

Przemysław WŁODARSKI¹

FRAKTALNE WŁAŚCIWOŚCI NATĘŻENIA RUCHU W SIECIACH TELEINFORMATYCZNYCH

W artykule omówiono podstawowe właściwości natężenia ruchu występującego w sieciach teleinformatycznych, który wykazuje cechy struktur fraktali – obiektów samopodobnych. Zjawisko to powoduje znaczący spadek efektywności działania sieci teleinformatycznych ze względu na większe opóźnienia oraz poziom strat pakietów. Oprócz wyników pomiarów przeprowadzonych w rzeczywistej sieci komputerowej omówione zostaną dwie metody estymacji poziomu samopodobieństwa, w ramach których zostanie zaproponowany sposób uśredniania prowadzący do bardziej stabilnych wyników. Ponadto zestawione zostaną ze sobą dwa modele – klasyczny Poissona oraz samopodobny model przełączanych strumieni on/off.

FRACTAL PROPERTIES OF NETWORK TRAFFIC IN DATA COMMUNICATION NETWORKS

The article discusses the basic properties of network traffic occurring in data communication networks, which exhibits the characteristic features of fractals – self-similar objects. This leads to a significant decrease in the network efficiency because of the increased latency and packet loss level. In addition to the results of measurements carried out in a real computer network, two methods for estimating the level of self-similarity will be discussed, in which the other averaging method, that leads to more stable results, is proposed. Furthermore two models will be compared: classical Poisson model and self-similar model of switching on/off streams.

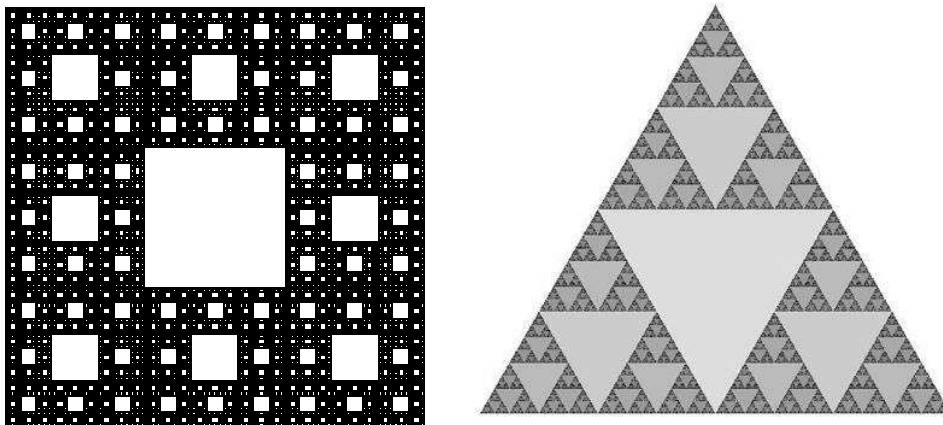
1. WSTĘP

Często zakłada się, że natężenie ruchu w sieciach teleinformatycznych ma charakter klasycznego strumienia Poissona. Podejście takie powoduje, że wszelkie analizy i wnioski odnośnie efektywności działania sieci są bardziej optymistyczne niż jest w rzeczywistości. Skorelowany ruch wprowadza fraktalną strukturę, która powoduje wzrost

¹Zachodniopomorski Uniwersytet Technologiczny w Szczecinie
Wydział Elektryczny, Katedra Przetwarzania Sygnałów i Inżynierii Multimedialnej
ul. 26 kwietnia 10, 71-126 Szczecin, tel. 91 449 5311, email: przemyslaw.wlodarski@zut.edu.pl

średniego czasu przebywania pakietu w systemie kolejkowym, jakim jest np. router, jak również wzrost poziomu strat pakietów. Zatem chcąc wykonać jakąkolwiek analizę, należy wziąć pod uwagę oprócz podstawowych parametrów strumienia wejściowego takich jak średnie natężenie oraz rozkład, także poziom samopodobieństwa [1, 2, 4].

Podstawową własnością fraktali jest brak zmiany podstawowego kształtu dla dowolnego powiększenia / pomniejszenia obserwowanej struktury. Przykłady takich struktur są przedstawione na rys. 1.



