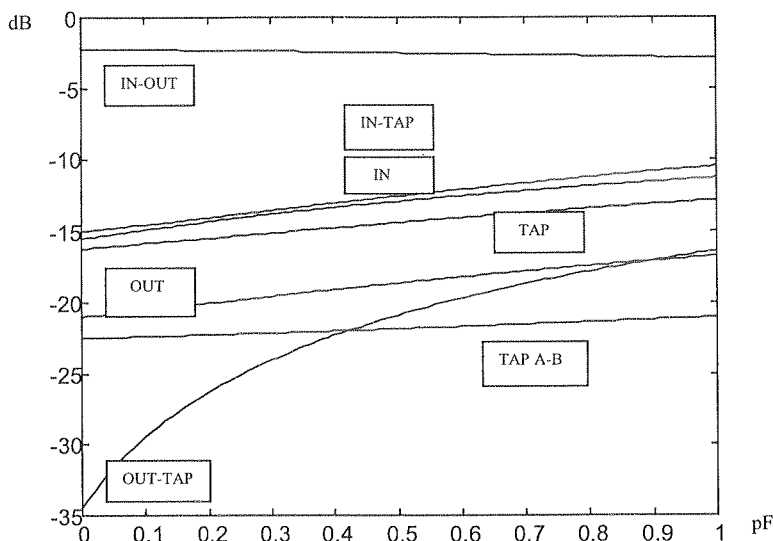
Rys. 13. Wpływ pojemności międzyzwojowej C_{4A} na charakterystyki transformatora w konfiguracji podstawowej dla częstotliwości 862 MHz

Fig. 13. Influence of winding capability C_{4A} to characteristics of the transformer in basic configuration for frequency 862 MHz



Rys. 14. Wpływ pojemności międzyzwojowej C_{3A} na charakterystyki transformatora w konfiguracji podstawowej dla częstotliwości 862 MHz

Fig. 14. Influence of winding capability C_{3A} to characteristics of the transformer in basic configuration for frequency 862 MHz



Rys. 15. Wpływ pojemności międzyzwojowej C_{34A} na charakterystyki transformatora w konfiguracji podstawowej dla częstotliwości 862 MHz

Fig. 15. Influence of winding capability C_{34A} to characteristics of the transformer in basic configuration for frequency 862 MHz

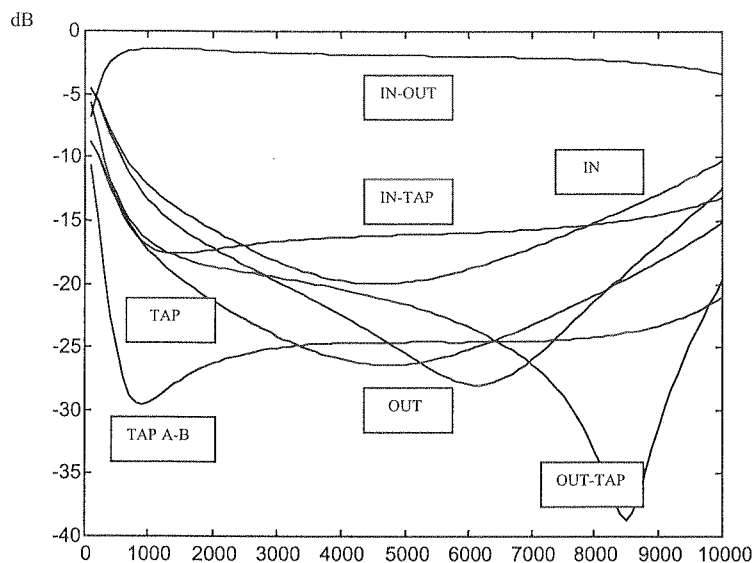
Fig.

2.4. WPŁYW PRZENIKALNOŚCI POCZĄTKOWEJ

Przenikalność początkowa μ_i rdzenia ferromagnetycznego na który nawinięte są uzwojenia transformatora ma decydujący wpływ na jego parametry. Rys. 16 obrazuje przebieg wpływu przenikalności na parametry transformatora dla częstotliwości 5 MHz, rys. 17, dla częstotliwości 862 MHz. Z rys. 16. widać, że najkorzystniejsze dla transformatora pracującego dla częstotliwości 5 MHz, są przenikalności z przedziału 3000 - 5000, chociaż ze względu na tłumienie przenikowe najlepsza jest przenikalność ok. 1000. Zbyt duże przenikalności nie są dla transformatora korzystne, gdyż tłumienie w linii głównej rośnie wraz ze wzrostem przenikalności początkowej. Ponieważ tłumienie w linii głównej i sprzężonej nie są stałe w funkcji przenikalności początkowej, a zatem przy projektowaniu takiego transformatora należy się liczyć ze zmianą tłumień w zależności od użytego rdzenia. Rys. 18 i 19 przedstawiają charakterystyki pomierzone i obliczone dla transformatora z wykorzystaniem rdzenia o przenikalności początkowej $\mu_i = 2500$.

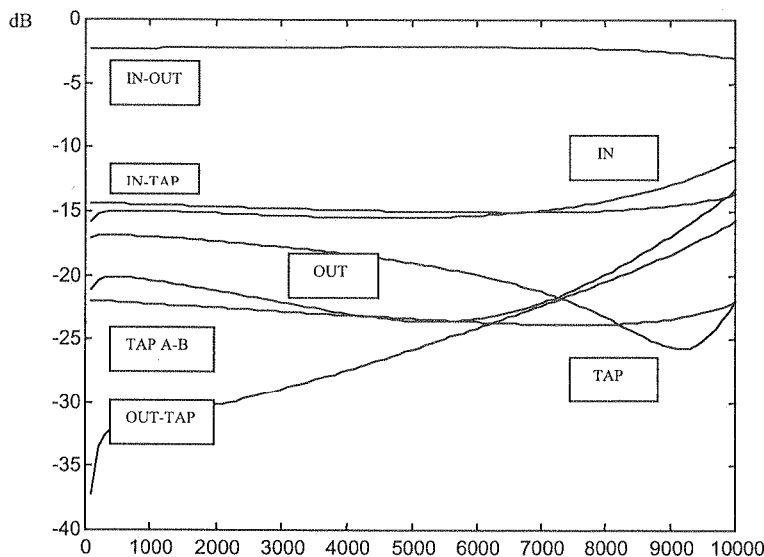
W stosunku do charakterystyk z rys. 6 i 7 (pomiaru dla przenikalności początkowej $\mu_i = 1400$), widoczna jest poprawa tłumienności niedopasowania na wszystkich wrotach, oraz tłumienie oddzielenia dla niskich częstotliwości. Jednak tłumienie przenikowe zgodnie z oczekiwaniami z rys. 16. uległo pogorszeniu. Wprawdzie nie odpowiada to dokładnie obliczeniom, ale jest zgodne z tendencją, jakie mają opisywane parametry w funkcji przenikalności z rys. 16.

Fig



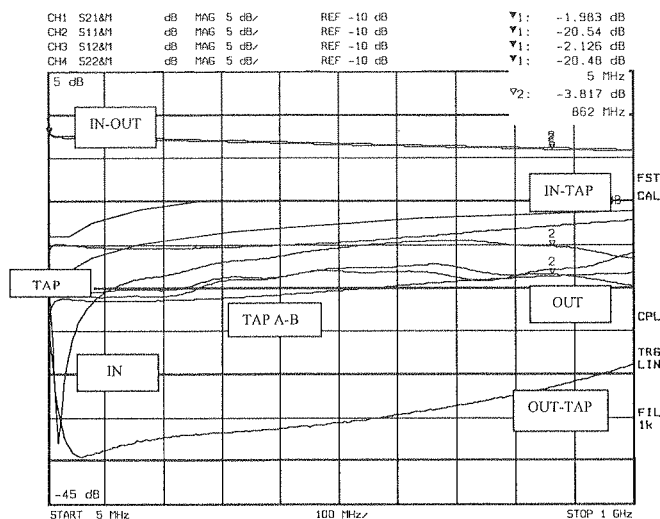
Rys. 16. Wpływ przenikalności początkowej na charakterystyki transformatora w konfiguracji podstawowej dla częstotliwości 5 MHz

Fig. 16. Influence of initial permeability to characteristics of the transformer in basic configuration for frequency 5 MHz



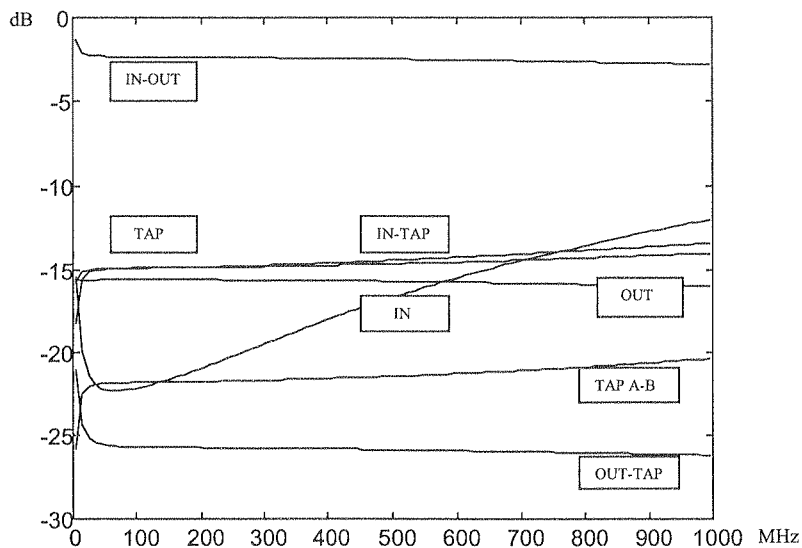
Rys. 17. Wpływ przenikalności początkowej na charakterystyki transformatora w konfiguracji podstawowej dla częstotliwości 862 MHz

Fig. 17. Influence of initial permeability to characteristics of the transformer in basic configuration for frequency 862 MHz



Rys. 18. Pomierzone charakterystyki transformatora dwukierunkowego w konfiguracji podstawowej dla przenikalności $\mu_i = 2500$

Fig. 18. Measured characteristics of the two-directional transformer in basic onfiguration for initial permeability $\mu_i = 2500$



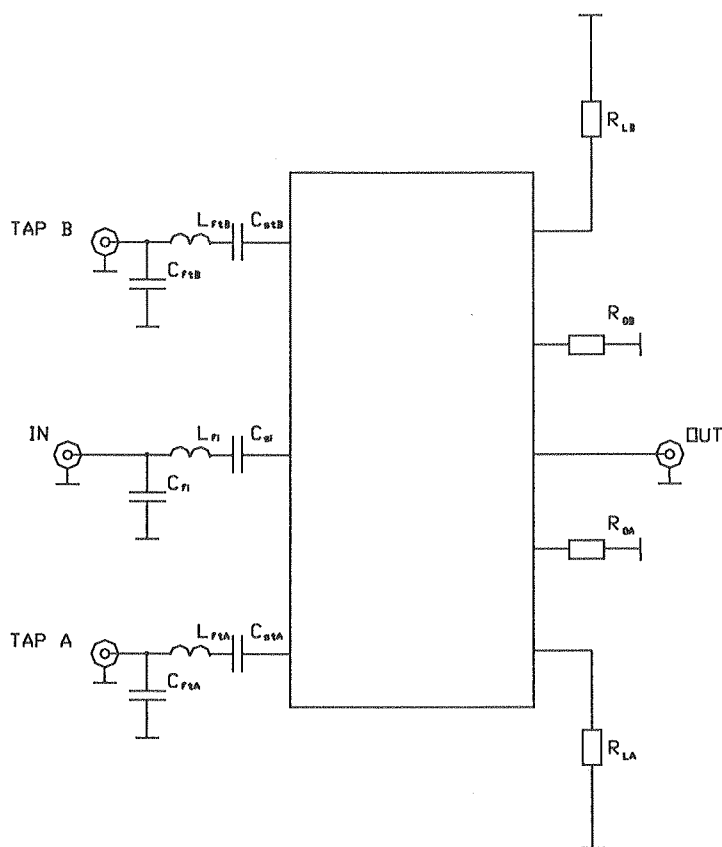
Rys. 19. Obliczone charakterystyki transformatora dwukierunkowego w konfiguracji podstawowej dla przenikalności $\mu_i = 2500$

Fig. 19. Calculated characteristics of the two-directional transformer in basic onfiguration for initial permeability $\mu_i = 2500$

Przebiegi parametrów transformatora w funkcji przenikalności początkowej są różne dla różnych częstotliwości. Rys. 17 przedstawia przebieg wpływu przenikalności na parametry transformatora dla częstotliwości 862 MHz. W tym przypadku optimum przenikalności przesunęło się na niższe jej wartości, chociaż trudniej jest ustalić przenikalność optymalną. Na przykład ze względu na tłumienie oddzielenia przenikalność $\mu_i = 1400$ wydaje się najkorzystniejsza, dlatego przebiegi tłumienia oddzielenia z rys. 6, 7 są tak dobre dla wysokich częstotliwości. Podobnie jak dla niskich częstotliwości, również w tym przypadku niekorzystne jest stosowanie zbyt wysokich przenikalności.

2.5. OPTYMALIZACJA TRANSFORMATORA

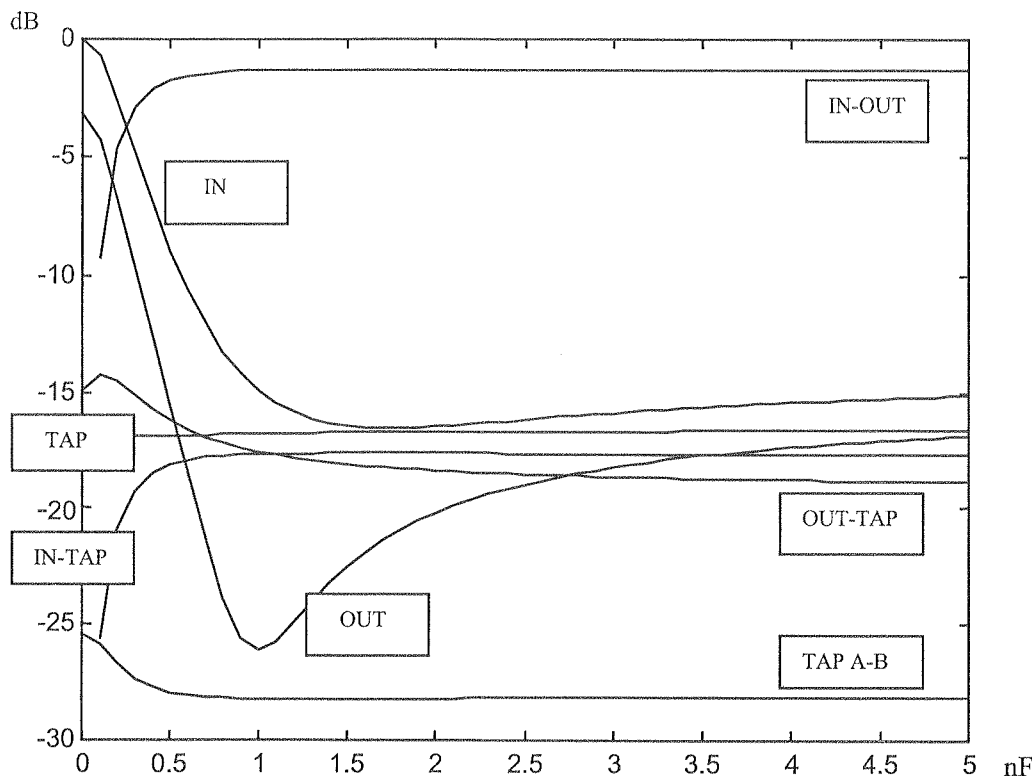
W niniejszym punkcie przedstawione zostały metody optymalizacji parametrów transformatora zarówno dla niskich, jak i wysokich częstotliwości. Poprawę parametrów uzyskuje się stosując proste układy korekcyjne pokazane na rys. 20.



Rys. 20. Układy korekcyjne transformatora dwukierunkowego

Fig. 20. Corrective arrangements of the two-directional transformer

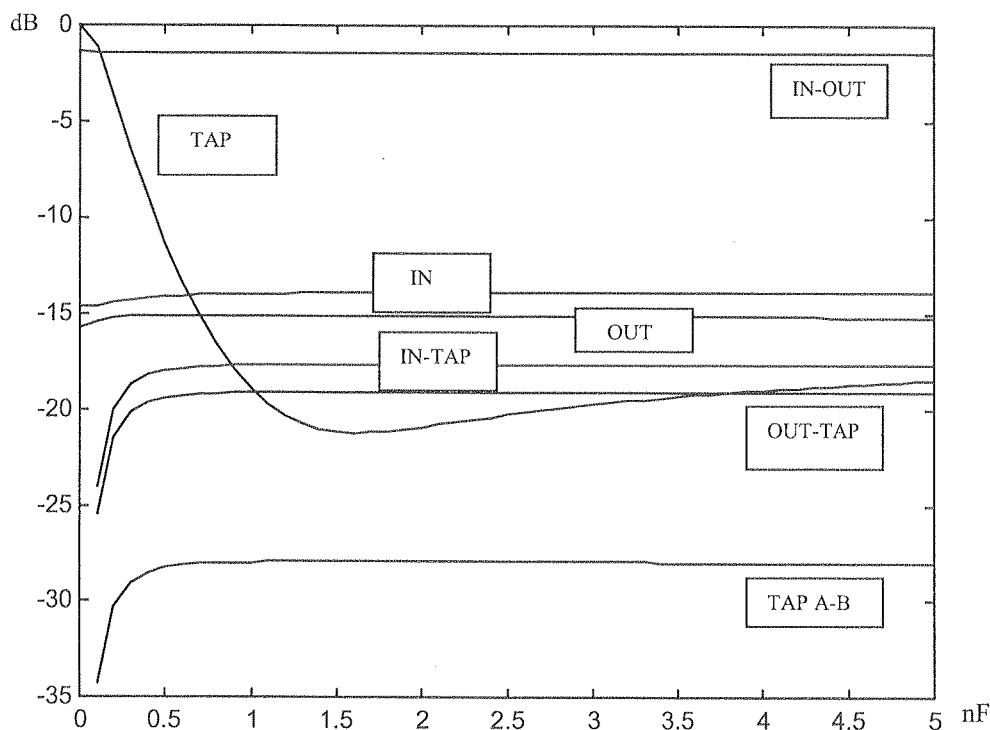
W celu poprawy parametrów na wejściu dla niskich częstotliwości zastosowano duży (rzędu kilku nanofaradów) kondensator szeregowy C_{si} . Efekt zastosowania kondensatora szeregowego ilustruje rys. 21. Kondensator ten ma wpływ na tłumienność niedopasowania na wejściu i (zwłaszcza) na wyjściu i na tłumienie oddzielenia. Nie ma on wpływu na tłumienia w linii głównej i sprzężonej, tłumienność niedopasowania na odgałęzieniu i na tłumienie przenikowe. Można dobrać tak wielkość kondensatora szeregowego C_{si} , aby uzyskać poprawę tych parametrów na które ten kondensator wpływa. Na przykład zastosowanie kondensatora $C_{si} = 1$ nF daje dużą poprawę tłumienności niedopasowania na wyjściu.



Rys. 21. Wpływ pojemności korekcyjnej C_{si} na charakterystyki transformatora dla częstotliwości 5 MHz

Fig. 21. Influence of corrective capability C_{si} to characteristics of the transformer for frequency 5 MHz

Analogicznie na odgałęzieniach można zastosować kondensator szeregowy C_{st} . On z kolei wpływa na tłumienność niedopasowania na odgałęzieniu, której przebieg jest analogiczny jak tłumienność niedopasowania na wyjściu (rys. 22). Na pozostałe parametry nie ma on wpływu. Zastosowanie kondensatorów szeregowych na wejściu i odgałęzieniu pozwala zatem na poprawę parametrów rozpatrywanego tu transformatora za wyjątkiem oczywiście tłumień w linii głównej i sprzężonej, co zresztą nie jest potrzebne.



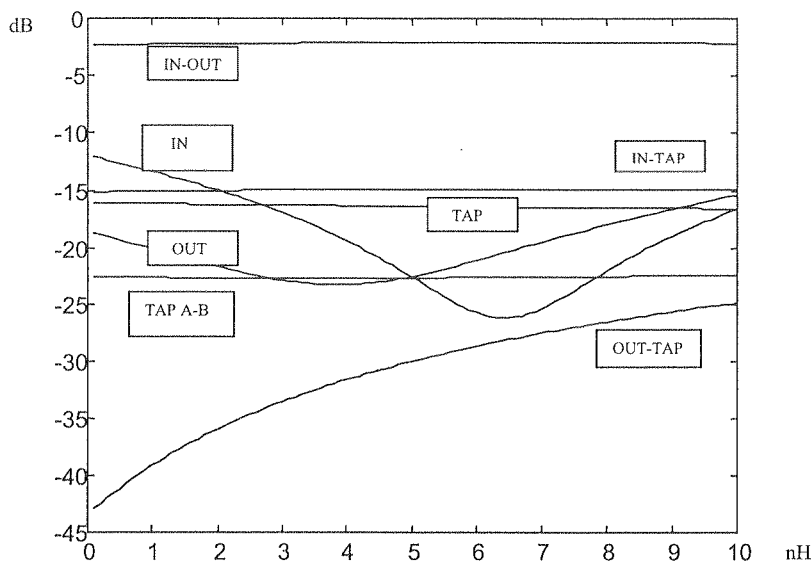
Rys. 22. Wpływ pojemności korekcyjnej C_{st} na charakterystyki transformatora dla częstotliwości 5 MHz

Fig. 22. Influence of corrective capability C_{st} to characteristics of the transformer for frequency 5 MHz

Aby poprawić parametry transformatora dla wysokich częstotliwości stosuje się na wejściu i odgałęzieniu układy korekcyjne LC. Na rys. 20 układy te stanowią odpowiednio $L_f C_{fi}$ na wejściu, oraz $L_{fi} C_{fi}$ na odgałęzieniu. Poprzez odpowiedni dobór pojemności i indukcyjności można uzyskać poprawę parametrów dla wysokich częstotliwości (rys. 23). Rys. 24 ilustruje efekty zastosowania układów korekcyjnych. Należy go porównać z rys. 6, który wprowadza ilustruje przebiegi dla optymalnej wartości rezystora balastowego, ale bez układów korekcyjnych. Dzięki zastosowaniu kondensatorów szeregowych C_{si} , C_{st} , uzyskano poprawę tłumienności niedopasowania na wejściu, wyjściu i odgałęzieniu, oraz poprawę tłumienia oddzielenia dla niskich częstotliwości.

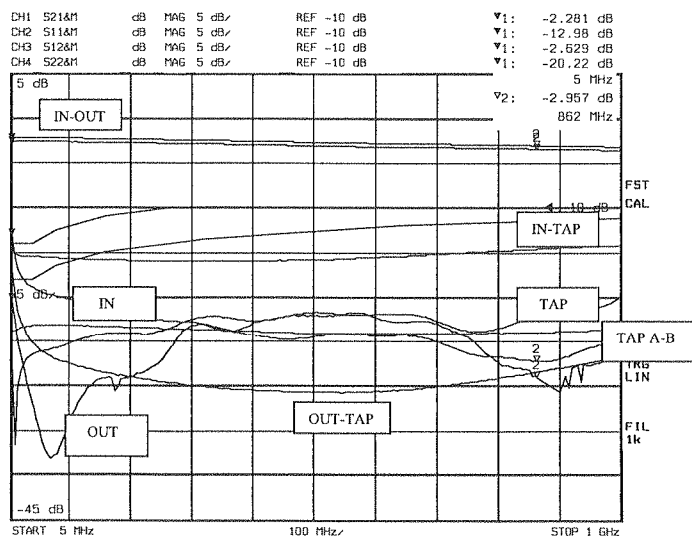
Dla wysokich częstotliwości również uzyskano poprawę wszystkich parametrów dzięki zastosowaniu korektorów LC. Przykładowe przebiegi parametrów w funkcji wielkości korekcyjnych dla wysokich częstotliwości przedstawione są na rys. 23.

Należy tu nadmienić, że optymalizację transformatora należy wspomóc badaniami praktycznymi, gdyż ze względu na 10% dokładność wzoru (24) uzyskanie optimum nie daje pełnej gwarancji uzyskania zamierzonego rezultatu. Optymalizacja transformatora dla wysokich częstotliwości jest trudniejsza z kilku względów. Otóż wartości kondensatorów korekcyjnych są rzędu ułamków pikofarada, natomiast cewek – rzędu



Rys. 23. Wpływ indukcyjności korekcyjnej L_{fi} na charakterystyki transformatora przy stałej pojemności korekcyjnej C_{fi} dla częst. 862 MHz

Fig. 23. Influence of corrective inductance L_{fi} to characteristics of the transformer for constant corrective capability C_{fi} for freq. 862 MHz



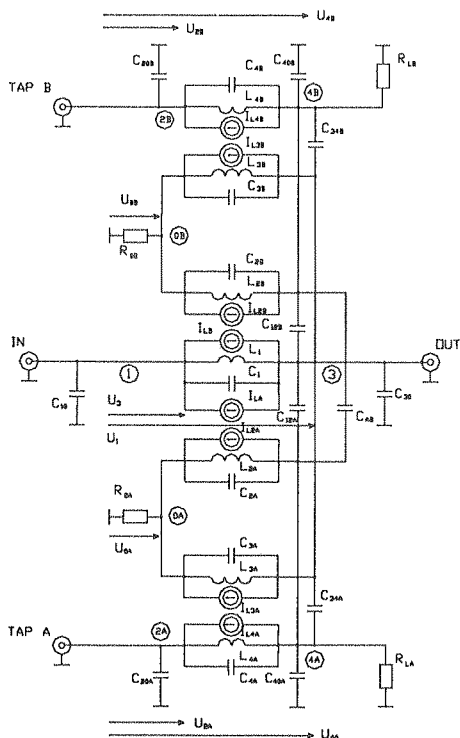
Rys. 24. Charakterystyki transformatora w konfiguracji podstawowej po optymalizacji

Fig. 24. Characteristics of the transformer in basic configuration after optimization

kilku nanohenrów. Ze względu na tak małe wartości, drukowane są one w miedzi na podstawie badań doświadczalnych. Dobór ich komplikuje obok niejednoznaczności z obliczeniami teoretycznymi, szereg dodatkowych czynników mających miejsce dla wysokich częstotliwości, o których wspomniano w punkcie 2.1.

3. TRANSFORMATOR W KONFIGURACJI ODWRÓCONEJ

Analizę transformatora w konfiguracji odwróconej przeprowadzono w oparciu o rys. 25. Badany transformator składa się podobnie jak poprzednio z transformatorów prostych, które tworzone są przez trójkę cewek L_1 , L_{2A} , L_{2B} , oraz pary cewek L_{3A} , L_{4A} i L_{3B} , L_{4B} .



Rys. 25. Model pełny transformatora dwukierunkowego w konfiguracji odwróconej

Fig. 25. Complete model of the two-directional transformer in inverse configuration

Dla trójki cewek L_1 , L_{2A} , L_{2B} , wyznaczamy równania napięciowe:

$$U_{4A} - U_{0A} = j\omega L_{2A} I_{L2A} - j\omega M_{2B} I_{L2B} - j\omega M_{1B} I_{L1} \quad (25)$$

$$U_1 - U_3 = -j\omega M_A I_{L2A} - j\omega M_B I_{L2B} + j\omega L_1 I_{L1} \quad (26)$$

$$U_{4B} - U_{0B} = -j\omega M_{2A} I_{L2A} + j\omega L_{2B} I_{L2B} - j\omega M_{1A} I_{L1} \quad (27)$$

gdzie elementy M to tak jak w p. 2.1 indukcyjności wzajemne, jakie występują w analizowanym układzie cewek; na podstawie równań (25)...(27) otrzymujemy równania prądowe:

$$I_{L1A} = (U_{4A} - U_{0A})\xi_{110} + (U_1 - U_3)\xi_{120} + (U_{4B} - U_{0B})\xi_{130} \quad (28)$$

$$I_{L1} = (U_{4A} - U_{0A})\xi_{210} + (U_1 - U_3)\xi_{220} + (U_{4B} - U_{0B})\xi_{230} \quad (29)$$

$$I_{L2B} = (U_{4A} - U_{0A})\xi_{310} + (U_1 - U_3)\xi_{320} + (U_{4B} - U_{0B})\xi_{330} \quad (30)$$

gdzie elementy ξ stanowią transkonduktancje układu cewek;

Analogicznie dla pary cewek L_{3A} , L_{4A} :

$$U_{2A} - U_{4A} = j\omega L_{4A} I_{L4A} - j\omega M_{43A} I_{L3A} \quad (31)$$

$$U_3 - U_{0A} = -j\omega L_{34A} I_{L4A} + j\omega L_{3A} I_{L3A} \quad (32)$$

a stąd:

$$I_{L4A} = (U_{2A} - U_{4A})\xi_{44A} + (U_3 - U_{0A})\xi_{43A} \quad (33)$$

$$I_{L3A} = (U_{2A} - U_{4A})\xi_{34A} + (U_3 - U_{0A})\xi_{33A} \quad (34)$$

i dla pary cewek L_{3B} , L_{4B} :

$$U_{2B} - U_{4B} = j\omega L_{4B} I_{L4B} - j\omega M_{43B} I_{L3B} \quad (35)$$

$$U_3 - U_{0B} = -j\omega L_{34B} I_{L4A} + j\omega L_{3B} I_{L3B} \quad (36)$$

stąd:

$$I_{L4B} = (U_{2B} - U_{4B})\xi_{44B} + (U_1 - U_{0B})\xi_{43B} \quad (37)$$

$$I_{L3B} = (U_{2B} - U_{4B})\xi_{34B} + (U_1 - U_{0B})\xi_{33B} \quad (38)$$

Analogicznie jak w p. 2.1 współczynniki w napięciowych równaniach (25), (26), (27), (31), (32), (35), (36) tworzą macierze impedancyjne pojedynczych transformatorów prostych. Macierze te zostały odwrócone. Na ich podstawie utworzono równania napięciowe (28), (29), (30), (33), (34), (37), (38).

Równania węzłowe transformatora przyjmują postać:

— dla węzła 1:

$$U_{1Yf} = U_1(\xi_{220} + j\omega C_1 + j\omega C_{10}) + U_3(\xi_{220} + j\omega C_1 + j\omega C_{30}) + \\ + U_{4A}\xi_{210} + U_{4B}\xi_{230} - U_{0A}\xi_{210} - U_{0B}\xi_{230} \quad (39)$$

— dla węzła 2A:

$$U_{2AYf} = U_{2A}(\xi_{44A} + j\omega C_{4A} + j\omega C_{20A}) + U_3\xi_{43A} - U_{4A}(\xi_{44A} + j\omega C_{4A}) \\ - U_{0A}\xi_{43A} \quad (40)$$

— dla węzła 2B:

$$U_{2B}y_f = U_{2B}(\xi_{44B} + j\omega C_{4B} + j\omega C_{20B}) + U_3\xi_{43B} - U_{4B}(\xi_{44B} + j\omega C_{4B}) - U_{0B}\xi_{43B} \quad (41)$$

— dla węzła 3:

$$\begin{aligned} U_{3y_f} = & -U_1(\xi_{220} + j\omega C_1) + U_{2A}\xi_{34A} + U_{2B}\xi_{34B} + \\ & + U_3(\xi_{220} + \xi_{33A} + \xi_{33B}) + j\omega C_1 + j\omega C_{3A} + j\omega C_{3B} + j\omega C_{30A} + j\omega C_{30B} + \\ & + j\omega C_{12A} + j\omega C_{12B} + j\omega C_{34A} + j\omega C_{34B} + j\omega C_{30} + \\ & - U_{4A}(\xi_{34A} + \xi_{210} - j\omega C_{12A} - j\omega C_{34A}) + \\ & - U_{4B}(\xi_{34B} + \xi_{230} - j\omega C_{12B} - j\omega C_{34B}) + \\ & + U_{0A}(\xi_{210} - \xi_{33A} - j\omega C_{3A}) - U_{0B}(\xi_{230} - \xi_{33B} - j\omega C_{3B}) \end{aligned} \quad (42)$$

— dla węzła 4A:

$$\begin{aligned} U_{4A}y_L = & U_1\xi_{120} - U_{2A}(\xi_{44A} + j\omega C_{4A}) - U_3(\xi_{43A} + \xi_{120} + j\omega C_{12A} + j\omega C_{34A}) + \\ & + U_{4A}(\xi_{44A} + \xi_{110} + j\omega C_{2A} + j\omega C_{4A} + j\omega C_{12A} + j\omega C_{34A} + j\omega C_{40A} + j\omega C_{AB}) + \\ & + U_{4B}(\xi_{130} - j\omega C_{AB}) + U_{0A}(\xi_{43A} - \xi_{110} - j\omega C_{2A}) - U_{0B}\xi_{130} \end{aligned} \quad (43)$$

— dla węzła 4B:

$$\begin{aligned} U_{4B}y_L = & U_1\xi_{320} - U_{2B}(\xi_{44B} + j\omega C_{4B}) - U_3(\xi_{43B} + \xi_{320} + j\omega C_{12B} + j\omega C_{34B}) + \\ & + U_{4A}(\xi_{310} - j\omega C_{AB}) + \\ & + U_{4B}(\xi_{44B} + \xi_{330} + j\omega C_{2B} + j\omega C_{4B} + j\omega C_{12B} + j\omega C_{34B} + j\omega C_{40B} + j\omega C_{AB}) + \\ & - U_{0A}\xi_{310} + U_{0B}(\xi_{43B} - \xi_{330} - j\omega C_{2B}) \end{aligned} \quad (44)$$

— dla węzła 0A:

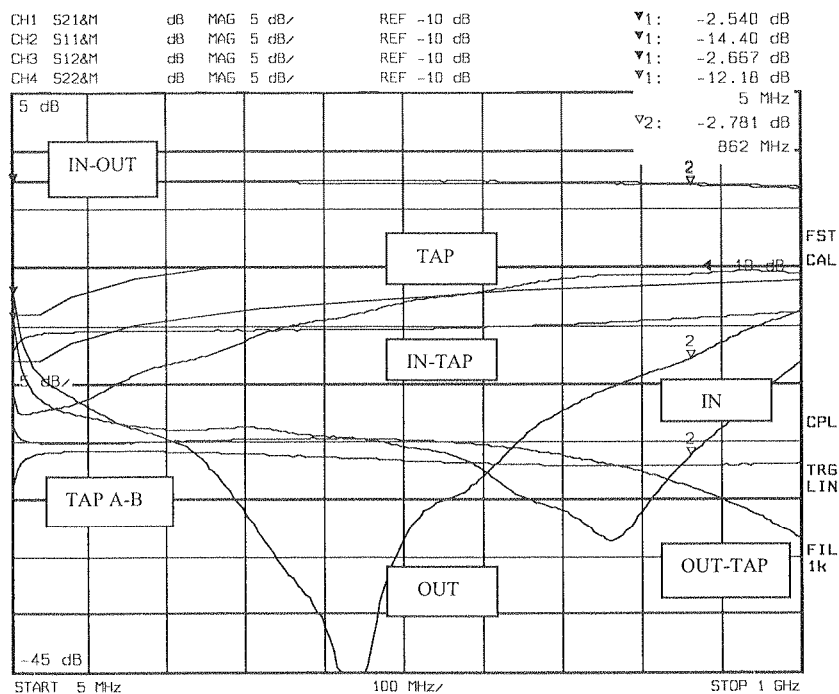
$$\begin{aligned} U_{0A}y_0 = & -U_1\xi_{120} - U_{2A}\xi_{34A} + U_3(\xi_{120} - \xi_{33A} - j\omega C_{3A}) + U_{4A}(\xi_{34A} - \xi_{110} - j\omega C_{2A}) + \\ & - U_{4B}\xi_{130} + U_{0A}(\xi_{33A} + \xi_{110} + j\omega C_{2A} + j\omega C_{3A}) + U_{0B}\xi_{130} \end{aligned} \quad (45)$$

— dla węzła 0B:

$$\begin{aligned} U_{0B}y_0 = & -U_1\xi_{320} - U_{2B}\xi_{34B} + U_3(\xi_{320} - \xi_{33B} - j\omega C_{3B}) - U_{4A}\xi_{310} + \\ & + U_{4B}(\xi_{34B} - \xi_{330} - j\omega C_{2B}) + U_{0A}\xi_{310} \\ & + U_{0B}(\xi_{33B} + \xi_{330} + j\omega C_{2B} + j\omega C_{3B}) \end{aligned} \quad (46)$$

Rys. 26 przedstawia charakterystyki transformatora pomierzone, a rys. 27 obliczone za pomocą powyższych wzorów. Wartości rezystancji balastowej i zerowej są tak dobrane, aby charakterystyki tłumienności niedopasowania na wejściu, odgałęzieniu i tłumienia oddzielenia były optymalne. Ma to miejsce dla $R_L = 330 \Omega$, oraz dla

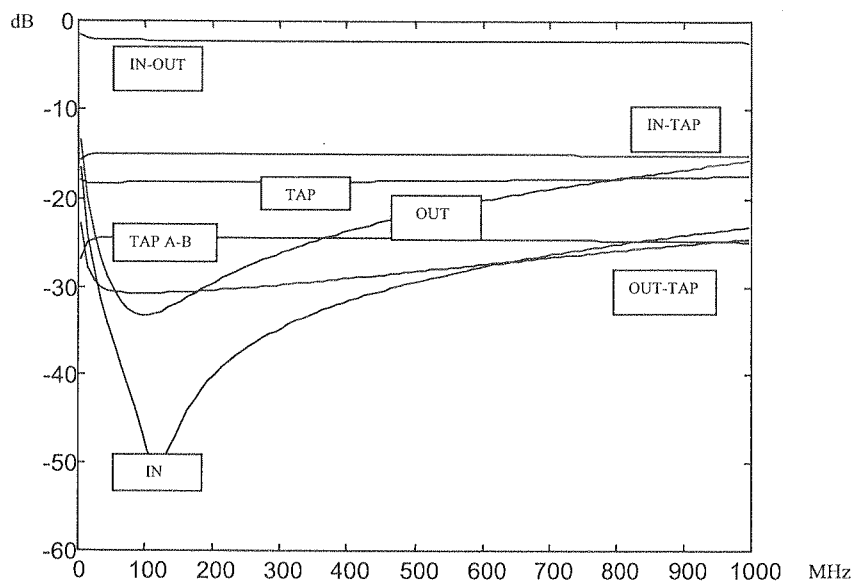
$R_0 = 51 \Omega$. Wpływ rezystancji balastowej ma analogiczny charakter jak dla transformatora w konfiguracji podstawowej. Przebiegi parametrów transformatora w funkcji rezystancji balastowej i zerowej są przedstawione na rys. 28 i 29. Należy jednak zwrócić uwagę, że optimum ze względu na tłumienie przenikowe ma miejsce dla $R_L = 120 \Omega$, a ze względu na tłumienność niedopasowania na odgałęzieniu dla $R_L = 150 \Omega$. Jednak dla tych wartości R_L tłumienie oddzielenia jest stosunkowo małe dlatego przyjęto $R_L = 330 \Omega$ jako optymalne. Rezystor balastowy nie ma wpływu na tłumienie w linii głównej i sprzężonej, co stanowi cechę pozytywną. Nie ma on również wpływu na tłumienność niedopasowania na wejściu i wyjściu. W przypadku rezystancji zerowej optimum tłumienności niedopasowania na odgałęzieniu ma miejsce dla $R_0 = 51 \Omega$. Wzrost rezystancji zerowej powoduje wprowadzić znaczną poprawę tłumienia przenikowego, ale jednocześnie następuje wzrost tłumienia w linii sprzężonej co stanowi cechę niekorzystną. Ponadto znacznie pogarsza się wówczas tłumienność niedopasowania na odgałęzieniu. Na pozostałe parametry rezystor zerowy ma mniejszy wpływ. Rezystor zerowy nie ma wpływu na tłumienie w linii głównej, co jest w tym przypadku cechą korzystną.



