Piotr KAWALEC
POLITECHNIKA WARSZAWSKA, WYDZIAŁ TRANSPORTU

Analiza probabilistyczno – czasowych przetworników informacji i ich implementacja w układach FPGA

Dr inż. Piotr KAWALEC

Ukończył studia na Wydziale Elektroniki Instytutu Transportu w Leningradzie w 1975 r., obronił pracę doktorską w 1979 r. w Instytucie Elektrotechnicznym w Leningradzie. Jest adiunktem na Wydziale Transportu Politechniki Warszawskiej. Jego zainteresowania naukowe to automatyka, elektronika, technika cyfrowa i ich zastosowanie w układach i systemach sterowania i teleinformatyki stosowanych w transporcie.

e-mail: pka@it.pw.edu.pl

Streszczenie

W artykule przedstawiono przetworniki informacji, w których zmienną maszynową jest pierwszy moment (wartość oczekiwana) stacjonarnego i ergodycznego binarnego ciągu losowego. Przetworniki te, zwane probabilistyczno-czasowymi, umożliwiają prostą realizację podstawowych operacji arytmetycznych na strumieniach losowych z wydaniem wyniku w postaci zdeterminowanej. Dla najtrudniejszej operacji dzielenia przedstawiono szczegółową analizę działania z określeniem dokładności przetwarzania dla wejściowych strumieni losowych o rozkładach dwumianowych oraz hipergeometrycznych.

Słowa kluczowe: przetworniki informacji, ciągi losowe, układy dzielące, dokładność przetwarzania, układy FPGA.

Analysis of probabilistic – times information converters and their implementation in FPGA device

Abstract

The article presents information converters in which the machine variable is the first moment (expected value) of stationary and ergodic binary random sequence. These converters, called probabilistic and time, enable us to conduct a simple implementation of basic arithmetical operations on random series giving the result in a determinantal form. For the most difficult division operation, a detailed analysis has been presented determining conversion accuracy for input random series of binomial and hypergeometrical distributions.

Keywords: information converters, random sequences, division systems, conversion accuracy, FPGA devices.

1. Wstęp

W stochastycznych modelach procesów transportowych, często występuje konieczność przeprowadzania operacji na ciągach losowych, których parametry (najczęściej pierwszy moment rozkładu – wartość oczekiwana) przenoszą istotne informacje opisujące modelowany proces. Jeśli modelowany proces transportowy (zgłoszenia pojazdów, odstępy między pojazdami w strumieniu itd.) będzie procesem ergodycznym i co najmniej przedziałami stacjonarny, to opisujące go ciągi losowe mogą być przetwarzane przy pomocy specjalizowanych układów i systemów cyfrowych nazywanych komputerami stochastycznymi, lub stochastycznymi przetwornikami informacji [1]. Zastosowanie w nich jako zmiennej maszynowej wartości oczekiwanej zmiennych losowych występujących w binarnych ciągach losowych w istotny sposób upraszcza wykonywanie operacji arytmetycznych. Przy tym zachowana jest cyfrowa postać danych, natomiast wyniki przetwarzania uzyskujemy w postaci wartości oczekiwanej ciągu wynikowego, bądź w postaci zdeterminowanej, jako estymatory wartości oczekiwanej.

Implementacja specjalizowanych układów arytmetyki stochastycznej w strukturach FPGA pozwala uzyskać jednoukładowe realizacje o maksymalnej częstotliwością taktowania określanej w setkach MHz. Dzięki temu przy przetwarzaniu ciągów losowych opisujących procesy transportowe, przebiegających, co najwyżej w zakresie milisekund, możliwa jest praca takich układów w czasie rzeczywistym z zapewnieniem wymaganej dokładności przetwarzania. Ogromną zaletą zastosowania stochastycznych przetworników informacji jest ich odporność na przemijające uszkodzenia, co w przypadku urządzeń sterowania i kierowania ruchem w transporcie ma zasadnicze znaczenie.

Z operacji arytmetycznych, wykonywanych na stacjonarnych procesach losowych, najbardziej złożonymi operacjami są ą operacje odwracania i dzielenia liczb przedstawionych w postaci prawdopodobieństw wystąpienia symboli o wartości 1 (lub 0) w synchronicznych binarnych ciągach losowych. Zwykle do tego celu wykorzystywane są cyfrowe integratory ze sprzężeniem zwrotnym.