Rys. 1. Dywan oraz trójkąt Sierpińskiego jako przykłady struktur samopodobnych

W przypadku analizy natężenia ruchu w sieciach teleinformatycznych struktury fraktalne powstają nie na płaszczyźnie, jak to przedstawiono na rys. 1, lecz w przestrzeni jednowymiarowej. Pojęcie powiększania i pomniejszania dotyczy tu nie kształtu, ale czasu obserwacji. W następnych rozdziałach zostanie przedstawione jak w prosty sposób określić ich wymiar fraktalny, a także jak zamodelować ruch o zadanym poziomie samopodobieństwa.

2. PODTAWOWE WŁAŚCIWOŚCI FRAKTALNYCH STRUMIENI PAKIETÓW

Podstawowe własności strumieni samopodobnych można przedstawić przyglądając się procesowi ułamkowego ruchu Browna [3], [7] z charakterystycznymi dla niego cechami:

- niezmienność w sensie rozkładu statystycznego:

$$Y(at) = a^H Y(t) \text{ dla } t, a \in R \text{ oraz } a > 0 \quad (1)$$

- posiadanie stacjonarnych Gaussowskich przyrostów o następującej funkcji gęstości prawdopodobieństwa:

$$f_G(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad (2)$$

- funkcja kowariancji zdefiniowana jako:

$$\text{Cov}(Y(s), Y(t)) = K_H(s, t) = \frac{\sigma^2}{2} (|t|^{2H} + |s|^{2H} - |t-s|^{2H}) \quad (3)$$

- dla $\sigma > 0$ oraz $0.5 \leq H < 1$ otrzymujemy:

$$E(Y^2(t)) = \sigma^2 |t|^{2H} \quad (4)$$

Z kolei proces przyrostowy $X_k = Y(k+1) - Y(k)$, $k \in N$, nazywany ułamkowym szumem Gaussa, posiada następujące właściwości będące podstawą konstrukcji wielu modeli natężenia ruchu:

- ciąg $\{X_k\}$ jest stacjonarny, $E(X_k) = 0$
- $E(X_k^2) = \sigma^2 = E(Y(1)^2)$
- autokowariancja procesu $\{X_k\}$ jest określona jako:

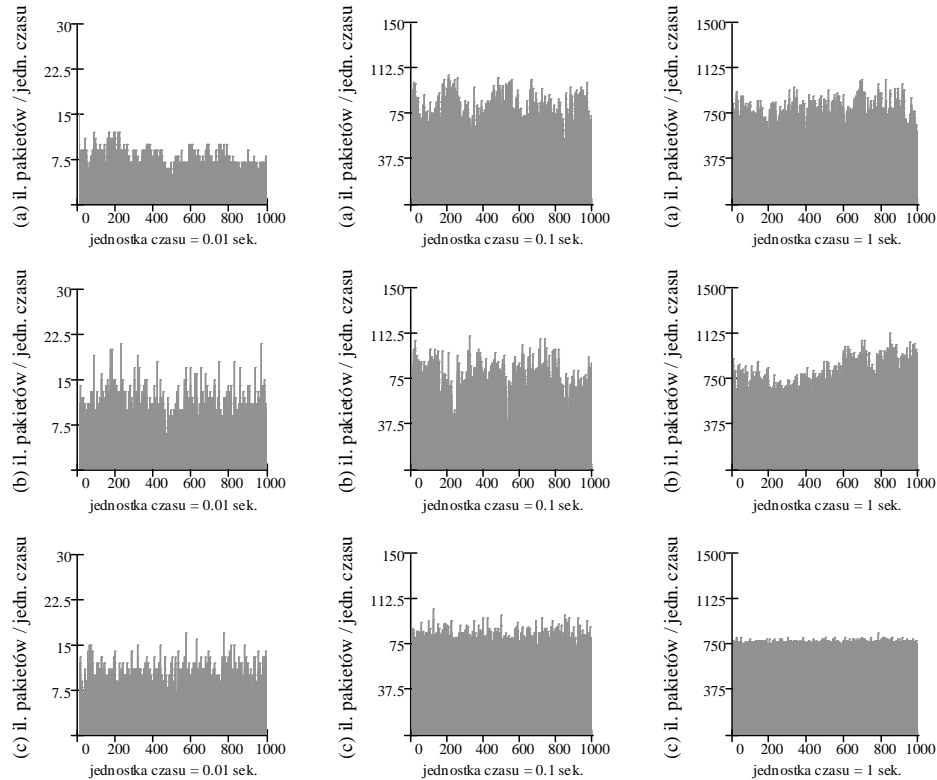
$$r(k) = \frac{\sigma^2}{2} [|k+1|^{2H} - 2|k|^{2H} + |k-1|^{2H}] = \frac{\sigma^2}{2} \delta^2(|k|^{2H}) \quad (5)$$