Rys. 26. Pomierzone charakterystyki transformatora dwukierunkowego w konfiguracji odwróconej

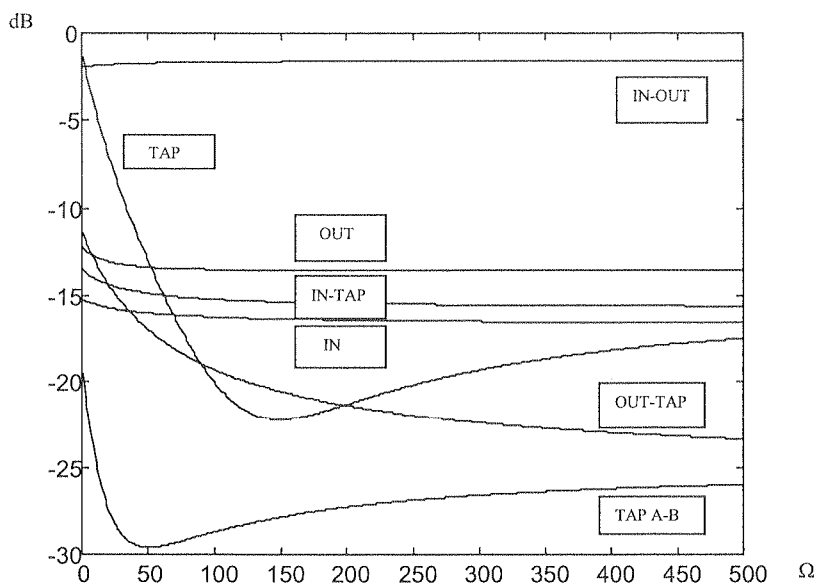
Fig. 26. Measured characteristics of the two-directional transformer in inverse configuration

trans-
funkcji
zwrócić
120 Ω ,
 Ω . Jed-
przyjęto
w linii
ywu na
zerowej
51 Ω .
zeniko-
i cechę
ania na
ezystor
u cechę



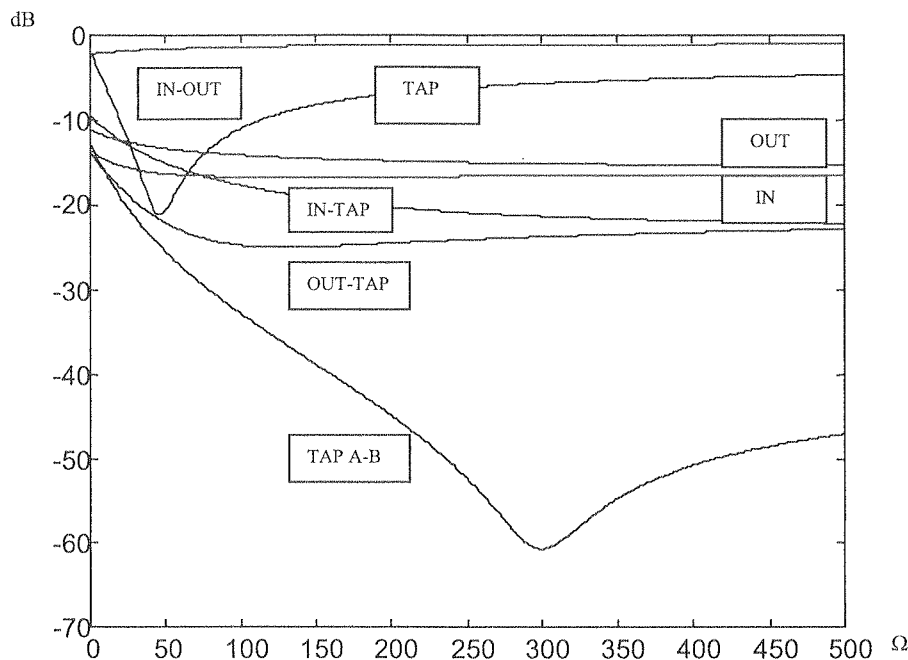
Rys. 27. Obliczone charakterystyki transformatora dwukierunkowego w konfiguracji odwróconej

Fig. 27. Calculated characteristics of the two-directional transformer in inverse configuration



Rys. 28. Wpływ rezystancji balastowej R_L na charakterystyki transformatora w konfiguracji odwróconej dla częstotliwości 5 MHz

Fig. 28. Influence of ballast resistance R_L to characteristics of the transformer in inverse configuration for frequency 5 MHz

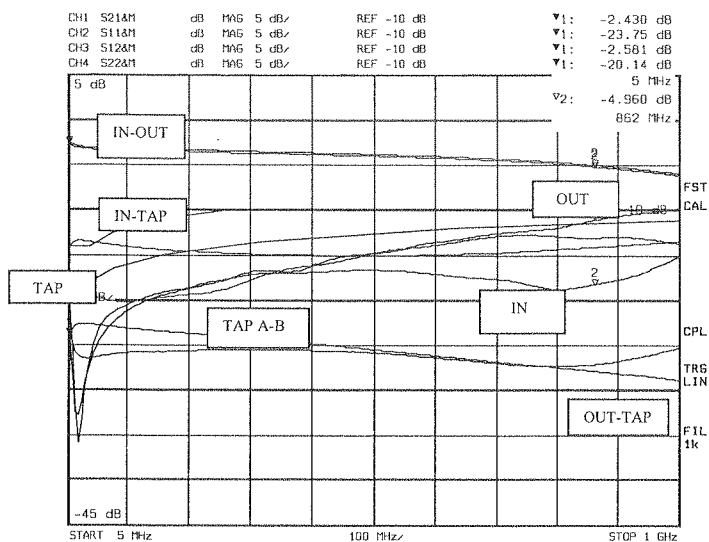


Rys. 29. Wpływ rezystancji zerowej R_0 na charakterystyki transformatora w konfiguracji odwróconej dla częstotliwości 5 MHz

Fig. 29. Influence of zero resistance R_0 to characteristics of the transformer in inverse configuration for frequency 5 MHz

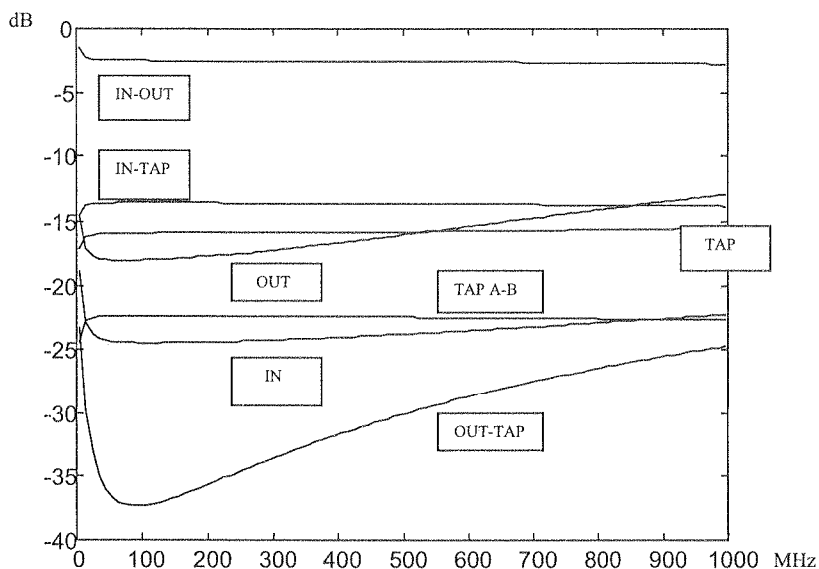
Wpływ przenikalności początkowej jest tu podobny jak w przypadku transformatora w konfiguracji podstawowej. Istotną różnicą jest przebieg tłumienia w linii sprzężonej w funkcji przenikalności początkowej, które bardzo znacząco wzrasta wraz ze wzrostem przenikalności (ok. 25 dB dla przenikalności rzędu 9000). Natomiast charakter zmienności pozostałych parametrów jest analogiczny do konfiguracji podstawowej. Rys. 30 i 31 przedstawiają przebiegi parametrów transformatora z zastosowanym rdzeniem o przenikalności początkowej wartości $\mu_i = 2500$. Obserwujemy poprawę tłumienia niedopasowania na wejściu i wyjściu dla m.cz. ale następuje pogorszenie tłumienia przenikowego — dla tego parametru najkorzystniejsza jest przenikalność ok. 1000. Również wpływ pojemności międzyzwojowych i montażowych ma analogiczny charakter jak w przypadku transformatora w konfiguracji podstawowej. Przebiegi są bardzo podobne jak do tych z rys. 11 — 15.

W konfiguracji odwróconej transformator można zoptymalizować jak na rys. 20. Zastosowanie pojemności szeregowych dla m.cz. i układów LC dla w.cz. daje podobne rezultaty jak poprzednio. Rys. 32 ilustruje przykładowe przebiegi parametrów transformatora przy zastosowaniu układu korekcyjnego LC w linii sprzężonej. Na rys. 33 pokazane są przebiegi transformatora po zastosowaniu układów optymalizujących. Jako pojemności szeregowe zarówno w linii głównej i sprzężonej zastosowano 2.2 nF.



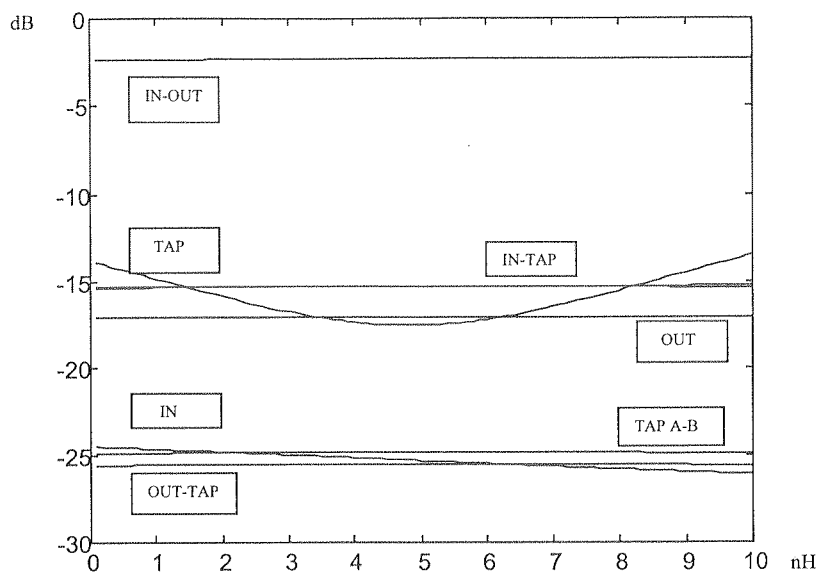
Rys. 30. Pomierzone charakterystyki transformatora dwukierunkowego w konfiguracji odwróconej dla przenikalności $\mu_i = 2500$

Fig. 30. Measured characteristics of the two-directional transformer in inverse configuration for initial permeability $\mu_i = 2500$



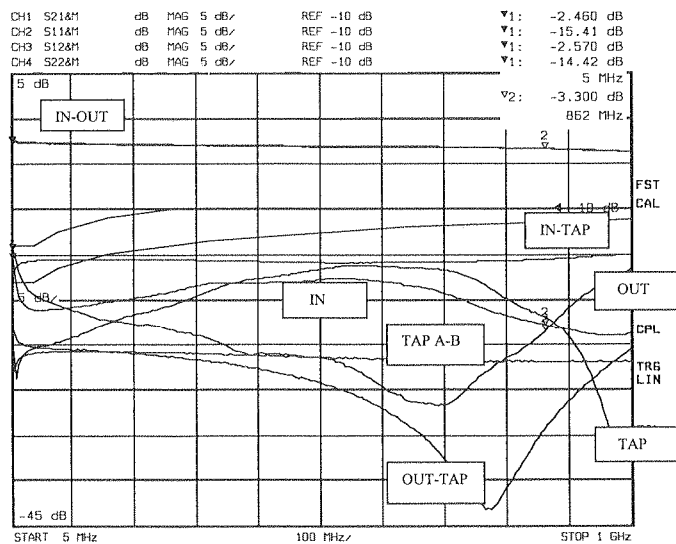
Rys. 31. Obliczone charakterystyki transformatora dwukierunkowego w konfiguracji odwróconej dla przenikalności $\mu_i = 2500$

Fig. 31. Calculated characteristics of the two-directional transformer in in-verse configuration for initial permeability $\mu_i = 2500$



Rys. 32. Wpływ indukcyjności korekcyjnej L_{ft} na charakterystyki transformatora przy stałej pojemności korekcyjnej C_{ft} dla częst. 862 MHz

Fig. 32. Influence of corrective inductance L_{ft} to characteristics of the transformer for constant corrective capacity C_{ft} for freq862



Rys. 33. Charakterystyki transformatora w konfiguracji odwróconej po optymalizacji

Fig. 33. Characteristics of the transformer in inverse configuration after optimization

4. PORÓWNANIE KONFIGURACJI TRANSFORMATORA

Obie konfiguracje transformatora dwukierunkowego zachowują kilka cech wspólnych. Dla obu wariantów wpływ przenikalności początkowej jest podobny. Ze względu na tłumienie w linii głównej i sprzężonej transformatory zachowują się również bardzo podobnie. W obu przypadkach najkorzystniejsze wyniki uzyskuje się dla przenikalności rzędu 3000-5000, natomiast dla wyższych przenikalności obserwujemy pogorszenie właściwości obu transformatorów. Również ze względu na pojemności transformatora, można co do charakteru zmian i w przybliżeniu wartości uznać, że ich wpływ jest zbliżony w obu konfiguracjach. Zatem rys. 11 — 15 można uznać za reprezentatywne. Wpływ rezystancji balastowych i zerowych również ma tu analogiczny charakter. Istnieją wartości dla których w przypadku rezystancji balastowych i zerowych tłumienność niedopasowania na odgałęzieniu osiąga ekstremum z tym, że dla różnych wartości (co zostało opisane wcześniej). Również dla obu konfiguracji korzystniej jest stosować jak największe rezystory balastowe i zerowe ze względu na tłumienie przenikowe.

Zasadniczymi wielkościami które różnią obie konfiguracje są stosunki ich tłumień przenikowych i tłumień oddzielenia. Różnice wynikają z faktu, że w konfiguracji podstawowej cewki L_{3A} , L_{3B} łączone są z wejściem, a w odwróconej z wyjściem. W pierwszym przypadku powoduje to silniejsze sprzężenie między wyjściami odgałęźnymi, a to sprawia, że gorsze jest tłumienie przenikowe. W drugim przypadku silniejsze sprzężenie występuje między wyjściem a odgałęzieniami co sprawia, że tłumienie oddzielenia jest gorsze niż w przypadku konfiguracji podstawowej. Zatem dla konfiguracji podstawowej uzyskujemy lepsze tłumienie oddzielenia, a gorsze przenikowe, zaś dla konfiguracji odwróconej — przeciwnie.

Ponadto na szczególną uwagę zasługuje nadmieniona różnica w przebiegu tłumienia w linii sprzężonej w funkcji przenikalności początkowej dla różnych konfiguracji transformatora.

5. PODSUMOWANIE

Podobnie jak w przypadku transformatora jednokierunkowego, bardzo ważną cechą transformatora dwukierunkowego niezależnie od konfiguracji jest stałość tłumień w linii głównej i sprzężonej. Taki stan rzeczy ma miejsce dzięki temu, że przenikalność magnetyczna rdzenia maleje w funkcji częstotliwości. W przeciwnym przypadku nie można byłoby zaprojektować szerokopasmowego transformatora. Zaletą ta jest tym cenniejsza, że w linii głównej stałość tłumienia jest utrzymana niezależnie od zmian balastu, rezystancji zerowej i do pewnego stopnia pojemności międzyzwojowych. Z linią sprzężoną jest trochę inaczej. Zmiany balastu, rezystancji zerowej i pojemności międzyzwojowych wywołują zmiany tłumienia w linii sprzężonej. Zmiany te występują dość wyraźnie również w przypadku zmian przenikalności początkowej. Te czynniki należy mieć na uwadze przy projektowaniu transformatora.

W pracy opisano możliwe konfiguracje transformatora. Zaprezentowane wykresy i próby optymalizacji pomyślane zostały o szerokopasmowym zastosowaniu transformatora. Przy uściśleniu zastosowań i przyjęciu określonych priorytetów można oczywiście uzyskać jeszcze bardziej zadowalające wyniki. Teoretyczne wyniki badań różnią się od praktycznych z kilku powodów. Zastosowanie do obliczeń wzoru (23), gwarantuje wg [2] dokładność 10%. Jest to główna przyczyna niedokładności dla niskich częstotliwości. Dla wysokich częstotliwości dochodzi ponadto niedoskonałość złączy F, wpływ układu pomiarowego, oraz układu przeniesienia masy. Jednak dla wartości bezwzględnych dokładność 10% została utrzymana. Przy projektowaniu transformatora w celu optymalizacji można posłużyć się do pewnego stopnia obliczeniami teoretycznymi, ale zgrubnie. W celu dokonania precyzyjnej optymalizacji, należy wspomóc się badaniami praktycznymi. Praca miała na celu prezentację nowej konstrukcji transformatora i wskazanie jego specyfiki. Na podstawie analizy i przedstawionych wykresów można wykorzystać zawarty w niej materiał do projektowania takich transformatorów i wykorzystania ich do różnych celów.

6. BIBLIOGRAFIA

1. D. I. K i m, M. T a k a h a s h i, K. A r a k i, Y. N a i t o: *Optimum design of the power dividers with ferrite toroids for CATV and/or MATV systems*. IEEE Trans. On Consumer Electronics, vol. CE-29, no1, Feb. 1983, pp. 27-38.
2. Y. N a i t o: *Formulation of frequency dispersion of permeability*. Trans. IECE, vol. 59-c, May 1976, pp. 297-304.
3. M. G. E l l i s: *RF Directional Couplers*. Electronics System Product a division of ANTEC 2002.
4. D. I. K i m, J. H. H w a n g M. T a k a h a s h i: *A new analysis method and broad-band design of tap-offs for CATV/DBS systems*. IEEE Trans. On Consumer Electronics, vol. 46, no. 1, Feb. 2000, pp. 171-178.
5. D. K r z e m i e n i e c k i: *Analiza i optymalizacja szerokopasmowego transformatora kierunkowego*. Kwartalnik Elektroniki i Telekomunikacji (złożono do druku).

D. KRZEMIENIECKI

ANALYSIS AND OPTIMIZATION OF A BROADBAND TWO-DIRECTIONAL TRANSFORMER

S u m m a r y

Object of present elaboration is complex analysis of symmetrical two-directional transformer (two-directional coupler). This is a new construction which may be use as a device of cable television (CATV), most often as tap. Executed they are on base of little, line 3×3 mm ferrite cores joined, on which wound are windings of the transformer.

Examined transformer can exist in two usable configurations named here: basic and inverse.

Work contains theoretical analysis and results of measurements and calculations of parameters of the transformer like: insertion losses of main and coupled line, isolation of output and tap, isolation of taps and return losses on input, output and tap in function of frequency in range 5 - 1000 MHz.

In work examined influence of initial permeability (in range 100-10000), on parameters of the transformer, and such factors as ballast-resistor, zero-resistor considering of obtainment of optimum parameters and its influence on winding and assembly capability of the transformer and its parameters for high frequencies. Besides one examined possibilities of use of corrective arrangements having in view improvement of parameters of the transformer for low and high frequencies. One introduced also effects of such optimization for two possible configuration of the transformer.

Taking into account optimization and influences of ballast and zero resistors one examined courses parameters of the transformer for its extreme frequencies of work: 5 and 862 MHz.

General analysis of the transformer passed in regard on broadband its use, but results of researches can also be use for more particular arrangements.

Work addressed is mostly to constructors and has in view help in projecting of arrangements taps.

To researches used 15-dB symmetrical transformer with use of cores of initial permeability 1400 and 2500 produced by Taiwan.

Keywords: two-directional transformer, basic configuration, inverse configuration

prob
przy
wiąz
w sw

10] c

Elektrotermiczny model transformatora impulsowego dla programu SPICE

KRZYSZTOF GÓRECKI, JANUSZ ZARĘBSKI

*Katedra Radioelektroniki Morskiej, Akademia Morska w Gdyni
ul. Morska 83, 81-225 Gdynia
e-mail: gorecki@am.gdynia.pl,
zarebski@am.gdynia.pl*

*Otrzymano 2004.06.02
Autoryzowano 2005.02.24*

W pracy zaproponowano nowy elektrotermiczny model transformatora impulsowego z rdzeniem ferromagnetycznym. Opracowany model uwzględnia nieliniowość charakterystyk magnesowania rdzenia ferromagnetycznego oraz wpływ samonagrzewania na ich przebieg, a także zależność rezystancji uzwojeń od temperatury i częstotliwości transformowanego sygnału. Przy opisie zjawiska samonagrzewania uwzględniono zróżnicowanie temperatur rdzenia i uzwojeń oraz straty energii występujące zarówno w rdzeniu jak i w uzwojeniach. Nowy model zaimplementowano w programie SPICE w postaci podukładu i porównano wyniki obliczeń z wynikami pomiarów dla transformatorów z różnymi typami rdzeni ferrytowych. Zaproponowano prostą procedurę wyznaczania wartości parametrów modelu na podstawie odpowiednich pomiarów lub danych katalogowych zastosowanych rdzeni ferrytowych oraz przewodów nawojowych.

Słowa kluczowe: transformator impulsowy, model elektrotermiczny, samonagrzewanie, SPICE

1. WPROWADZENIE

W wielu układach elektronicznych, np. w zasilaczach impulsowych, występuje problem przenoszenia energii między źródłem napięcia wejściowego a obciążeniem przy zachowaniu ich separacji galwanicznej. Typowo stosowanym elementem, umożliwiającym separację galwaniczną obwodów jest transformator impulsowy, zawierający w swej strukturze rdzeń ferrytowy oraz co najmniej dwa uzwojenia.

W literaturze dużo uwagi poświęca się własnościom materiałów ferrytowych [1 – 10] oraz wpływowi oddziaływań elektromagnetycznych, zachodzących w transformato-

rze, na jego charakterystyki oraz parametry robocze [2, 7, 11-13], natomiast stosunkowo mało prac dotyczy problemu modelowania tego elementu [12, 14, 15, 16]. Z drugiej strony, coraz powszechniej stosowane do analizy układów elektronicznych programy komputerowe, wśród których standardem jest obecnie program SPICE, zawierają stosunkowo proste, wyidealizowane modele transformatora. Jak zaznaczono w pracy [17], użytkownik programu SPICE może zamodelować transformator jako sprzężone cewki, przy czym ich indukcyjność wzajemna może być stała (transformator liniowy) lub też zależna od prądów tych cewek (transformator nieliniowy). Nieliniowość charakterystyk transformatora może być uwzględniona poprzez uzależnienie indukcyjności cewek, stanowiących poszczególne uzwojenia, od płynącego przez nie prądu za pomocą funkcji kwadratowej lub wykorzystanie nieliniowego modelu rdzenia ferromagnetycznego. Wbudowane w programie SPICE modele rdzenia, a wśród nich model Jilesa-Athertona [1, 18] uwzględniający najwięcej zjawisk fizycznych, istotnych z punktu widzenia opisu charakterystyki magnesowania, nie uwzględniają wpływu temperatury na charakterystyki materiału ferromagnetycznego, a jak wynika między innymi z prac [4, 6, 8, 16, 19, 20], wpływ ten może być bardzo istotny.

Należy podkreślić, że pod pojęciem temperatury należy rozumieć temperaturę wnętrza rozważanego elementu, różniącą się od temperatury otoczenia na skutek zjawiska samonagrzewania. Zjawisko to wynika z zamiany wydzielanej w transformatorze energii elektrycznej na ciepło przy nieidealnym chłodzeniu tego elementu. Ze względu na różne wartości energii wydzielanych w rdzeniu i w uzwojeniach, temperatury obu wymienionych części transformatora mogą różnić się między sobą.

Uwzględnienie w analizach komputerowych zjawiska samonagrzewania w transformatorze wymaga zastosowania modeli specjalnego rodzaju, nazywanych elektrotermicznymi modelami rozważanego elementu. Jednak, jak wynika z dostępnych źródeł obejmujących publikacje naukowe, zasoby bibliotek dołączonych do programu SPICE oraz strony internetowej producentów rdzeni ferrytowych, jak dotąd nie sformułowano elektrotermicznego modelu transformatora impulsowego dla programu SPICE.

W pracy przedstawiono opracowany przez autorów elektrotermiczny model transformatora impulsowego dla programu SPICE, uwzględniający zjawiska elektryczne, magnetyczne i termiczne, zachodzące w tym elemencie. Zaproponowano także algorytm estymacji wartości parametrów modelu, a poprawność tego modelu zweryfikowano doświadczalnie.

2. POSTAĆ MODELU

Opracowany przez autorów model transformatora dla programu SPICE bazuje na wzajemnych oddziaływaniach elektromagnetycznych i elektrotermicznych, zachodzących w rdzeniu i w uzwojeniach oraz między nimi. Model transformatora składa się z pięciu submodeli reprezentowanych przez:

- równania opisujące charakterystyki materiału ferrytowego, z którego wykonano rdzeń,

- zależność natężenia pola magnetycznego w rdzeniu od prądów uzwojeń,
- równania indukcji elektromagnetycznej dla wszystkich uzwojeń,
- zależność strat energii w uzwojeniach od temperatury, prądu uzwojeń i częstotliwości,
- zależności temperatury rdzenia oraz uzwojeń od napięć i prądów zaciskowych transformatora, stanowiących submodel termiczny.

Model materiału ferrytowego zaczerpnięto z pracy [18], uzupełniając go o temperaturowe zależności występujących w nim parametrów. Szczegółowe równania modelu rdzenia podano w pracy [19]. Są to równania różniczkowo-algebraiczne o postaci

$$f(M_a, M, H, \dot{M}_a, \dot{M}, \dot{H}, T_r, \underline{par}) = 0 \quad (1)$$

gdzie M oznacza magnetyzację w rdzeniu, M_a — magnetyzację wyznaczoną z krzywej pierwotnego magnesowania, H — natężenie pola magnetycznego, T_r — temperaturę rdzenia, $\dot{M}$, $\dot{M}_a$, $\dot{H}$ oznaczają czasowe pochodne odpowiednio wielkości M , M_a oraz H , natomiast symbol $\underline{par}$ oznacza wektor parametrów modelu rdzenia. W wektorze $\underline{par}$ występują parametry (tabela 1) opisujące właściwości rdzenia w warunkach izotermicznych, tzn. dla ustalonej temperatury rdzenia.

Opisany model uwzględnia istnienie pętli histerezy i zależność podstawowych parametrów katalogowych rdzenia, tzn. indukcji nasycenia B_s , indukcji remanencji B_r oraz pola koercji H_c , od temperatury. W modelu uwzględniono także temperaturę Curie T_C i temperaturową zależność początkowej przenikalności magnetycznej μ_i .

Pole magnetyczne H wytwarzane w rdzeniu jest funkcją prądów oraz liczby zwojów wszystkich indukcji, zgodnie z zależnością

$$H = \frac{\sum_{k=1}^N z_k \cdot i_k}{l} \quad (2)$$

gdzie z_k oznacza liczbę zwojów k -tego uzwojenia, i_k — prąd wpływający do początku k -tego uzwojenia, natomiast l to długość drogi magnetycznej w rdzeniu.

Siła elektromotoryczna e_k indukowana w k -tym uzwojeniu opisana jest wzorem

$$e_k = m_k \cdot z_k \cdot S \cdot \frac{dB}{dt} \quad (3)$$

gdzie m_k oznacza współczynnik sprzężenia dla k -tego uzwojenia, S — pole przekroju czynnego rdzenia, natomiast B — indukcję magnetyczną w rdzeniu.

W opisie modelu uzwojeń uwzględniono straty energii wynikające z rezystywności przewodu oraz z efektu naskórkowości za pomocą rezystancji R opisanej wzorem

$$R = \rho \cdot (1 + \alpha \cdot (T_u - T_0)) \cdot \frac{l_p}{S_p} + \frac{l_p}{\sqrt{8 \cdot S_p}} \cdot \sqrt{\rho \cdot \mu \cdot (1 + \alpha \cdot (T_u - T_0))} \cdot f \quad (4)$$

Tabela 1

Parametry modelu rdzenia

Parameters of the core model

Nazwa parametru	Symbol
przenikalność magnetyczna próżni	μ_0
przenikalność początkowa materiału rdzenia	μ_i
temperatura Curie	T_C
pole koercji w temperaturze odniesienia T_0	H_{C0}
indukcja remanencji w temperaturze T_0	B_{r0}
indukcja nasycenia wyznaczona w temperaturze T_0 przy natężeniu pola magnetycznego H_S	B_{S0}
współrzędne punktu leżącego na krzywej pierwotnego magnesowania	H_X, M_X
temperaturowy współczynnik zmian pola koercji	α_r
temperaturowy współczynnik zmian indukcji remanencji	α_c
temperaturowy współczynnik zmian indukcji nasycenia	α_s
temperaturowy współczynnik zmian przenikalności początkowej	α_u

gdzie S_p oznacza pole przekroju poprzecznego przewodu, l_p — długość tego przewodu, ρ — rezystywność przewodu, α — temperaturowy współczynnik zmian rezystywności, μ — przenikalność magnetyczną przewodu, f — częstotliwość prądu uzwojenia, natomiast T_u — temperaturę uzwojenia.

Straty energii w rdzeniu wynikają głównie z istnienia pętli histerezy i są proporcjonalne do jej powierzchni. W modelu zastosowano empiryczną zależność, opisującą całkowite straty w rdzeniu p_r , o postaci [2]

$$p_r = V_e \cdot A \cdot f^a \cdot B_m^b \cdot [1 - D \cdot (T_r - T_m)^2] \quad (5)$$

gdzie B_m oznacza amplitudę składowej zmiennej indukcji, V_e — ekwiwalentną objętość rdzenia, D — kwadratowy współczynnik temperaturowy zmian gęstości mocy strat p_v w rdzeniu, T_m to temperatura odpowiadająca minimum funkcji $p_v(T)$, A — współczynnik proporcjonalności we wzorze (5), natomiast a oraz b są wykładnikami we wzorach opisujących zależność strat w rdzeniu od częstotliwości i amplitudy indukcji. W modelu przyjęto typowe wartości parametrów a oraz b równe odpowiednio 1 oraz 2 [20].

Tabela 1

Symbol
μ_0
μ_i
T_C
H_{C0}
B_{r0}
B_{s0}
H_X, M_X
α_r
α_c
α_s
α_u

Z kolei, moc wydzielana w uzwojeniach dana jest wzorem

$$p_u = \sum_{j=1}^N R_j \cdot i_j^2 \quad (6)$$

gdzie R_j oznacza rezystancję j-tego uzwojenia zależną od temperatury i częstotliwości, natomiast i_j jest prądem płynącym przez to uzwojenie.

W submodelu termicznym transformatora występują dwie temperatury: temperatura rdzenia T_r oraz temperatura uzwojeń T_u . Submodel ten opisany jest za pomocą układu równań o postaci

$$T_r = T_a + \int_0^t Z'_{rr}(x) \cdot p_r(t-x) \cdot dx + \int_0^t Z'_{ru}(x) \cdot p_u(t-x) \cdot dx \quad (7)$$

$$T_u = T_a + \int_0^t Z'_{uu}(x) \cdot p_u(t-x) \cdot dx + \int_0^t Z'_{ur}(x) \cdot p_r(t-x) \cdot dx \quad (8)$$

gdzie p_r oraz p_u oznaczają odpowiednio moce wydzielane w rdzeniu oraz w uzwojeniach, $Z'_{rr}(t)$ oraz $Z'_{uu}(t)$ to czasowe pochodne przejściowej impedancji termicznej rdzenia oraz uzwojeń, natomiast $Z'_{ru}(t)$ oraz $Z'_{ur}(t)$ to pochodne wzajemnych przejściowych impedancji termicznych między rdzeniem a uzwojeniem. Wszystkie z wymienionych przejściowych impedancji termicznych modelowane są za pomocą sumy funkcji eksponencjalnych o postaci

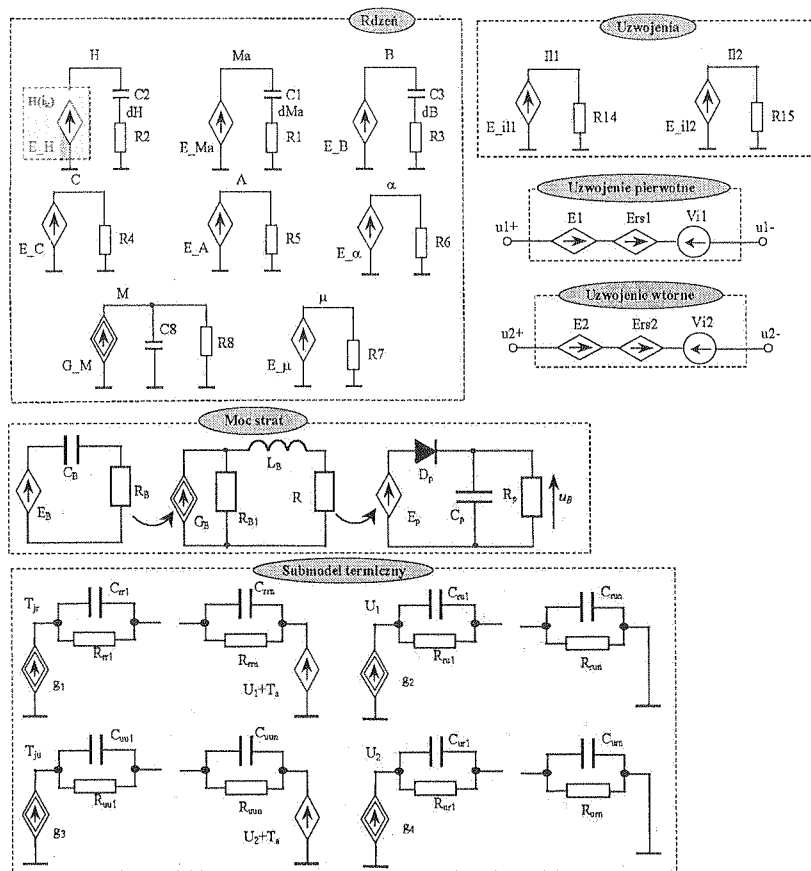
$$Z(t) = R_{th} \cdot \left[1 - \sum_{i=1}^N a_i \cdot \exp\left(-\frac{t}{\tau_{thi}}\right) \right] \quad (9)$$

gdzie R_{th} oznacza rezystancję termiczną, τ_{thi} reprezentuje i-tą termiczną stałą czasową, a_i to współczynniki wagowe, których suma wynosi 1, natomiast N symbolizuje ilość termicznych stałych czasowych w modelu $Z(t)$. Wzór (9) dotyczy zarówno własnych przejściowych impedancji termicznych rdzenia $Z_{rr}(t)$ oraz uzwojenia $Z_{uu}(t)$, jak i wzajemnych przejściowych impedancji termicznych między rdzeniem a uzwojeniem $Z_{ru}(t)$ oraz $Z_{ur}(t)$. W oznaczeniach poszczególnych parametrów modelu $Z(t)$ wystąpią dodatkowo indeksy identyczne jak w symbolu odpowiedniej przejściowej impedancji termicznej. Przykładowo, własna rezystancja termiczna rdzenia jest oznaczona symbolem R_{thrr} , natomiast wzajemna rezystancja termiczna między rdzeniem a uzwojeniem — symbolem R_{thru} .

3. IMPLEMENTACJA MODELU W PROGRAMIE SPICE

Program SPICE umożliwia użytkownikowi tworzenie własnych modeli (makromodeli) elementów elektronicznych w postaci podukładów, poprzez łączenie odpowiednich elementów, reprezentowanych za pomocą ich modeli wbudowanych w tym programie.

Model transformatora impulsowego opisany w rozdziale drugim zaimplementowano w postaci podukładu, którego reprezentację obwodową przedstawiono na rys.1.



Rys. 1. Reprezentacja obwodowa modelu transformatora impulsowego

Fig. 1. The network representation of the pulse transformer model

W podukładzie tym wyróżniono bloki stanowiące obwody główne transformatora (uzwojenie pierwotne i uzwojenie wtórne) oraz poszczególne submodele transformatora. W obwodach głównych występują napięciowe źródła sterowane, modelujące zjawisko indukcji elektromagnetycznej w obu uzwojeniach (źródła E1 i E2) oraz spadek napięcia na rezystancjach szeregowych (źródła Ers1 i Ers2). Niezależne źródła napięciowe o zerowej wydajności V11 oraz V12 służą do wyznaczania prądów w uzwojeniach, wykorzystywanych w opisie wydajności źródeł sterowanych zawartych w submodelach rdzenia i uzwojeń. Źródło napięciowe E_H modeluje zależność pola magnetycznego w rdzeniu od prądów uzwojeń. Z kolei, w submodelu rdzenia występują m.in. obwody

elemento-
na rys.1.

realizujące czasowe pochodne magnetyzacji dla próbki nienamagnesowanej M_a , magnetyzacji M , indukcji B oraz natężenia pola magnetycznego H . W obwodach tych zastosowano klasyczne układy różniczkujące RC (np. R2-C2). Zastosowane podejście, pomimo ograniczenia zakresu jego słuszności dla dużych częstotliwości pozwala na wykorzystanie zaproponowanego modelu również w analizie małosygnałowej, stosowanej np. przy badaniu stabilności układów impulsowych [24]. Takich możliwości nie daje standardowa funkcja różniczkowania DDT (ang. time derivative), która przyjmuje wartości różne od zera tylko w analizie stanów przejściowych [21].

Obwody złożone ze sterowanych źródeł napięciowych oraz rezystorów służą do wyznaczania wartości zmiennych pomocniczych, występujących w modelu materiału ferromagnetycznego, które ze względu na ograniczenie długości wyrażenia opisującego wydajność źródła sterowanego, nie mogą być wprowadzone bezpośrednio do wyrażeń opisujących zmienne B , H , M oraz M_a .