2. Probabilistyczno – czasowe przetworniki informacji

Najprostszą realizację operacji odwracania liczb uzyskamy, rozpatrując binarny ciąg losowy będący produktem zastosowania schematu Bernoulliego. W tym przypadku pojawiające się w kolejnych taktach w binarnym ciągu losowym wartości 1 oraz 0 występują w sposób niezależny, z prawdopodobieństwem odpowiednio p oraz $1– p$. Zakładając, że w trakcie m taktów w losowym ciągu binarnym wystąpi N wartości 1, można wyznaczyć wartość oczekiwaną liczby taktów m

$$\mathrm{E(m)} = \frac{\mathrm{N}}{\mathrm{p}}. \tag{1}$$

Tak więc operacja odwracania liczb zostaje sprowadzona do odmierzenia przedziału czasowego $\tau = mf^{-1}$ (gdzie f – częstotliwość taktowania) w trakcie którego zostanie zliczone w binarnym ciągu losowym N wartości 1. Stąd określenie takich przetworników, jako przetworników probabilistyczno – czasowych.

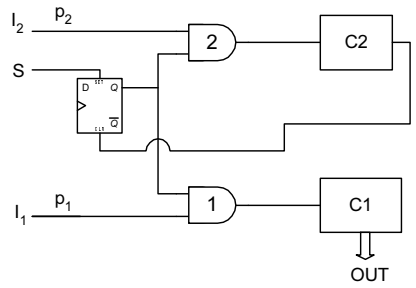
Sprzętowa realizacja tej operacji wymaga zastosowania dwóch liczników, w jednym z nich zostanie zliczone N wartości 1, natomiast w drugim zostanie zliczona liczba taktów m. Jeśli licznik zliczający wartości 1 w binarnym ciągu losowym będzie miał pojemność N, można sygnałem jego przepełnienia sterować procesem zliczania liczby taktów m w drugim liczniku.

Wariancja zmiennej losowej m określa dokładność wykonania operacji odwracania liczby p

$$\mathrm{Var(m)} = \frac{\mathrm{N(1-p)}}{\mathrm{p}}. \tag{2}$$

Zależności (1) i (2) są prawdziwe tylko wtedy, gdy kolejne produkcje symboli tworzące binarny ciąg losowy są wzajemnie niezależne.

Przedstawioną koncepcję przetworników probabilistyczno – czasowych do odwracania liczb, można wykorzystać do realizacji operacji dzielenia, podając na wejście zliczające liczbę taktów m binarne ciągi losowe o zadanym rozkładzie. Probabilistyczno – czasowy przetwornik działający jako układ dzielący zawiera dwa liczniki C1 i C2, elementy bramkujące AND 1 i 2, oraz przerzutnik z wejściami asynchronicznymi (rys. 1).



Rys. 1. Schemat probabilistyczno – czasowego układu dzielącego
Fig. 1. Diagram of probabilistic – time division system

Sygnałem S uruchamia się równoczesne podawanie z wejść I_1 i I_2 binarnych ciągów losowych (z prawdopodobieństwem wystąpienia wartości 1 odpowiednio p_1 i p_2) na liczniki C1 i C2. Sygnał przeniesienia z licznika C2 odcina podawanie ciągów losowych na liczniki, ustalając w ten sposób wynik dzielenia na wyjściu licznika C1.

Operacja odwracania liczb jest szczególnym przypadkiem dzielenia, gdy na wejście I_1 podawane są sygnały synchronizacji przyjmujące w każdym takcie wartość 1 z prawdopodobieństwem p_1 =1. Szybkość i dokładność działania przetworników probabilistyczno – czasowych w istotny sposób zależy od rodzaju binarnych strumieni losowych podawanych na jego wejście I_1.

3. Analiza układu dzielącego sterowanego ciągiem losowym o rozkładzie dwumianowym

Rozkład dwumianowy jest modelem losowania niezależnego, które można przeprowadzić porównując kod binarny stałoprzecinkowej wartości liczbowej z niezależnymi liczbami losowymi uzyskanymi z generatora o rozkładzie równomiernym, stosując np. generatory układowe na rejestrach przesuwających [2].

Uogólnimy rezultaty otrzymane dla operacji odwracania liczby na operację dzielenia p_1/p_2, zakładając, że na wejście I_1 podawany jest binarny ciąg losowy o rozkładzie dwumianowym.