gdzie δ^2 oznacza operator różnicowy drugiego rzędu
dla $k \neq 0$: $r(k) = 0$ jeżeli $H = 0.5$, $r(k) > 0$ jeżeli $0.5 < H < 1$

Wartość H ($0.5 \leq H < 1$) jest zwana wykładnikiem Hursta i jest miarą stopnia samopodobieństwa. Dla $H = 0.5$ ułamkowy ruch Browna jest zwykłym ruchem Browna, gdzie $r(k) = 0$, $k \geq 1$ oraz $\text{Cov}(Y(s), Y(t)) = K_{0.5}(s, t) = \min(s, t)$.

3. ESTYMACJA POZIOMU SAMOPODOBIEŃSTWA

Jedne z najprostszych i często używane metody estymacji poziomu samopodobieństwa natężenia ruchu to działające w dziedzinie czasu metody: *variance-time* (VT) oraz *index of despiration for counts* (IDC). Dla obu zostanie zaproponowana metoda uśredniania prowadząca do bardziej stabilnych wyników. Wyniki pomiarów dotyczą rzeczywistego ruchu sieciowego zarejestrowanego na Politechnice Szczecińskiej. Na rys. 2 pokazano zestawienie struktur natężenia ruchu przy różnych skalach czasowych. Widać, że dla modelu Poissona przy coraz większej skali czasowej „poszarpanie” wykresu maleje, natomiast dla rzeczywistego ruchu oraz modelu on/off poziom „poszarpania” pozostaje wskazując na ceche charakterystyczna fraktali – niezmiennosc struktury przy powiększaniu / pomniejszaniu.



Rys. 2. a) model on/off, b) rzeczywisty ruch, c) model Poissona

3.1. Metoda variance-time

Metoda analizy wariancyjno-czasowej (variance-time, VT) charakteryzuje się tym, że agregowany jest proces przyrostowy $Y(t)$: $X(i) = Y(i) - Y(i-1)$ w następujący sposób [6]:

$$X^{(m)}(i) = \frac{1}{m} \sum_{t=m(i-1)+1}^{mi} X(t), \quad (6)$$

dzięki czemu otrzymujemy szereg czasowy $X(t)$ podzielony na niezachodzące na siebie bloki o rozmiarze m . Wartości te następnie są uśredniane tworząc $X^{(m)}(i)$, gdzie i jest wskaźnikiem danego bloku. Aby uprościć zapis, przyjmijmy $X = X(1)$ oraz $X^{(m)} = X^{(m)}(1)$, wtedy $X^{(m)}$ jest uśrednioną zmienną losową o stopniu agregacji rzędu m :

$$X^{(m)} = \frac{1}{m} \sum_{t=1}^m X(t) = m^{H-1} X. \quad (7)$$

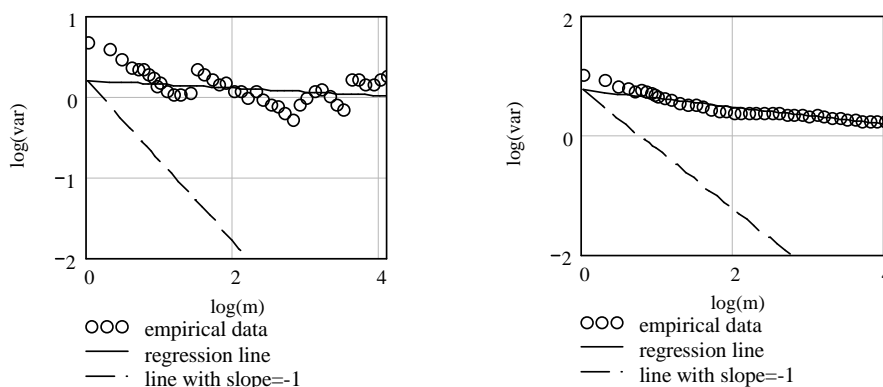
Jeżeli $X^{(m)}$ jest zmienną losową niezależną, wtedy jej wariancja $\text{var}(X^{(m)})$ redukuje się do $\sigma^2 m^{-1}$ w przypadku gdy $H=0.5$, ale jeżeli $0.5 < H < 1$, wtedy:

$$\text{var}(X^{(m)}) = \sigma^2 m^{-\beta}, \tag{8}$$

gdzie $0 < \beta < 1$ oraz $H = 1 - \beta/2$. Wariancja zagregowanej zmiennej losowej o strukturze fraktalnej zanika bardziej powoli, niż dla zmiennej losowej niezależnej, nieskorelowanej, dla której współczynnik nachylenia prostej na wykresie logarytmicznym wynosi -1 (rys. 3). Ze względu na duże rozproszenie wyników wokół prostej regresji dla metody według wzoru (7), proponuje się inny sposób uśredniania prowadzący do bardziej stabilnych wyników:

$$X'^{(m)}(i) = \frac{1}{m} \sum_{t=i-M}^{i+M} X(t) \tag{9}$$

Powyższe uśrednianie różni się od (7) tym, że przeprowadzone jest od razu na pełnym zakresie próbek, aż do momentu zapełnienia bloku o rozmiarze M .



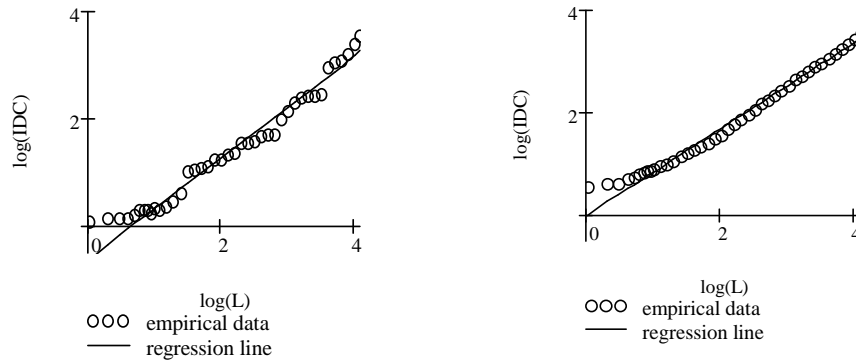
Rys. 3. Wyniki metody VT dla zarejestrowanego ruchu w ZUT: po lewej metoda uśredniania wg (7), po prawej metoda wg (9)

3.2. Metoda index of dispersion for counts

Kolejną metodą estymacji wykładnika Hursta, a tym samym poziomu samopodobieństwa, jest metoda index of dispersion for counts (IDC), działająca, podobnie jak metoda VT, w dziedzinie czasu. W tej metodzie, dla danego przedziału czasu L , wskaźnik $IDC(L)$ definiowany jest jako stosunek wariancji zagregowanej zmiennej losowej do jej wartości oczekiwanej [6]:

$$IDC(L) = \frac{\text{var}\left(\sum_{j=1}^L X_j\right)}{E\left(\sum_{j=1}^L X_j\right)} \approx cL^{2H-1}, \tag{10}$$

gdzie c jest dodatnią wartością niezależną od L . Wykreślając zależność $\log(IDC(L))$ od $\log(L)$, można dopasować prostą regresji, dla której współczynnik nachylenia wynosi $2H - 1$. Podobnie jak dla metody VT, sposób uśredniania zaproponowany w (9) daje bardziej stabilne rezultaty (rys. 4).



Rys. 4. Wyniki metody IDC dla zarejestrowanego ruchu w ZUT: po lewej metoda uśredniania wg (7), po prawej metoda wg (9)