Do wyznaczania mocy strat związanych z histerezą w rdzeniu zastosowano układ złożony ze źródła napięciowego E_B o wydajności równej indukcji B w rdzeniu, kondensatora C_B oraz rezystora R_B . Napięcie na rezystorze R_B , odpowiadające składowej zmiennej indukcji B , steruje wydajnością źródła prądowego G_B , obciążonego cewką L_B oraz rezystorem R . Układ ten umożliwia wyznaczenie częstotliwości zmian indukcji B , natomiast detektor szczytowy złożony z elementów E_P , D_P , C_P oraz R_P służy do wyznaczenia amplitudy indukcji B_m , wykorzystywanej w opisie strat w rdzeniu.

Submodel termiczny transformatora skonstruowano analogicznie do modelu termicznego dwóch elementów półprzewodnikowych sprzężonych termicznie [22], np. osadzonych na wspólnym podłożu lub radiatorze. W modelu tym występują cztery obwody złożone ze źródeł prądowych o wydajnościach odpowiadających mocom wydzielanym w rdzeniu (g_1 oraz g_2) i w uzwojeniach (g_3 oraz g_4) oraz z sieci RC modelujących odpowiednie własne oraz wzajemne przejściowe impedancje termiczne. Z każdym dwójnikiem RC związana jest jedna termiczna stała czasowa τ_{th} . Napięcie na źródle g_1 jest równe temperaturze rdzenia T_r , natomiast napięcie na źródle g_3 — temperaturze uzwojeń T_u .

W bloku „uzwojenia” są wyliczane wartości spadku napięcia na poszczególnych uzwojeniach, wynikające ze zjawiska naskórkowości (źródła E_{il1} oraz E_{il2}). Wydajności tych źródeł opisano za pomocą funkcji LAPLACE.

4. WYZNACZANIE WARTOŚCI PARAMETRÓW MODELU

Parametry modelu transformatora można podzielić na sześć grup. Są to:

- stałe fizyczne (μ_0) i materiałowe (ρ , α)
- parametry materiału ferrytowego, z którego wykonano rdzeń (B_{r0} , B_{S0} , H_{c0} , H_S , H_X , M_X , μ_{i0} , α_r , α_C , α_S , α_U , T_C),
- parametry geometryczne rdzenia (V_e , l , S),
- parametry geometryczne uzwojeń (l_w , S_p , l_p , z_1 , z_2),
- parametry opisujące straty w rdzeniu (A , D , T_m , a , b),

- parametry termiczne modelu (R_{thuu} , R_{thrr} , R_{thru} , R_{thur} , a_{irr} , a_{iru} , a_{iuu} , a_{iur} , τ_{thirr} , τ_{thiuu} , τ_{thiru} , τ_{thiur} , N_{rr} , N_{uu} , N_{ru} , N_{ur})

Wartości stałych fizycznych i materiałowych podane są w tablicach fizycznych. Wynoszą one: $\mu_0 = 4\pi \cdot 10^{-7} \text{H/m}$, $\rho = 1,724 \cdot 10^{-8} \Omega \cdot \text{m}$ oraz $\alpha = 4,577 \cdot 10^{-4} \text{K}^{-1}$.

W celu wyznaczenia wartości parametrów materiału ferrytowego należy:

1. Odczytać z danych katalogowych wartości parametrów μ_i , H_{c0} , B_{r0} , B_{S0} , H_S dla temperatury odniesienia $T_0 = 25^\circ\text{C}$ oraz wartość temperatury Curie T_C ,
2. Odczytać wartości parametrów H_C , B_r oraz B_S dla temperatury $T_1 = 100^\circ\text{C}$,
3. Wyliczyć wartości temperaturowych współczynników zmian indukcji nasycenia α_S , pola koercji α_C oraz indukcji remanencji α_r ze wzoru o postaci

$$\alpha_X = \frac{X_1/X_0 - 1}{T_1 - T_0} \quad (10)$$

gdzie α_X jest określonym współczynnikiem temperaturowym, natomiast X_0 oraz X_1 oznaczają wartości określonego parametru w temperaturach odpowiednio T_0 i T_1 .

4. Z katalogowego przebiegu pętli histerezy odczytać współrzędne punktu $A_1(H_X, B_X)$ leżącego na wierzchołku pętli histerezy
5. Wyliczyć wartość parametru M_X ze wzoru

$$M_X = B_X/\mu_0 - H_X \quad (11)$$

Z kolei, aby wyznaczyć wartości parametrów opisujących straty w rdzeniu należy:

1. Na katalogowej zależności gęstości mocy strat w rdzeniu od temperatury $p_V(T)$, odpowiadającej danym wartościom amplitudy indukcji w rdzeniu B_m oraz częstotliwości f , wyznaczyć wartość temperatury T_m , przy której charakterystyka ta posiada minimum (punkt A_2)
2. Dla tego samego punktu $A_2(T_m, p_{V0})$ na charakterystyce $p_V(T)$ odczytać wartość p_{V0} i wyliczyć wartość parametru A ze wzoru

$$A = \frac{p_{V0}}{f \cdot B_m^2} \quad (12)$$

3. Wybierając punkt $A_3(T_1, p_{V1})$ na tej samej charakterystyce $p_V(T)$, wyznaczyć wartość parametru D ze wzoru

$$D = 1 - \frac{p_{V1}}{A \cdot f \cdot B_m^2 \cdot (T_1 - T_m)^2} \quad (13)$$

Należy przyjąć możliwie dużą różnicę $T_1 - T_m$, zapewniającą zadawalającą dokładność wyznaczenia parametru D .

Wartości parametrów geometrycznych rdzenia, tzn. długość drogi magnetycznej l , pole przekroju czynnego rdzenia S oraz jego objętość ekwiwalentną V_e należy odczytać z danych katalogowych rdzenia.

Takie parametry geometryczne uzwojenia jak liczba zwojów uzwojenia pierwotnego z_1 i wtórnego z_2 oraz średnica przewodu nawojowego d_l są podawane przez projektanta transformatora.

Długości uzwojenia pierwotnego l_p i wtórnego l_w należy wyliczyć na podstawie wymiarów geometrycznych karkasu i liczby zwojów poszczególnych uzwojeń ze wzorów

$$l_p = z_1 \cdot 2 \cdot \pi \cdot r_{szp} + l_{wyp} \quad (14)$$

$$l_w = z_2 \cdot 2 \cdot \pi \cdot (r_{szp} + d_l) + l_{wyp} \quad (15)$$

gdzie r_{szp} oznacza promień karkasu, natomiast l_{wyp} stanowi całkowitą długość wyprowadzeń. We wzorze (15) promień karkasu powiększono o średnicę przewodu nawojowego, ponieważ typowo uzwojenie wtórne nawinięte jest na uzwojeniu pierwotnym.

Parametry modelu termicznego należy wyznaczyć na podstawie pomiaru czasowego przebiegu temperatury rdzenia $T_r(t)$ oraz uzwojenia $T_u(t)$, np. za pomocą urządzenia do ultraszybkich pomiarów temperatury firmy Vigo System [23], przy pobudzeniu transformatora uskokiem mocy o wysokości P .

Własne oraz wzajemne przejściowe impedancje termiczne, zgodnie z definicją, opisane są wzorami

$$Z_{thuu} = \frac{T_u(t) - T_a}{P} \quad (16)$$

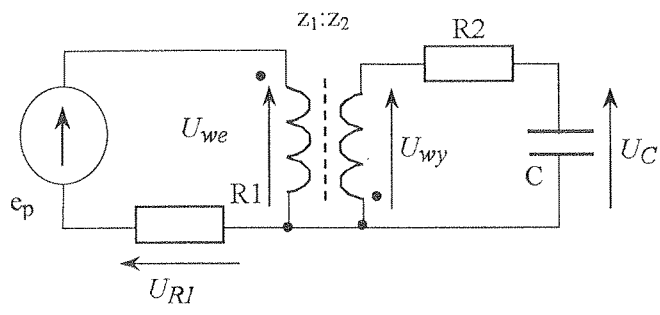
$$Z_{thur} = \frac{T_r(t) - T_a}{P} \quad (17)$$

Parametry modelu tych impedancji (R_{thuu} , R_{thrr} , R_{thur} , R_{thru} , a_{thiuu} , a_{thirr} , a_{thiru} , a_{thiur} , τ_{thuu} , τ_{thrr} , τ_{thur} , τ_{thru} , N_{uu} , N_{rr} , N_{ru} , N_{ur}) należy wyznaczać przy wykorzystaniu metody z pracy [25].

5. WERYFIKACJA POPRAWNOŚCI MODELU

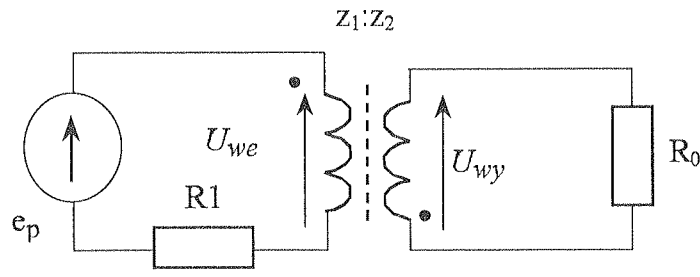
W celu zbadania poprawności zaproponowanego modelu transformatora przeprowadzono obliczenia i pomiary jego charakterystyk zaciskowych. W badaniach wykorzystano dwa układy pomiarowe. Pierwszy z nich (rys.2), przeznaczony jest do wyznaczania charakterystyk rdzenia ferrytowego, natomiast drugi (rys.3) jest klasycznym układem pracy transformatora z obciążeniem rezystancyjnym.

Przedstawiono wyniki badań dwóch transformatorów. Pierwszy z nich, oznaczany dalej symbolem Tr1 zawiera rdzeń B65701-D-R26 firmy EPCOS z materiału N26 [11]. Transformator ten posiada dwa uzwojenia zawierające po 35 zwojów, wykonane poczwórnym przewodem miedzianym w emalii o średnicy 0,28 mm. Z kolei, drugi transformator, oznaczany dalej symbolem Tr2 wykonano przy wykorzystaniu rdzenia B65671-W-R67 tej samej firmy z materiału N67, na którym nawinięto dwa identyczne uzwojenia po 15 zwojów każde przewodem miedzianym w emalii o średnicy 0,63 mm. Wartości parametrów modelu dla obu transformatorów, uzyskane zgodnie z metodą przedstawioną w rozdziale 4, przedstawiono w tabeli 2.



Rys. 2. Układ do pomiaru charakterystyk rdzenia

Fig. 2. The circuit for the measurements of the core characteristic



Rys. 3. Układ do pomiaru charakterystyk transformatora

Fig. 3. The circuit for the measurements of the transformer characteristics

Wartości parametrów modelu dla transformatorów Tr1 i Tr2

Parameter values of the model for transformers Tr1 and Tr2

Parametr	T_C	T_0	B_{r0}	H_{C0}	α_r	α_C	α_U
Tr1	420 K	300 K	140mT	23 A/m	-1,9e-3 K ⁻¹	-1e-3 K ⁻¹	4,33e-3 K ⁻¹
Tr2	490 K	298 K	200 mT	20 A/m	-3,3e-3 K ⁻¹	-4e-3 K ⁻¹	11,4e-3K ⁻¹
Parametr	B_{S0}	α_S	H_S	B_X	H_X	μ_i	V_e
Tr1	380 mT	-4.2e-3 K ⁻¹	1200 A/m	345 mT	190 A/m	1950	6,1e-6 m ³
Tr2	480 mT	-2,8e-3 K ⁻¹	1200 A/m	410 mT	180 A/m	2100	4,32e-6 m ³
Parametr	S	l	z ₁	z ₂	l _p	l _w	S _p
Tr1	1,36e-4 m ²	0,045 m	35	35	1,1 m	1,3 m	0,246 mm ²
Tr2	1,08e-4 m ²	0,04 m	15	15	0,6 m	0,65 m	0,312 mm ²
Parametr	A	D	T _m	R _{thrr}	R _{thuu}	R _{thru}	R _{thur}
Tr1	150 m/H	0,2·10 ⁻³ K ⁻¹	353 K	28,2 K/W	28,2 K/W	23 K/W	23 K/W
Tr2	150 m/H	0,2 · 10 ⁻³ K ⁻¹	373 K	32,4 K/W	32,4 K/W	20,13 K/W	20,13 K/W

TOM 5

U
je cha
rezysta
moc w
jenia s
zrealiz
rezysta
iz efek
Na
wyjści
temper
sinuso

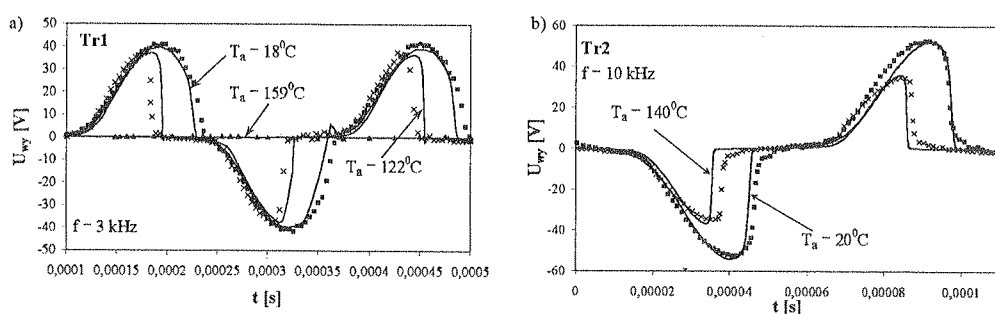
a) 50
40
30
20
10
0
-10
-20
-30
-40
-50
0,0

U_{wy} [V]

Rys
Fig.
Jak
ze wzr
wyjści
nawet
Tr1 prz
skutek
Na
(punkty
wości f
widmow
Jak
transfor
tego sy
rystyki
zgodno
W
szy niż

Uzyskane wyniki badań można podzielić na dwie grupy. Pierwsza z nich obejmuje charakterystyki transformatora pracującego w układzie z rys. 2 przy dużej wartości rezystancji R2 w szerokim zakresie zmian temperatury otoczenia. W tym przypadku moc wydzielana w transformatorze jest niewielka, a zatem temperatury rdzenia i uzwojenia są praktycznie równe temperaturze otoczenia. Z kolei, druga grupa badań została zrealizowana w układzie z rys. 3 w temperaturze pokojowej przy małych wartościach rezystancji obciążenia, dla których moce wydzielane w transformatorze są na tyle duże, iż efekt samonagrzewania istotnie wpływa na charakterystyki transformatora.

Na rys. 4 porównano obliczone (linie) i zmierzone (punkty) przebiegi napięcia wyjściowego transformatora w stanie ustalonym odpowiadające różnym wartościom temperatury otoczenia T_a dla obu transformatorów przy pobudzeniu układu napięciem sinusoidalnym.



Rys. 4. Obliczone i zmierzone przebiegi napięcia wyjściowego transformatorów: a) Tr1, b) Tr2

Tabela 2

Fig. 4. The calculated and measured transients of the transformer output voltage: a) Tr1, b) Tr2

Jak widać, uzyskano zadawalającą zgodność wyników pomiarów i obliczeń. Wraz ze wzrostem temperatury otoczenia zauważono wzrost poziomu zniekształceń napięcia wyjściowego oraz spadek amplitudy tego napięcia, który dla transformatora Tr2 wynosi nawet 30% przy zmianie temperatury z 20°C na 140°C, natomiast dla transformatora Tr1 przy temperaturze $T_a = 159^\circ\text{C}$ napięcie wyjściowe jest praktycznie równe zero na skutek przekroczenia przez temperaturę rdzenia wartości temperatury Curie.

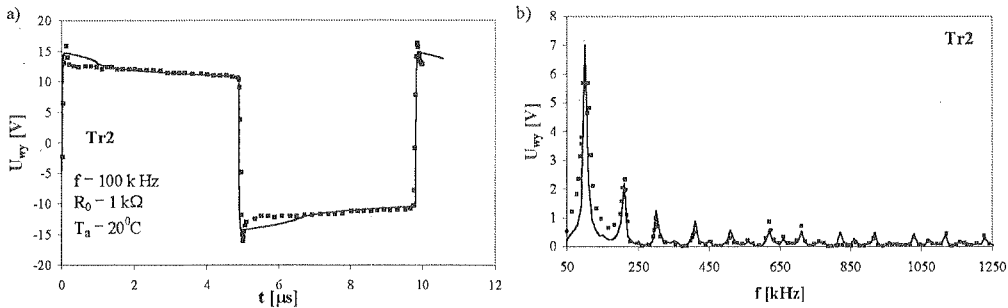
Na rys. 5a porównano przebiegi napięcia wyjściowego otrzymane z pomiarów (punkty) oraz z obliczeń (linie) przy pobudzeniu sygnałem prostokątnym o częstotliwości $f = 100\text{ kHz}$, zaś na rys. 5b porównano obliczone i zmierzone charakterystyki widmowe napięcia wyjściowego.

Jak widać, sygnał prostokątny o rozważanej częstotliwości jest przenoszony przez transformator przy małych zniekształceniach, ponieważ najważniejsze składowe widma tego sygnału znajdują się w paśmie przenoszenia transformatora. Zarówno charakterystyki czasowe, jak i widmowe uzyskane z pomiarów i obliczeń wykazują dobrą zgodność.

W rozważanym układzie z rys. 2, prąd uzwojenia pierwotnego był znacznie większy niż prąd uzwojenia wtórnego, co skutkowało tym, że o wartości natężenia pola

α_U
$33\text{e-}3\text{ K}^{-1}$
$1,4\text{e-}3\text{ K}^{-1}$
V_e
$5,1\text{e-}6\text{ m}^3$
$3,32\text{e-}6\text{ m}^3$
S_p
246 mm^2
312 mm^2
R_{thur}
23 K/W
$0,13\text{ K/W}$

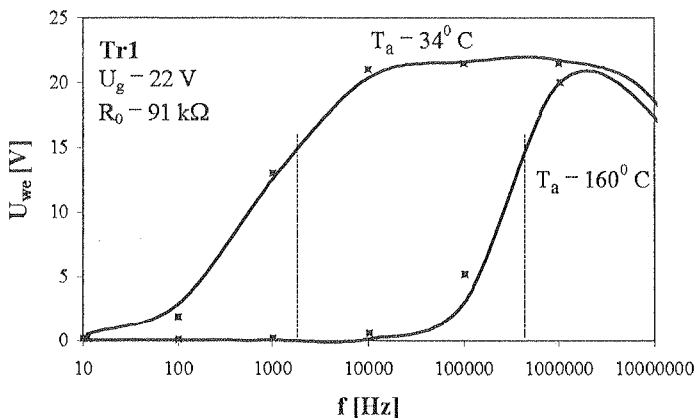
magnetycznego w rdzeniu decydował tylko prąd tego uzwojenia. Z kolei, w układzie z rys. 3, przez oba uzwojenia płynie prąd o zbliżonej wartości, lecz o przeciwnych znakach powodując, że w tym przypadku transformator wchodzi w nasycenie dla znacznie większych wartości prądów uzwojeń.



Rys. 5. Obliczone i zmierzone przebiegi napięcia wyjściowego transformatora Tr2 przy pobudzeniu sygnałem prostokątnym: a) w dziedzinie czasu, b) w dziedzinie częstotliwości

Fig. 5. The calculated and measured transients of the output voltage of the Tr2 transformer excited by the rectangular train: a) in the time domain, b) in the frequency domain

Na rys. 6 przedstawiono obliczoną i zmierzoną w układzie z rys. 3 zależność napięcia wejściowego transformatora od częstotliwości przy rezystancji obciążenia $R_0 = 91$ kΩ dla dwóch temperatur otoczenia.



Rys. 6. Obliczona i zmierzona charakterystyka częstotliwościowa napięcia wejściowego transformatora

Fig. 6. Calculated and measured frequency characteristics of the transformer input voltage

Z przeprowadzonych pomiarów i symulacji widać ograniczenie pasma przenoszenia transformatora zarówno od dołu, jak i od góry. Dolna częstotliwość graniczna transformatora rośnie wraz ze wzrostem temperatury. Wartości tych częstotliwości wynoszą

około
Dodatk
wartość
napięc

Z
między
oraz T

Jak
spadek
rezysta
o około
15 Ω d

Z

mocą p

T_u obu

rezysta

przy cz

świadcz

Tr1 war

tempera

a) 70

65

60

55

50

45

40

0

U_{wep} [V]

Tr

T_a

f

Pod

którego

ściowym

graficzne

Jak wyn

rozkład t

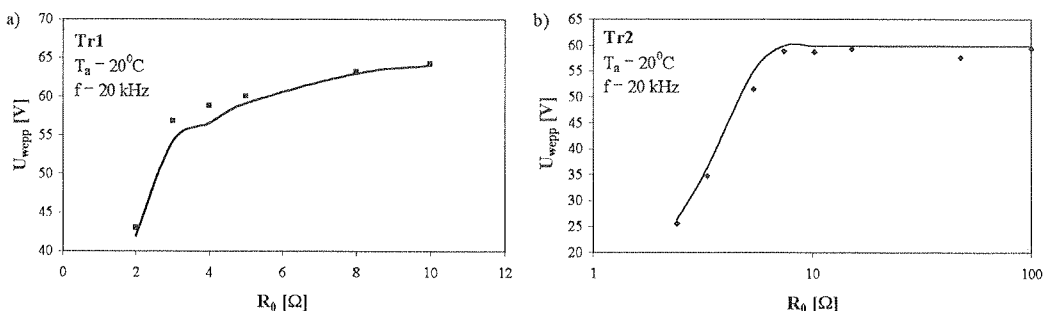
wierzchn

około 1,5 kHz przy temperaturze $T_a = 34^\circ\text{C}$ oraz około 500 kHz przy $T_a = 160^\circ\text{C}$. Dodatkowo warto podkreślić, że wraz ze spadkiem częstotliwości nie tylko maleje wartość napięcia na wejściu transformatora, ale także występują silne zniekształcenie napięcia wyjściowego [16].

Z kolei na rys. 7 przedstawiono wyniki pomiarów i obliczeń zależności wartości międzyszczytowej napięcia na uzwojeniach pierwotnych transformatorów Tr1 (rys. 7a) oraz Tr2 (rys. 7b) od rezystancji obciążenia.

Jak widać, zmniejszenie wartości rezystancji obciążenia transformatora powoduje spadek wartości napięcia wejściowego. W przypadku transformatora Tr1, zmiana tej rezystancji w przedziale od $10\ \Omega$ do $2\ \Omega$ spowodowała spadek napięcia wejściowego o około 30%, natomiast dla transformatora Tr2, zmiana rezystancji w przedziale od $15\ \Omega$ do $2,4\ \Omega$ spowodowała spadek napięcia aż o 50%.

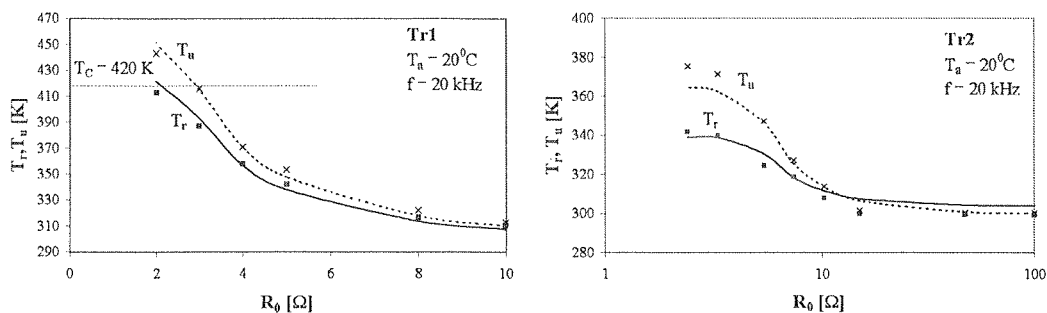
Z kolei, na rys. 8 przedstawiono obliczone (linie) i zmierzone (punkty) za pomocą pirometru Optex PT — 3S zależności temperatury rdzenia T_r oraz uzwojenia T_u obu transformatorów od rezystancji obciążenia. Jak widać, wraz ze zmniejszaniem rezystancji obciążenia rośnie wartość temperatury zarówno rdzenia, jak i uzwojenia, przy czym w rozważanym przypadku większą wartość ma temperatura uzwojenia, co świadczy o tym, że straty energii występują głównie w uzwojeniu. Dla transformatora Tr1 wartość temperatury rdzenia odpowiadająca rezystancji $R_0 = 2\ \Omega$ jest bardzo bliska temperatury Curie, równej $420\ \text{K}$.



Rys. 7. Zależność napięcia na uzwojeniach pierwotnych transformatorów Tr1 oraz Tr2 od rezystancji obciążenia

Fig. 7. The dependence of the primary winding voltage of Tr1 and Tr2 transformers on the load resistance

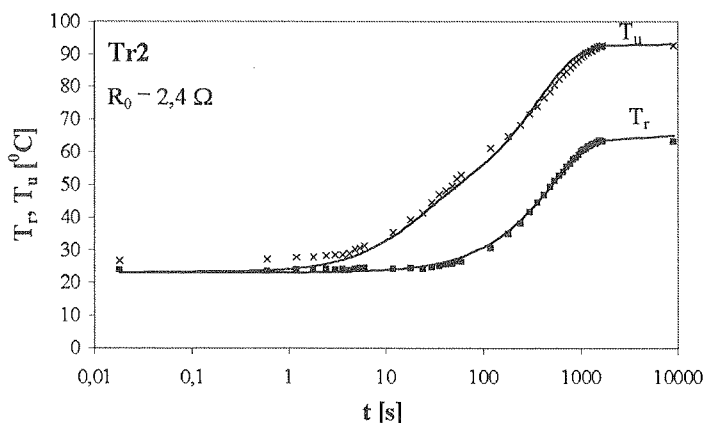
Podobne jakościowo charakterystyki są obserwowane dla transformatora Tr2, dla którego zmierzono także rozkład temperatury w stanie ustalonym, przy napięciu wejściowym $u_{we} = 10,3\ \text{V}$ i rezystancji obciążenia $R_0 = 2,4\ \Omega$ za pomocą kamery termograficznej V-20 oraz urządzenia do ultraszybkich pomiarów temperatury firmy Vigo. Jak wynika z tych pomiarów, rozkład temperatury na powierzchni rdzenia, jak również rozkład temperatury uzwojenia są praktycznie równomierne (różnice temperatur na powierzchni rdzenia nie przekraczają kilku kelwinów), co potwierdza słuszność przyjęcia



Rys. 8. Zależność temperatury rdzenia i uzwojeń transformatorów Tr1 i Tr2 od rezystancji obciążenia

Fig. 8. The dependence of the core and windings temperatures of Tr1 and Tr2 transformers on the load resistance

w modelu termicznym transformatora jednej temperatury rdzenia i jednej temperatury uzwojeń.



Rys. 9. Czasowe przebiegi zmian temperatury rdzenia T_r i uzwojeń T_u transformatora Tr2 przy $R_0 = 2,4 \Omega$

Fig. 9. The core T_r and the winding T_u temperatures of the transformer Tr2 for the resistance $R_0 = 2,4 \Omega$

Czasowe przebiegi temperatury rdzenia T_r i uzwojeń T_u przy pobudzeniu transformatora Tr2 uskokiem mocy o wartości około 2,5 W przedstawiono na rys. 9. Jak widać, temperatura uzwojeń zaczyna w widoczny sposób narastać po upływie kilku sekund od załączenia zasilania, natomiast temperatura rdzenia — po upływie kilkudziesięciu sekund. W stanie ustalonym, zaobserwowanym po upływie około 1000 s od chwili włączenia pobudzenia, różnica temperatur rdzenia i uzwojeń wynosi około 25°C . A zatem, słuszne jest przyjęcie w modelu dwóch różnych temperatur rdzenia i uzwojeń. W rozpatrywanym przypadku, na skutek wzajemnych sprzężeń termicznych między

TOM 3

rdzenia
strat

V

mu S
na ch
odnie
wych
nym c
zniek
tor pr
wzras
wyjść

P

elektr
chodz
miczn
ny m
napię
funkc
granic

1. D
- and
2. R
- we
3. J.
- Fe
- Co
4. R.
- Pro
5. R.
- ter
- Ele
6. SI
7. B.
- Ele
- Ele
8. W
- mi

Tr2
 $T_a = 20^\circ\text{C}$
 $f = 20\text{ kHz}$

obciążenia

on the load

temperatury

a Tr2

 $R_0 = 2,4\ \Omega$

u transfor-
Jak widać,
ku sekund
dziesięciu
od chwili
o 25°C . A
i uzwojeń.
ch między

rdzeniem a uzwojeniem temperatura rdzenia wzrosła aż o 40°C pomimo zerowych strat mocy w rdzeniu.

6. UWAGI KOŃCOWE

W pracy zaproponowano nowy elektrotermiczny model transformatora dla programu SPICE. Przedstawione w pracy przykłady potwierdzają silny wpływ temperatury na charakterystyki transformatora i poprawność zaproponowanego modelu zarówno w odniesieniu do charakterystyk częstotliwościowych transformatora, przebiegów czasowych napięć na jego zaciskach przy pobudzeniu sygnałem harmonicznym i prostokątnym oraz widma tych napięć. Wzrost temperatury rdzenia skutkuje wzrostem poziomu zniekształceń napięcia wejściowego, a po przekroczeniu temperatury Curie transformator praktycznie nie przenosi energii z wejścia na wyjście. Na skutek samonagrzewania wzrasta temperatura rdzenia i uzwojeń oraz widoczne są zmiany w przebiegu napięcia wyjściowego transformatora w funkcji rezystancji obciążenia.

Przedstawione w pracy przykłady potwierdzają, że zaproponowany przez autorów elektrotermiczny model transformatora impulsowego poprawnie modeluje zjawiska zachodzące w tym elemencie oraz dowodzą znaczenia prowadzenia analiz elektrotermicznych układów zawierających te elementy. W szczególności, stosując przedstawiony model elektrotermiczny można wyznaczyć zakres dopuszczalnych obciążeń oraz napięć wejściowych transformatora, przy których będzie on poprawnie spełniał swoją funkcję, a temperatury rdzenia i uzwojeń nie będą przekraczały katalogowych wartości granicznych.

7. BIBLIOGRAFIA

1. D. C. Jiles, D. L. Atherton: *Theory of Ferromagnetic Hysteresis*. Journal of Magnetism and Magnetic Materials, Vol. 61, 1986, p. 48.
2. R. Ericson, D. Maksimovic: *Fundamentals of Power Electronics*. Norwell (USA), Kluwer Academic Publisher, 2001.
3. J. Fuzi, E. Helerea, A. Ivanyi: *Experimental Construction of Preisach Models for Ferromagnetic Cores*. PCIM 1998 Europe Official Proc. of the 40th International Power Electronics Conference, Nurnberg, 1998, p. 661.
4. R. H. Coit: *Designing with Magnetic Cores at High Temperatures*. PCIM 2001 Europe Official Proc. of the 43th International Power Electronics Conference, Nurnberg, 2001, p. 607.
5. R. Lucke, S. Ahne, S. Plutzer, J. Wrba: *The Present and Future of Ferrite Materials for Power Applications*. PCIM 2001 Europe Official Proc. of the 43th International Power Electronics Conference, Nurnberg, 2001, p. 83.
6. SIFERRIT Materials. Materiały firmowe EPCOS AG, Munchen, 2000.
7. B. Grzesik, M. Stępień: *New Construction of High Frequency Transformer for Power Electronics Appliances*. V Krajowa Konferencja Naukowa Sterowanie w Energoelektronice i Napędzie Elektrycznym, Łódź, 2001, p. 167.
8. W. Dąbrowski, J. Król, M. Sobóń: *Materiały i rdzenie ferrytowe, ferryty magnetycznie miękkie ferroxyd*. Warszawa, Wydawnictwa Przemysłu Maszynowego "Wema", 1976.

9. R. Kurdziel: *Obwody magnetyczne*. Warszawa, PWN, 1962.
10. Z. Celiński: *Materiałoznawstwo elektrotechniczne*. Warszawa, Oficyna Wydawnicza Politechniki Warszawskiej, 1998.
11. N. Mohan, T. M. Undeland, W. P. Robbins: *Power Electronics: Converters, Applications, and Design*. New York, John Wiley & Sons, 1995.
12. A. Fery: *Overview of modeling of the main SMPS topologies with leakage inductance*. Infineon Technologies AG, 2001.
13. A. Van den Bossche, V. Valchev, J. Melkebeek: *Improved Thermal Modelling of Magnetic Components for Power Electronics*. EPE Journal, Vol. 12, No. 2, 2002, p. 7.
14. J. Porębski, P. Korohoda: *SPICE, program analizy nieliniowej układów elektronicznych*. Warszawa, WNT, 1995.
15. W. Pietrenko: *Układowy model transformatora*. I Konferencja Elektroniki, Kołobrzeg, 2002, s. 697.
16. K. Górecki, J. Zarębski: *Modelowanie wpływu temperatury na charakterystyki transformatora impulsowego w programie SPICE*. XXVI Międzynarodowa Konferencja z Podstaw Elektrotechniki i Teorii Obwodów IC-SPETO'2003, Niedzica, 2003, Vol. 2, s. 335.
17. K. Górecki, J. Zarębski: *Modelowanie elementów indukcyjnych w programie SPICE*. Elektronika konstrukcje, technologie, zastosowania, Not-Sigma, Nr 1, 2004, s. 40.
18. J. Izydorczyk: *Pspice. Komputerowa symulacja układów elektronicznych*. Gliwice, Wydawnictwo Helion, 1993.
19. K. Górecki: *Modelowanie cewki z rdzeniem ferrytowym w programie SPICE z uwzględnieniem samonagrzewania*. Kwartalnik Elektroniki i Telekomunikacji, Nr 3, 2003, s. 389.
20. K. Górecki, J. Zarębski, J. Krupa: *Wpływ temperatury na właściwości cewki z rdzeniem ferromagnetycznym*. Elektronizacja, Not-Sigma, Nr 11, 2002, s. 21.
21. K. Górecki, J. Zarębski: *Możliwości zastosowania funkcji programu Pspice do modelowania elementów półprzewodnikowych*. Elektronizacja, Not-Sigma, Nr 1, 2002, s. 7.
22. K. Górecki: *Elektrotermiczny makromodel tranzystora Darlingtona do analizy układów elektronicznych*. Praca doktorska, Politechnika Gdańska, 1999.
23. *Instrukcja obsługi urządzenia do ultraszybkich pomiarów temperatury*. Vigo- System, Warszawa 2001.
24. Ch. P. Basso: *Switch-Mode Power Supply SPICE Cookbook*. McGraw- Hill, 2001.
25. J. Zarębski, K. Górecki: *Thermal Transients in Bipolar Transistors; Measurements and Simulations*. Int. Conf. on Information, Systems Methods Applied to Engineering Problems, Malta 1993, V. 2, p. 111.

K. GÓRECKI, J. ZARĘBSKI

THE ELECTROTHERMAL MODEL OF THE PULSE TRANSFORMER FOR SPICE

Summary

The pulse transformer is a commonly used electronic device, which makes it possible the transformation of the electrical energy between separated circuits. Any pulse transformer consists of a ferrite core and at least two windings. Because of the nonlinear core characteristics and non-ideal parameters of windings, the characteristics of the real pulse transformer differ from the characteristics of the ideal one.

On the other hand, more and more commonly used softwares for the analysis of electronic circuits, e.g. the SPICE software, include the relatively simple, idealised transformer models. In this models the temperature influencing on the transformer terminal characteristics and selfheating are omitted. In the paper a new electrothermal model (ETM) of the pulse transformer is proposed. In the ETM the description of the

nonlinear core characteristics, the influence of selfheating on the terminal characteristics, the dependence of the windings resistances on the temperature as well as on frequency of the transformed signal are taken into account. In the thermal part of the considered ETM, the two different temperatures are distinguished: the core temperature and the windings one. The power losses in the core and the windings, as well as the mutual thermal interactions between both the parts of the transformer, are included in the proposed model. This model consists of five submodels representing:

- the dependencies describing the properties of the ferrite material used for the construction of the core,
- the dependencies of the magnetic field strength in the core on the windings currents,
- the equation of the electromagnetic induction for all windings,
- the dependencies of power losses in windings on the temperature, winding current and frequency,
- the dependencies of the core and windings temperatures on the transformer terminal currents and voltages.

The new ETM was implemented in SPICE software as the subcircuit, including the current and voltage sources, resistors and capacitors. The simple method of the model parameters estimation on the basis of the catalogue data of the core, the ferrite material and the coil wire is also proposed. The correctness of the elaborated ETM was verified experimentally in the wide range of the ambient temperature, the frequency of the transformed signal, the input voltage and the load resistance, for two transformers of the different types of the ferrite cores and different numbers of the coils in both windings. In the all cases the good agreement between measurement and calculation results have been obtained.

Keywords: pulse transformer, electrothermal model, selfheating, SPICE

E
Z

C
Z
Z
Z
L
P
Z
S

W
Coraz
menta
siebie
przem
pomia
przetw

ELEKTRONICZNY MODUŁ DZIESIĘCIOKROTNEGO ZWIĘKSZANIA DOKŁADNOŚCI PRZETWARZANIA SYGNAŁÓW OPTOELEKTRONICZNEGO PRZETWORNIKA POŁOŻENIA

ZBIGNIEW SZCZEŚNIAK

*Politechnika Świętokrzyska, Wydział Elektrotechniki, Automatyki i Informatyki
Al. Tysiąclecia PP7, Kielce, Polska
e-mail: Zbigniew.Szczesniak@poczta.fm*

*Otrzymano 2005.01.10
Autoryzowano 2005.03.31*

W artykule przedstawiono układ zwiększenia dokładności i wyróżnienia kierunku ruchu optoelektronicznego przetwornika. W układzie tym wytworzono napięcia fazowe o założonym przesunięciu wynoszącym 36° . Umożliwiają one po odpowiednim przetworzeniu 2,5-krotne zwiększenie częstotliwości w stosunku do sygnałów podstawowych wynikających z liniału pomiarowego przetwornika. W układzie zliczania impulsów oraz rozróżniania ich kolejności, a tym samym rozróżniania kierunku ruchu przetwornika, możliwe jest czterokrotne zwiększenie jego dokładności. Przedstawione układy umożliwiają dziesięciokrotne zwiększenie dokładności przetwarzania przetwornika.