Jako zmienną losową Z oznaczymy zawartość licznika C1. Wielkość ta jest sumą m zmiennych losowych z_i (z_i =1 albo 0 w ciągu losowym p_1, i – numer taktu). Ponieważ m jest również zmienną losową, więc wartość oczekiwana zmiennej losowej Z

$$E(Z) = E\Big(\sum_{i=1}^{m} z_i\Big) = E(m)E(Z_i) = N\frac{p_1}{p_2}, \tag{3}$$

gdzie N – pojemność licznika C2.

Wariancję zmiennej losowej Z wyznaczymy z zależności

$$\begin{aligned} Var(Z) &= E\Big\{\Big(\sum_{i=1}^{m} z_i\Big)^2\Big\} - E^2(Z) = \sum_m p_m\Big\{m(m-1)p_1^2 + mp_1 + 2\sum_{i\neq j} K_{ij}\Big\} - E^2(Z) = \\ &= \sum_m p_m m^2 p_1^2 - \sum_m p_m m p_1^2 + \sum_m p_m m p_1 + \sum_m p_m\Big(2\sum_{i\neq j} K_{ij}\Big) - E^2(Z) = \\ &= p_1^2\{Var(m) + E^2(m)\} - p_1^2 E(m) + p_1 E(m) + 2\sum_{i\neq j} K_{ij} - E^2(Z) = N\frac{p_1}{p_2}\Big(\frac{p_1}{p_2} + 1 - 2p_1\Big) + 2\sum_{i\neq j} K_{ij}, \end{aligned} \tag{4}$$

gdzie K_{ij} – moment korelacyjny zmiennych losowych z_i , z_j.

Błąd względny dzielenia liczb wyniesie:

$$\delta(Z) = \lambda_p \frac{\sqrt{Var(Z)}}{E(Z)} = \lambda_p \sqrt{\frac{1}{N}\Big(\frac{p_2}{p_1} + 1 - 2p_2\Big)}\ 100\ [\%], \tag{5}$$

gdzie λ_p – kawntyl rozkładu normalnego.

Przykładowo dla $p_1 = 0{,}1$; $p_2 = 0{,}2$ dla uzyskania błędu względnego $\delta(Z) < 5\%$ z prawdopodobieństwem 0,95 (λ_p = 1,645), licznik C2 powinien być licznikiem co najmniej 14 pozycyjnym. Ogólnie, z zależności (5) wynika, że k krotne zwiększenie dokładności dzielenia, wydłuża czas obliczeń k^2 razy.

4. Analiza układu dzielącego sterowanego ciągiem losowym o rozkładzie hipergeometrycznym

Rozkład hipergeometryczny jest modelem losowania bezzwrotnego, które można przeprowadzić porównując kod binarny zadający prawdopodobieństwo wystąpienia jedności z ciągiem $L = 2^l$ niepowtarzających się liczb losowych. W tym przypadku wartość oczekiwana wyników symbolu v określana jest ilością N_1 liczb losowych porównywanych z symbolem v = 1 i równa jest średniej arytmetycznej wielkości v za $m = L$ taktów

$$E(v) = \frac{N_1}{L} = \frac{1}{L}\sum_{i=1}^{L} v_i.$$

Wartość oczekiwana i wariancja zmiennej losowej

$$V = \sum_{i=1}^{m} v_i$$

wyrażają się wzorami [3]

$$\begin{aligned} E(V) &= m\frac{N_1}{L} = mp; \\ Var(V) &= mpq\Big(\frac{L-m}{L-1}\Big) \end{aligned} \tag{6}$$

Porównując charakterystyki rozkładu hipergeometrycznego z charakterystykami rozkładu dwumianowego, można stwierdzić, że w wyrażeniu (6) dla wariancji $Var(V)$ występuje dodatkowy czynnik $\Big(\frac{L-m}{L-1}\Big)$. Z tego powodu przy m=L, gdy wartość średnia równa jest wartości oczekiwanej, wariancja $Var(V)$ przyjmuje wartość zero. Przy liczności próbki m > L błąd oceny $E(V)$ określa się wariancją zmiennej losowej $\sum_{i=1}^{m'} v_i$, gdzie m' – reszta z dzielenia m na cykle po L taktów. W ten sposób zastosowanie krótkich ciągów losowych o rozkładzie hipergeometrycznym umożliwia uzyskanie oceny wartości oczekiwanej z mniejszą wariancją niż w przypadku losowania niezależnego.