4. MODEL ON/OFF

Jednym z najbardziej naturalnych modeli działania interfejsu sieciowego jest model on/off, który m.in. jest w stanie wygenerować natężenie ruchu o strukturze fraktalnej. Działa on w ten sposób, że w okresie 'on' generowany jest strumień o stałej intensywności λ_{on} , natomiast w okresie 'off' intensywność ta wynosi 0, otrzymując szereg czasowy $\{W(t), t \geq 0\}$, który przybiera na zmianę wartości λ_{on} oraz 0. Superpozycja M takich źródeł $\{W^{(m)}(t), m = 1, 2, \dots, M\}$ w czasie $T \cdot t$, gdzie T jest wartością skalującą, wynosi [1]:

$$W'_M(Tt) = \int_0^{Tt} \left(\sum_{m=1}^M W^{(m)}(u) \right) du \quad (11)$$

Jeżeli przyjrzymy się rys. 2. dla modelu on/off, widać, że nieregularności zostają zachowane dla różnych skal czasowych. Wynika to z tego, że do jego konstrukcji zostały użyte rozkłady Pareto o nieskończonej wariancji i parametrach $\alpha_1 = \alpha_2 = \alpha$ z przedziału $1 < \alpha < 2$. Okresy 'on' i 'off' występujące naprzemiennie posiadają rozkłady Pareto, których dystrybuanta oraz wartości oczekiwane wynoszą odpowiednio:

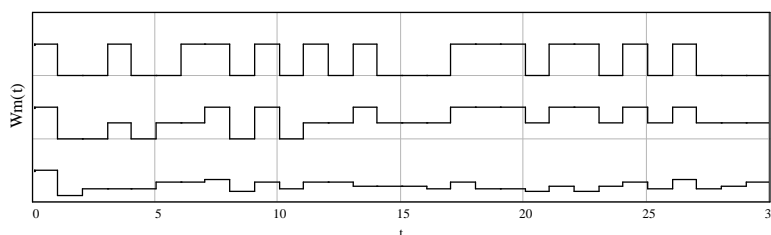
$$F(x) = 1 - \left(\frac{b}{x} \right)^\alpha, \quad E(x) = \frac{\alpha b}{\alpha - 1} \quad \text{dla } 1 < \alpha < 2 \text{ oraz } x > b \quad (12)$$

Zatem średnia intensywność pojedynczego źródła modelu on/off wynosi:

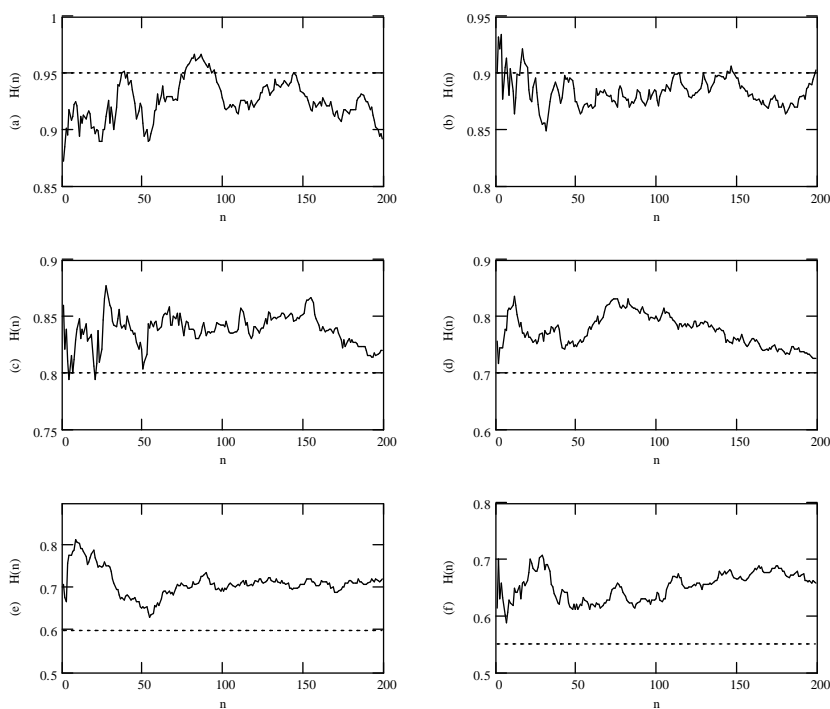
$$\lambda_s = \frac{t_{on}}{t_{on} + t_{off}} \lambda_{on} \tag{13}$$

W naszym przypadku $\lambda_s = 0.5 \cdot \lambda_{on}$. Tworzymy następnie proces składający się z M strumieni on/off, dla którego wartość oczekiwana intensywności wynosi:

$$\lambda_s^M = E\left(\sum_{m=1}^M W^{(m)}(t)\right) = \frac{Mt_{on}\lambda_{on}}{t_{on} + t_{off}} \tag{14}$$