Słowa kluczowe: wyróżnienie kierunku ruchu zwiększenie, dokładności przetwornika optoelektronicznego

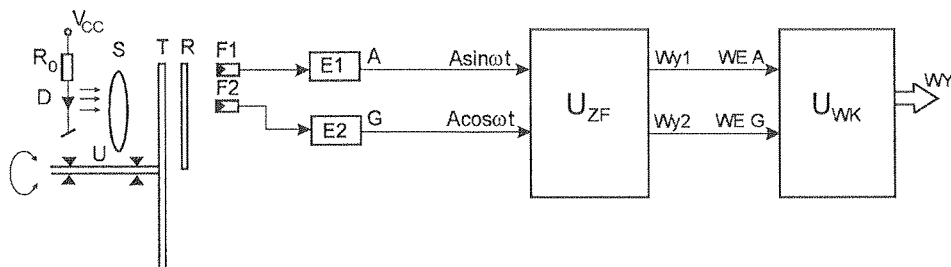
1. WPROWADZENIE

W pomiarach położenia przywiązywana jest duża waga do dokładności pomiaru. Coraz częściej stosuje się optoelektroniczne przetworniki położenia, kodowe lub inkrementalne. Te ostatnie wytwarzają na jego wyjściu dwa sygnały przesunięte względem siebie o $1/4$ okresu jego liniału pomiarowego. Zmniejszenie przedziału kwantowania przemieszczenia obiektu można dokonać na drodze precyzyjniejszej budowy liniału pomiarowego przetwornika lub na drodze odpowiedniego przetwarzania sygnałów z przetwornika optoelektronicznego. Tendencja stosowania elektronicznych metod prze-

tworzania sygnałów znajduje coraz większą liczbę zwolenników z racji prostszej technologii wykonania. W artykule przedstawiono, opracowaną metodę i elektroniczny układ zwielokrotnienia częstotliwości sygnałów optoelektronicznego przetwornika (interpolacja) oraz przekształcania ich na postać cyfrową (digitalizacja). Układ ten wraz z układem wyróżniania kierunku ruchu przetwornika, umożliwia zwiększenie dokładności optoelektronicznego przetwornika położenia.

Ogólny schemat blokowy współpracy optoelektrycznego przetwornika położenia z układem przetwarzania sygnałów przedstawiono na rys. 1.

Obecnie najczęściej stosowane są przetworniki fotoelektryczne obrotowe, składające się z tarczy T zawierającej na obwodzie sekwencję pól przezroczystych i nieprzezroczystych dla światła, tworząc liniał pomiarowy. Tarcza pomiarowa mocowana jest do ułożyskowanego wału U sprzęgającego przetwornik z obiektem. Jednostka skanująca przetwornika składa się ze źródła światła D , soczewek kondensujących S służących do otrzymania równoległej wiązki światła, skanującej siatki rozdzielczej R z siatką indeksową tarczy T i krzemowych ogniw fotoelektrycznych $F1, F2$. Gdy skala przesuwana jest względem jednostki skanującej, linie skali zbiegają się na przemian z liniami lub odstępami na siatce indeksowej. Okresowe fluktuacje natężenia światła będące tego rezultatem, zamieniane są przez ogniwa fotoelektryczne w sygnały elektryczne.



Rys. 1. Schemat ogólny układu przetwarzania sygnałów optoelektronicznego przetwornika położenia

Fig. 1. General diagram of the processing system of photoelectric position transducer signals

Liczba impulsów jest proporcjonalna do przemieszczenia kąowego tarczy.

Fotoelementy są przesunięte względem siebie o $1/4$ okresu liniału pomiarowego. Umożliwia to uzyskanie na wyjściu A układu elektroniki $E1$, sygnału przetwornika o fali $Asin\omega t$, natomiast na wyjściu G układu elektroniki $E2$, sygnału przetwornika o fali $Acos\omega t$. Okres sygnału jest równy okresowi siatki podziałki skali /liniału/ tego przetwornika.

Odpowiednie przetwarzanie tych sygnałów w układzie U_{ZF} , umożliwia wytworzenie sygnałów o zwielokrotnionej częstotliwości w stosunku do sygnałów wejściowych tego modułu. Wytworzone sygnały $Wy1, Wy2$ to sygnały o przebiegu prostokątnym i współczynniku wypełnienia 0,5, wzajemnie przesunięte o $1/4$ okresu. Na bazie tych sygnałów, w układzie U_{WK} , określany jest kierunek obrotu optoelektronicznego przetwornika.

2. METODA I UKŁAD ZWIELOKROTNIANIA CZĘSTOTLIWOŚCI SYGNAŁÓW POMIAROWYCH OPTOELEKTRONICZNEGO PRZETWORNIKA POŁOŻENIA

Wyjściowe sygnały z fotoelektrycznego przetwornika położenia to dwa sygnały sinusoidalne, przesunięte w fazie względem siebie o $1/4$ okresu [2]. Okres sygnału jest równy okresowi siatki podziałki skali /liniału/ tego przetwornika. Sygnały skanujące z przetworników są po pierwsze wzmacniane, a następnie interpolowane. Opracowana metoda interpolacji wykorzystuje sieć rezystorów, która wytwarza przesunięte fazowo sygnały z dwóch sinusoidalnych sygnałów skanujących poprzez wektorowe ich sumowanie zgodnie z niżej przedstawionym opisem. Sygnały napięciowe z czujnika

$$x = A \sin \varphi \quad ; \quad y = A \cos \varphi \quad (1)$$

podawane są na dzielnik rezystancyjny wytwarzając na nim napięcie równe

$$U_d = A(\sin \varphi R_1 + \cos \varphi R_2) \quad (2)$$

gdzie: $\varphi = 2\pi y/T$ przy czym A — amplituda sygnału; y — przemieszczenie; T — okres sygnału.

W powyższej zależności rezystancje R_1 , R_2 określono jako :

$$R_1 = \frac{R}{|\sin \alpha|} \quad \text{oraz} \quad R_2 = \frac{R}{|\cos \alpha|} \quad (3)$$

gdzie: α — kąt założonego przesunięcia fazowego, R — przyjęta rezystancja odniesienia dla sygnału.

Po uwzględnieniu powyższych zależności i przekształceniu otrzymuje się

$$U_d = A_1 \sin(\varphi + \alpha) \quad (4)$$

przy czym A_1 — amplituda wytworzonego sygnału o założonym przesunięciu fazowym α .

Dla przejrzystości opisu metody założono, że przebiegi wyjściowe z przetwornika ($\sin$, $\cos$) są o stałych częstotliwościach (okresach). Przypadek taki odzwierciedla współpracę przetwornika z obiektem przemieszczającym się ze stałą prędkością /Rys.2,Rys.3/.

W opracowanej metodzie, w pierwszym etapie sygnały są wzmacnianie i odwracane tworząc cztery przebiegi: $\sin$, $\cos$, $-\sin$, $-\cos$, po czym odpowiednie dwa z nich poddawane są interpolacji. Istotne jest aby sygnały brane do interpolacji miały chwilowe wartości napięcia zawsze przeciwnych znaków w interesującej nas ćwiartce ich okresu. Dla I ćwiartki ($0^\circ - 90^\circ$) wykorzystywane są sygnały $\sin$ i $-\cos$, dla II ćwiartki ($90^\circ - 180^\circ$) $\sin$ i $\cos$, dla ćwiartki III ($180^\circ - 270^\circ$) sygnały $-\sin$ i $\cos$, oraz $-\sin$ i $-\cos$ dla ćwiartki IV ($270^\circ - 360^\circ$).

Sygnały trafiają do dzielników rezystancyjnych o tak dobranych wartościach, aby na ich wyjściach powstały przebiegi sinusoidalne przesunięte względem siebie o założony

kąt. W przedstawianej metodzie wynosi on 36° (zgodnie z zależnościami 2,3,4). Dla tak przyjętego przesunięcia fazowego, dla kątów 0, 180, stopni interpolacja nie występuje i do dalszej obróbki brane są sygnały: $\sin$, $-\sin$. Wytworzone sygnały podawane są do jednego z dziesięciu komparatorów zmieniających jego kształt na sygnał prostokątny, przy czym zmiana stanu następuje w miejscu przejścia sygnału wejściowego przez zero. W ten sposób powstaje dziesięć sygnałów prostokątnych przesuniętych względem siebie o $1/10$ okresu (36°). Sygnały rozdzielane są na dwie grupy w taki sposób, aby w grupie każdy kolejny sygnał przesunięty był o $1/5$ okresu względem poprzedniego. Powstaną dwie grupy pięciu sygnałów przesuniętych o $1/5$ okresu w każdej grupie i o $1/10$ okresu względem drugiej grupy /Rys.2,Rys.3/.

Kolejnym etapem jest zróżniczkowanie przebiegów tak, aby z każdego przebiegu powstał tylko jeden impuls w okresie, czyli z jednej grupy powstanie pięć impulsów dzielących jeden okres wejściowego sygnału odniesienia na pięć równych części przy założeniu, że częstotliwość sygnałów wejściowych jest stała. Impulsy z dwóch grup są sumowane w dwa przebiegi impulsowe o częstotliwości pięciokrotnie większej od przebiegu wejściowego (rys.2, rys.3). Służą one jako sygnały zegarowe dla przerzutników JK ustawionych w taki sposób, aby każdy impuls wejściowego sygnału zegarowego powodował zmianę stanu na wyjściu przerzutnika na przeciwny. Tworzone są w ten sposób dwa symetryczne przebiegi prostokątne o częstotliwości 2,5-krotnie większej o częstotliwości wejściowej, czyli na jeden okres wejściowego sygnału sinusoidalnego przypada 2,5 okresów wyjściowego sygnału prostokątnego. Sygnały te przesunięte są względem siebie o $1/4$ okresu, a kierunek ich przesunięcia zależy od kierunku przesunięcia sygnałów wejściowych uzależnionych od kierunku ruchu przetwornika (rys.3.)

Mimo tego, że w module tym występuje znaczna ilość układów cyfrowych i część operacji wykonywana jest na sygnałach cyfrowych, to istotą działania układu jest możliwość współpracy z analogowymi sygnałami wejściowymi i interpolacji tych sygnałów metodą analogową. Część cyfrowa przygotowuje jedynie sygnał do współpracy z następnymi modułami [1],[2],[3].

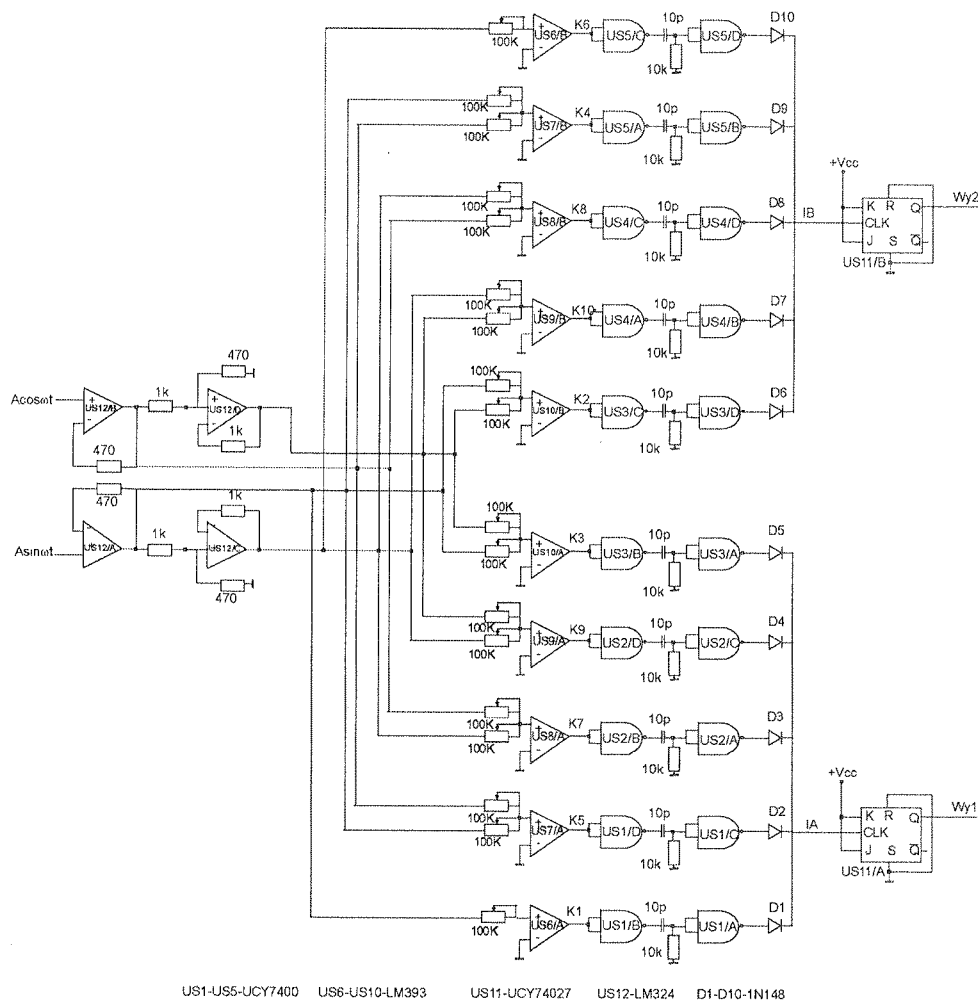
Obwody wejściowe (rys.2) umożliwiają współpracę z sygnałami sinusoidalnymi symetrycznymi. Składają się na nie 4 wzmacniacze operacyjne zawarte w jednej obudowie układu LM324. Dwa z nich pracują jako wtórники, aby układ nie obciążał przetwornika pomiarowego. Dwa kolejne wzmacniacze odwracają sygnał o 180° . Odpowiednie dwa sygnały trafiają do rezystorów interpolujących te przebiegi tworząc trzeci o założonym przesunięciu fazowym. Dla kątów: 0, 180 stopni nie występuje interpolacja i do dalszej obróbki brane są odpowiednio sygnały: $\sin$, $-\sin$.

Wyliczone wartości rezystorów, są precyzyjnie ustawione na potencjometrach montażowych. W ten sposób powstaje 10 przebiegów sinusoidalnych przesuniętych względem siebie o 36° ($1/10$ okresu), które trafiają do wejść komparatorów. Funkcje komparatorów spełniają układy LM393. Prostokątne sygnały z wyjść komparatorów podawane są na wejścia układów różniczkujących zbudowanych z bramek NAND, kondensatorów

Acos

Asin

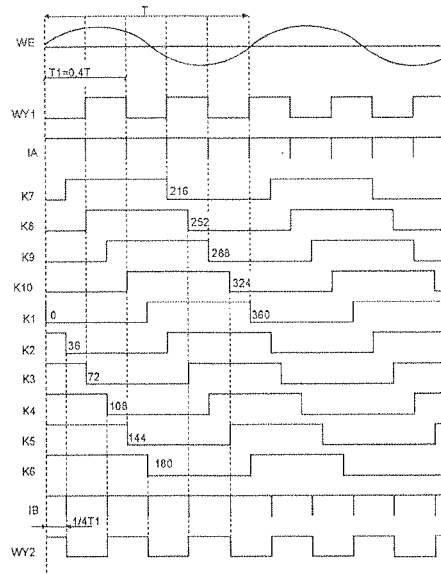
i rez
wanc
Sygn
zsum
przer
Mod
mont



Rys. 2. Układ 2,5- krotnego zwielokrotnienia częstotliwości sygnałów przetwornika

Fig. 2. System for transducer signals frequency increment for 2.5 times

i rezystorów. W związku z występowaniem sygnałów o wartościach do 15V zastosowano układy 4011 serii CMOS zawierające po 4 bramki NAND. Sygnały z wyjść układów różniczkujących zostały poprzez diody 1N4148 odpowiednio zsumowane tworząc dwa sygnały zegarowe dla przerzutników JK (4027). Na wyjściach przerzutników powstają o przesunięciu fazowym 1/4 okresu względem siebie. Moduł ten wykonany został na druku dwustronnym z elementami półprzewodnikowymi montowanymi w podstawkach.



Rys. 3. Metoda 2,5-krotnego zwielokrotnienia częstotliwości sygnałów przetwornika

Fig. 3. Method for transducer signals frequency increment for 2.5 times

3. METODA ROZRÓŻNIANIA KIERUNKU RUCHU OPTOELEKTRONICZNEGO PRZETWORNIKA ZREALIZOWANA W OPARCIU O FUNKCJE LOGICZNE SYGNAŁÓW TEGO PRZETWORNIKA I IMPULSÓW RUCHU WYGENEROWANYCH W UKŁADACH RC

W metodzie tej, (rys. 4, rys. 5.) poprzez realizację sumy odpowiednich iloczynów sygnałów $A, \bar{A}, G, \bar{G}$ i sygnałów powstałych z generacji ich impulsów o czasie τ (od ich zbocza narastającego) można zliczać impulsy w liczniku rewersyjnym w zależności od kierunku ruchu przetwornika. W czasie ruchu w prawo na wyjściu WY2 funkтора NOR otrzymuje się stan logiczny "1", a na wyjściu WY1 funkтора NOR szereg impulsów powtarzających się z okresem impulsów jednego kanału przetwornika:

$$WY1 = \bar{A} \cdot \bar{IG} + A \cdot IG + G \cdot \bar{IA} + \bar{G} \cdot IA, \quad (5)$$

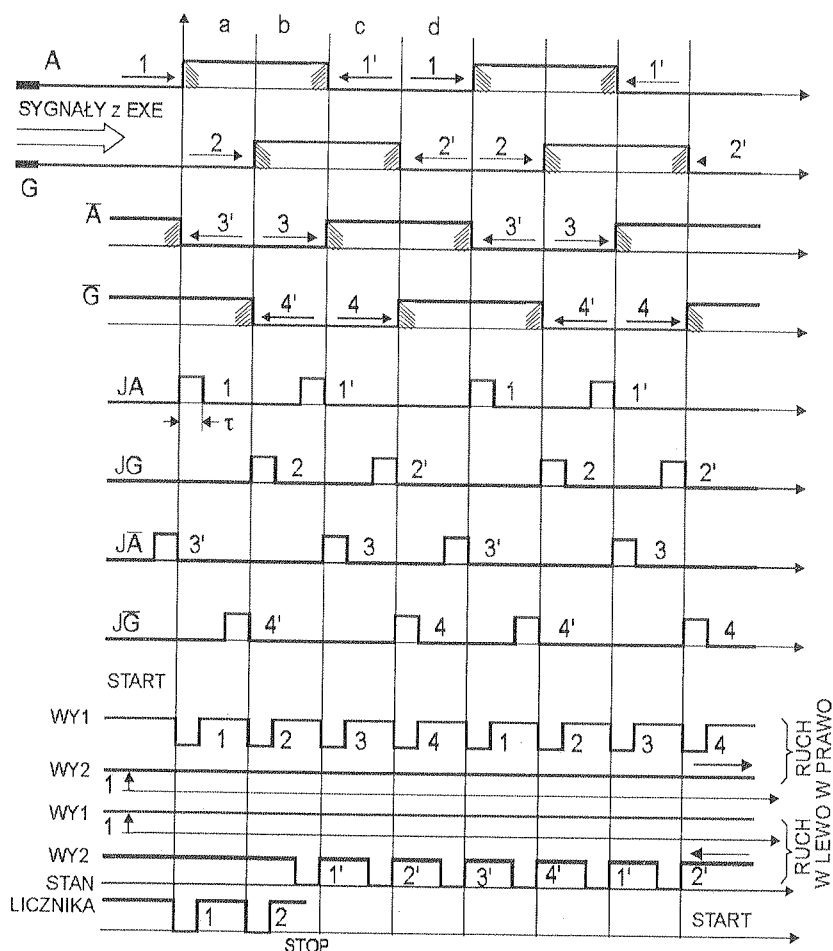
gdzie: IA, IG — impulsy od zbocza sygnału A, G ; $\bar{IA}, \bar{IG}$ — impulsy od zbocza sygnału $\bar{A}, \bar{G}$.

Natomiast w czasie ruchu w kierunku przeciwnym, na wyjściu WY1 funkтора NOR otrzymujemy stan logiczny "1", a na wyjściu WY2 funkтора NOR otrzymujemy szereg impulsów podobnie jak przy ruchu w prawo:

$$WY2 = \bar{A} \cdot \bar{IG} + A \cdot IG + G \cdot \bar{IA} + \bar{G} \cdot IA, \quad (6)$$

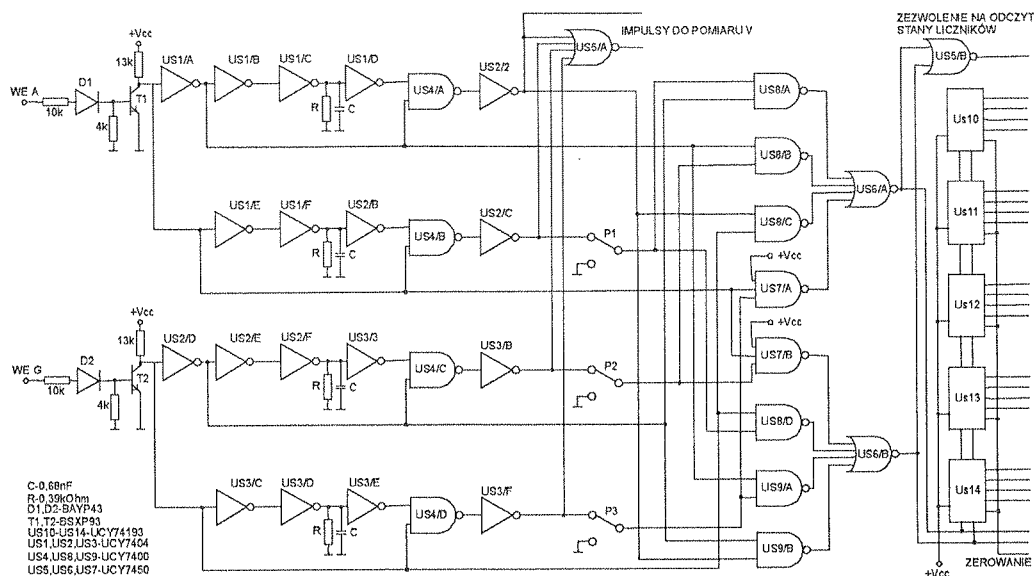
które są odejmowane od stanu licznika rewersyjnego podczas ruchu w prawo.

Zmiana kierunku ruchu może nastąpić w czterech różnych stanach a, b, c i d, co zobrazowane jest na rys.4. W każdym z tych stanów układ musi reagować identycznie, a więc największy błąd wyróżniania kierunku obrotu wynosi $\frac{1}{4}T$ (T — okres impulsów jednego kanału przetwornika). Należy podkreślić, że układ posiada możliwość regulowania ilości zliczanych impulsów przypadających na jeden okres sygnału wejściowego (1, 2, 4) w zależności od nastaw przełączników P1, P2, P3 (rys. 5), czyli istnieje możliwość nastawiania dokładności pozycjonowania mechanizmu 1,2 lub 4 krotnej.



Rys. 4. Metoda rozróżniania kierunku ruchu optoelektronicznego przetwornika zrealizowana w oparciu o funkcje logiczne sygnałów tego przetwornika i impulsów ruchu wygenerowanych w układach RC

Fig. 4. Method of photoelectric transducer motion direction discrimination based on logical functions of the transducer signals and motion pulses generated in the RC circuits



Rys. 5. Układ rozróżniania kierunku ruchu optoelektronicznego przetwornika zrealizowany w oparciu o funkcje logiczne sygnałów tego przetwornika i impulsów ruchu wygenerowanych w układach RC

Fig. 5. System of the photoelectric transducer motion direction discrimination based on logical functions of the transducer signals and motion pulses generated in the RC circuits

4. PODSUMOWANIE

Analiza opracowanego układu prowadzi do następujących uwag i wniosków:

1. Zmniejszenie przedziału kwantowania przemieszczenia napędu współpracującego z przetwornikiem optoelektrycznym można dokonać na drodze precyzyjniejszej budowy przetwornika /jego liniału pomiarowego/ lub na drodze odpowiedniego przetwarzania sygnałów z przetwornika, co technologicznie jest prostsze.
2. Opracowana metoda elektronicznego zwielokrotnienia częstotliwości sygnałów (interpolacja) i przekształcania na postać cyfrową (digitalizacja), umożliwia zwiększenie dokładności fotoelektrycznego przetwornika. Podana metoda 2,5-krotnego zwiększenia częstotliwości jest uniwersalna i można ją wykorzystać do innej skali powielania częstotliwości sygnału z przetwornika.
3. Opracowana metoda zliczania impulsów oraz rozróżniania ich kolejności, a tym samym kierunku ruchu przetwornika, umożliwia czterokrotne zwiększenie jego dokładności.
4. Przedstawione metody interpolacji i digitalizacji oraz zliczania impulsów umożliwiają dziesięciokrotne zwiększenie dokładności przetwarzania przetwornika.

5. BIBLIOGRAFIA

1. Z. Szcześniak, K. Sikora, A. Pizoń, I. Smolewski: *Wykonanie prototypu elektronicznego układu pomiaru wysokości odkuwki dla prasy o nacisku 30 MN w Hucie Warszawa oraz nadzór autorski nad budową i wdrożeniem tego układu*. Opracowanie wykonane w Politechnice Krakowskiej im.T. Kościuszki dla Huty Warszawa. Etap II, 1989r
2. Z. Szcześniak: *Metoda interpolacji sygnałów elektrycznych optoelektronicznych przetworników w pomiarach położenia*. Elektronika nr 1. Warszawa 2005. ss. 74-76.
3. Z. Szcześniak, K. Sikora, A. Pizoń, T. Stefański: *Projekt Badawczy KBN, „Napędy hydrauliczne z zaworami o działaniu ciągłym sterowane mikroprocesorowo*. Opracowanie wykonane w Politechnice Świętokrzyskiej, Kielce Etap II, 1993r.

Z. SZCZEŚNIAK

ELECTRONIC MODULE FOR 10-TIMES INCREMENT OF PROCESSING ACCURACY OF PHOTOELECTRIC TRANSDUCER SIGNALS

Summary

A system for accuracy increment and motion direction discrimination of the photoelectric transducer has been presented in the paper. In the system phase voltages /with the assumed phase shift/ are generated, what makes possible to increase the frequency for 2.5 times in relation to the basic signals resulting from the transducer measurement bar. In the system of pulse counting and sequence discrimination and this way discrimination of the transducer motion direction it is possible to increase the accuracy for 4 times. Presented systems make possible to increase the transducer processing accuracy for ten times.

Keywords: motion direction discrimination increasing of the photoelectric transducer accuracy

ON
CO

r
r
t
c
t
n
r
e

K

Th
sign an
This te
high sp
i.e. a v
connect
packets
of trans
delay v

ON THE ROBUSTNESS OF A NONLINEAR ABR FLOW CONTROLLER FOR MULTI-SOURCE ATM NETWORKS

ANDRZEJ BARTOSZEWICZ

*Institute of Automatic Control, Technical University of Łódź,
18/22 Stefanowskiego St., 90-924 Łódź, Poland
Email: andpbart@p.lodz.pl*

*Otrzymano 2004.05.21
Autoryzowano 2004.10.14*

In this paper the problem of flow control in fast, connection oriented communication networks supporting traffic generated by multiple sources is considered. The network is modelled as a dynamic system with different delay times. A nonlinear strategy governing the behaviour of the sources is proposed. The strategy combines Smith principle with the conventional relay type (i.e. continuous with respect to time and discontinuous with respect to its output signal) controller. The strategy guarantees full bottleneck link utilisation and no cell loss in the controlled network. These favourable properties are obtained even if the round trip times of the connections contributing to the bottleneck link queue are not known exactly, but can only be estimated with some degree of accuracy.

Keywords: Congestion control, ATM networks, time delay systems, Smith principle

1. INTRODUCTION

The asynchronous transfer mode (ATM) technology plays the crucial role in the design and implementation of broadband integrated services digital networks (B-ISDN). This technology is well suited for video, voice and data transmission through the high speed telecommunication networks. The ATM networks are connection oriented, i.e. a virtual circuit (VC) is established between each source and destination for the connection lifetime. After setting up a VC, data is sent in relatively short, fixed size packets, usually called cells. Each data cell is 53 bytes long and consists of 48 bytes of transmitted information and a 5 byte long header. The small fixed cell size reduces delay variation which could be particularly harmful for multimedia traffic.

In order to properly serve diverse needs of different users the ATM Forum defines five service categories. Constant bit rate (CBR) service category provides the bandwidth which is always available to its user. This category is used by real time service. Typical examples are television, telephone, etc. Variable bit rate (VBR) is designed for both real and non-real time applications. An example of such real time application is video conferencing, and multimedia email is an example of non-real time VBR service. Available bit rate (ABR) is a service category whose rate depends on the available bandwidth. Users should adjust their flow rates according to the feedback information received from the network. Electronic mail is an example of this service category. Unspecified bit rate (UBR) category is used to send data on the first in first out (FIFO) basis, using the capacity not consumed by other services. No initial commitment is made to a UBR source and no feedback concerning congestion is provided. This type of service can be used for background file transfer. Guaranteed frame rate (GFR) is a service intended for non-real time applications with little rate requirements. No feedback control protocol is applied in this service category. An example of this service is frame relay interworking.

As stated above ABR is the only service category using feedback information to control source flow rate. Therefore, ABR control is particularly important for congestion avoidance and full resource utilisation. This problem, i.e. the ABR flow control in ATM networks is considered in the paper.

The difficulty of the ABR flow control is mainly caused by long propagation delays in the network. If congestion occurs at a specific node, information about this condition must be conveyed to all the sources transmitting data cells through the node. Transferring this information involves feedback propagation delays. After this information has been received by the particular source, it can be used to adjust the flow rate of this source. However, the adjusted flow rate will start to affect the congested node only after forward propagation delay.

ABR flow rate control has recently been studied in several papers [1-4, 6-16]. A valuable survey of earlier congestion control mechanisms is given in [9]. Furthermore, Izmailov [7] considered a single connection controlled by a linear regulator whose output signal is generated according to the several states of the buffer measured at different time instants. Asymptotic stability, nonoscillatory system behaviour and locally optimal rate of convergence have been proved. Chong et. al. proposed and thoroughly studied the performance of a simple queue length based flow control algorithm with dynamic queue threshold adjustment [3]. Lengliz and Kamoun [12] introduced a proportional plus derivative controller which is computationally efficient and can be easily implemented in ATM networks. Imer et. al. [6] gave a brief, excellent tutorial exposition of the ABR control problem and presented new stochastic and deterministic control algorithms. Another interesting approach to the problem of flow rate control in communication networks has been proposed by Quet et al. In the recent paper [16] the authors considered a single bottleneck multi-source ATM network and applied minimisation of an H-infinity norm to the design of a flow rate controller. The proposed

contro
delays
been
time
ABR
[8] sh
loop s

D
perform
commu
single
He use
functio
was ap
standar
net Pro
In the
support
control
expon
oversho
principl
which I
dwidth.
On the
the ABR
principl

In t
is simila
we prop
the con
with res
allocatio
in the n
maximu
algorithm
of the pr
controlle
time, the
estimates

The
used thro

controller guarantees stability robustness to uncertain and time-varying propagation delays in various channels. Adaptive control strategies for ABR flow regulation have been proposed by Laberteaux and Rohrs [11]. Their strategies reduce convergence time and improve queue length management. Also a neural network controller for ABR service in ATM networks has recently been proposed. Jagannathan and Talluri [8] showed that their neural network controller can guarantee stability of the closed loop system and the desired quality of service (QoS).

Due to the significant propagation delays which are critical for the closed loop performance, several researchers applied the Smith principle to control ABR flow in communication networks [1, 2, 4, 13–15]. In the paper [14] Mascolo considered the single connection congestion control problem in a general packet switching network. He used the deterministic fluid model approximation of packet flow and applied transfer functions to describe the network dynamics. The designed continuous time controller was applied to the ABR traffic control in ATM network and compared with ERICA standard. Furthermore, Mascolo showed that Transmission Control Protocol / Internet Protocol (TCP/IP) implements a Smith predictor to control network congestion. In the next paper [15] the same author applied the Smith principle to the network supporting multiple ABR connections with different propagation delays. The proposed control algorithm guarantees no cell loss, full and fair network utilisation, and ensures exponential convergence of queue levels to stationary values without oscillations or overshoots. Gomez-Stern et. al. further studied the ABR flow control using Smith principle [4]. They proposed a continuous time proportional-integral (PI) controller which helps reduce the average queue level and its sensitivity to the available bandwidth. Saturation issues in the system were handled using anti-wind up techniques. On the other hand, in the papers [1, 2] linear, discrete-time flow control strategies for the ABR service in ATM networks have been proposed. The strategies combine Smith principle with the discrete time proportional controllers.

In this paper the ABR flow control in ATM networks is considered. Our approach is similar to that introduced in [1, 2, 4, 13 – 15], however as opposed to those papers we propose a nonlinear control strategy. The strategy combines Smith principle with the conventional relay type (i.e. continuous with respect to time and discontinuous with respect to its output signal) controller. The strategy guarantees equal resource allocation between various users, full bottleneck node link utilisation and no cell loss in the network. As a result, the need of cell retransmission is eliminated and the maximum throughput is achieved. Furthermore, transmission rates generated by the algorithm are nonnegative and limited. These properties allow direct implementation of the proposed strategy in the network environment. Since the Smith principle based controllers are often sensitive with respect to the inaccurate estimation of the delay time, the effect of the possible mismatch between the actual round trip times and their estimates is thoroughly studied in the paper.

The remainder of this paper is organised as follows. The model of the network used throughout the paper is introduced in section 2. Then the nonlinear ABR flow

control algorithm and the properties of the system under the assumption that the round trip times of all connections are known precisely are described in section 3. Section 4 is devoted to the analysis of the system robustness. In that section, it is shown how the desired properties of the proposed control system are preserved in the presence of the mismatch between the actual round trip times and their estimates used by the controller. Section 5 presents a simulation example, and finally Section 6 comprises conclusions of the paper.

2. NETWORK MODEL

The network considered in this paper consists of data sources, nodes and destinations — all of them interconnected via bi-directional links. Each node of the network (i.e. switch) maintains one queue per output port. When a new data cell arrives at an input port of the node, it is directed to the appropriate output buffer, stored and forwarded to the next node on the first in first out (FIFO) basis. Similarly as in the papers [2, 3, 4, 8, 10, 11, 12, 15, 16] the case of a single bottleneck link shared by n sources is considered. Our purpose is to design the ABR flow controller, to be implemented at the node, which will assure full bottleneck link utilisation and no buffer overflow.

The rate of cell outflow from the bottleneck buffer depends on the available bandwidth modelled as an a priori unknown, bounded function of time $d(t)$ where

$$0 \leq d(t) \leq d_{max} \quad (1)$$

This is motivated by the fact that the ABR service dynamically uses the bandwidth temporarily left not consumed by rt-VBR and nrt-VBR which typically support unpredictable traffic.

The sources send data cells (at the rate determined by the controller) and resource management (RM) cells. The RM cells are processed by the nodes on the priority basis, i.e. they are not queued but sent to the next node without delay. These cells carry information about the network conditions. After reaching the destination they are immediately sent back to the source, along the same path they arrived. The information carried by the RM cells is used to adjust the source rates.

Further in this paper t denotes time and RTT_j represents the round trip time of the j -th ($j = 1, 2, \dots, n$) virtual circuit contributing to the bottleneck queue under control. This time is equal to the sum of forward and backward propagation delays denoted as T_{fj} and T_{bj} respectively

$$RTT_j = T_{fj} + T_{bj} \quad (2)$$

Furthermore, $x(t)$ denotes the bottleneck queue length at time t , and x_d the demand value of $x(t)$. The virtual connections are numbered in such a way that

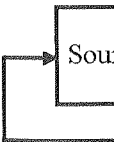
$$RTT_1 \leq RTT_2 \leq \dots \leq RTT_{n-1} \leq RTT_n \quad (3)$$

Before se

On the ot

where a_j
is actual
 $j = 1, 2,$
 $x(t)$ is gr
Otherwis
node. In
any time

The
in Figure
The
and equa
Thus the



Before setting up the connection, the bottleneck buffer is empty, i.e.

$$x(t < 0) = 0 \quad (4)$$

On the other hand for $t \geq 0$

$$x(t) = \sum_{j=1}^n \int_0^t a_j(\tau - T_{fj}) d\tau - \int_0^t h(\tau) d\tau \quad (5)$$

where $a_j(t)$ is the j -th source rate at time t , and $h(t)$ represents the bandwidth which is actually consumed by the bottleneck link at time t . It is assumed that for any $j = 1, 2, \dots, n$ $a_j(t < 0) = 0$. Furthermore, if the queue length at the bottleneck link $x(t)$ is greater than zero, then the entire available bandwidth is consumed $h(t) = d(t)$. Otherwise, i.e. when $x(t) = 0$, then $h(t)$ is determined by the rate of data arrival at the node. In this case the available bandwidth may not be fully utilised. Consequently, for any time t

$$0 \leq h(t) \leq d(t) \leq d_{\max} \quad (6)$$

The block diagram of the flow control system considered in this paper is shown in Figure 1.

The total source rate $\tilde{a}$ is generated by the controller implemented at the switch, and equally allocated between all the sources contributing to the bottleneck queue. Thus the j -th source rate

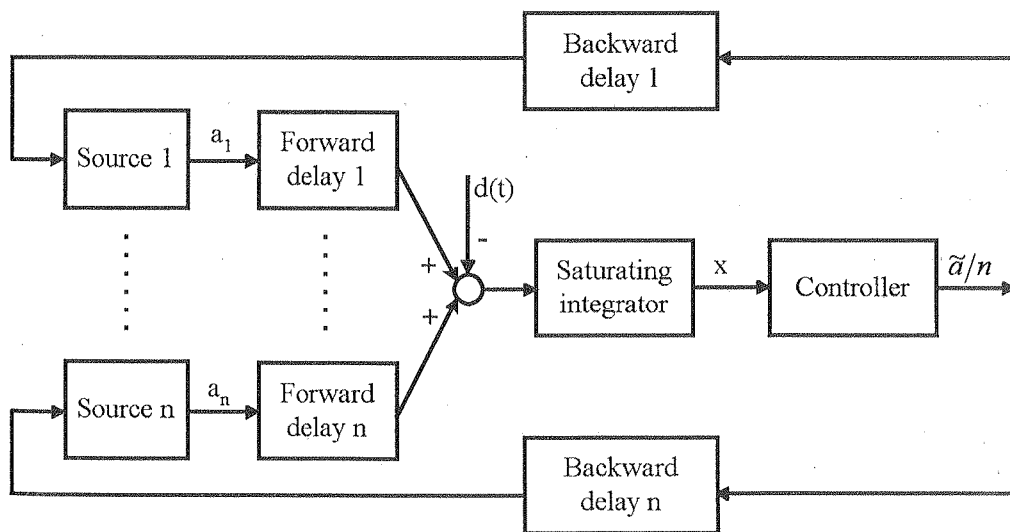


Fig. 1. Control system

Rys. 1. Układ sterowania

$$a_j(t) = \frac{1}{n} \tilde{a}(t - T_{bj}) \quad (7)$$

The source rate should be determined in such a way that the bottleneck buffer overflow is avoided and the node has always enough data to send. The first condition implies that data cells are not lost and there is no need for their retransmission, while the latter one assures full bottleneck link utilisation which is highly desirable for economic reasons. In the next section a nonlinear controller which ensures that the two conditions are satisfied is proposed.