Technicznie otrzymanie ciągu z rozkładem hipergeometrycznym jest stosunkowo proste, bowiem generatorem cyklicznego ciągu $L= 2^l$ binarnych liczb pseudolosowych może być l –bitowy rejestr przesuwający z nieliniowym sprzężeniem zwrotnym (nLFSR) [4].

Analizując operację dzielenia p_1/p_2 binarnych ciągów losowych podawanych na wejścia I_1 i I_2 probabilistyczno-czasowego przetwornika, założymy, że wejściowe ciągi losowe są niezależne i mają jednakowy okres L. Warunek ten jest łatwy do spełnienia jeśli do konwersji binarnych ciągów losowych zastosujemy dwa rejestry przesuwające z nieliniowym sprzężeniem zwrotnym (nLFSR) o różnych funkcjach charakterystycznych, lecz o jednakowym okresie $L=2^l$.

W przypadku dzielenia liczb dla wartości oczekiwanej zmiennej losowej Z odwzorowującej zawartość licznika C1 w chwili wystąpienia przeniesienia w liczniku C2 określa się z zależności

$$E(Z) = kp_1L + E\left(\sum_{i=1}^{m'} z_i\right) = kp_1L + E(z_i)\cdot E(m') = kp_1L + p_1\frac{Q}{p_2} = N\frac{p_1}{p_2} \quad (7)$$

Wariancję *Var(Z)* losowego czasu zliczania *m'* wyznaczymy przy pomocy parametrów (6) warunkowego rozkładu hipergeometrycznego zmiennej losowej *Z* (przy stałej *m'*), korzystając ze znanych *E(m')* i $D^2(m')$

$$\mathrm{Var}(Z) = \sum_{m'} p_{m'}E(Z^2/m') - E^2(Z) = \sum_{m'} p_{m'}[E^2(Z/m') + \mathrm{Var}(Z/m')] - E^2(Z) =$$

$$= \sum_{m'} p_{m'}\left[p_1(m')^2 + m'p_1q_1\left(\frac{L-m'}{L-1}\right)\right] - E^2(Z) = \left(p_1^2 - \frac{p_1q_1}{L-1}\right)\sum_{m'} p_{m'}(m')^2 +$$

$$+p_1q_1\frac{L}{L-1}\sum_{m'} p_{m'}m' - E^2(Z) = \left(p_1^2 - \frac{p_1q_1}{L-1}\right)E(m')^2 + p_1q_1\frac{L}{L-1}E(m') - p_1^2E^2(m') =$$

$$= \left(p_1^2 - \frac{p_1q_1}{L-1}\right)\mathrm{Var}(m') + \frac{p_1q_1}{L-1}[LE(m') - E(m')^2] = Q\left(\frac{N_1-Q}{N_1+1}\right)\frac{p_1}{p_2}\left[\frac{p_1q_2}{p_2}\cdot\frac{N_1-1}{N_1+1} +\right.$$

$$\left. -\frac{q_1q_2}{p_2}\cdot\frac{N_1-1}{N_1+1}\cdot\frac{1}{L-1} + q_1\frac{N_1-1}{N_1}\cdot\frac{L}{L-1}\right] < Q\left(\frac{N_1-Q}{N_1-1}\right)\frac{p_1}{p_2}\left(1-2p_1+\frac{p_1}{p_2}\right). \quad (8)$$

Błąd względny operacji dzielenia w tym przypadku wyniesie

$$\delta(Z) = \lambda_p\frac{\sqrt{\mathrm{Var}(Z)}}{E(Z)} < \frac{\lambda_p}{N}\sqrt{Q\left(\frac{N_1-Q}{N_1-1}\right)\left(\frac{p_2}{p_1}+1-2p_2\right)}\cdot 100[\%] \quad (9)$$

Przykładowo dla p_1=0,1; p_2=0, 2; N = 1024; L = 128, $\delta(Z)$< 0,5% z prawdopodobieństwem 0,95.

Tak znaczne zwiększenie dokładności dzielenia w tym przypadku w porównaniu ze sterowaniem układu binarnymi ciągami losowymi o rozkładzie dwumianowym, można uzasadnić tym, że wariancja wyniku *Var(Z)* zależy nie od pojemności *N* licznika C2, lecz od okresu L cyklicznego ciągu losowego. Przy czym zwiększenie pojemności *N* prowadzi do proporcjonalnego zmniejszenia błędu *δ(Z)*, co wynika z wrażenia (9). Kolejną przyczyną zmniejszenia błędu jest zbliżenie czasu wystąpienia nadmiaru w liczniku C2 do wielkości *k·L*, co we wzorze (9) wyraża się niską wartością współczynnika