3. PROPOSED CONTROLLER

In this section we propose the following control algorithm. The j -th source rate is determined by equation (7). For any time $t \geq 0$, the total source rate signal $\tilde{a}(t)$ in this equation is generated by the controller (placed at the bottleneck node) according to the following formula

$$\tilde{a}(t) = \frac{1}{2}a_{\max} + \frac{1}{2}a_{\max} \operatorname{sgn} \left[x_d - x(t) - \sum_{j=1}^n \frac{1}{n} \int_{t-RTT_j}^t \tilde{a}(\tau) d\tau \right] \quad (8)$$

where $a_{\max}$ is a positive constant greater than $d_{\max}$, and the $\operatorname{sgn}(\cdot)$ function denotes the sign of its argument. It is assumed that for any time $t < 0$, the total source rate signal $\tilde{a}(t) = 0$. Notice that when the proposed strategy is applied the j -th source does not send any data for the time smaller than T_{bj} , i.e.

$$\forall j \forall t < T_{bj}, a_j(t) = 0 \quad (9)$$

Consequently, for any time smaller than or equal to RTT_1 the queue length $x(t) = 0$.

The strategy proposed in this section combines a generalised continuous time Smith predictor with the nonlinear on-off controller. In the sequel two theorems presenting important properties of the proposed control strategy are introduced.

Theorem 1. If the proposed strategy is applied, then the bottleneck link queue length is always upper bounded by its demand value, i.e.

$$x(t) \leq x_d \quad (10)$$

Proof: As it has already been mentioned, for any time smaller than or equal to RTT_1 the queue length $x(t) = 0$. Therefore, in order to prove the theorem, it is necessary to show that the queue length will not exceed its demand value x_d at any time t greater than RTT_1 .

Let

$$\varphi(t) = x(t) + \sum_{j=1}^n \frac{1}{n} \int_{t-RTT_j}^t \tilde{a}(\tau) d\tau \quad (11)$$

(7)

ck buffer
condition
while the
conomic
conditions

This function represents the sum of three terms:
i) the number of cells currently waiting in the bottleneck buffer;
ii) the number of 'in flight' cells, i.e. those which have already been sent by the sources but not arrived at the bottleneck node yet;
iii) and the number of those cells which will inevitably be sent by the sources (this is because the controller already sent out the command to do so to all the sources).
It follows from equations (5) and (7), that for any $t \geq 0$ the queue length

$$\begin{aligned} x(t) &= \sum_{j=1}^n \int_0^t \frac{1}{n} \tilde{a}(\tau - RTT_j) d\tau - \int_0^t h(\tau) d\tau = \\ &= \sum_{j=1}^n \int_0^{t-RTT_j} \frac{1}{n} \tilde{a}(\tau) d\tau - \int_0^t h(\tau) d\tau \end{aligned} \quad (12)$$

source rate
al $\tilde{a}(t)$ in
according

Consequently, the function $\varphi(t)$ can be expressed as

(8)

$$\varphi(t) = \sum_{j=1}^n \int_0^t \frac{1}{n} \tilde{a}(\tau) d\tau - \int_0^t h(\tau) d\tau = \int_0^t \tilde{a}(\tau) d\tau - \int_0^t h(\tau) d\tau \quad (13)$$

notes the
te signal
does not

Taking into account relation (6) it can be easily concluded that this function increases only if $a(t) = a_{\max}$. Together with equation (8) this implies that

$$\forall t \geq 0, \quad \varphi(t) \leq x_d \quad (14)$$

(9)

$\varphi(t) = 0$.
ne Smith
representing

On the other hand, since $\tilde{a}(t)$ is always non-negative, one concludes that

$$x(t) = \varphi(t) - \sum_{j=1}^n \frac{1}{n} \int_{t-RTT_j}^t \tilde{a}(\tau) d\tau \leq \varphi(t) \leq x_d \quad (15)$$

ue length

This conclusion ends the proof.

(10)

equal to
necessary
 t greater

Another desired property of a properly designed flow control system is full link utilisation. If the queue length is greater than zero, then the link bandwidth is fully used. The next theorem shows how the buffer capacity should be chosen in order to ensure the strictly positive queue length and as a consequence full bottleneck link bandwidth utilisation.

Theorem 2. If the bottleneck link buffer capacity is greater than or equal to the demand value of the queue length x_d and the following inequality holds

(11)

$$x_d > \left(\frac{1}{n} \sum_{j=1}^n RTT_j \right) a_{\max} \quad (16)$$

then for any $t > RTT_n$ the queue length is greater than zero.

Proof: At the initial time $t = 0$ the function $\varphi(t)$, defined by equation (11), equals zero. Then (since by the definition $a_{\max}$ is greater than $d_{\max}$) the function $\varphi(t)$ increases until its value becomes x_d . At the time instant $t = RTT_n$ either $\varphi(RTT_n) = x_d$ or $\varphi(RTT_n)$ is smaller than x_d and increasing. In the first case $\varphi(t) = x_d$ for any time $t \geq RTT_n$. Thus, taking into account relations (11) and (16), one can directly conclude that for any $t \geq RTT_n$

$$x(t) = \varphi(t) - \sum_{j=1}^n \frac{1}{n} \int_{t-RTT_j}^t \tilde{a}(\tau) d\tau \geq x_d - \left(\sum_{j=1}^n \frac{1}{n} RTT_j \right) a_{\max} > 0 \quad (17)$$

On the other hand, the latter case implies that for any time $t \leq RTT_n$ $\tilde{a}(t) = a_{\max}$. Consequently, the number of cells accumulated in the network (i.e. these cells which have actually been sent and those which will inevitably be sent by the sources because an appropriate control signal has already been generated) at the time instant $t = RTT_n$ can be expressed as

$$\begin{aligned} \lambda(RTT_n) = a_{\max} & \left[\frac{1}{n} (RTT_n - RTT_{n-1}) + \frac{2}{n} (RTT_{n-1} - RTT_{n-2}) + \dots + \right. \\ & \left. + \frac{n-1}{n} (RTT_2 - RTT_1) + RTT_1 \right] = a_{\max} \left(\frac{1}{n} \sum_{j=1}^n RTT_j \right) \end{aligned} \quad (18)$$

and

$$\begin{aligned} \varphi(RTT_n) &= \lambda(RTT_n) + x(RTT_n) = \\ &= \left(\frac{1}{n} \sum_{j=1}^n RTT_j \right) a_{\max} + x(RTT_n) \geq \left(\frac{1}{n} \sum_{j=1}^n RTT_j \right) a_{\max} \end{aligned} \quad (19)$$

As it has already been mentioned, in this case $\varphi(t)$ is increasing at $t = RTT_n$ and non-decreasing for any $t > RTT_n$. Consequently, for any time t greater than RTT_n , $\varphi(t)$ is strictly greater than $\left(\frac{1}{n} \sum_{j=1}^n RTT_j \right) a_{\max}$. Thus, it follows from equation (11) that for any $t > RTT_n$

$$x(t) = \varphi(t) - \sum_{j=1}^n \frac{1}{n} \int_{t-RTT_j}^t \tilde{a}(\tau) d\tau \geq \varphi(t) - \left(\sum_{j=1}^n \frac{1}{n} RTT_j \right) a_{\max} > 0 \quad (20)$$

Finally, taking into account inequalities (17) and (20) one concludes that for any time t greater than RTT_n the queue length is strictly positive. This ends the proof of theorem 2.

Theorem 2 shows that using the strategy proposed in this paper one can always assure full link utilisation, provided the bottleneck node buffer capacity is greater than

the fol
buffer c

Ho
any tim
setting

The
principle
usually
in this
 RTT_j a
for exam
only on
slightly
In c
written

where t
($j = 1$,
actual A
their es
hold an
formula

Theore
always

where

the following product $\left(\frac{1}{n} \sum_{j=1}^n RTT_j\right) a_{\max}$. In fact this property can be achieved if the buffer capacity satisfies a weaker condition, i.e.

$$x_d > \left(\frac{1}{n} \sum_{j=1}^n RTT_j\right) d_{\max} \quad (21)$$

However, in this case the queue length is not guaranteed to be strictly positive for any time t greater than RTT_n , but only after the elapse of some longer time since setting up the controlled connections.

4. ROBUSTNESS ISSUES

The control strategy presented in the previous section effectively exploits the Smith principle. However, it is well known that the Smith principle based controllers are usually sensitive with respect to the inaccurate estimation of the delay time. Therefore, in this section the effect of the possible mismatch between the actual round trip times RTT_j and their estimates, further denoted as $\overline{RTT}_j$, will be studied. The mismatch may for example occur if two or more RM cells arrive at a node at the same time. Then only one of them can be forwarded immediately and the round trip times of the others slightly increase.

In order to analyse the effect of imperfect RTT_j evaluation, formula (8) can be rewritten as follows

$$\tilde{a}(t) = \frac{1}{2} a_{\max} + \frac{1}{2} a_{\max} \operatorname{sgn} \left[x_d - x(t) - \sum_{j=1}^n \frac{1}{n} \int_{t-\overline{RTT}_j}^t \tilde{a}(\tau) d\tau \right] \quad (22)$$

where the parameter $\overline{RTT}_j$ determined while setting up the j -th virtual connection ($j = 1, 2, \dots, n$) and used by the controller is an estimate of the (exactly) unknown actual RTT_j time. The possible discrepancies between the actual round trip times and their estimates cause that the two theorems proved in the previous section do not hold anymore. However, the properties of the proposed control strategy can now be formulated as stated in the next two theorems.

Theorem 3. If the strategy (22) is applied, then the bottleneck link queue length is always limited

$$x(t) \leq x_d + \Delta_{\max} \quad (23)$$

where

$$\Delta_{\max} = \frac{a_{\max}}{n} \sum_{j: RTT_j > \overline{RTT}_j} (RTT_j - \overline{RTT}_j) \quad (24)$$

Proof: Let

$$\bar{\varphi}(t) = x(t) + \sum_{j=1}^n \frac{1}{n} \int_{t-\overline{RTT}_j}^t \tilde{a}(\tau) d\tau \quad (25)$$

This function is an estimate of $\varphi(t)$ and can be calculated by the controller on-line. The relation between the $\varphi(t)$ function and its estimate $\bar{\varphi}(t)$ is given by

$$\varphi(t) = \bar{\varphi}(t) + \sum_{j=1}^n \int_{t-\overline{RTT}_j}^{t-\overline{RTT}_j} \frac{1}{n} \tilde{a}(\tau) d\tau = \bar{\varphi}(t) + \Delta(t) \quad (26)$$

where

$$\Delta(t) = \sum_{j=1}^n \int_{t-\overline{RTT}_j}^{t-\overline{RTT}_j} \frac{1}{n} \tilde{a}(\tau) d\tau \quad (27)$$

It can be easily noticed that $\Delta(t)$ is bounded by the following two inequalities

$$\Delta(t) \geq \frac{a_{\max}}{n} \sum_{j: RTT_j < \overline{RTT}_j} (RTT_j - \overline{RTT}_j) = -\Delta_{\min} \quad (28)$$

and

$$\Delta(t) \leq \frac{a_{\max}}{n} \sum_{j: RTT_j > \overline{RTT}_j} (RTT_j - \overline{RTT}_j) = \Delta_{\max} \quad (29)$$

Both $\Delta_{\max}$ and $\Delta_{\min}$ in relations (28) and (29) are nonnegative real numbers. With this notation one gets

$$\bar{\varphi}(t) - \Delta_{\min} \leq \varphi(t) \leq \bar{\varphi}(t) + \Delta_{\max} \quad (30)$$

On the other hand, it follows from relations (6) and (13) that this function increases if and only if $\bar{\varphi}(t) < x_d$. Since at the initial time $\varphi(0) = 0$, then inequality (30) implies that $\varphi(t)$ cannot become greater than $x_d + \Delta_{\max}$. Consequently, taking into account equation (11) one can easily notice that the queue length $x(t)$ never exceeds $x_d + \Delta_{\max}$. This conclusion ends the proof.

Theorem 3 shows that if the bottleneck buffer capacity is selected appropriately, then data cells do not get lost even when the actual round trip times of various connections and their estimates do not match exactly. In the sequel it will be shown that under these circumstances the other desirable property of the proposed strategy, i.e. full bottleneck link utilisation can also be preserved.

Theorem 4. If the bottleneck buffer capacity is greater than or equal to the reference value x_d and

$$x_d > \left(\frac{1}{n} \sum_{j=1}^n RTT_j \right) a_{\max} + \Delta_{\min} \quad (31)$$

then for any time $t > RTT_n$ the queue length $x(t)$ is strictly positive.

Proof: Inequality (31) is equivalent to

(25)

$$x_d = \left(\frac{1}{n} \sum_{j=1}^n RTT_j \right) a_{\max} + \Delta_{\min} + \delta \quad (32)$$

line.

where δ is a positive constant. On the other hand it follows from equation (26) that

(26)

$$\bar{\varphi}(t) = \varphi(t) - \Delta(t) \quad (33)$$

Thus if

$$\varphi(t) < x_d - \Delta_{\min} \quad (34)$$

then

(27)

$$\bar{\varphi}(t) < x_d - \Delta_{\min} - \Delta(t) \quad (35)$$

and it follows from relations (28) and (35) that if inequality (34) is satisfied, then

(28)

$$\bar{\varphi}(t) < x_d \quad (36)$$

Therefore, if $\varphi(t) < x_d - \Delta_{\min}$, which taking into account assumption (31) is equivalent to $\varphi(t) < \delta + a_{\max} \sum_{j=1}^n RTT_j/n$, then $\varphi(t)$ increases. This conclusion follows directly

(29)

from the fact that $a_{\max}$ is greater than $d_{\max}$. Furthermore, it can be noticed that:

with this

- for any $t < RTT_1$, $h(t) = 0$,
- for any $t \in [RTT_1, RTT_2)$, $h(t) \leq a_{\max}/n$,
- for any $t \in [RTT_2, RTT_3)$, $h(t) \leq 2a_{\max}/n$,
- ,
- for any $t \in [RTT_{n-1}, RTT_n)$, $h(t) \leq (n-1)a_{\max}/n$.

(30)

ases if
implies
account
+ $\Delta_{\max}$.

Consequently, it follows from relation (13) that if $\varphi(t) < x_d - \Delta_{\min}$, i.e. if inequality

$$\varphi(t) < \delta + a_{\max} \sum_{j=1}^n RTT_j/n \text{ is satisfied, then:}$$

ropriately,
ous con-
own that
egy, i.e.

- for any $t < RTT_1$, $\varphi(t)$ increases at the rate $a_{\max}$,
- for any $t \in [RTT_1, RTT_2)$, $\varphi(t)$ increases at least at the rate $a_{\max}(n-1)/n$,
- for any $t \in [RTT_2, RTT_3)$, $\varphi(t)$ increases at least at the rate $a_{\max}(n-2)/n$,
- ,
- for any $t \in [RTT_{n-1}, RTT_n)$, $\varphi(t)$ increases at least at the rate $a_{\max}/n$.

reference

Furthermore, if for any time $t > RTT_n$, function $\varphi(t)$ is smaller than $x_d - \Delta_{\min}$, then its value increases at least at the rate equal to the difference between $a_{\max}$ and $d_{\max}$.

Consequently, if $\varphi(t) < x_d - \Delta_{\min}$, i.e. if $\varphi(t) < \delta + a_{\max} \sum_{j=1}^n RTT_j/n$, then

(31)

$$\varphi(t) \geq f(t) \quad (37)$$

Finally, relations (41) and (42) imply

$$\forall t > RTT_n, \quad x(t) > 0 \quad (43)$$

This conclusion ends the proof.

Theorem 4 shows that the full bottleneck link utilisation can actually be preserved in the presence of imperfect knowledge of the round trip times in the controlled virtual connections.

5. SIMULATION EXAMPLE

In order to present the properties of the control strategy proposed in this paper computer simulations of a wide area ATM network with $n = 3$ sources were performed. The connections of the network are characterised by the following round trip times: $RTT_1 = 20$ ms, $RTT_2 = 30$ ms, and $RTT_3 = 70$ ms. The estimates of these round trip times available for the controller are: $\overline{RTT}_1 = 22$ ms, $\overline{RTT}_2 = 34$ ms, and $\overline{RTT}_3 = 67$ ms. The maximum bandwidth $d_{max} = 4$ Mb/s, the source sending rate $a_{max} = 10000$ cells/s $\approx 1.01 \cdot d_{max}$ and the bandwidth actually available for the controlled connections at the bottleneck node is shown in figure 2. In order to assure full link utilisation, the demand value of the queue length must be greater than 420 cells. To satisfy this condition $x_d = 425$ cells has been chosen. Furthermore, in order to avoid cell loss the bottleneck buffer capacity should be greater than or equal to 435 cells. The queue length in the buffer of the bottleneck link is presented in figure 3. It can be clearly seen from this figure that the length never exceeds the bottleneck buffer capacity. This implies no cell loss in the bottleneck node and no need for data retransmission. Furthermore, for any time greater than $T_f = 70$ ms the queue length is strictly positive, which leads to the conclusion that the bottleneck link bandwidth is fully utilised.

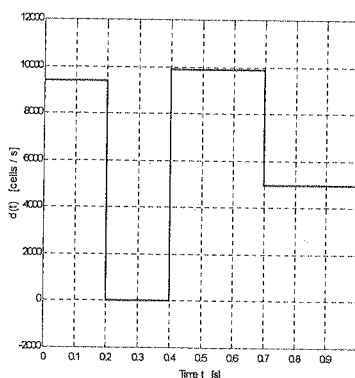


Fig. 2. Available bandwidth

Rys. 2. Dostępne pasmo

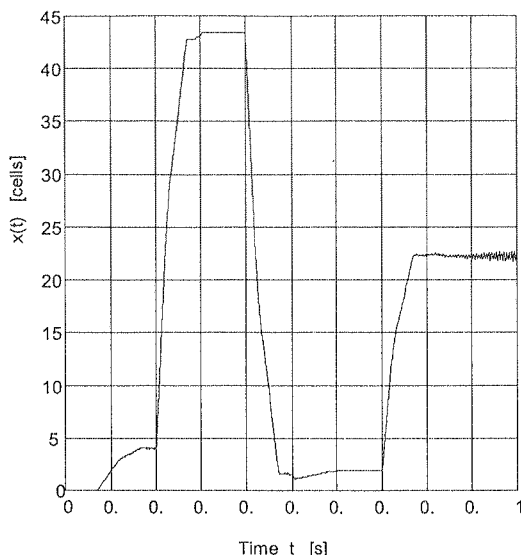


Fig. 3. Queue length

Rys. 3. Długość kolejki danych

6. CONCLUSIONS

In this paper a new nonlinear flow control strategy for multi-source, single-bottleneck ATM networks has been proposed. The strategy effectively combines the Smith predictor with the conventional on-off controller. Since the Smith predictors are known to be sensitive with respect to imperfect delay estimation, the properties of the proposed strategy in the presence of the mismatch between the actual round trip times and their estimates available for the controller are analysed. Furthermore, sufficient conditions for full bottleneck link utilisation and no cell loss in the controlled virtual circuits are derived.

7. ACKNOWLEDGEMENTS

This work has been supported by the Polish State Committee for Scientific Research (KBN) grant no. 3T11C 043 26.

8. REFERENCES

1. A. Bartoszewicz, M. Karbowanczyk: *Sampled Time Flow Control Algorithm For Fast Connection Oriented Communication Networks*. Journal of Applied Computer Science, 2003, Vol. 11, No. 1, pp. 5 - 16.
2. A. Bartoszewicz, T. Molik: *ABR traffic control over multi-source single-bottleneck ATM networks* Journal of Applied Mathematics and Computer Science. 2004, Vol. 14, No. 1, pp. 43-51.

- TOM 51
3. S. .
mic
1999
 4. F. .
traff
 5. A. .
Wro
 6. O. .
cont
 7. R. .
twor
 8. S. .
sour
 9. R. .
surv
 10. L. .
IEEE
 11. K. .
expl
pp. 9
 12. I. L. .
Com
 13. S. M. .
ed .
Dieg
 14. S. M. .
Auto
 15. S. M. .
action
 16. P. F. .
Rate-
multi
 17. The
 18. The
ftp://

AN.

W ar
wych sieci
Zapropono
artykule r

3. S. Chong, R. Nagarajan, Y. T. Wang: *First-order rate-based flow control with dynamic queue threshold for high-speed wide-area ATM networks*. Computer Networks and ISDN Systems, 1998, Vol. 29, No. 17-18, pp. 2201-2212.
4. F. Gómez-Stern, J. M. Fornés, F. R. Rubio: *Dead-time compensation for ABR traffic control over ATM networks*. Control Engineering Practice, 2002, Vol. 10, No. 5, pp. 481-491.
5. A. Grzech: *Sterowanie ruchem w sieciach teleinformatycznych*. Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 2002.
6. O. C. Imer, S. Compans, T. Basar, R. Srikant: *Available bit rate congestion control in ATM networks*. IEEE Control Systems Magazine, 2001, pp. 38-56.
7. R. Izmailov: *Adaptive feedback control algorithms for large data transfers in high-speed networks*. IEEE Transactions on Automatic Control, 1995, Vol. 40, No. 8, pp. 1469-1471.
8. S. Jagannathan, J. Talluri: *Predictive congestion control of ATM networks: multiple sources/single buffer scenario*. Automatica, 2002, Vol. 38, No. 5, pp. 815-820.
9. R. Jain: *Congestion control and traffic management in ATM networks: recent advances and a survey*. Computer Networks and ISDN Systems, 1996, Vol. 28, No. 13, pp. 1723-1738.
10. L. A. Kulkarni, S. Li: *Performance analysis of a rate-based feedback control scheme*. IEEE/ACM Transaction on Networking, 1998, Vol. 6, No. 6, pp. 797-810.
11. K. P. Laberteaux, Ch. E. Rohrs, P. J. Antsaklis: *A practical controller for explicit rate congestion control*. IEEE Transactions on Automatic Control, 2002, Vol. 47, No. 6, pp. 960-978.
12. I. Lengli, F. Kamoun: *A rate-based flow control method for ABR service in ATM networks*. Computer Networks, 2000, Vol. 34, No. 1, pp. 129-138.
13. S. Mascolo: *Smith's principle for congestion control in high speed ATM networks*. Proceedings of the 36th IEEE Conference on Decision and Control, San Diego, 1997, pp. 4595-4600.
14. S. Mascolo: *Congestion control in high-speed communication networks using the Smith principle*. Automatica, 1999, Vol. 35, No. 12, pp. 1921-1935.
15. S. Mascolo: *Smith's principle for congestion control in high-speed data networks*. IEEE Transactions on Automatic Control, 2000, Vol. 45, No. 2, pp. 358-364.
16. P. F. Quet, B. Ataslar, A. Iftar, H. Özbay, S. Kalyanaraman, T. Kang: *Rate-based flow controllers for communication networks in the presence of uncertain time-varying multiple time-delays*. Automatica, 2002, Vol. 38, No. 6, pp. 917-928.
17. The ATM Forum: *Beginner's Overview*. available at <http://www.atmforum.com/aboutatm/guide.html>.
18. The ATM Forum: *Traffic Specification*. ver. 4.1, available at <ftp://ftp.atmforum.com/pub/approved-specs/af-tm-0121.000.pdf>.

A. BARTOSZEWICZ

ANALIZA ODPORNOŚCI NIELINIOWEGO REGULATORA PRĘDKOŚCI PRZEPŁYWU DANYCH W SIECIACH ATM O WIELU ŹRÓDŁACH

Streszczenie

W artykule zaproponowano nową strategię sterowania prędkością nadawania danych w połączeniowych sieciach teleinformatycznych, w których występuje wiele źródeł i pojedynczy węzeł wąskiego gardła. Zaproponowana strategia wykorzystuje elementy zaczerpnięte z klasycznej teorii sterowania. Opisany w artykule regulator stanowi połączenie predyktora Smitha oraz nieliniowego elementu dwupołożeniowego.

go. Zastosowanie przedstawionego regulatora zapewnia uzyskanie niezerowej długości kolejki danych w buforze wąskiego gardła, a równocześnie gwarantuje, że długość tej kolejki nie przekroczy zadanej wartości. Dzięki temu możliwe jest pełne wykorzystanie łącza przy równoczesnym wyeliminowaniu niebezpieczeństwa gubienia przesyłanych pakietów. Pozwala to zapobiec konieczności retransmisji danych i zagwarantować stuprocentowe wykorzystanie zasobów sieci.

Ponieważ układy sterowania wykorzystujące predyktor Smitha często są wrażliwe na niedokładność oszacowania czasu opóźnienia, w artykule dokładnie przeanalizowano wpływ nieprecyzyjnego określenia czasu RTT pełnego obiegu komórki RM na działanie zaproponowanego układu regulacji. Pokazano jakie warunki muszą być spełnione aby zapewnić pełne wykorzystanie dostępnej przepustowości i nie dopuścić do przepełnienia bufora wąskiego gardła nawet wtedy, gdy czasy RTT nie są znane dokładnie, a jedynie mogą być oszacowane z pewnym przybliżeniem. Sformułowane twierdzenia zostały zilustrowane przykładem symulacyjnym.

Słowa kluczowe: sterowanie przepływem danych, sieci ATM, układy z opóźnieniem, predyktor Smitha

Acti

pro
ari
lin
the
is l
ran
an
5th

Ke

CM
ny analo
filters an
converte
One of t
rential p
makes it
This tech

ki danych
ry zadanej
waniu nie-
i danych i

okładność
określenia
zako jakie
nie dopu-
okładnie, a
lustrowane

Smitha

Active-error feedforward technique for linearization of cmos transconductance amplifier

STANISŁAW SZCZEPAŃSKI, SŁAWOMIR KOZIEL

*Faculty of Electronics, Telecommunications and Informatics,
Gdańsk University of Technology, 80-952 Gdańsk, Poland
e-mail: stanisla@pg.gda.pl, koziel@ue.eti.pg.gda.pl*

Otrzymano 2003.07.21

Autoryzowano 2004.10.24

In the paper, a novel highly linear operational transconductance amplifier (OTA) is proposed. The circuit is developed using simple differential pair transconductors and linearization technique based on an active-error feedforward concept. As a result, improved linearity of the single-input two-output OTA is obtained. SPICE simulations show that for the circuit working with a $\pm 1.25\text{V}$ power supply, total harmonic distortion (THD) at 0.6V_{pp} is less than 0.38% in comparison to 10.0% without linearization. Moreover, the input voltage range of linear operation is increased. Power consumption of the overall circuit is 0.3mW. As an example, the linearized OTA is used to design a highly linear G_m -C filter that implements 5th order Butterworth low-pass transfer function.

Keywords: operational transconductance amplifiers, feedforward linearization, CMOS feedforward amplifiers

1. INTRODUCTION

CMOS transconductance elements are useful building blocks for the design of many analog and analog-digital signal processing systems. Such applications (e.g. active filters and transconductance multipliers) usually require very linear voltage-to-current converters (transconductors) or operational transconductance amplifiers (OTAs) [1]-[3]. One of the simplest and most widely used transconductors is a source-coupled differential pair [4]. It is commonly known that application of the feedforward technique makes it possible to improve performance of many analog signal processing circuits. This technique is widely used to reduce nonlinear distortion in amplifiers [5]-[12]. Mo-

reover, it is successfully employed to frequency compensation of operational amplifiers and OTAs [13]-[16].

In this work, a novel highly linear operational transconductance amplifier (OTA) is proposed. The circuit uses, as basic building blocks, simple differential pair transconductors. The linearization follows by employing active-error feedforward scheme. The error signal is generated using an additional differential pair transconductor and a resistor which is assumed to be linear. This resistor can be an external one or it can be implemented as an monolithic element e.g. using a high resistive poly technology. Moreover, this resistor can be tuned digitally using an appropriate matrix of switches and reference resistors.

The paper is organized as follows. In Section 2, a representation of nonlinear transfer characteristic of transconductors using power series expansion is described. Section 3 gives a description of the proposed linearization method. In Section 4 the circuit realization of the linearized differential pair OTA is presented. In Section 5 the SPICE simulation results of the OTA are discussed, including THD analysis. Simulation results of the 5th order low-pass Butterworth Gm-C filter using linearized differential pair CMOS OTAs are discussed in Section 6. Section 7 concludes the paper.

2. BEHAVIORAL MODELING OF TRANSCONDUCTORS

For the purpose of the subsequent analysis we will represent the nonlinear transfer characteristic of transconductor using power series expansion. Let i_G denotes the output current of the transconductor, while v_{IN} is the differential input voltage of the transconductor as shown in Fig. 1. Then we have

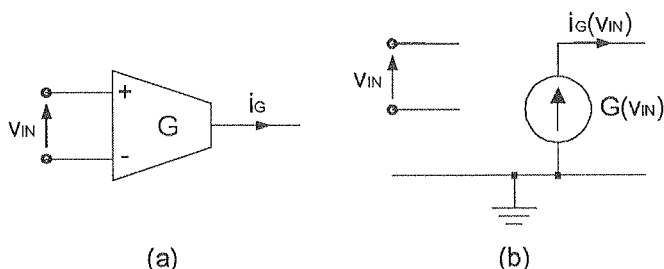


Fig. 1. Nonlinear model of transconductance element: symbolic representation (a), AC-equivalent model (b)

Rys. 1. Nieliniowy model elementu transkonduktancyjnego: reprezentacja symboliczna (a), odpowiadający model zmiennoprądowy (b)

$$i_G(v_{IN}) = G(v_{IN}) = \sum_{n=1}^{\infty} g_n v_{IN}^n(t) \quad (1)$$

amplifiers

er (OTA)

pair trans-

d scheme.

ector and a

or it can

chnology.

switches

nonlinear

described.

tion 4 the

tion 5 the

imulation

differential

r.

near trans-

enotes the

age of the

where coefficients g_n are defined as

$$g_n = \frac{1}{n!} \left. \frac{d^n G(v_{IN})}{dv_{IN}^n} \right|_{v_{IN}=0} \quad (2)$$

By definition, coefficient g_1 is the transconductance g_m of the amplifier.

Consider the following simple CMOS differential pair transconductor shown in Fig. 2. It can be shown, using square-law MOS transistor modeling (e.g. [17]) that its normalized transfer characteristic around zero is:

$$i_G(x) = 2I_S x \sqrt{1 - x^2} \quad (3)$$

where x is a normalized input voltage defined as $x = v_{IN}/2(V_{GS} - V_T)$, with v_{IN} being a differential input voltage, V_{GS} and V_T — gate-source DC voltage and threshold voltage, respectively; I_S is a bias current (see Fig. 2). Actually, formula (3) is valid for $|x| \leq \sqrt{2}/2$. For larger x the transfer characteristic saturates. The corresponding power series expansion is of the form:

$$i_G(x) = I_S \left(x - \frac{1}{2}x^3 - \frac{1}{8}x^5 - \frac{1}{16}x^7 - \frac{1}{128}x^9 \dots \right) \quad (4)$$

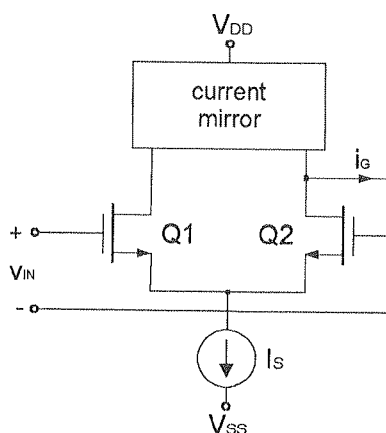


Fig. 2. Simple CMOS differential pair transconductor

Rys. 2. Prosty transkonduktor oparty na parze różnicowej CMOS

3. DESCRIPTION OF FEEDFORWARD LINEARIZATION METHOD

Fig. 3. shows the concept of transconductance amplifier linearization based on active-error feedforward method. All amplifiers $G^{<1>}$, $G^{<2>}$, $G^{<3>}$, modeled as in Fig. 1.

(1)

are assumed to be identical. Their transfer characteristics are described by the power series expansion (1). Moreover, it is assumed that resistor R in Fig. 3 is perfectly linear and equal to $1/g_m$. In practice, e.g. in integrated circuit implementations some technologies offer high resistive poly which can be used to realize resistor R . Alternatively, such a resistor can be treated as an external (discrete) element.

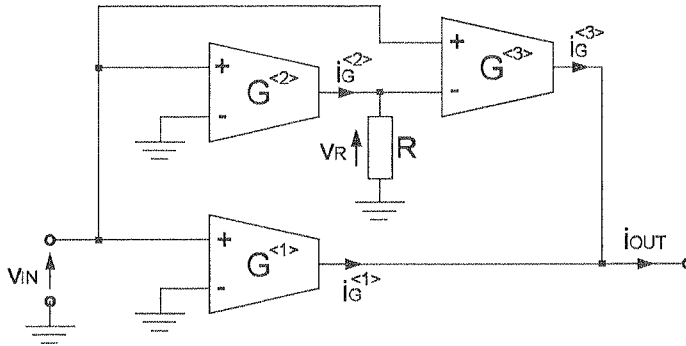


Fig. 3. Three-block feedforward transconductance amplifier

Rys. 3. Trójblokowy wzmacniacz transkonduktancyjny ze sprzężeniem typu „feedforward”

Using (1), the output current i_{OUT} of the overall circuit in Fig. 3. can be written as follows:

$$i_{OUT}(t) = \sum_{n=1}^{\infty} g_n [v_{IN}(t)]^n + \sum_{n=1}^{\infty} g_n [v_{IN}(t) - v_R(t)]^{n+} \quad (5)$$

where

$$v_R(t) = g_1^{-1} \sum_{n=1}^{\infty} g_n [v_{IN}(t)]^n \quad (6)$$

This means that the voltage at the input of the transconductor $G^{<3>}$ (working as an error amplifier) equals $v_{IN}(t) - v_R(t) = v_{IN}(t) - g_1^{-1} \sum_{n=1}^{\infty} g_n [v_{IN}(t)]^n = -g_1^{-1} \sum_{n=2}^{\infty} g_n [v_{IN}(t)]^n$.

Hence, we obtain

$$i_{OUT}(t) = \sum_{n=1}^{\infty} g_n [v_{IN}(t)]^n + \sum_{n=1}^{\infty} g_n \left[-g_1^{-1} \sum_{k=2}^{\infty} g_k [v_{IN}(t)]^k \right]^n \quad (7)$$

Normally, $v_{IN}(t) - v_R(t)$ is much smaller than the input voltage of transconductors $G^{<1>}$ and $G^{<2>}$, which allows us to neglect the higher order terms in the output current of $G^{<3>}$. This leads to the following approximation:

$$i_{OUT}(t) \cong \sum_{n=1}^{\infty} g_n [v_{IN}(t)]^n - \sum_{n=2}^{\infty} g_n [v_{IN}(t)]^n = g_1 v_{IN}(t) \quad (8)$$

power
linear
techno-
actively,

which shows the perfect cancellation of nonlinearities of the overall transconductance amplifier in Fig. 3.

For the purpose of illustration consider again the simple differential pair transconductor in Fig. 2. Consider the following approximation of the power series expansion (4)

$$\bar{i}_G(x) \cong I_S \left(x - \frac{1}{2}x^3 \right) \quad (9)$$

The error of this approximation is less than 1% for $x < 0.17$. THD for the considered transconductor described by equation (9) can be calculated assuming $x(t) = x \cos \omega t$. Then we have

$$\bar{i}_G = I_S x \cos \omega t - 0.5 I_S x^3 \cos^3 \omega t = I_S \left[\left(x - \frac{3}{8}x^3 \right) \cos \omega t - \frac{1}{8}x^3 \cos 3\omega t \right] \quad (10)$$

Thus, THD can be calculated as

$$THD_1 = \sqrt{\frac{x^6/64}{(x - 3x^3/8)^2 + x^6/64}} \quad (11)$$

Now, employing the active-error feedforward scheme gives (see (7))

$$\bar{i}_{OUT}(t) = I_S x - 0.5 I_S x^3 + I_S \cdot 0.5 I_S x^3 - 0.5 I_S (0.5 x^3)^3 = I_S x - 2^{-4} I_S x^9 \quad (12)$$

Again, assuming $x(t) = x \cos \omega t$ one gets

$$\bar{i}_{OUT}(t) = I_S x \cos \omega t - 2^{-4} I_S \cos^9 \omega t \quad (13)$$

which can be rewritten (using some elementary trigonometric formulas) as

$$\begin{aligned} \bar{i}_{OUT}(t) = I_S \left[\left(x - 126Ax^9 \right) \cos \omega t - 84Ax^9 \cos 3\omega t - 36Ax^9 \cos 5\omega t - \right. \\ \left. - 9Ax^9 \cos 7\omega t - Ax^9 \cos 9\omega t \right] \end{aligned} \quad (14)$$

with $A = 2^{-12}$. Having (14) one can now easily calculate THD_2 of the linearized circuit, which is

$$THD_2 = \sqrt{\frac{A^2 x^{18} (84^2 + 36^2 + 9^2 + 1^2)}{(x - 126Ax^9)^2 + A^2 x^{18} (84^2 + 36^2 + 9^2 + 1^2)}} \quad (15)$$

One can easily note that linearized circuit has not only third harmonic component but also 5th, 7th and 9th. However, at least for x low enough, they are much smaller than the third harmonic of the original circuit. For example, for $x = 0.1$ we have the following values of the third harmonics: $(HD_3)_1 = 1.25 \cdot 10^{-4}$ (original circuit) and $(HD_3)_2 = 2.05 \cdot 10^{-11}$. Respective THD are $THD_1 = 1.3 \cdot 10^{-1}\%$ and

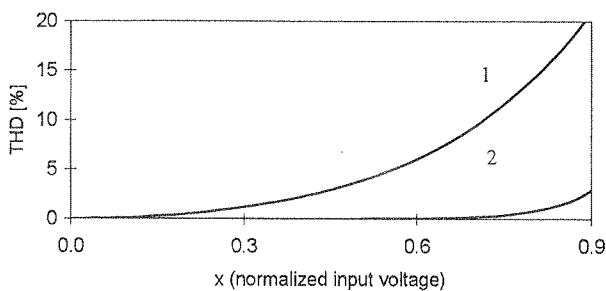


Fig. 4. Theoretical THD characteristics for simple differential pair transconductor (1), and linearized transconductor (2)

Rys. 4. Teoretyczne charakterystyki THD dla prostego transkonduktora (1), oraz transkonduktora linearyzowanego (2)

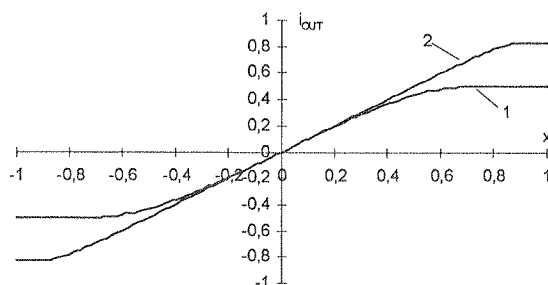


Fig. 5. Theoretical transfer characteristics for simple differential pair transconductor (1), and linearized transconductor (2)

Rys. 5. Teoretyczne charakterystyki przejściowe dla prostego transkonduktora (1), oraz transkonduktora linearyzowanego (2)

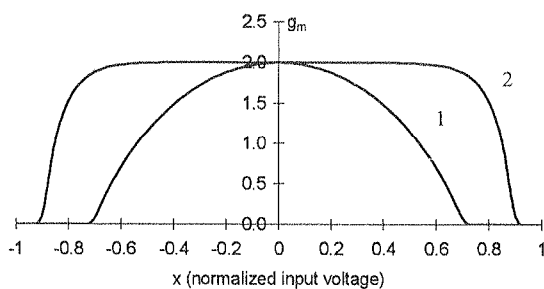


Fig. 6. Theoretical transconductance characteristics for simple differential pair transconductor (1), and linearized transconductor (2)

Rys. 6. Teoretyczne charakterystyki transkonduktancji dla prostego transkonduktora (1), oraz transkonduktora linearyzowanego (2)

$\text{THD}_2 = 7.3 \cdot 10^{-7} \%$ (for $x = 0.2$ we have $(\text{HD}_3)_1 = 1.0 \cdot 10^{-3}$, $(\text{HD}_3)_2 = 1.2 \cdot 10^{-8}$, $\text{THD}_1 = 5.1 \cdot 10^{-1}$, $\text{THD}_2 = 6.1 \cdot 10^{-6}$).

One can also calculate THD for both original and linearized circuit using the exact transfer characteristic (3). In this case, calculations have to be performed numerically, however, there is no restriction for the normalized input voltage x . Fig. 4 shows theoretical THD characteristics for the considered circuits. Figs. 5 and 6 show theoretical transfer and transconductance characteristics, respectively. It is worth noticing that using the active-error feedforward technique one can obtain not only significant reduction of THD but also considerable increase of linear range of operation (recall that the transfer characteristic for differential pair transconductor saturates for $x \approx 0.7$; for linearized circuit it happens for $x \approx 0.9$).

4. LINEARIZED DIFFERENTIAL PAIR OTA IMPLEMENTATION

In Fig. 7 a circuit implementation of the active-error feedforward linearization concept discussed in Section 3 is shown. Note that the circuit in Fig. 7 is a two-output OTA, which follows from the fact that such a configuration is more suitable for filtering applications. Thus, this is a slight modification of the concept presented in Fig. 3. Transistors $Q_{1,2,3,4}$ form classical differential source-coupled pairs with current sink

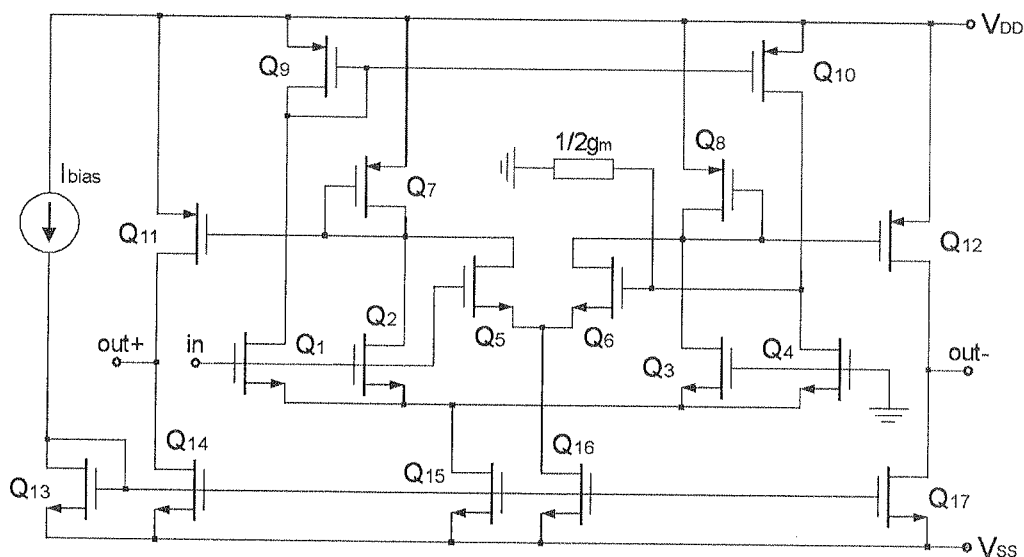


Fig. 7. Complete diagram of the linearized CMOS OTA

Rys. 7. Kompletny schemat linearyzowanego wzmacniacza transkonduktancyjnego CMOS

realized by transistor Q_{15} . Actually, the pair Q_2 , Q_3 with transistors Q_7 , Q_{11} and Q_8 , Q_{12} acting as simple current mirrors implement transconductor $G^{<1>}$ in Fig. 3. The

only difference is that implementation in Fig. 7 is two-output one. Transistors Q_1 , Q_2 with current mirror Q_9 , Q_{10} implement transconductor $G^{<2>}$ loaded by resistor $1/2g_m$. Differential pair Q_5 , Q_6 with current sink Q_{16} realize error amplifier $G^{<3>}$. Note that in order to change the transconductance of the circuit in Fig. 7, the bias current I_{bias} and resistor $1/2g_m$ have to be adjusted simultaneously.

5. SIMULATION RESULTS

The complete circuit in Fig. 7 was designed for the 0.35- μm AMS process and simulated using SPICE. All p-channel and n-channel transistors have their bulks connected to V_{DD} and V_{SS} , respectively. The W/L ratios for transistors in Fig. 7 are: 4/1 for $Q_{1,2,3,4,5,6}$, 20/1 for $Q_{7,8,11,12,13,14,16,17}$, 10/1 for $Q_{9,10}$ and 40/1 for Q_{15} . Channel length is $1\mu\text{m}$ for all transistors. The bias current was set to $I = 20\mu\text{A}$. The circuit was simulated with $V_{DD} = -V_{SS} = 1.25\text{V}$. The power consumption of the circuit is about 0.3mW.

Figs. 8 and 9 show the simulated transfer characteristics and transconductance characteristics of the circuit in Fig. 7 and the reference simple differential pair transconductor in Fig. 2, respectively. It is seen that the improved linearity of overall transconductance element is obtained without deteriorating the transconductance of the circuit. This is not the case for the feedback technique, where the improvement of linearity is obtained at the expense of the reduced transconductance of the overall element. Moreover, we can observe a significant increase of the linear input voltage range for the linearized circuit, which confirms the theoretical predictions presented in Section 3.

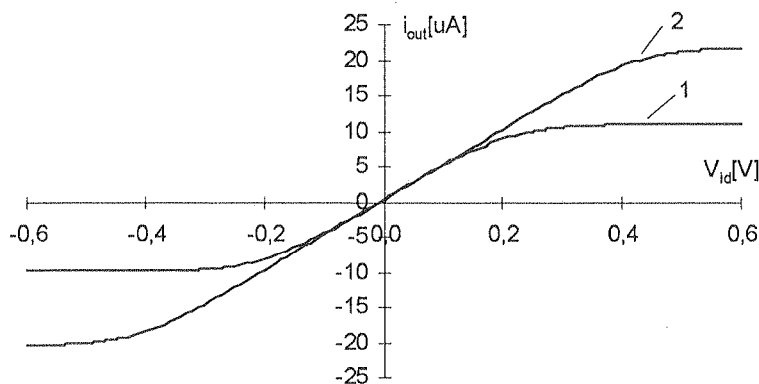


Fig. 8. SPICE simulated transfer characteristics of the circuit in Fig. 7: without linearization (1), and with linearization (2)

Rys. 8. Symulowane charakterystyki przejściowe dla układu z Rys. 7: bez linearyzacji (1), z linearyzacją (2)

Q_1, Q_2
 $1/2 g_m$.
 Note that
 ent I_{bias}

ess and
 ks con-
 $4/1$ for
 length is
 mulated
 mW.

uctance
 ir trans-
 overall
 ance of
 vement
 overall
 voltage
 entered in

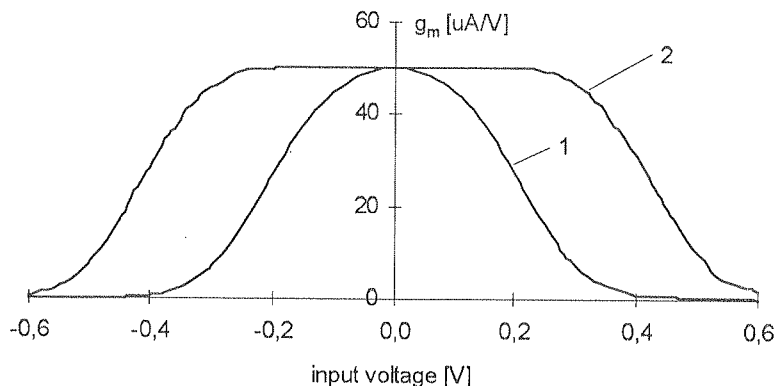


Fig. 9. SPICE simulated transconductance characteristics of the circuit in Fig. 7: without linearization (1), and with linearization (2)

Rys. 9. Symulowane charakterystyki transkonduktancji dla układu z Rys. 7: bez linearyzacji (1), z linearyzacją (2)

In order to have a quantitative measure of linearity improvement, the total harmonic distortion has been calculated for the circuit in Fig. 7 assuming 1MHz sine wave input. The THD characteristics of the circuit with and without linearization are shown in Fig. 10. We can observe a big difference between both curves in Fig. 10. For example, THD at $0.4V_{pp}$ input signal for linearized circuit is less than 0.11% while THD for the circuit without linearization is about 3.9%. For input signal $0.6V_{pp}$ the figure is 0.38% and 10.0%, respectively. These results also confirm the theoretical predictions of Section 3.

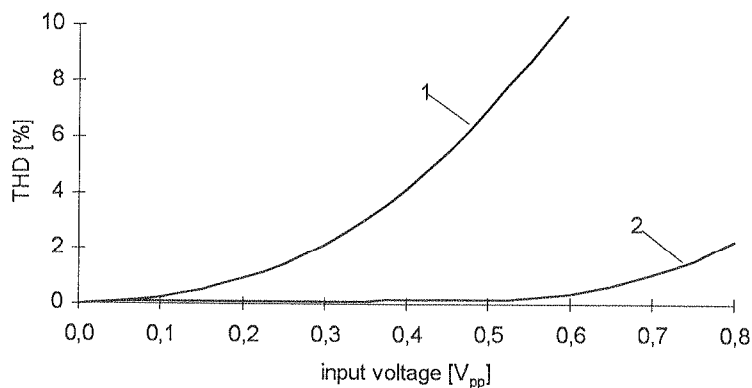


Fig. 10. SPICE simulated THD characteristics of the circuit in Fig. 7 with 1MHz sine wave: without linearization (1), and with linearization (2)

Rys. 10. Symulowane charakterystyki THD dla układu z Rys. 7 dla pobudzenia sygnałem o częstotliwości 1MHz: bez linearyzacji (1), z linearyzacją (2)

6. FILTER EXAMPLE

As an example, a 5th order low-pass Butterworth G_m -C filter in Leap-Frog structure [18] has been designed and simulated. The general structure of the filter is shown in Fig. 11. We have assumed 7.7MHz cut-off frequency. Element values are: $g_i = -g_{fi} = 50\mu\text{A/V}$, $i = 1, 2, 3, 4, 5$, $C_1 = 1.55\text{pF}$, $C_2 = 1.69\text{pF}$, $C_3 = 1.38\text{pF}$, $C_4 = 0.89\text{pF}$, $C_5 = 0.31\text{pF}$.

We have considered two cases. In the first one we have simulated the filter with conventional differential pair transconductors in Fig. 2. It follows that THD of the output signal of the filter with 1MHz sine wave input equals 5.4% (12.1%) for $0.2V_{pp}$ ($0.3V_{pp}$) input signal amplitude. One can easily note that THD is relatively high despite the fact that input signal amplitude is not too large. The reason is that the filter is designed with all transconductances equal, which is convenient for the automatic frequency tuning [1]. At the same time, due to employing equal transconductances, signal levels at some internal nodes of the filter can significantly exceed input voltage amplitude, which causes the aforementioned THD growth.

The filter in Fig. 11 has been also implemented using linearized transconductors in Fig. 7. The simulation results for the filter with linearized OTAs show that THD of its output signal with 1MHz sine wave input equals 0.52% (0.98%) for $0.2V_{pp}$ ($0.3V_{pp}$) input signal amplitude. Thus, application of linearized transconductors gives a significant reduction of the nonlinearity distortion of the whole filter circuit. Due to the increase of the linear input voltage range of the circuit in Fig. 7, the filter can now work with much larger input signals. For example, THD level of 5.4% (12.1%), which was obtained for the filter with reference differential pair transconductors with $0.2V_{pp}$ ($0.3V_{pp}$) of input signal, is now attained for input signal amplitude equal to $0.9V_{pp}$ ($1.15V_{pp}$).

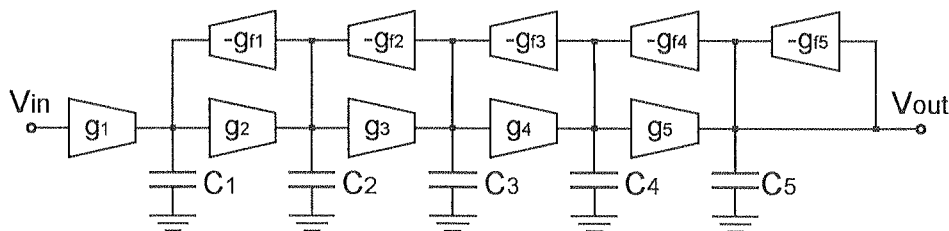


Fig. 11. Diagram of the 5th order low-pass G_m -C filter in Leap-Frog structure

Rys. 11. Schemat dolnoprzepustowego filtra G_m -C piątego rzędu w strukturze Leap-Frog

7. CONCLUSIONS

An effective linearization method based on the active-error feedforward concept has been developed for realizing a very linear CMOS OTA. The proposed circuit technique combines a classical source-coupled differential pair transconductor with a simple error amplifier in feedforward path. Theoretical considerations have been performed using power series expansion of the transfer characteristic of differential pair transconductor showing that both linearization of the transfer characteristic of the circuit as well as increasing of the linear input voltage range actually takes place. The complete linearized OTA in single-input structure has been simulated via SPICE using the 0.35- μm AMS process. For power supply $\pm 1.25\text{V}$, total harmonic distortion at 0.6V_{pp} is less than 0.38%. The obtained simulation results confirm that the linearity of the overall transconductance element is significantly improved in comparison to the reference circuit (i.e. simple differential pair transconductor). Moreover, a significant increase of the linear input voltage range for the linearized circuit was observed. The presented linearization technique can be successfully employed in high-frequency analog filtering. This was confirmed by the simulation results of the 5th order low-pass G_m -C filter. Its THD was significantly reduced while working with linearized OTAs in comparison to the same filter working with conventional differential pair transconductors.

This work was supported in part by the State Scientific Research Committee, Poland, under Grant 4T11B01625

8. REFERENCES

1. R. Schaumann, M. S. Ghausi, K. R. Laker: *Design of Analog Filters, Passive, Active RC, and Switched Capacitor*. Englewood Cliff, NJ: Prentice-Hall, 1990.
2. T. Deliyannis, Y. Sun, J. K. Fidler: *Continuous-time active filter design*. CRC Press, USA, 1999.
3. G. Han, E. Sanchez-Sinencio: *CMOS Transconductance Multipliers: A Tutorial*. IEEE Trans. Circuits and Systems II, Vol. 45, No. 12, 1998, pp. 1550-1562.
4. S. T. Dupie, M. Ismail: *High Frequency CMOS Transconductors, in Analogue IC Design: The Current-Mode Approach*. 1990, C. Toumazou, F. J. Lidgley, D. G. Haigh, Eds., London: Peter-Peregrinus.
5. R. G. Meyer, R. Eschenbach, W. Edgerley: *A Wide-Band Feedforward Amplifier*. IEEE J. Solid-State Circuits, Vol. SC-9, No. 6, 1974, pp. 422-428.
6. S. Szczepanski, Z. P. Tomaszewski: *Linearization Technique Using the Mixed Feedforward-Feedback Concept*. in Proc. European Conference on Circuit Theory and Design, 1980, ECCTD, Vol. 1, pp. 214-219.
7. S. Narahashi, T. Nojima: *Extremely low-distortion multi-carrier amplifier self-adjusting feedforward amplifier*. Proc. of IEEE Int. Communications Conf., 1991, pp. 1485-1490.
8. J. K. Cavers: *Adaptation Behavior of a Feedforward Amplifier Linearizer*. IEEE Trans. Vehicular Tech., vol. 44, no. 1, 1995, pp. 31-40.
9. E. E. Eid, F. M. Ghannouchi, F. Beaugerard: *Optimal Feedforward Linearization System Design*. Microwave Journal, Vol. 38, No. 11, 1995, pp. 78-86.

10. P. B. Kenington, D. W. Bennett: *Linear distortion correction using a feedforward system*. IEEE Trans. on Vehicular Tech., vol. 45, no. 1, 1996, pp. 74-81.
11. Q. Cheng, C. Yiyuan, Z. Xiaowei: *A 1.9 GHz Adaptive Feedforward Power Amplifier*. Microwave Journal, Vol. 41, No. 11, 1998, pp. 86-96.
12. A. Katz: *Linearization: Reducing Distortion in Power Amplifiers*. IEEE Microwave Magazine, 2:4, 2001, pp. 37-49.
13. A. B. Greene: *Bipolar and MOS Analog Integrated Circuit Design*. New York, John Wiley & Sons, Inc., 1984.
14. W. Sansen, Z. Y. Chang: *Feedforward compensation techniques for high-frequency CMOS amplifier*. IEEE J. Solid-State Circuits, Vol. 25, No. 6, 1990, pp. 1590-1595.
15. S. Setty, C. Toumazou: *Feedforward compensation technique in the design of low voltage opamps and OTAs*. IEEE ISCAS, Vol. 1, 1998, pp. 464-467.
16. B. K. Thandri, J. Silva-Martinez: *A Robust Feedforward Compensation Scheme for Multistage Operational Transconductance Amplifiers With No Miller Capacitors*. IEEE J. Solid-State Circuits, Vol. 38, No. 2, 2003, pp. 237-243.
17. P. Wambacq, W. Sansen: *Distortion Analysis of Analog Integrated Circuits*. Kluwer Academic Publishers, 1998.
18. S. Kozieł, S. Szczepański: *Sensitivity Performance of All-pole Canonical Low-pass G_m -C Filters*. Bull. Pol. Ac.: Technical Sciences, 2002, Vol. 50, No. 4, pp. 313-340.

S. SZCZEPAŃSKI, S. KOZIEL

LINEARYZACJA WZMACNIACZA TRANSKONDUKTANCYJNEGO CMOS METODĄ AKTYWNEGO SPRZĘŻENIA TYPU „FEEDFORWARD”

Streszczenie

W pracy przedstawiono koncepcję układową operacyjnego wzmacniacza transkonduktancyjnego CMOS o zwiększonej liniowości charakterystyk przejściowych. Proponowana struktura wzmacniacza transkonduktancyjnego opiera się na wykorzystaniu prostych transkonduktorów realizowanych za pomocą par różnicowych. Linearyzacja charakterystyki przejściowej układu następuje dzięki zastosowaniu w ścieżce „feedforward” wzmacniacza błędu. Wzmacniacz ten sterowany jest sygnałem będącym różnicą pobudzenia układu oraz sygnału wyjściowego podstawowego (nieliniowego) transkonduktora przetworzonego do postaci napięciowej. Różnica ta jest w istocie proporcjonalna do sumy składników nieliniowych w sygnale wyjściowym transkonduktora. Efekt ten warunkowany jest użyciem, jako elementu odwzorowującego, liniowego rezystora, który może być zrealizowany zarówno jako element dyskretny jak też monolitycznie, wewnątrz wzmacniacza operacyjnego (jest to możliwe przy zastosowaniu odpowiednich technologii, tzw. „high-resistive poly”). Sygnał wyjściowy wzmacniacza błędu zostaje następnie odjęty od sygnału wyjściowego wzmacniacza toru głównego. Przy założeniu liniowości wzmacniacza błędu powoduje to całkowite usunięcie składników nieliniowych w sygnale wyjściowym wzmacniacza głównego. W praktyce, wzmacniacz błędu realizowany jest w identycznej strukturze jak wzmacniacz główny. Jego nieliniowość nie odgrywa jednak większej roli z uwagi na fakt, że wzmacniacz błędu sterowany jest sygnałem, którego poziom jest znacznie niższy niż sygnał pobudzający wzmacniacz główny.

Wzmacniacz pracujący zgodnie z opisaną wyżej koncepcją został zaprojektowany z uwzględnieniem parametrów technologii $0.35\mu\text{m}$ AMS dla napięcia zasilania $\pm 1.25\text{V}$. Wyniki symulacji komputerowej pokazują, że współczynnik zawartości harmonicznych (THD) dla napięcia wejściowego 0.6V_{pp} jest mniejszy niż 0.38% w porównaniu do 10% otrzymanych w przypadku układu nielinearyzowanego. Ponadto, układ

linearyzowany wykazuje poszerzenie zakresu liniowej pracy (co należy rozumieć jako zwiększenie progu napięciowego, w którym charakterystyka przejściowa wchodzi w zakres nasycenia) w porównaniu do układu podstawowego transkonduktora. Pobór mocy układu wynosi zaledwie 0.3mW, co czyni prezentowany układ atrakcyjnym dla wielu zastosowań niskomocowych, np. w radiokomunikacji bezprzewodowej. Jako przykład, linearyzowany wzmacniacz transkonduktancyjny zastosowano do zaprojektowania filtra G_m -C implementującego maksymalnie płaską charakterystykę amplitudową 5-go rzędu. Wyniki symulacji potwierdziły, że filtr wykorzystujący wzmacniacz linearyzowany posiada znacznie lepszą liniowość niż taki sam filtr wykorzystujący podstawowy układ transkonduktora.

Słowa kluczowe: operacyjne wzmacniacze transkonduktancyjne, linearyzacja typu „feedforward”, wzmacniacze CMOS ze sprzężeniem „feedforward”

i
f
t
s
l
f
b
f

A

Co
fiers and
current
types and
ted cir
filters
subject
for ap
comput
ugh th

NOISE ANALYSIS AND OPTIMIZATION OF CONTINUOUS-TIME OTA-C FILTERS

ŚLAWOMIR KOZIEŁ

*Faculty of Electronics, Telecommunications and Informatics,
Gdańsk University of Technology, 80-952 Gdańsk, Poland
e-mail: koziel@ue.eti.pg.gda.pl*

*Otrzymano 2004.03.16
Autoryzowano 2004.11.15*

In the paper, a general approach to noise analysis in continuous-time OTA-C filters is presented. Based on a matrix description of a general OTA-C filter topology, universal formulas for evaluating noise in any OTA-C filter of arbitrary order are derived. The presented formulas can be easily implemented and used in computer-aided analysis/optimization software. The accuracy of the proposed method is confirmed by comparison to SPICE simulation. The examples of application for finding the minimum-noise 5th order multiple-loop feedback filters implementing Butterworth and Bessel transfer functions, and for optimal biquad sequencing and gain distribution in cascade realization of 8th order Butterworth filter are given.

Keywords: OTA-C filters, noise analysis, filter optimization

1. INTRODUCTION

Continuous-time analog filters and equalizers based on transconductance amplifiers and capacitors (OTA-C) constitute suitable solutions to various voltage-mode and current-mode signal processing tasks. Many synthesis and design methods for different types and architectures of this class of filters have been reported [1]-[9]. Since integrated circuit technologies offer possibilities to realize fully integrated continuous-time filters using the OTA-C technique, in recent years there has been a great interest in this subject. Integrated continuous-time OTA-C filters have received substantial attention for applications such as in hard disc drives, video filters, wireless communications, computer systems, biomedicine, control and instrumentation systems [10]-[21]. Although these filters offer excellent high frequency performance, their other properties in

terms of low supply voltage, power consumption, parasitic effects, sensitivity, noise and dynamic range, etc., still need improvements [5], [22].

In this paper we deal with noise in OTA-C filters. It is important for filter design purposes to develop efficient tools for performing noise analysis of OTA-C filters, especially tools which can be embedded into automated filter design systems. There have been several attempts to solve this problem described in the literature [23]-[27]. In this paper we propose a general approach to noise analysis in OTA-C filters based on the matrix description of OTA-C filters developed in [8]. The derived formulas can be applied to any known OTA-C filter architecture. They can also be easily implemented and used in computer-aided analysis/optimization software.

The paper is organized as follows. In Section 2, we present a general OTA-C topology along with its matrix description that is based on separate treatments of the passive and active parts of the circuit. The model is sufficiently general to comprise any known OTA-C filter architecture. The matrix description permits the transfer function of any OTA-C filter to be calculated as a special case of the general structure. In Section 3, the noise analysis of the general OTA-C filter model is carried out, which leads to explicit and compact formulas for output and input referred noise spectrum of any OTA-C filter. Verification of the proposed approach is given in Section 4 by comparing theoretical results to SPICE simulation. Section 4 also gives a few examples of application of the discussed approach to noise optimization of OTA-C filters. Remarks and conclusions are contained in Section 5.

2. GENERAL STRUCTURE OF OTA-C FILTER

Consider a general structure of a voltage-mode OTA-C filter shown in Fig. 1. It is easily seen that any OTA-C filter is a particular case of this structure, since any particular filter topology can be obtained from the general one by removing appropriate number of elements. The structure in Fig. 1 contains n internal nodes denoted as x_i , $i = 1, \dots, n$, an active network consisting of n input transconductors G_{mbi} , a set of feedforward and feedback transconductors G_{mij} , $i, j = 1, \dots, n$, an output summer consisting of transconductors G_{mci} , G_{mo} and a feedforward transconductor G_{md} , and a passive network including input capacitors C_{bi} , $i = 1, \dots, n$ as well as grounded and floating capacitors C_{ij} , $1 \leq i < j \leq n$.

In the sequel we shall denote the voltage at the i -th node x_i also by x_i . A general structure of the OTA-C filter in Fig. 1 can be described by the following matrix equations:

$$\begin{aligned} sT_c X &= GX + B^T u_i \\ u_o &= CX + Du_i \end{aligned} \quad (1)$$

noise

design

filters,

There

[27].

based

s can

ented

TA-C

of the

any

action

re. In

which

spectrum

on 4

a few

TA-C

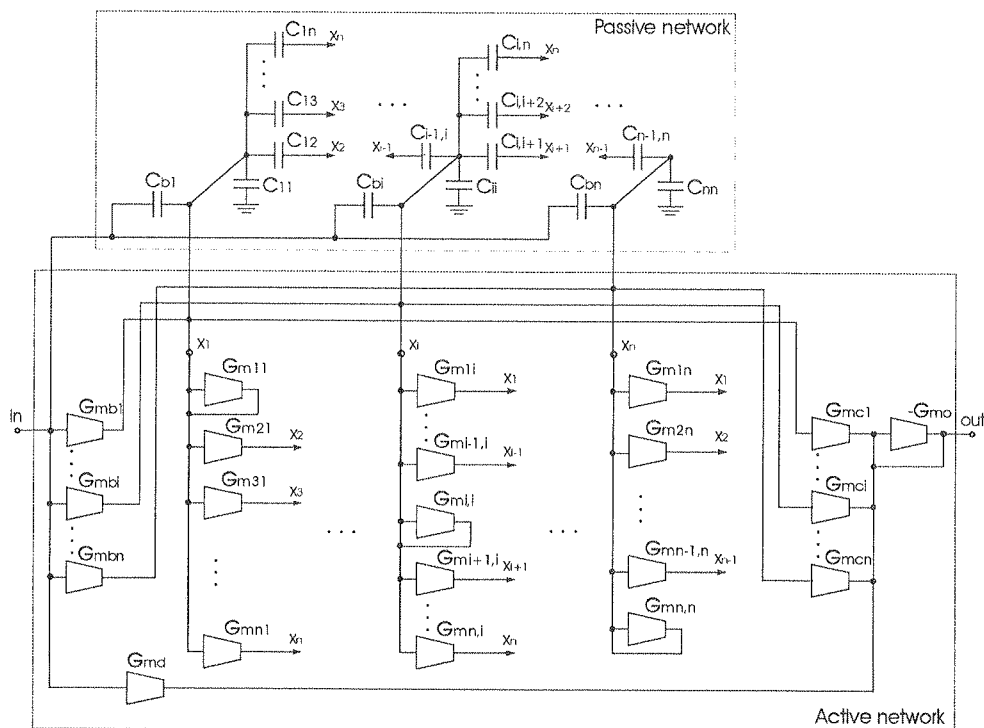


Fig. 1. General structure of a voltage-mode OTA-C filter

Rys. 1. Ogólna struktura filtru napięciowego OTA-C

where u_i , u_o are the input and output voltages, respectively, and

$$T_c = \begin{bmatrix} C_{b1} + \sum_{j=1}^n C_{1j} & -C_{12} & \cdots & -C_{1n} \\ -C_{12} & C_{b2} + \sum_{j=1}^n C_{2j} & \cdots & -C_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -C_{1n} & -C_{2n} & \cdots & C_{bn} + \sum_{j=1}^n C_{1n} \end{bmatrix},$$

$$G = \begin{bmatrix} G_{m11} & G_{m12} & \cdots & G_{m1n} \\ G_{m21} & G_{m22} & \cdots & G_{m2n} \\ \vdots & \vdots & \ddots & \vdots \\ G_{mn1} & G_{mn2} & \cdots & G_{mnn} \end{bmatrix}, \quad X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix},$$

(2)

$$B = [G_{mb1} + sC_{b1} + \cdots + G_{mbn} + sC_{bn}],$$

$$C = [c_1 \cdots c_n], \quad c_i = G_{mci}/G_{mo}, \quad i = 1, 2, \dots, n$$

$$D = d, \quad d = G_{md}/G_{mo}$$

(1)

On the basis of (1) we can calculate the filter transfer function:

$$H(s) = \frac{u_o(s)}{u_i(s)} = C(T - G)^{-1} B^T + D \quad (3)$$

Now, let us denote adjoint matrix of $sT_c - G$ as $\tilde{A}$ where

$$\tilde{A}(s) = \text{adj}(sT_c - G) = [\tilde{A}_{ij}(s)]_{i,j=1}^n \quad (4)$$

This allows us to rewrite transfer function in the form:

$$H(s) = \frac{1}{\det(sT_c - G)} \sum_{i,j=1}^n c_i (G_{mbj} + sC_{bj}) \tilde{A}_{ij}(s) + d \quad (5)$$

Note that many filter structures have only one input transconductor (i.e. no input signal distribution), a trivial output summer (i.e. one of the internal nodes is the output of the filter), and no input capacitors. This means that $B = [0 \cdots 0 \ G_{mbk} \ 0 \cdots 0]$, that $C = [0 \cdots 0 \ 1 \ 0 \cdots 0] - 1$ at l -th position and $C_{bi} = 0$ for $i = 0, 1, \dots, n$. In such case, expression (5) reduces to the form:

$$H(s) = \frac{G_{mbk} \tilde{A}_{lk}(s)}{\det(sT_c - G)} \quad (6)$$

Similar expression can be written for slightly more general case, with no input capacitors, input signal distribution (i.e. $B = [G_{mbl} \cdots G_{mbn}]$), and a trivial output summer ($C = [0 \cdots 0 \ 1 \ 0 \cdots 0] - 1$ at l -th position). Then, we have

$$H(s) = \frac{1}{\det(sT_c - G)} \sum_{i,j=1}^n G_{mbj} \tilde{A}_{ij}(s) \quad (7)$$

On the basis of the above expressions one can easily calculate the transfer function of any particular structure of OTA-C filter. In a similar way one can treat current-mode structures as well as state-space filters (see [8] for more details). It follows that the order n_H of transfer function of general filter structure in Fig. 1 is not necessarily equal to the number of internal nodes. It can be shown that it essentially depends on the passive network. In particular, n_H is not larger than the rank of the matrix T_C corresponding to the filter.

3. NOISE ANALYSIS IN GENERAL OTA-C FILTERS

The output noise of any OTA-C filter is a combination of the noise contributions of its all transconductors. The noise in CMOS amplifier with transconductance g_m can

be described in terms of an equivalent input referred noise voltage source v_n as shown in Fig. 2. Spectral density $S_n(f)$ of the noise source can be modeled as [23]:

$$S_n(f) = \frac{S_t}{g_m} + \frac{S_f}{f} \quad (8)$$

where both S_t (thermal noise component) and S_f (flicker noise component) depend on amplifier topology. We shall assume that noise sources associated to different OTAs are statistically independent.

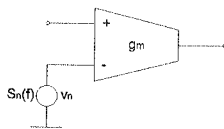


Fig. 2. Equivalent input voltage noise source representation of noise in OTA

Rys. 2. Reprezentacja szumów OTA w postaci napięciowego źródła szumów odniesionych do wejścia

Our immediate goal is to obtain the explicit formula for output (and/or input) noise spectrum of the general OTA-C filter in Fig. 1. In order to do this, one has to consider what is the contribution of the noise of each individual transconductance amplifier to the output noise spectrum of the filter. This can be modeled as shown in Fig. 3. Let

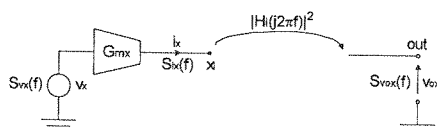


Fig. 3. Noise contribution of individual filter transconductor to the total output noise of the filter

Rys. 3. Wkład pojedynczego transkonduktora do całkowitego szumu wyjściowego filtru

G_{mx} denote one of the filter transconductors (e.g. G_{mbi} , G_{mij} , etc.), which is connected to one of the nodes, say x_i (if the filter contains non-trivial output summer then it has an additional — output — node which will be denoted as x_0). Denote by v_x the input referred noise voltage of the noise source corresponding to G_{mx} , whose spectral density is $S_{vx}(f)$. Transconductor G_{mx} injects its noise current $i_x = v_x G_{mx}$ into the node x_i . Spectral density $S_{ix}(f)$ of this current is given by

$$S_{ix}(f) = G_{mx}^2 S_{vx}(f) \quad (9)$$

The corresponding output noise voltage v_{ox} can be calculated as

$$v_{ox} = i_x H_i = G_{mx} H_i v_x \quad (10)$$

where H_i is the current-to-voltage transfer function from node x_i to the output of the filter. Corresponding spectral density $S_{vox}(f)$ is given by the formula:

$$S_{vox}(f) = S_{ix}(f) |H_i(j2\pi f)|^2 = G_{mx}^2 S_{vx}(f) |H_i(j2\pi f)|^2 \quad (11)$$

It can be easily shown using the matrix equation (1) that the transfer functions $H_i(s)$, $i = 1, 2, \dots, n$ are components of the $1 \times n$ vector $\mathbf{H}_{cv}$ of rational functions in s defined as follows

$$\mathbf{H}_{cv}(s) = \mathbf{C}(s\mathbf{T}_c - \mathbf{G})^{-1} \quad (12)$$

If non-trivial output summer is present (cf. Fig.1) then we also need the current-to-voltage transfer function H_0 from output node to itself, which is

$$H_0 = G_{mo}^{-1} \quad (13)$$

Thus, each filter transconductor injects its noise current (whose spectrum density can be calculated using (9)) into one of the internal nodes of the filter (or directly into the output node if the filter possesses nontrivial output summer). Subsequently, this current is converted into output noise voltage according to (11). In order to calculate the total output noise voltage of the filter we can employ our assumption of statistical independence of noise sources of filter transconductors and simply add all the corresponding noise spectra. In general, the outputs of one input transconductor G_{mbi} , and n transconductors G_{mij} , $j = 1, \dots, n$ are connected to each internal node x_i . In the presence of non-trivial output summer we have an additional output node x_0 with outputs of transconductors G_{mcj} , $j = 1, \dots, n$, G_{md} , and G_{mo} . Let us define auxiliary matrices

$$\begin{aligned} \mathbf{S}_t &= [S_{t,ij}]_{i,j=1}^n, & \mathbf{S}_f &= [S_{f,ij}]_{i,j=1}^n, \\ \mathbf{S}_{tb} &= [S_{tbn} \cdots S_{tbn}]^T, & \mathbf{S}_{fb} &= [S_{f,1n} \cdots S_{f,1n}]^T, \\ \mathbf{S}_{tc} &= [S_{tc,1} \cdots S_{tc,n}]^T, & \mathbf{S}_{fc} &= [S_{f,c,1} \cdots S_{f,c,n}]^T, \\ \mathbf{S}_{td} &= S_{td}, & \mathbf{S}_{fd} &= S_{fd}, \\ \mathbf{S}_{to} &= S_{to}, & \mathbf{S}_{fo} &= S_{fo} \end{aligned} \quad (14)$$

representing the thermal noise (subscript t) and $1/f$ noise (subscript f) of transconductors G_{mij} , G_{mbi} , G_{mci} , G_{md} and G_{mo} , respectively. Let us introduce the following notation:

$$\begin{aligned} \bar{\mathbf{G}} &= [G_{mij}]_{i,j=1}^n, \\ \bar{\mathbf{B}} &= [G_{mb1} \cdots G_{mbn}]^T, \\ \bar{\mathbf{C}} &= [G_{mc1} \cdots G_{mcn}], \\ \bar{\mathbf{D}} &= G_{md}, & \bar{\mathbf{O}} &= G_{mo} \end{aligned} \quad (15)$$

Denote by $\circ$ the Hadamard product of two matrices, i.e. for $\mathbf{P} = [p_{ij}]_{i,j=1}^n$ and for $\mathbf{Q} = [q_{ij}]_{i,j=1}^n$ we have for $\mathbf{P} \circ \mathbf{Q} = [p_{ij}q_{ij}]_{i,j=1}^n$ (the same definition holds, with obvious changes for $n \times 1$, $1 \times n$ and 1×1 matrices) [29]. Define function $F(\mathbf{P}, \mathbf{Q}, \mathbf{R})(x)$, where $\mathbf{P}$, $\mathbf{Q}$, and $\mathbf{R}$ are matrices of the same dimension and x is a real variable:

$$F(\mathbf{P}, \mathbf{Q}, \mathbf{R})(x) = \mathbf{P} \circ (\mathbf{Q} + (2\pi/x)\mathbf{P} \circ \mathbf{R}) \quad (16)$$

It follows from (8) and (9) that spectral densities $S_i(\omega)$ of the total noise current injected into the nodes x_i , $i = 1, \dots, n$ can be expressed, using (16), as components of the current spectral density vector S , given by the formula

$$S(\omega) = \begin{bmatrix} S_1(\omega) \\ \vdots \\ S_n(\omega) \end{bmatrix} = F(\bar{G}, S_t, S_f)(\omega) \cdot \hat{I} + F(\bar{B}, S_{tb}, S_{fb})(\omega) \quad (17)$$

where $\hat{I} = [1 \dots 1]^T$ is $n \times 1$ unit vector. Spectral density $S_0(\omega)$ of the noise current injected into the node x_0 (if there is a non-trivial output summer) is given by

$$S_0(\omega) = F(\bar{C}, S_{tc}, S_{fc})(\omega) \cdot \hat{I} + F(\bar{D}, S_{td}, S_{fd})(\omega) + F(\bar{O}, S_{to}, S_{fo})(\omega) \quad (18)$$

The spectrum density $S_{n0}(\omega)$ of the total output noise voltage u_{no} can be then calculated as

$$S_{n0}(\omega) = |H_{cv}(\omega)|^2 S(\omega) + H_0^2 S_0(\omega) \quad (19)$$

where $|H_{cv}(\omega)|^2 = H_{cv}(j\omega) \circ H_{cv}(-j\omega)$, with H_{cv} and H_0 given by (12) and (13), respectively. In general, $S_0(\omega)$ is a rational function of ω with numerator and denominator of order not larger than $2n + 1$. Formula (19) allows us to calculate the output noise spectrum of any OTA-C filter. In order to get the output noise voltage one needs to integrate (19) over the suitable frequency range. The equivalent input noise spectrum $S_{ni}(\omega)$ can be obtained by dividing (19) by the square of the transfer function of the filter given by (3). This, especially if there is no output summer in the filter, leads to simplification of formulas because determinant of matrix $sT_C - G$ is canceled out. It is worth noting that because of matrix formulation, the presented formulas are particularly convenient to be implemented in a computer program which will allow to carry out the noise analysis of arbitrary OTA-C filter.