$$Q\left(\frac{N_1-Q}{N_1-1}\right).$$

Do rozważenia pozostaje zależność współczynnika $Q\left(\frac{N_1-Q}{N_1-1}\right)$ od okresu *L* cyklicznego ciągu binarnego o rozkładzie hipergeometrycznym. W tym celu znajdziemy maksimum funkcji

$$Q\left(\frac{N_1-Q}{N_1-1}\right). \quad (10)$$

Początkowym argumentem tej funkcji jest prawdopodobieństwo $p = \frac{N_1}{L}$. Reszta z dzielenia *N* na N_1, wynosząca Q = $N - kN_1$, (gdzie $k = ent(N/N_1)$ też jest funkcją *p*. W zależności od *p* wielkość *k* może przyjmować następujący szereg wartości $k = 2^n, 2^{n-1}, \ldots, 2^{n-l}$, podstawiając $Q = (2^{n-l} - kp)L$ do zależności (10), po przekształceniach otrzymamy

$$Q\frac{N_1-Q}{N_1-1} = \frac{(2^{n-l}kp)[(k+1)p - 2^{n-l}]}{p}\cdot L\frac{N_1}{N_1-1} \quad (11)$$

Analizując zależność

$$w = w(p) = \frac{(2^{n-l}kp)[(k+1)p - 2^{n-l}]}{p} \quad (12)$$

można stwierdzić, że absolutne maksimum funkcja *w(p)* posiada na odcinku, gdzie $k = k_{min} = 2^{n-l}$. Podstawiając tę wartość do (12), a następnie różniczkując *w* po *p* i porównując pochodną do zera otrzymujemy równanie kwadratowe

$$p^2 = \frac{2^{n-l}}{2^{n-l}+1}$$

Podstawiając pierwiastki tego równania do (12) otrzymujemy wyrażenie

$$w_{max} = \frac{2^{n-l}}{1+2^{n-l+1}+2\sqrt{2^{n-l}(2^{n-l}+1)}},$$

którego granica

$$\lim_{(n-l)\to\infty} w_{max} = \frac{1}{4}.$$

W ten sposób wyznaczyliśmy zależność współczynnika $Q\left(\frac{N_1-Q}{N_1-1}\right)$ od okresu *L* cyklicznego binarnego ciągu losowego o rozkładzie hipergeometrycznym

$$Q\left(\frac{N_1-Q}{N_1-1}\right) < \frac{1}{4}L.$$

Podstawiając wyznaczone wartości graniczne współczynników do równania (9) przy λ_p = 2,576 (dla prawdopodobieństwa 0,995) otrzymamy $n = 1{,}5l$. Tak więc dla przyjętych założeń liczba pozycji *n* licznika C2 powinna być 1, 5 krotnie większa od pozycyjności *l* rejestru generującego okresowy ciąg losowy.

Aby otrzymać taką samą dokładność dzielenia przy zastosowaniu ciągów losowych o rozkładzie dwumianowym, wymagana liczba pozycji licznika C2 wyniesie $n' = 2l + 2$.

Współczynnik $\gamma = \frac{2^{n'}}{2^n} = 2^{0,5l+2}$ określa względne zwiększenie szybkości pracy układu dzielącego, uzyskane dzięki zastosowaniu, jako sygnałów wejściowych, okresowych binarnych ciągów losowych o rozkładzie hipergeometrycznym. Np. przy l = 8, uzyskujemy 64 krotne przyśpieszenie wykonywania operacji dzielenia.

Praca naukowa finansowana ze środków na naukę w latach 2007 – 2009 jako projekt badawczy.

5. Literatura

[1] Gaines B.R.: Stochastic computing systems. Advances in Information Systems Science, New York, 1969, pp. 37 – 172.

[2] Kawalec P.: Synteza i weryfikacja wielokanałowego generatora liczb pseudolosowych zaimplementowanego w układzie FPGA. Materiały IV Krajowej Konferencji Naukowej „Reprogramowalne układy cyfrowe RUC'2001", Szczecin, 2001, str. 291 – 298.

[3] Krysicki W.: Rachunek prawdopodobieństwa. PWN, Warszawa, 2004.

[4] Janicka-Lipska I., Stokłosa J.: Boolean feedback functions for full-length nonlinear shift registers. Journal of Telecommunications and Information Technology, No. 4, 2004, pp. 28 − 30.

Artykuł recenzowany