In practice, it is a common situation, when the transconductance value for all filter transconductors are the same. Then, each of the matrices S in (14) is proportional to the unit matrix (i.e. the matrix with all entries equal to 1) $\tilde{I}$ of the suitable dimension. For example, we have $S_t, S_i, \tilde{I}, S_f = S_f \tilde{I}$, where S_t, S_f are noise parameter of transconductor and $\tilde{I}$ is $n \times n$ unit matrix. Then, equations (17) and (18) take the form

$$S(\omega) = (\bar{G}\hat{I} + \bar{B})S_t + \frac{2\pi}{\omega}(\bar{G} \circ \bar{G}\hat{I} + \bar{B} \circ \bar{B})S_f \quad (20)$$

$$S_0(\omega) = (\bar{C}\hat{I} + \bar{D} + \bar{O})S_t + \frac{2\pi}{\omega}(\bar{C} \circ \bar{C}\hat{I} + \bar{D} \circ \bar{D} + \bar{O} \circ \bar{O})S_f \quad (21)$$

For the sake of illustration, consider the second-order low-pass OTA-C filter shown in Fig. 4. We shall assume that $g_1 = g_2 = -g_3 = -g_4 = g$, and noise parameters are S_t and S_f . Corresponding matrices T_C , G , B , C , and D are:

$$T_C = \begin{bmatrix} C_1 & 0 \\ 0 & C_2 \end{bmatrix}, \quad G = \begin{bmatrix} 0 & -g \\ g & -g \end{bmatrix}, \quad B = \begin{bmatrix} g \\ 0 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & 1 \end{bmatrix}, \quad D = 0 \quad (22)$$

Vector $\mathbf{H}_{cv}$ has the form of

$$\mathbf{H}_{cv}(s) = \begin{bmatrix} \frac{g}{D(s)} & \frac{sC_1}{D(s)} \end{bmatrix} \quad (23)$$

where $D(s) = C_1C_2s^2 + C_1gs + g^2$ is denominator of the filter transfer function. Noise vector $\mathbf{S}$ can be calculated as in (20)

$$\mathbf{S}(\omega) = 2g \left(S_t + (2\pi/\omega) g S_f \right) \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad (24)$$

Now, it follows from (19) that the spectrum density $S_{no}(\omega)$ of the total output noise voltage of the filter is given by

$$S_{no}(\omega) = \frac{2g(g^2 + \omega^2C_1^2)(S_t + 2\pi g S_f/\omega)}{|D(\omega)|^2} = \frac{2g(g^2 + \omega^2C_1^2)(S_t + 2\pi g S_f/\omega)}{C_1^2C_2^2\omega^4 + (C_1 - 2C_2)C_1g^2\omega^2 + g^4} \quad (25)$$

Corresponding spectral density $S_{ni}(\omega)$ of the input noise voltage can be calculated as

$$S_{ni}(\omega) = 2(g^2 + \omega^2C_1^2) \left(\frac{S_t}{g} + \frac{2\pi S_f}{\omega} \right) \quad (26)$$

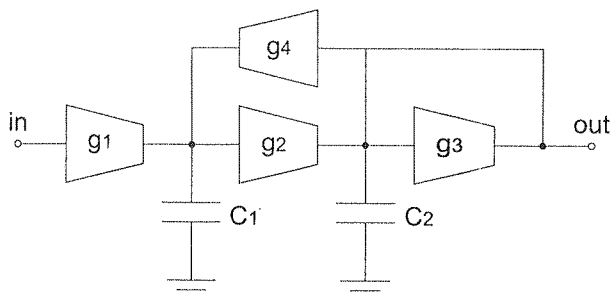


Fig. 4. Diagram of second-order low-pass OTA-C filter

Rys. 4. Schemat dolnoprzepustowego filtru OTA-C drugiego rzędu

4. VERIFICATION AND APPLICATION EXAMPLES

In this section we discuss some application examples of the general OTA-C filter noise analysis presented in Section 3. Due to its generality and matrix formulation, the presented approach can be used to solve many optimization tasks such as proper choice of filter topology with respect optimal noise performance or direct optimization of filter noise performance in various settings.

We start, however, from a verification of the presented approach by comparing the theoretical results of Section 3 to the SPICE simulation. For our comparison we use a simple differential-pair transconductor shown in Fig. 5. The circuit was implemented in standard $0.35\mu\text{m}$ AMS technology and simulated using SPICE. The OTA noise parameters (see (8)) extracted from the simulations are $S_i = 5.2 \cdot 10^{-16}$ V. A/Hz, and $S_f = 2.3 \cdot 10^{-10}$ V². Transconductance of the circuit is $g_m = 100\mu\text{A/V}$.

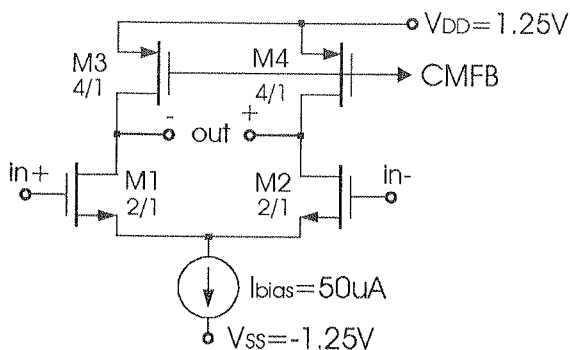


Fig. 5. Simple differential-pair transconductor (Common-Mode Feedback circuit not shown)

Rys. 5. Schemat prostego transkonduktora CMOS

The OTA circuit in Fig. 5 was used to implement two Butterworth low-pass filters in a leap-frog (LF) structure: third- and eight-order ones. Fig. 6 shows a general structure of n^{th} order LF structure.

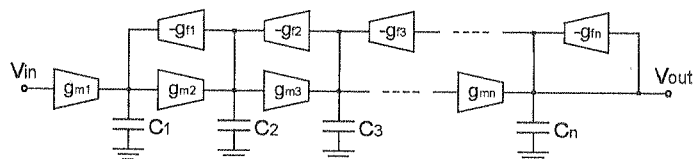


Fig. 6. Diagram of n^{th} order leap-frog (LF) filter

Rys. 6. Schemat filtru typu leap-frog (LF) n -tego rzędu

Actual filters were implemented in fully differential structures. 3dB frequency of the filters is 10MHz. Figs. 7 and 8 show input and output noise spectrum versus frequency for 3rd order and 8th order filter respectively. The agreement between theoretical results (line) and SPICE simulation results (points) is very good. The little discrepancy for output noise spectrum in Fig. 8 follows not from the limited accuracy of the model but from the fact of finite DC gain of the OTA (in this case about 35dB) which gives DC gain of the filter equal to -1.3dB (instead of ideal value of 0dB). It should be emphasized that the effect of finite DC gain of OTAs can also be incorporated into the model presented in Section 2, which would result in even more accurate predictions (cf. [28]).

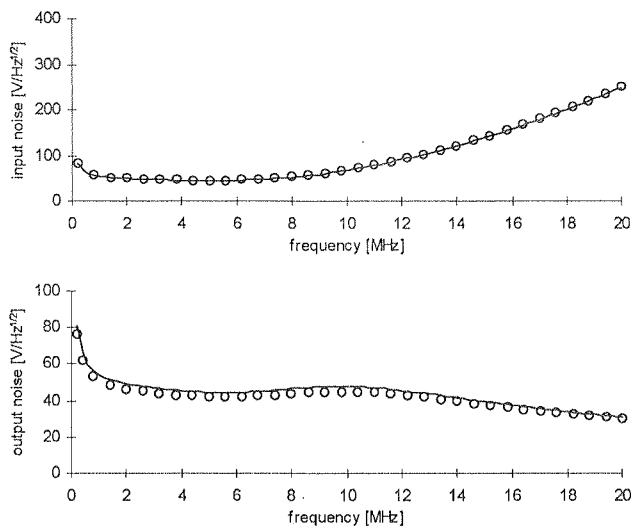


Fig. 7. Input and output noise spectrum vs. frequency for 3rd order Butterworth filter; theoretical data (line), and simulation (points)

Rys. 7. Odniesione do wejścia i wyjścia widmo szumów filtru Butterwortha trzeciego rzędu; wyniki teoretyczne (linia ciągła) oraz symulacje (punkty)

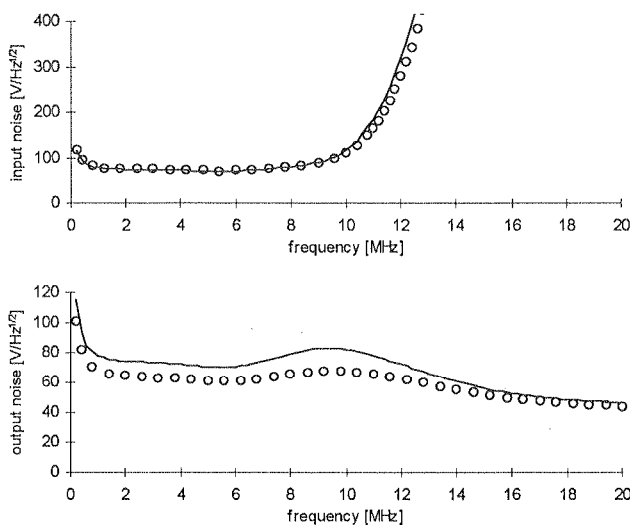


Fig. 8. Input and output noise spectrum vs. frequency for 8th order Butterworth filter; theoretical data (line), and simulation (points)

Rys. 8. Odniesione do wejścia i wyjścia widmo szumów filtru Butterwortha ósmego rzędu; wyniki teoretyczne (linia ciągła) oraz symulacje (punkty)

We now turn to the examples of noise optimization. As a first example let us consider the optimal choice of the multiple-loop feedback topology for the fifth-order low-pass filter. Fig. 9 shows the general structure of n -th order all-pole canonical low-pass OTA-C filter. The feedback signal emerging at node x_i can be taken from node x_{kj} , where $kj \in \{i+1, \dots, n\}$ for $i < n$, and $kn = n$. Two most popular multiple-loop

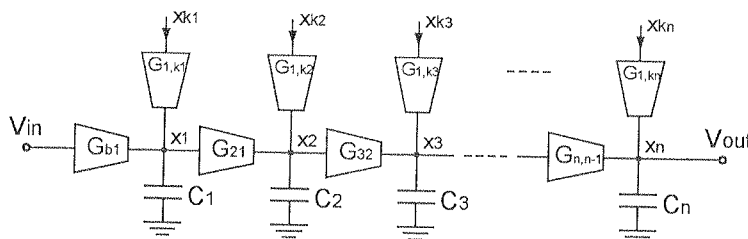


Fig. 9. A general structure of an n -th order all-pole canonical low-pass OTA-C filter

Rys. 9. Ogólna struktura kanonicznego filtru dolnoprzepustowego OTA-C n -tego rzędu bez zer transmisyjnych

feedback topologies — leap-frog (LF) and inverse follow-the-leader (IFLF) can be considered as two ‘extreme’ canonical structures of all-pole filters [8]. In particular, for LF structure the feedback signal at intrinsic node x_i is taken from the node x_{i+1} , $i = 1, 2, \dots, n-1$ and the node x_n has the inner feedback loop. In IFLF structure all feedback signals at nodes x_i , $i = 1, 2, \dots, n$ are taken from the node x_n . In general, the feedback signal emerging at node x_i can be taken from node x_j , where $j = i+1, \dots, n$. Thus, we have in total $(n-1)!$ canonical multiple-loop feedback structures of order n . The matrix approach proposed in Section 3 makes it possible to perform an exhaustive search through all topologies because it is very fast and can be easily automated.

To simplify further description, we introduce the following notation: let $F_{n|j1,j2,\dots,jn}$ denote the canonical all-pole low-pass structure of n -th order, for which the feedback signal emerging at node x_i is taken from node x_{j_i} . For example, n -th order LF structure will be denoted as $F_{n|2,3,4,\dots,n,n}$, while n -th order IFLF structure as $F_{n|n,n,\dots,n}$.

Here, we consider 5th order Butterworth and Bessel filters. According to the discussion above, there are 24 different canonical structures that realize this transfer function. They can be implemented with the same value of transconductances for all filter transconductors. Table 1 shows integrated input noise for all the structures and both approximations. The noise is integrated over 3dB bandwidth and normalized to the noise of LF filter. Only thermal noise has been considered in this example.

It follows from the results in Table 1 that the difference of integrated noise between best and worst topology is as large as 7dB for Butterworth filters and 4.3dB for Bessel filters. This difference tells us about the improvement in noise performance we can get by proper choice of the filter topology only. Note also that the lowest-noise filter topology is transfer function dependent, which means that it has to be determined individually for different approximations.

Table 1

Integrated input noise comparison for 5th order all-pole low-pass Butterworth and Bessel filters; best [worst] topologies shaded horizontally [vertically]; noise values normalized with respect to LF filter

Porównanie scałkowanego szumu odniesionego do wejścia dla dolnoprzepustowych filtrów Butterworth'a i Bessela piątego rzędu. Dane dla najlepszych [najgorszych] struktur zacieniowano liniami pionowymi [poziomymi]

Filter structure	Normalized noise [dB]		Filter structure	Normalized noise [dB]	
	Butterworth	Bessel		Butterworth	Bessel
$F_{S[2,3,4,5,5]}$ (LF)	0.0	0.0	$F_{S[4,3,4,5,5]}$	2.6	1.1
$F_{S[2,3,5,5,5]}$	-0.2	-0.4	$F_{S[4,3,5,5,5]}$	3.2	1.3
$F_{S[2,4,4,5,5]}$	2.4	0.4	$F_{S[4,4,4,5,5]}$	5.7	2.4
$F_{S[2,4,5,5,5]}$	2.9	-0.1	$F_{S[4,4,5,5,5]}$	7.1	3.0
$F_{S[2,5,4,5,5]}$	2.4	-0.5	$F_{S[4,5,4,5,5]}$	5.3	1.3
$F_{S[2,5,5,5,5]}$	3.0	-1.2	$F_{S[4,5,5,5,5]}$	6.8	2.1
$F_{S[3,3,4,5,5]}$	3.4	1.8	$F_{S[5,3,4,5,5]}$	3.6	1.9
$F_{S[3,3,5,5,5]}$	3.1	1.2	$F_{S[5,3,5,5,5]}$	2.9	1.0
$F_{S[3,4,4,5,5]}$	6.0	3.1	$F_{S[5,4,4,5,5]}$	6.2	2.9
$F_{S[3,4,5,5,5]}$	6.7	2.8	$F_{S[5,4,5,5,5]}$	6.7	2.5
$F_{S[3,5,4,5,5]}$	5.5	2.1	$F_{S[5,5,4,5,5]}$	5.9	1.9
$F_{S[3,5,5,5,5]}$	6.2	1.7	$F_{S[5,5,5,5,5]}$ (IFLF)	6.3	1.3

As the second example let us consider noise optimization of 8th order Butterworth low-pass filter implemented as cascade of four second-order sections. Each section was implemented using the filter topology in Fig. 4 with simple OTAs in Fig. 5. The transfer function of the filter is

$$H(s) = \prod_{i=1}^4 H_i(s) \quad (27)$$

where

$$H_i(s) = \frac{K_i}{s^2 + 2 \sin((2i-1)\pi/16) + 1} \quad (28)$$

We assume two degrees of freedom in noise optimization process. On the one hand, exhaustive search through all possible permutations ($4! = 24$) of second-order section is carried out. On the other hand, the optimal gain distribution for each permutation is found. We assume unity gain setting for the whole filter. The gain of individual block is adjusted by changing the transconductance value of its input transconductor (g_1 in Fig. 5). Note that noise parameters S_t and S_f of transconductor are transconductance dependent so we have $S_t = S_t(g_1)$ and $S_f = S_f(g_1)$ for input transconductors. Here, we model this dependence using polynomial interpolation. We assume that g_1 is changed from $70.7 \mu\text{A/V}$ to $141.4 \mu\text{A/V}$, which allows us to adjust block gain in the range $\pm 3\text{dB}$. Let us denote by B_i the second order section implementing the transfer function H_i

(see (impos
0[dB]
Butte
is the
follow
reduc
may
to the

V

Biqu
sequen
 $B_1 B_2$
 $B_1 B_2$
 $B_1 B_3$
 $B_1 B_3$
 $B_1 B_4$
 $B_1 B_4$
 $B_2 B_1$
 $B_2 B_1$
 $B_2 B_3$
 $B_2 B_3$
 $B_2 B_4$
 $B_2 B_4$
 $B_3 B_1$
 $B_3 B_1$
 $B_3 B_2$
 $B_3 B_2$
 $B_3 B_4$
 $B_3 B_4$
 $B_4 B_1$
 $B_4 B_1$
 $B_4 B_2$
 $B_4 B_2$
 $B_4 B_3$
 $B_4 B_3$

(see (28)), $i = 1, 2, 3, 4$, and by K_i the gain of respective block. Then, unity gain setting imposes the following constraint for the noise optimization process: $K_1 + K_2 + K_3 + K_4 = 0[\text{dB}]$. Table 2 shows the results of noise optimization of the considered 8th order Butterworth filter. As expected, the optimal gain distribution for all biquad permutations is the maximum gain for the first two blocks and minimum gain for the last two. It follows from the data in Table 2 that optimal gain distribution typically results in noise reduction of about 25%. On the other hand, the optimal choice of biquad sequencing may give us additional and significant (up to 20%) reduction of noise in comparison to the average noise integral over all 24 biquad permutations (222 μV).

Table 2

Noise optimization results for 8th order Butterworth filter in cascade realization

Wyniki optymalizacji szumów dla filtru Butterwortha ósmego rzędu w realizacji kaskadowej

Biquad sequencing	Optimal gain distribution				Integrated input noise [μV] (optimal gain distribution)	Integrated input noise [μV] ($K_1 = K_2 = K_3 = K_4 = 0\text{dB}$)
	$K_1[\text{dB}]$	$K_2[\text{dB}]$	$K_3[\text{dB}]$	$K_4[\text{dB}]$		
$B_1 B_2 B_3 B_4$	+3	+3	-3	-3	176	221
$B_1 B_2 B_4 B_3$	+3	+3	-3	-3	179	226
$B_1 B_3 B_2 B_4$	+3	+3	-3	-3	178	228
$B_1 B_3 B_4 B_2$	+3	+3	-3	-3	193	246
$B_1 B_4 B_2 B_3$	+3	+3	-3	-3	183	237
$B_1 B_4 B_3 B_2$	+3	+3	-3	-3	194	250
$B_2 B_1 B_3 B_4$	+3	+3	-3	-3	184	237
$B_2 B_1 B_4 B_3$	+3	+3	-3	-3	188	241
$B_2 B_3 B_1 B_4$	+3	+3	-3	-3	199	270
$B_2 B_3 B_4 B_1$	+3	+3	-3	-3	266	348
$B_2 B_4 B_1 B_3$	+3	+3	-3	-3	207	287
$B_2 B_4 B_3 B_1$	+3	+3	-3	-3	270	358
$B_3 B_1 B_2 B_4$	+3	+3	-3	-3	198	263
$B_3 B_1 B_4 B_2$	+3	+3	-3	-3	211	279
$B_3 B_2 B_1 B_4$	+3	+3	-3	-3	209	289
$B_3 B_2 B_4 B_1$	+3	+3	-3	-3	274	362
$B_3 B_4 B_1 B_2$	+3	+3	-3	-3	245	354
$B_3 B_4 B_2 B_1$	+3	+3	-3	-3	293	406
$B_4 B_1 B_2 B_3$	+3	+3	-3	-3	210	284
$B_4 B_1 B_3 B_2$	+3	+3	-3	-3	220	295
$B_4 B_2 B_1 B_3$	+3	+3	-3	-3	224	315
$B_4 B_2 B_3 B_1$	+3	+3	-3	-3	283	381
$B_4 B_3 B_1 B_2$	+3	+3	-3	-3	251	364
$B_4 B_3 B_2 B_1$	+3	+3	-3	-3	298	415

5. CONCLUSIONS

In the paper, an efficient procedure for evaluating noise in OTA-C filters has been proposed based on the general model of OTA-C filter and its matrix description. The procedure has been verified by comparing with SPICE simulation proving its accuracy. The derived formulas can be applied to any known OTA-C filter architecture and can be easily implemented and used in computer-aided analysis/optimization systems. For the sake of illustration, two examples of noise optimization tasks for 5th and 8th order filters have been solved using the software based on the presented noise evaluation formulas. It should be emphasized that the presented approach can be effectively used to solving any noise-related optimization problems in OTA-C filters, which is especially because of the algebraic formulation of the underlying filter model.

This work was supported in part by the State Scientific Research Committee, Poland, under Grant 4T11B01625

6. REFERENCES

1. R. Schaumann, M. S. Ghausi, K. R. Laker: *Design of Analog Filters, Passive, Active RC, and Switched Capacitor*. Englewood Cliff, NJ: Prentice-Hall, 1990.
2. T. Deliyannis, Y. Sun, J. K. Fidler: *Continuous-time active filter design*. CRC Press, USA, 1999.
3. R. L. Geiger, E. Sánchez-Sinencio: *Active filter design using operational transconductance amplifiers: A tutorial*. IEEE Circuit and Devices Mag. Vol. 1, 1985, pp. 20-32.
4. B. Nauta: *Analog CMOS filters for very high frequencies*. Kluwer Academic Publishers, 1993.
5. Y. Sun (Editor): *Design of high frequency integrated analogue filters*. The Institution of Electrical Engineers, London, 2002.
6. E. Sánchez-Sinencio, J. Silva-Martinez: *CMOS transconductance amplifiers, architectures and active filters: a tutorial*. IEE Proc.-Circuits Dev. Syst., Vol. 147, No. 1, 2000, pp. 3-12.
7. S. Kozieł, S. Szczepański: *Structure Generation and Performance Comparison of Canonical Elliptic G_m -C Filters*. Proc. Int. Conf. Electron. Circuits, Syst., ICECS, Vol. I, Dubrovnik, Croatia, 2002, pp. 157-160.
8. S. Kozieł, S. Szczepański, R. Schaumann: *General Approach to Continuous-Time G_m -C Filters*. Int. J. Circuit Theory Appl., Vol.31, July/Aug. 2003, pp. 361-383.
9. Y. P. Tsividis: *Integrated continuous-time filter design — an overview*. IEEE J. Solid-State Circuits, vol. 29, 1994, pp. 166-176.
10. P. K. D. Pai, and A. A. Abidi: *A 40-mW 55 Mb/s CMOS Equalizer for Use in Magnetic Storage Read Channels*. IEEE J. Solid-State Circuits, vol. 29, Mar. 1994, pp. 489-499.
11. H. Khorramabadi, M. J. Tarsie, and N. S. Woo: *Baseband filters for IS-95 CDMA receiver applications featuring digital automatic frequency tuning*. in Int. Solid State Circuits Conf., San Francisco, 1996, pp. 172-173.
12. W. Dehaene, M. S. J. Steyaert, and W. Sansen: *A 50-MHz Standard CMOS Pulse Equalizer for Hard Disk Read Channels*. IEEE J. Solid-State Circuits, vol. 32, July 1997, pp. 977-988.
13. A. Hassan, K. Sharaf, H. El-Ghitani, H. F. Ragai: *The Design and Implementation of a Bandpass G_m -C Filter for Bluetooth*. Proc. Midwest Symp. on Circuits Syst., MWSCAS, Vol. 2, 2002, pp. 629-632.

14. S. Mehrmanesh, M. Atarodi: *A High Dynamic Range CMOS Variable Gain Filter for ADSL*. Proc. Int. Symp. Circuits Syst. Circuits and Systems, ISCAS, Vol. 4, 2002, pp. 257-260.
15. S. Mehrmanesh, H. A. Aslanzadeh, M. B. Vahidfar, M. Atarodi: *A 1.8V High Dynamic Range CMOS G_m -C Filter for Portable Video Systems*. Proc. Int. Conf. Microelectronics, ICM, 2002, pp. 38-41.
16. W. Jendernalik, S. Szczepański: *A CMOS OTA-C Channel-Select Filter for Mobile Receiver*. Proc. Int. Conf. Circuits Syst. Comm., ICCSC, 2002, pp. 50-53.
17. C. D. Salthouse, R. Sarpeshkar: *A Practical Micropower Programmable Bandpass Filter for Use in Bionic Ears*. IEEE J. Solid-State Circuits, Vol. 38, No. 1, 2003, pp. 63-70.
18. A. A. Emira, E. Sanchez-Sinencio: *A Pseudo Differential Complex Filter for Bluetooth With Frequency Tuning*. IEEE Trans. Circuits Syst. II, Vol. 50, No. 10, 2003, pp. 742-754.
19. S. D'Amico, A. Baschirotto: *0.18 μ m CMOS gmc Digitally Tuned Filter for Telecom Receivers*. Proc. Int. Conf. Circuits Syst., ISCAS, Vol. 1, 2003, pp. 493-496.
20. J. Shin, S. Min, S. Kim, J. Choi, S. Lee, H. Park, J. Kim: *3.3-V Baseband G_m -C Filters for Wireless Transceiver Applications*. Proc. Int. Conf. Circuits Syst., ISCAS, Vol. 1, 2003, pp. 457-460.
21. S. Mehrmanesh, M. B. Vahidfar, H. A. Aslanzadeh, M. Atarodi: *An Ultra Low-Voltage G_m -C Filter for Video Applications*. Proc. Int. Conf. Circuits Syst., ISCAS, Vol. 1, 2003, pp. 561-564.
22. Y. Palaskas, Y. Tsvividis: *Dynamic Range Optimization of Weakly Nonlinear, Fully Balanced, G_m -C Filters With Power Dissipation Constraints*. IEEE Trans. Circuits Syst.-II, Vol. 50, No. 10, Oct. 2003, pp. 714-727.
23. A. Brambilla, G. Espinosa, F. Montecchi, E. Sánchez-Sinencio: *Noise optimization in operational transconductance amplifier filters*. in Proc. Int. Symp. Circuits Syst. ISCAS, Vol. 1, 1989, pp. 118-121.
24. G. Ejthivoulidis, L. Toth, Y. P. Tsvividis: *Noise in G_m -C Filters*. IEEE Trans. Circuits Syst., Vol. 45, No. 3, Mar. 1998, pp. 295-302.
25. A. S. Korotkov, R. Unbehauen: *Calculation and Estimation of Noise in All-Pole G_m -C Filters*. Int. J. Electron. Commun., Vol. 53, No. 3, 1999, pp. 151-154.
26. K. A. Mezher, P. Bowron: *Noise-Flow-Graph Analysis of OTA-C Filters*. Proc. Int. Symp. Circuits Syst. ISCAS, Vol. 1, 2001, pp. 695-698.
27. S. Kozieł, S. Szczepański: *Dynamic Range Comparison of Voltage-Mode and Current-Mode State-Space G_m -C Biquad Filters in Reciprocal Structures*. IEEE Trans. Circuits Syst.-I, Vol. 50, No. 10, 2003, pp. 1245-1255.
28. S. Kozieł, S. Szczepański: *General G_m -C Filters with Finite-Band Transconductors*. Proc. European Conf. Circuit Theory and Design, ECCTD, Vol. II, 2003, pp. 293-296.
29. T. Kaczorek: *Wektory i macierze w automatyce i elektrotechnice*. Wydawnictwa Naukowo-Techniczne, Warszawa, 1998.

S. KOZIEŁ

ANALIZA I OPTYMALIZACJA SZUMÓW FILTRÓW OTA-C CZASU CIĄGŁEGO

Streszczenie

W pracy przedstawiono efektywną procedurę wyznaczania szumów w filtrach OTA-C czasu ciągłego. Procedura wyprowadzona jest w oparciu o ogólny model filtru OTA-C oraz jego opis macierzowy. Podano jawne formuły pozwalające określić odniesione do wejścia widmo szumów dowolnego filtru rozważanej

klasy. Jako ilustrację, przedstawiono „ręczną” analizę szumów filtru dolnoprzepustowego drugiego rzędu opartą o wyprowadzone formuły.

Otrzymane wyniki teoretyczne zostały zweryfikowane poprzez porównanie z wynikami otrzymywanymi za pomocą symulatora SPICE dla wybranych przykładów filtrów dolnoprzepustowych trzeciego oraz ósmego rzędu. Rolę elementu aktywnego w filtrach pełnił prosty transkontaktor CMOS wykorzystujący parę różnicową. Pokazano, że charakterystyki szumowe badanych filtrów otrzymane przy użyciu rozważanego ogólnego modelu filtru OTA-C są zgodne z wynikami symulacji komputerowych na poziomie tranzystorowym.

Sformułowanie modelu w języku algebry liniowej umożliwia bezpośrednią implementację wyprowadzonych formuł i użycie we wspomaganej komputerowo optymalizacji szumów filtrów OTA-C. Jako ilustrację, przedstawiono zastosowanie modelu do rozwiązania wybranych problemów: znajdowania struktur filtrów z wielopętlowym sprzężeniem zwrotnym o minimalnych szumach na przykładzie filtrów realizujących funkcje przenoszenia Butterworth’a oraz Bessela piątego rzędu oraz optymalizacji szumów filtru Butterworth’a ósmego rzędu w realizacji kaskadowej.

Przedstawiony model może być zastosowany w komputerowych systemach wspomagania projektowania filtrów klasy OTA-C do rozwiązywania problemów związanych z optymalizacją parametrów szumowych oraz zakresu dynamiki.

Słowa kluczowe: filtry OTA-C, analiza szumowa, optymalizacja filtrów

Software Implemented Fault Detection And Fault Tolerance Mechanisms — PART II: Experimental evaluation of error coverage

PIOTR GAWKOWSKI, JANUSZ SOSNOWSKI

*Institute of Computer Science, Warsaw University of Technology,
ul. Nowowiejska 15/19 Warsaw Poland,
Email: jss@ii.pw.edu.pl*

Otrzymano 2004.12.10

Autoryzowano 2005.03.11

The paper deals with the problem of evaluating the impact of hardware faults on program execution. For this purpose we use software implemented fault injector supplemented with various statistical tools. It simulates faults by disturbing CPU registers and memory cells in the analysed system. Using this tool we analyse fault effects in ordinary applications and in applications with embedded fault detection and fault tolerance mechanisms. The paper presents experimental results for a wide spectrum of applications and different fault hardening techniques. It outlines some critical problems and discusses the effectiveness of the proposed solutions.

Keywords: fault injection, fault detection, fault tolerance, fault coverage

1. INTRODUCTION

Many applications of digital systems require high dependability [2]. In part I of the paper [13], we discussed software methods of achieving this goal. Another important issue is the evaluation of the effectiveness of fault handling procedures. For this purpose various tools have been proposed which rely either on fault (physical or logical) simulation in the system model or in the real system [1,7,15,22]. Amongst them, the most universal are software implemented fault injectors (SWIFI).

Various fault injection experiments have been performed to demonstrate the effectiveness of fault hardening methods (e.g. [6-9,16,18-20,23] and the references therein). Unfortunately, most of them were limited to a few simple applications and the presented reports were not sufficiently detailed to reveal many problems appearing in real world.

While analysing effectiveness of fault tolerant techniques, most authors give experimental results averaged over all injected faults. Unfortunately, our experiments have shown that this approach does not reveal weak points of these techniques, which is especially important in critical applications. In this paper, we show the impact of various program features which influence fault susceptibility (e.g. resource activity, compiler specificity, operational and test profiles). In the performed experiments, we have used a software implemented fault injector [21] developed by us (section 2), which delivers various original statistics and fault tracing procedures unavailable in other tools. These capabilities were very useful in the performed analysis. We have analysed many applications, checked their natural susceptibility to faults (section 3) and implemented various software fault detection and fault tolerance mechanisms (according to the concepts described in part I [13]). Several fault-hardened applications were subjected to fault injection experiments described in section 4. These experiments were targeted not only for the dependability evaluation but also for tracing fault effects in the system. This last issue has allowed us to get better insight into real problems and to find weak points in fault handling strategies. We have also analysed fault susceptibility of the additional code in improved applications (section 5). This is especially crucial in methods with frequent fault checking (finer-grained approaches), as we have found that special care is needed at the implementation level to avoid the risk of unnoticed incorrect results.

2. EXPERIMENTAL DEPENDABILITY ASSESSMENT

In fault injection experiments we have verified the effectiveness of various approaches to improve system fault robustness with software procedures. In all experiments we use our fault injector FITS [21]. FITS simulates faults in the Windows environment by disturbing the states of processor registers or RAM locations (storing program code and data). A fault type (bit flip, bit setting, resetting and bridging), the duration, triggering moments and location can be specified explicitly or generated pseudorandomly. The presented experimental results deal with transient faults (pseudorandomly generated bit flips) which are most critical in real applications [6,14,16].

FITS compares the generated states of registers with those found during the golden run time. Also, it registers the exit code, results and all generated events and exceptions. In general, we distinguish 5 classes of test results: C — correct result, INC — incorrect result, S — fault detected by the system, T — time-out, U — user messages (generated by the program if an error is detected). More detailed qualification of test results is also possible. System exceptions (S) are mostly generated by special hardware mechanisms embedded in contemporary COTS (commercial off-the-shelves) systems. Microprocessors signal such exceptions as: access violation (within RAM), in page error, array bounds exceeded, data type misalignment (wrong word boundaries), illegal instruction, etc. [21]. FITS delivers also various statistics related to distribution of fault injections, system resource activity etc.

In fault injection experiments, it is important to specify test scenarios which cover fault distribution in time and space (fault localisation and triggering), fault types, input data profiles etc. The experiments performed by us were targeted for transient faults (bit flips) injected into registers (specified CPU or FPU registers, or all of them), the code or data area of the memory. By concentrating on specified system resources (or code segments), we can perform deeper analysis and tune appropriate fault handling mechanisms. The number of injected faults is sufficiently large (typically $10^4 - 10^6$ faults) to assure statistical significance of the obtained results. For each application, we choose a representative set of input data, to assure high coverage of the code, decisions etc. This selection is based on the analysis of some coverage measures (block, decision, c-use coverage etc. [5,9]).

To get a better insight into the analysed problems, we concentrate on three classes of experiments: checking natural fault susceptibility for a wide spectrum of applications, checking the effectiveness of fault hardened applications and checking fault susceptibility of fault handling procedures. These experiments allow us to identify various features influencing fault susceptibility as well as "fault leakage" sources.

3. EXPERIMENTS WITH NON-HARDENED APPLICATIONS

Having analysed many applications, we found that they differ quite significantly in fault susceptibility which in particular, depends upon the activity of used resources, input data, compiler features etc. In tab.1 we give fault injection results for different applications and different fault locations (register, data or code area). For each category of results (C, INC, S and T), we highlight the minimal (underlined) and the maximal (in bold) values. The applications A1-A4 relate to various sorting algorithms (simple bubble, by selection, modified bubble and optimal bubble), A5 — calculation of the CRC code, A6-A8 cryptography algorithms (IDEA, MD5, SHA1), A9 — benchmark of operations on double linked ring. The presented results relate to single bit flips injected into CPU (Pentium) registers, the memory data area (used by the application) and the code area.

Tab. 1 shows quite big dispersion (defined by underlined and bold figures in each column) of fault effects depending upon application. This is mostly related to resource activity and code specificity. Higher fault robustness (correct results) is observed for faults injected into registers and data (it will be explained in the sequel). Exceptions are mostly generated by faults injected into some registers or the code and most of them are related to access violation. Timeouts practically appear for faults in the code only. More detailed analysis of the fault propagation and sources of natural fault robustness are given in [12].

Using different algorithms for the same computational problem (e.g. sorting in A1-A4) have led us to different results (correct results in the range 1.92-17.8% for data faults). It is worth noting that even the same algorithm implemented by different programmers or with the use of different languages may differ significantly in fault

susceptibility. Similarly, the same source code translated with different compilers also gives different results. For example, the faults injected into the data area of the quick sort algorithm compiled in ADA95, Borland C++ 5.5 and MsVC++6.0 resulted in 8%, 38% and 50% of incorrect results, respectively (for a fixed input data). In many applications the used input data set may influence program flow resulting in different fault susceptibility. For the above-mentioned quick sort program in MsC++6.0 faults injected into data area generated 38-50% of incorrect results depending upon the input data. In general, numerical applications are more susceptible to faults than. non-numerical

Table 1

Distribution of test results for selected applications [%]
Rozkład wyników testów dla wybranych aplikacji [%]

	Correct results — C			Incorrect results — INC			System exceptions — S			Time-out —T		
	Data	Code	Reg.	Data	Code	Reg.	Data	Code	Reg.	Data	Code	Reg.
A1	17.8	7.83	61	82.2	29.9	<u>4.13</u>	<u>0</u>	50.3	34.9	<u>0</u>	12	<u>0</u>
A2	<u>1.92</u>	5.1	43.9	97.5	24.5	11.9	0.6	66.2	44.2	<u>0</u>	4.26	<u>0</u>
A3	2.71	5.58	41.54	96.5	26.3	10.86	0.8	65.9	47.6	<u>0</u>	2.26	<u>0</u>
A4	10.58	9.52	42.46	86.7	<u>23.8</u>	10.24	2.6	65.5	47.3	<u>0</u>	1.13	<u>0</u>
A5	41.4	<u>3.36</u>	<u>27.5</u>	58.6	41.4	20.5	<u>0</u>	53.8	52.0	<u>0</u>	1.43	<u>0</u>
A6	74.7	14.9	39.5	21.8	35.1	14.8	3.5	47.1	45.6	<u>0</u>	2.93	0.1
A7	76.2	11.1	33.2	21.2	50	35.7	2.2	28.2	30.7	<u>0</u>	0.74	0.37
A8	67.9	7.33	35.1	28.8	39	27.8	3.4	53.6	37.1	<u>0</u>	<u>0.1</u>	<u>0</u>
A9	71.1	16.7	48.9	11.3	57	37	15	<u>16.1</u>	<u>13.1</u>	2.43	10.2	1.02

ones.

Table 2

Register susceptibility to faults (C — correct results, S — system exceptions) [%]
Wrażliwość rejestrów na błędy (C — wyniki poprawne, S — wyjątki systemowe) [%]

Applic.	EAX	EBX	ECX	EDX	ESP	EBP	ESI	EDI	EIP
LZW C	95.6	100	96.2	96.2	26.5	5.7	96.0	95.6	9.1
S	3.5	0	3.6	3.7	69.9	93.3	3.8	4.1	87.0
MIX1 C	60.1	82.7	92.3	98.1	15.8	3.5	53.9	100	3.8
S	0	00	0	1.5	73.9	96.1	0	0	90.8

In most publications authors give fault insertion results averaged over all injected faults e.g. [6,8,16,20]. Sometimes they give averaged figures for faults injected into the code, the data area or registers. It is worth noting that result averaging may hide many real problems. By inspection of the results from tab.1, we can find big differences in test results depending upon the area of fault injection (code, registers, data area). Table 2 shows register susceptibility to faults for the data compression algorithm LZW and a mixture of integer calculations (MIX1). Here faults injected into some registers strongly influence the program control flow (e.g. ESP, EBP, EIP) and generate large percentage of system exceptions. The same also holds for segment addressing registers (not shown in tab. 2).

Register sensitivity to faults depends upon its activity and the type of calculations performed. The register activity (AR) is the percentage of the execution time during which the register holds useful data (from its loading to the last read operation). The disturbed result (DR) ratio is defined as $DR = INC + S + T + U$. For calculation oriented applications with low redundancy we have good correlation between DR and AR, practically in the range of a few percent. In many applications DR is lower than AR due to more important impact of data relations than their values and algorithm specificity. For example, in the quick sort algorithm, EBX holds the index of the left most element of the table to sort and the activity ratio is 91% as opposed to 10% of the program disturbance ratio.

The utilization of some resources may be quite low which is typical e.g. for the floating point unit (FPU). For a sample of programs using FPU, the faults injected into the control CWD and status (SWD and TWD) registers resulted in 80-95% of correct results, incorrect results constituted 5-15% and system exceptions appeared mostly for faults injected into the CWD register (~ 5%). In simple algorithms only few stack FPU data registers are active, so, on average, in more than 95% cases injected faults had no impact on the results (correct). In complex calculation of an integral, we have obtained 50-99% of correct results for faults injected into FPU data. Also, stack registers closer to the bottom of the stack are more susceptible to faults because they store important data for a longer time period (high activity ratio).

A detailed knowledge of the fault susceptibility facilitates optimisation of hardening procedures. On the other hand, the averaged figures could give some estimation of the general fault susceptibility features. The correlation of these features with the physical fault susceptibility is more complex. For some faults (e.g. transient faults caused by cosmic rays), it can be assumed that faults are equally distributed in chip space and time. Some mapping can be found between experiment fault location distribution and chip areas. This may give better estimation of the real fault susceptibility (compare [6]). However, we should be conscious that SWIFIs do not disturb directly some resources like hidden registers (related to CPU microstructure and not accessible directly from the machine code). Similar problem arises for complex combinational logic. So the mapping from the physical fault distribution into the RTL program level model is more complex and needs further research (compare [6,14]).

Another problem relates to the dynamic use of resources in the application. In many applications some program segments are executed more frequently than others, so equal distribution of faults in memory address space may not well characterise the real fault susceptibility which mostly depends on the most frequently used program segments. This situation is encountered especially in various embedded applications [11]. Hence it is important to explicitly specify what kind of fault distribution is used: in space or in time; e.g. uniform distribution into dynamically executed instructions. Independent stressing (with faults) of different program segments, registers etc. enables deeper analysis of the fault propagation etc.

4. EXPERIMENTS WITH HARDENED APPLICATIONS

We have analysed the effectiveness of various software mechanisms increasing dependability of a system. For this purpose, we have embedded these mechanisms in many applications and performed fault injection experiments. Here we present selected results related to fault detection and fault tolerance mechanisms. We concentrate mostly on universal and systematic approaches. Nevertheless some atypical (application dependent) solutions will also be commented.

Table 3

Test results for applications with enhanced fault detection [%]

Wyniki testów dla aplikacji z wbudowanymi mechanizmami detekcji błędów [%]

	Correct results			Incorrect results			System exceptions			User exceptions		
	Data	Code	Reg.	Data	Code	Reg.	Data	Code	Reg.	Data	Code	Reg.
R1	44.5	19.4	45.6	19.1	8.7	2.7	0.0	45.7	39.1	36.3	25.9	38.4
R2	78.86	17.9	45.6	0.12	2.86	6.7	2.21	40.9	31.7	18.8	37.3	15.9
R3	52.1	21.4	63.7	15.6	0.8	0.0	0.25	51	31.8	32	25.4	4.5
ED1	45.7	6.8	56.7	0.8	0.0	1.8	1.8	50.6	28.3	51.7	37.6	13.1
ED2	265.6	13	57.2	3.8	1.7	0.0	0.65	46.3	28.8	30.0	30.8	14.0
AS	23.9	18.8	60.9	54.1	8.4	0.9	8.7	58.5	35.0	13.3	4.8	3.3

In tab. 3 we present test results for applications with software implemented fault detection mechanisms. The applications R1, R2 and R3 relate, respectively, to bubble sorting, Ar tgh calculation ($Ar\ tgh = 0.5 \ln 2 \log_2[(1+x)/(1-x)]$ with $\log_2$ calculated according to Mitchell method [17]) and data compression according to the LZW algorithm. All these applications were enhanced by modifying the primary programs according to fine-grained approach: variable duplication rules (VDR, see section 2,

Table 4

Test results for applications with embedded fault tolerance mechanisms [%]

Wyniki testów dla aplikacji z wbudowanymi mechanizmami tolerowania błędów [%]

	Correct results			Incorrect results			System exceptions			User exceptions		
	Data	Code	Reg.	Data	Code	Reg.	Data	Code	Reg.	Data	Code	Reg.
T1	100	82.4	78.9	0.0	0.4	1.0	0.0	12.8	16.1	0.0	4.3	3.9
T2	97.6	93.8	90.0	2.4	2.0	1.1	0.0	3.9	8.9	0.0	0.0	0.0
T3	99.7	98.2	93.0	0.0	0.0	0.0	0.3	1.8	7.0	0.0	0.0	0.0
RC	99.3	28.4	70.5	0.5	0.19	0.2	0.1	42	26.9	0.0	25.7	2.4
MA	36.1	2.1	28.4	0.0	0.5	0.0	6.3	72.8	59.1	57.6	22.8	12.5
MB	51.6	86.5	88.2	0.4	1.0	0.0	0.4	5.5	11.0	47.6	7.0	0.8

part I [13]). The applications ED1 and ED2 relate to bubble sorting and Ar tgh calculations enhanced according to algorithm ED⁴I (section 2 — part I, coarse-grained approach). Finally, AS is a quick sort program with embedded assertions on correct table indices and final monotonicity of ordered output data.

Deeper analysis of the results revealed some important aspects which should be taken into account in the design process, in particular, the problems of fine- vs. coarse-grained fault handling, dealing with exceptions and recovery. The coarse-grained error detection assures the lowest percentage of incorrect results. The fine-grained approach based on VDR rules is less effective for faults injected into code or registers (due to higher overhead of fault detection code — susceptible to faults). Quite high percentage of incorrect results (faults injected in data area) for applications R1 and R2 is caused by not checking the final result. Obviously, this can be easily improved by adding such checks (not included in the primary rules). The results for the application AS show that a simple intuitively generated assertion may not be effective. In all applications, significant percentage of faults has been detected by system exceptions (mostly access violation), especially for faults injected into the code and registers, this relates to our observations for not hardened applications (section 2).

While analysing the effectiveness of fault detection mechanisms, it is interesting to evaluate the attenuation ratio of incorrect results ($A_{INC} = (\text{incorrect results for the improved application})/(\text{incorrect results for the primary application})$). For the applications considered, we have obtained: $A_{INC}(R1) = (0.23, 0.29, 0.65)$, $A_{INC}(R2) = (0.01, 0.01, 0.45)$, $A_{INC}(R3) = (0.91, 0.00, 0.0)$, $A_{INC}(ED1) = (0.015, 0.0, 0.12)$, $A_{INC}(ED2) = (0.30, 0.057, 0.00)$, $A_{INC}(AS) = (0.82, 0.53, 0.1)$. The figures in the parentheses define attenuation ratio A_{INC} for faults injected into data, code and registers, respectively.

More correct results could be assured by embedding fault tolerance mechanisms, like those presented in tab. 4:

- simple triple modular redundancy (TMR in the code and data) based on the final result voting, i.e. the coarse-grained approach (the applications T1 and T2 related to the CRC calculation and the cryptographic algorithm DES);
- more sophisticated TMR with three processing threads and retries according to the idea described in part I [13] (the application T3 with CRC calculation);
- modification of the primary application according to error detection and error correction rules (ECR) described in section 2 (the application RC is LZW data compressing);
- application specific error correction (MA is the matrix multiplication with checksums in rows and columns and error correction, MB is enhanced MA by inclusion of the code retry on detected errors).

All these techniques assure relatively low percentage of incorrect results, but their fault tolerance capabilities are different. The highest percentage of correct results is assured by TMR coarse-grained redundancy. Nevertheless, for a short application T1, the faults injected into code are tolerated only at 82.4% (due to relatively high contribution of the voting code). Similarly for fine-grained fault tolerance technique used in the RC application, we obtained very low percentage of correct results for faults injected into the code. The best fault tolerance has been achieved in the application T3, wherein we applied efficient fault handling procedures described in part I [13]. In the application MA, fault tolerance is based on correcting the results by taking into account the row and column checksums. This technique, considered in the literature as efficient in hardware implementation, does not increase fault tolerance capabilities if implemented in software (especially for faults injected into the code). This can be improved by including the code to perform retries initiated by detected faults (version MB). In all cases we observe some limitation on the fraction of correct results for faults injected into registers. This is because the faults injected into some registers (e.g. EIP, ESP, segment registers) generate system exceptions which could be used to correct errors, provided that a system is equipped with a more effective mechanism of fault handling.

Table 4 shows the average ratio of correct results in the code, data and register areas. More detailed analysis shows that different resources have different fault tolerance capabilities. In particular, table 5 provides the distribution of correct results for registers in two applications: the data compression LZW hardened with ECR rules and a mixture of integer calculations (MIX1) with coarse-grained TMR. It seems that faults in some registers (ESP, EBP and EIP) cause errors difficult to correct (they generate mostly system exceptions, compare section 3). To analyse the effectiveness of fault tolerance capabilities, we should compare the ratio (FR) of correct results for the primary application and its fault-hardened version. For the applications considered, we have achieved the following values of fault tolerance ratio $FR(A,L)$, where A specifies the application, L specifies the fault location: D, C and R respectively for data, code

and registers:

$$\begin{aligned} \text{FR}(\text{T1}, \text{D}) &= 2.42, \text{FR}(\text{T1}, \text{C}) = 24.5, \text{FR}(\text{T1}, \text{R}) = 2.87; \\ \text{FR}(\text{T2}, \text{D}) &= 1.4, \text{FR}(\text{T2}, \text{C}) = 13.43, \text{FR}(\text{T2}, \text{R}) = 2.66; \\ \text{FR}(\text{T3}, \text{D}) &= 2.35, \text{FR}(\text{T3}, \text{C}) = 59.16, \text{FR}(\text{T3}, \text{R}) = 2.66; \\ \text{FR}(\text{RC}, \text{D}) &= 1.21, \text{FR}(\text{RC}, \text{C}) = 6.05, \text{FR}(\text{RC}, \text{R}) = 1.07; \\ \text{FR}(\text{MA}, \text{D}) &= 1.12, \text{FR}(\text{MA}, \text{C}) = 0.8, \text{FR}(\text{MA}, \text{R}) = 0.92; \\ \text{FR}(\text{MB}, \text{D}) &= 1.64, \text{FR}(\text{MB}, \text{C}) = 6.05, \text{FR}(\text{MB}, \text{R}) = 2.81; \end{aligned}$$

More detailed comparison needs inclusion of the fault distribution characteristics and the fault hardening overhead. For example, if we assume the equal distribution of faults in the circuit cross section area, we should add to the non hardened version correct results related to the space overhead of the hardened version.

It is important to note that, in general, fine-grained fault tolerance mechanisms show lower fault correction capabilities than their coarse-grained counterparts. It is due to relatively high code overhead of voting and fault handling. For the application MIX1 with fine-grained TMR, we achieved 92.5%, 57.9% and 71.3% of correct results for faults injected into the data, code and register area, respectively. For the coarse-grained TMR, they increased to 95.6%, 82.2% and 82.7%.

Table 5

Register susceptibility to faults in fault tolerant applications [%]

Wrażliwość rejestrów na błędy w aplikacjach z tolerancją błędów [%]

Applic.	EAX	EBX	ECX	EDX	ESP	EBP	ESI	EDI	EIP
LZW C	76.4	100	99.4	99.4	86.3	0.4	100	100	7.2
INC	0.17	0	0.0	0.1	0.1	0.0	0	0	0.6
MIX1 C	87.1	94.4	99.2	100	22.8	46.9	84.9	100	54.4
INC	12.7	5.6	0.7	0	13.9	2.8	15.1	0	5.8

5. EXPERIMENTS WITH FAULT HANDLING PROCEDURES

The fault insertion results given in tab.1 show that quite significant number of faults may cause incorrect results that go undetected. This situation can be improved with various fault detection and fault tolerance mechanisms (tab. 4-5). However, for some applications or mechanisms used, we can still observe "fault leakage" effect (not detected). Embedded fault handling procedures may occupy a significant part of the primary application, so their fault robustness is of great importance. This is especially critical, if these procedures (e.g. comparison, voting, checksum or code verification)

are activated frequently (fine granularity). To have better insight in this problem, we have arranged special experiments targeted to evaluate these mechanisms. Here we will present results for three of them: TMR with voting, CRC calculation and fault discrimination.

For software implemented voting procedure, extensive tests with different input data gave 65-71% and 54-57% correct results for faults injected into registers and the code, respectively. The percentage of generated exceptions was 38-41%. For a typical CRC checking procedure (used in many applications to detect errors), we found on average 40%, 52% and 15% incorrect results for faults injected into the data, code and registers, respectively. After hardening this procedure with the TMR scheme, significant improvements were observed: 99.7%, 98.2% and 93% of correct results for faults injected into the data, code and registers, respectively.

Table 6

Compiler impact on test results [%]

Wpływ kompilatora na wyniki testów [%]

Fault location		Test results			
		C	INC	S	TO
Registers	A	84.9	1.6	13.0	0.5
	B	94.2	0.4	4.9	0.5
Data	A	96.4	2.5	0.8	0.4
	B	94.1	3.6	1.2	1.1
Code	A	70.7	18.1	8.6	2.6
	B	89.1	0.8	5.4	4.7

In systems using software redundancy (TMR, DMR), one must be careful with compilers. In particular, compiling with optimisation may eliminate redundant functions or code segments, resulting in lower fault tolerance. We illustrate this for a program finding the minimum for a complex function, hardened by using the TMR structure (about 1KB of code). Table 6 comprises the fault injection results for the optimal (A) and non-optimal (B) compilation of this TMR program. Due to elimination of about 15% of the code redundancy in version A, we have got significant percentage of incorrect results (18%) for faults injected into the code (for version B, it was only 0.8%). The percentage of correct results can be improved by more efficient recovery of system exceptions and time-outs. Typically, the critical fault tolerance mechanisms contribute only a few percent of the code. So, if they are hardened, we can reduce incorrect results to 0-0.5%.

In handling faults, it is important to identify permanent, transient and intermittent faults. In [4] (compare part I [13]), quite an efficient algorithm (double threshold) has been proposed. We have analysed its fault robustness [10] for various test scenarios.

The injected faults into registers, data and instructions caused 8-13%, 1-25% and 13-30% incorrect results, respectively. We have enhanced this algorithm by making use of C++ class that triplicates internal data, duplicating code and exception handling. An exception switches the application to the redundant code. The enhanced version developed by us practically does not produce incorrect results (less than 1%), and the correct results were increased to the level of 82%, 30% and 55% for faults injected into the data, code and registers. Not recovered detected errors were signalled by system and user messages. Although the software hardening process could be applied iteratively, eventually we will reach some limits resulting from the critical code and data areas, which will be always sensitive to injected faults — not covered by the hardening mechanisms used.

In part I [13], we have outlined the problem of critical situations related to disturbing the program stack by faults. To lower the probability of such situations, we can perform triplicated pushing registers ESP, EBP and EIP on the stack and then check the popped values and select the correct ones by voting. For some test application, we were able to increase correct results for faults injected into the stack from 54.1 to 81.2% (incorrect results dropped down from 1.34 to 0.0%).

Table 7

Handling code errors [%]			
Obsługa błędów kodu [%]			
Test Result	Program version		
	V1	V2	V3
C	12.1; 0.1	53.2; 6.1	83.2; 91.6
C?	0; 0	21.1; 43.1	0; 0
C1	0; 0	16.1; 3.8	10.9; 2.7
U	0, 0	0; 38.7	0; 0
S+T	56.3; 89.5	9.5; 8.3	5.8; 5.7

Another problem relates to detection and recovering from code errors. In systems with error detection or error correction codes, significant amount of such errors can be handled directly by reloading the code or correcting it. If the code is protected only in the cache or RAM memory, the problem is more complicated, similarly as in the case of no protection. In such a situation, we have to rely on code checksums at the software level. We have considered two approaches: the vertical and horizontal parity checksum generated for every code block (e.g. 32 bytes) and the CRC checksum for each block. In the first case, all single errors can be corrected. In the second case, the correction possibilities are a little bit higher (if not successful, we have to reload the code). The effectiveness of these approaches is illustrated in tab. 7 for an application realizing quite complex iterative calculation (finding the n-th prime number). The faults have

been injected into the program code. The results for single and multiple bit flip faults are presented using roman and italic font (after semicolon), respectively. They provide a good insight into the anatomy of error handling processes.

The table shows results for three versions of the program: version V1 does not comprise any code recovery; version V2 corrects errors detected by horizontal and vertical parity bits; version 3 relies on program checksum and recovery by code re-loading. In the experiments, we have identified three classes of correct results: C, C? and C1. C relates to correct results, obtained directly due to the natural fault resistivity of the application (section 3) or due to error corrections (versions V2, V3). C? relates to correct results but with signalled parity errors caused by disturbed parity checking procedures. C1 corresponds to errors corrected by checksum verification at the end of the execution of the program. A relatively high value of C? for version V2 results from significant time overhead in parity checking. Lowering the frequency of parity checking (at the cost of higher error latency) can decrease this. The low value of direct correct results (C) for multiple faults is due to the fact that the parity bits allow to correct single faults only. Some exceptions (S) skip *try-catch* mechanisms (see part I [13]) and are collected at the system level. Such errors can also be corrected by reinitiating the calculation from the system level, but however, it could be quite complicated (like in Windows). Practically no undetected incorrect results appeared for versions V2 and V3. For version V2 incorrect results were signalled as user messages (U) in 38.7% cases of multiple faults. For the basic version V1 we have got 31.7% and 10% of not signalled incorrect results for single and multiple faults. In practice, we can encounter systems with parity embedded in caches or RAM memories. An error detected in caches can be easily corrected by cache invalidation (only the specified line or the whole cache). This scheme is effective for instruction and data write through caches. Hence it is possible to optimise time losses during recovery procedures.

6. CONCLUSION

The paper shows that the effectiveness of COTS error detection mechanisms depends upon various application features: resource activity, code specificity, algorithms and compilers used, etc. It can be improved using software techniques, so that a significant percentage of faults can be corrected or masked. It is important to take over system exceptions in the application code (embedded fault-handling procedures). Some operating systems (e.g. Windows) impose various restrictions, so an additional recovery procedure may be needed at the system level.

Our analysis of fault susceptibility of error handling procedures has revealed some possibilities of making them more robust. This can be achieved by optimisation of resource activities at the compiler level as well. The paper confirms that detailed analysis of fault effects performed using fault injector (supplemented with various statistics) facilitates the optimisation of fault hardening procedures. Many software techniques cannot detect faults that cause infinite loops, system hang-ups etc. (disturbed

program flow may skip embedded fault detection procedures). Such faults need some additional hardware detectors like watchdog timers or more sophisticated watchdog processors.

Since fault injection experiments performed on complex applications could be time consuming, some optimisations could be required. We can limit experiments to fault handling procedures or inject faults only during the active periods of resources (or even just before the read operations [3]). To get the overall dependability evaluation, we have to scale the obtained fault injection results by taking into account the ratio of activity periods etc.

This work was supported by KBN grant no. 4T11C049 25.

7. REFERENCES

1. J. Arlat, Y. Crouzet, et al.: *Comparison of physical and software implemented fault injection techniques*. IEEE Trans. on Computers, vol. 52, no. 9, 2003, pp. 1115-1135.
2. A. Avizienis, J.-C. Laprie, B. Randel, and C. Landwehr: *Basic concepts and taxonomy of dependable and secure computing*. IEEE Trans. on Dependable and Secure Computing, vol.1, no. 2, Jan.- Mar. 2004, pp. 11-33.
3. L. Berrojo, et al.: *New techniques for speeding up fault injection campaigns*. Proc. of IEEE Conference DATE 2002, pp. 847-852.
4. A. Bondavalli, et al.: *Threshold-based mechanisms to discriminate transient from intermittent faults*. IEEE Trans. on Computers, vol. 49, no. 3, March 2000, pp. 230-245.
5. L. C. Briand and D. Pfahl: *Using simulation for assessing the real impact of test coverage on defect coverage*. IEEE Trans. on Reliability, vol. 49, no. 1, March 2000, pp. 60-70.
6. G. C. Cardarilli, F. Kaddour, et al.: *Bit flip injection in processor based architectures: a case study*. Proc. of 6th IEEE On-Line Testing Workshop, 2002, pp. 117-128.
7. J. Carreira, H. Madeira and J. G. Silva: *Xception: a technique of the experimental evaluation of dependability in modern computers*. IEEE Trans. on Software Engineering, vol. 24, no. 2, February 1998, pp. 125-136.
8. P. Cheynet, et al.: *Experimentally evaluating an automatic approach for generating safety critical software with respect to transient errors*. IEEE Trans. on Nuclear Science, vol. 47, no. 6, Dec. 2000, pp. 231-236.
9. P. Folkesson and J. Karlsson: *Considering workload input variations in error coverage considerations*. Proc. of EDCC-3, Springer Verlag, LNCS, vol. 1667, 1996, pp. 171- 188.
10. P. Gawkowski and J. Sosnowski: *Experimental evaluation of fault handling mechanisms*. Proc. of 20th Int. Conference SAFECOMP, Springer Verlag, LNCS, vol. 2187, 2001, pp. 109-118.
11. P. Gawkowski and J. Sosnowski: *Evaluation of fault effects in programmable micro-controllers*. Proc. of IFAC Workshop on Programmable Devices and Systems, Pergamon Press, 2001, pp. 127-132.
12. P. Gawkowski and J. Sosnowski: *Evaluation of transient fault susceptibility in micro-processor systems*. Proc. of Digital System Design Euromicro Symposium, IEEE Comp. Soc. 2004, 2004, pp. 432-439.
13. P. Gawkowski and J. Sosnowski: *Software Implemented Fault Detection and Fault Tolerance Mechanisms, Part I: Concepts and algorithms*. Kwartalnik Elektroniki i Telekomunikacji, nr 2, 2005, ss. 291-303.

14. S. Kim and A. K. Somani: *Soft Error Sensitivity Characterisation for Microprocessor Dependability Enhancement Strategy*. Proc. of IEEE Int. Conference on Dependable Systems and Networks, 2002, pp. 416-428.
15. R. Leveugle: *Fault injection in VHDL description and emulation*. Proc. of IEEE DFT Symposium, 2000, pp. 414-419.
16. H. Madeira, R. R. Some, F. D. Costa and D. Rennels: *Experimental evaluation of a COTS system for space applications*. Proc. of IEEE Int. Conference on Dependable Systems and Networks, 2002, pp. 325-330.
17. J. N. Mitchell: *Computer multiplication and division using binary logarithms*. IRE Trans. on Electronic Computers, August 1962, pp. 512- 517.
18. B. Nicolescu, R. Velazco and M. S. Reorda: *Effectiveness and limitations of various software techniques for soft error detection, a comparative study*. Proc. of 7th IEEE Int. On-Line Testing Workshop, 2001, pp. 172-177.
19. M. Rebaudengo and M. S. Reorda: *Evaluating the fault tolerance capabilities of embedded systems via BDM*. Proc. of IEEE VLSI Test Symposium, 1999, pp. 452-459.
20. M. Rebaudengo, M. S. Reorda and M. Violante: *A new software based technique for low cost fault tolerant application*. Proc. of Annual Reliability and Maintainability Symposium, 2003, pp. 25- 28.
21. J. Sosnowski, P. Gawkowski and A. Lesiak: *Software implemented fault inserters*. Proc. of IFAC PDS2003 Workshop, Pergamon Press, 2003, pp. 293-298.
22. F. Vargas, R. D. R. Fagundes, D. Barros and D. R. Brum: *Briefing a new approach to improve EMI immunity of DSP systems*. Proc. of IEEE Asian Test Symposium, 2003, pp. 468- 471.
23. R. Velazco and S. Rezgui: *Assessing the soft error rate of digital architectures devoted to operate in radiation environment: a case study*. Journal of Electronic Testing, 19, 2003, pp. 83-90.

P. Gawkowski, J. Sosnowski

PROGRAMOWE MECHANIZMY DETEKCJI I TOLEROWANIA BŁĘDÓW SPRZĘTU — CZĘŚĆ II: EKSPERYMENTALNA ANALIZA POKRYCIA BŁĘDÓW

Streszczenie

Praca dotyczy analizy wpływu błędów sprzętu na wykonywanie się aplikacji programowych. Do tego celu użyto specjalnego symulatora błędów oraz odpowiednich modułów statystycznych. Symulator generuje błędy w badanym systemie poprzez modyfikację stanu rejestrów i komórek pamięci. Zachowanie się systemu zakłócanego błędami jest porównywane z zachowaniem wzorcowym (bez błędów). Symulator wykorzystano do charakteryzacji odporności na błędy aplikacji standardowych oraz ich wersji rozbudowanych o mechanizmy detekcji i obsługi błędów. Przedstawiono wyniki dla szerokiego spektrum aplikacji z różnymi mechanizmami uodparniania na błędy oraz wskazano na pewne krytyczne problemy i metody ich rozwiązania. Rozpatrzono głównie programowe mechanizmy detekcji oraz tolerowania błędów i systemowe mechanizmy generacji wyjątków. Mechanizmy te były omawiane w części pierwszej pracy.

Słowa kluczowe: symulator błędów, detekcja i tolerowanie błędów, pokrycie błędów

INFORMACJE DLA AUTORÓW

Redakcja przyjmuje do publikowania prace oryginalne, przeglądowe i monograficzne wchodzące w zakres szeroko pojętej elektroniki. Ponieważ KWARTALNIK ELEKTRONIKI I TELEKOMUNIKACJI jest czasopismem Komitetu Elektroniki i Telekomunikacji Polskiej Akademii Nauk, w związku z tym na jego łamach znajdują się prace naukowe dotyczące podstaw teoretycznych i zastosowań z zakresu elektroniki, telekomunikacji, mikroelektroniki, optoelektroniki, radiotechniki i elektroniki medycznej.

Artykuły powinny charakteryzować oryginalne ujęcie zagadnienia, własna klasyfikacja, krytyczna ocena (teorii lub metod), omówienie aktualnego stanu, lub postępu danej gałęzi techniki oraz omówienie perspektyw rozwojowych. Artykuły publikowane w innych czasopismach nie mogą być kierowane do druku w Kwartalniku Elektroniki i Telekomunikacji w drugiej kolejności zgłoszenia.

Objętość artykułu nie powinna przekraczać 30 stron po około 1800 znaków na stronie, w tym rysunki i tabele.

Wymagania podstawowe.

Artykuły należy nadsyłać na wyraźnym, jednostronnym, czarno-białym wydruku komputerowym. Wydruk w formacie A4 powinien mieć znormalizowaną liczbę wierszy i znaków w wierszu (30 wierszy po 60 znaków w wierszu), w dwóch egzemplarzach, w języku polskim lub angielskim wybranym przez autora. Do wydruku powinna być dołączona dyskietka z elektronicznym tekstem artykułu. Preferowane edytory to WORD 6 lub 8. Układ artykułu (w wersji podstawowej) musi być następujący:

- Tytuł.
- Autor (imię i nazwisko autora/ów).
- Miejsce pracy (nazwa instytucji, miejscowość, adres. + ew. adres elektroniczny (e-mail)).
- Zwięzłe streszczenie powinno być w języku takim, w jakim jest pisany artykuł (wraz ze słowami kluczowymi).
- Tekst podstawowy powinien mieć następujący układ:

1. WPROWADZENIE
2. np. TEORIA
3. np. WYNIKI NUMERYCZNE
- 3.1.
- 3.2.
4.
5.
6. PODSUMOWANIE
7. ew. PODZIĘKOWANIA
8. BIBLIOGRAFIA

- Układ streszczenia w dodatkowej wersji językowej powinien być następujący:

AUTOR (inicjał imienia i nazwisko).

TYTUŁ (w języku angielskim – o ile artykuł pisany jest w języku polskim i na odrwót).

Obszerne do 3600 znaków streszczenia (wraz z słowami kluczowymi) w języku:

a. angielskim, gdy artykuł pisany jest w języku polskim.

b. polskim, gdy artykuł pisany jest w języku angielskim.

Streszczenie to powinno pozwolić czytelnikowi na uzyskanie istotnych informacji zawartych w pracy. Z tego względu w streszczeniu tym mogą być cytowane numery istotnych wzorów, rysunków i tabel zawartych w podstawowej wersji językowej.

- Wszystkie strony muszą mieć numerację ciągłą.

Sposób pisania tekstu.

Tekst powinien być pisany bez używania wyróżnień, a w szczególności nie dopuszcza się spacjiowania, podkreślania i pisania tekstu dużymi literami z wyjątkiem wyrazów, które umownie pisze się dużymi literami (np. FORTRAN). Proponowane wyróżnienia Autor może zaznaczyć w maszynopisie zwykłym ołówkiem za pomocą przyjętych znaków adjustacyjnych, np. podkreślenie linią przerywaną oznacza spacjiowanie (rozstrzelenie), podkreślenie linią ciągłą – pogrubienie, podkreślenie wężymkiem — kursywa.

Tekst powinien być napisany z podwójnym odstępem między wierszami, tytuły i podtytuły małymi literami. Marginesy z każdej strony powinny mieć około 35 mm. Wielkość czcionki wydruku powinna być zbliżona co najmniej co wielkości czcionki maszyny do pisania (minimum 12 punktów). Przy podziale pracy na rozdziały i podrozdziały cyfrowe ich oznaczenia nie powinny być większe niż II stopnia (np. 4.1.1.).

Sposób pisania tabel.

Tabele powinny być pisane na oddzielnych stronach. Tytuły rubryk pionowych i poziomych powinny być napisane małymi literami z podwójnym odstępem między wierszami. Przypisy (notki) dotyczące tabel należy pisać bezpośrednio pod tabelami. Tabele należy numerować kolejno liczbami arabskimi, u góry każdej tabeli podać tytuł dwujęzyczny. W pierwszej kolejności w podstawowej wersji językowej, a później w dodatkowej wersji językowej. Tabele umieścić na końcu maszynopisu. Przyjmowane są tabele algorytmów i programy na wydrukach komputerowych. W tym przypadku zachowany jest ich oryginalny układ. Tabele powinny być cytowane w tekście.

Sposób pisania wzorów matematycznych.

Rozmieszczenie znaków, cyfr, liter i odstępów powinno być zbliżone do rozmieszczenia elementów druku. Wskaźniki i wykładniki potęg powinny być napisane wyraźnie i być prawidłowo obniżone lub podwyższone w stosunku do linii wiersza podstawowego. Znaki nad literami i cyframi, całkami i in. symbolami (strzałki, linie, kropki, daszki) powinny być umieszczone dokładnie nad tymi elementami, do których się odnoszą. Numery wzorów cyframi arabskimi powinny być kolejne i umieszczone w nawiasach okrągłych z prawej strony. Nazwy jednostek, symbole literowe i graficzne powinny być zgodne z wytycznymi IEE (International Electrotechnical Commission) oraz ISO (International Organization of Standardization).

Powołania.

Powołania na publikacje powinny być umieszczone na ostatnich stronach tekstu pod tytułem „Bibliografia”, opatrzone numeracją kolejną bez nawiasów. Numeracja ta powinna być zgodna z odnośnikami w tekście artykułu. Przykłady opisu publikacji:

- periodycznej 1. F. Valdoni: A new milimetre wave satelite. E.T.T. 1990, vol. 2, no 5, pp. 141–148
- nieperiodyczne 2. K. Andersen: A resource allocation framework. XVI International Symposium, Stockholm (Sweden), may 1991, paper A 2,4
- książki 3. Y.P. Tvidis: Operation and modeling of the MOS transistors. New York, McGraw-Hill, 1987, p. 553

Materiały ilustracyjne.

Rysunki powinny być wykonane wyraźnie, na papierze gładkim lub milimetrowym w formacie nie mniejszym niż 9 × 12 cm. Mogą być także w postaci wydruku komputerowego (preferowany edytor Corel Draw). Fotografie lub diapozytywy przyjmowane są raczej czarno-białe w formacie nie przekraczającym 10 × 15 cm. Na marginesie każdego rysunku i na odwrocie fotografii powinno być napisane ołówkiem imię i nazwisko Autora oraz skrót tytułu artykułu, do którego są przeznaczone oraz numer rysunku. Spis podpisów pod rysunki i fotografie powinny być umieszczone na oddzielnej stronie. Podpisy pod rysunkami (fotografiami) powinny być dwujęzyczne: w pierwszej kolejności w podstawowej wersji językowej, a później w dodatkowej wersji językowej. Rysunki powinny być cytowane w tekście.

Uwagi końcowe.

Na odrębnej stronie powinny być podane następujące informacje:

- adres do korespondencji z kodem pocztowym (domowy lub do miejsca pracy),
- telefon domowy i/lub do miejsca pracy,
- adres e-mailowy (jeśli autor posiada).

Autorowi przysługuje bezpłatnie 20 odbitek artykułu. Dodatkowe egzemplarze odbitek, lub cały zeszyt Autor może zamówić u wydawcy na własny koszt.

Autora obowiązują korekta autorska, którą powinien wykonać w ciągu 3 dni od daty otrzymania tekstu z Redakcji oraz zwrócić osobiście lub listownie pod adres Redakcji. Korekta powinna być naniesiona na przekazanych Autorowi szpaltach na marginesach ew. na osobnym arkuszu w przypadku uzupełnień tekstu większych niż dwa wiersze. W przypadku nie zwrócenia korekty w terminie, korektę przeprowadza Redakcja Techniczna Wydawcy. Redakcja prosi Autorów o powiadomienie ją o zmianie miejsca pracy i adresu prywatnego.

INFORMATION FOR AUTHORS OF K.E.T.

The editorial staff will accept for publishing only original monographic and survey papers concerning widely understood electronics. Because of the fact that KWARTALNIK ELEKTRONIKI I TELEKOMUNIKACJI is a journal of the Committee for Electronics and Telecommunications of Polish Academy of Science, it presents scientific works concerning theoretical bases and applications from the field of electronics, telecommunications, microelectronics, optoelectronics, radioelectronics and medical electronics.

Articles should be characterised by original depiction of a problem, its own classification, critical opinion (concerning theories or methods), discussion of an actual state or a progress of a given branch of a technique and discussion of development perspectives.

An article published in other magazines can not be submitted for publishing in K.E.T. The size of an article can not exceed 30 pages, 1800 character each, including figures and tables.

Basic requirements

The article should be submitted to the editorial staff as a one side, clear, black and white computer printout in two copies. The article should be prepared in English or Polish. Floppy disc with an electronic version of the article should be enclosed. Preferred wordprocessors: WORD 6 or 8.

Layout of the article.

– Title.

– Author (first name and surname of author/authors).

– Workplace (institution, address and e-mail).

– Concise summary in a language article is prepared in (with keywords).

– Main text with following layout:

- Introduction
- Theory (if applicable)
- Numerical results (if applicable)
- Paragraph 1
- Paragraph 2
-
-
- Conclusions
- Acknowledgements (if applicable)
- References

– Summary in additional language:

- Author (first name initials and surname)
- Title (in Polish, if article was prepared in English and vice versa)
- Extensive summary, however not exceeding 3600 characters (along with keywords) in Polish, if article was prepared in English and vice versa. The summary should be prepared in a way allowing a reader to obtain essential information contained in the article. For that reason in the summary author can place numbers of essential formulas, figures and tables from the article.

Pages should have continuous numbering.

Main text

Main text can not contain formatting such as spacing, underlining, words written in capital letters (except words that are commonly written in capital letters). Author can mark suggested formatting with pencil on the margin of the article using commonly accepted adjusting marks.

Text should be written with double line spacing with 35 mm left and right margin. Titles and subtitles should be written with small letters. Titles and subtitles should be numbered using no more than 3 levels (i.e. 4.1.1.).

Tables

Tables with their titles should be placed on separate page at the end of the article. Titles of rows and columns should be written in small letters with double line spacing. Annotations concerning tables

should be placed directly below the table. Tables should be numbered with Arabic numbers on the top of each table. Table can consist algorithm and program listings. In such cases original layout of the table will be preserved. Table should be cited in the text.

Mathematical formulas

Characters, numbers, letters and spacing of the formula should be adequate to layout of main text. Indexes should be properly lowered or raised above the basic line and clearly written. Special characters such as lines, arrows, dots should be placed exactly over symbols which they are attributed to. Formulas should be numbered with Arabic numbers placed in brackets on the right side of the page. Units of measure, letter and graphic symbols should be printed according to requirements of IEE (International Electrotechnical Commission) and ISO (International Organisation of Standardisation).

References

References should be placed at the end of the main text with the subtitle „References“. References should be numbered (without brackets) adequately to references placed in the text. Examples of periodical [1], non-periodical [2] and book [3] references:

1. F. Valdoni: A new millimetre wave satellite. E.T.T. 1990, vol. 2, no 5, pp. 141–148
2. K. Anderson: A resource allocation framework. XVI International Symposium (Sweden). May 1991. paper A 2.4
3. Y.P. Tvidis: Operation and modeling of the MOS transistors. New York. McGraw-Hill. 1987. p. 553

Figures

Figures should be clearly drawn on plain or millimetre paper in the format not smaller than 9 × 12 cm. Figures can be also printed (preferred editor – CorelDRAW). Photos or diapositives will be accepted in black and white format not greater than 10 × 15 cm. On the margin of each drawing and on the back side of each photo author's name and abbreviation of the title of article should be placed. Figure's captions should be given in two languages (first in the language the article is written in and then in additional language). Figure's captions should be also listed on separate page. Figures should be cited in the text.

Additional information

On the separate page following information should be placed:

- mailing address (home or office),
- phone (home or/and office),
- e-mail.

Author is entitled to free of charge 20 copies of article. Additional copies or the whole magazine can be ordered at publisher at the author's expense.

Author is obliged to perform the author's correction, which should be accomplished within 3 days starting from the date of receiving of the text from the editorial staff. Corrected text should be returned to the editorial staff personally or by mail. Correction marks should be placed on the margin of copies received from the editorial staff or if needed on separate pages. In the case when the correction is not returned in said time limit, correction will be performed by technical editorial staff of the publisher.

In case of changing of workplace or home address Authors are asked to inform the editorial staff.