Rys. 5. Przykładowe strumienie on/off składające się z różnej liczby źródeł (1, 2 i 10)



Rys. 6. Wartości wykładnika Hursta w zależności od liczby źródeł (n)

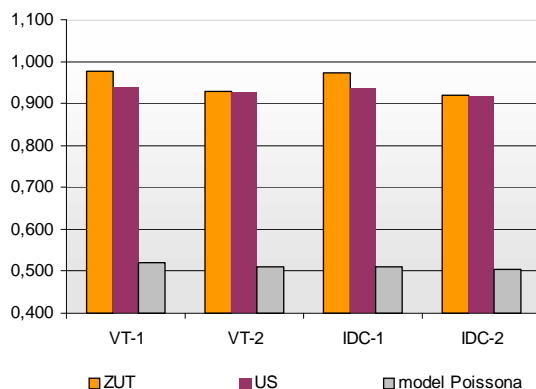
W naszym przypadku $\lambda_s^M = 0.5 \cdot M\lambda_{on}$ oraz $\lambda_s^M = \lambda_r = 7.434$ (średnia intensywność ruchu zmierzonego w sieci). Zależność $H(n)$ zbadano dla liczby źródeł $n_{max} = 200$ biorąc pod uwagę planowane wartości wykładnika Hursta kolejno (a-f na rys. 6): 0.95, 0.9, 0.8, 0.7, 0.6, 0.55. Otrzymane wyniki odbiegają zatem nieco od wartości teoretycznych (tab. 1).

Tab. 1. Zestawienie uśrednionych wyników modelu on/off dla różnych wartości H

Wartość teoretyczna wykładnika Hursta	Średnia wartość wykładnika Hursta ($n=200$)	Odchyłka od wartości teoretycznej
0.95	0.925	-0.025
0.9	0.884	-0.016
0.8	0.838	0.038
0.7	0.774	0.074
0.6	0.711	0.111
0.55	0.653	0.103

5. WNIOSKI

Porównując wyniki estymacji poziomu samopodobieństwa dla ruchu rzeczywistego z wynikami jakie daje model Poissona można zauważyć dużą różnicę w poziomie samopodobieństwa (rys. 7). Założenie, że strumień danych ma charakter fraktalny, o wcześniej ustalonym, wyestymowanym (np. metody VT, IDC) poziomie samopodobieństwa wyrażonym wartością wykładnika Hursta, jest zatem bardziej właściwe, niż założenie, że strumień ma charakter procesu Poissona. Ma to ogromne znaczenie w analizie systemów kolejkowych takich jak routery czy switchy, gdzie wyższy poziom samopodobieństwa skutkuje niższym poziomem jakości usług (Quality of Service – QoS) w stosunku do klasycznego podejścia wykorzystującego model Poissona.



Rys. 7. Wyniki estymacji poziomu samopodobieństwa metodami VT oraz IDC

Jeżeli chodzi o model on/off mający odzwierciedlać rzeczywisty ruch w sieciach teleinformatycznych, to jego konstrukcja naturalnie wynika z zasady działania interfejsów sieciowych, które w okresie 'on' generują serię pakietów, a w okresie 'off' nasłuchują. Jednak, mimo że użyta została stosunkowo duża liczba źródeł (200) tego modelu, zależność $H(n)$ nie wykazuje zbieżności do wartości teoretycznej. Pomijając tę wadę, model on/off zdecydowanie lepiej przybliża rzeczywisty ruch o strukturze fraktalnej, wypierając klasyczny model Poissona i stanowi dobre narzędzie do symulacji związanych np. z szacowaniem opóźnień czy poziomów strat w sieci.

6. BIBLIOGRAFIA

- [1] M. Taqqu, W. Willinger, R. Scherman. "Proof of a Fundamental Result in Self-Similar Traffic Modeling". *Computer Communication Review*, 27 (1997) 5-23
- [2] A.J. Field, U. Harder, P.G. Harrison: *Measurement and Modelling of Self-similar Traffic in Computer Networks*. IEE Proc-Commun., 151 (4), Aug 2004, 355-38
- [3] Beran, J. *Statistics for Long-Memory Processes*. Chapman and Hall, New York, 1994.
- [4] Li, M., Zhao, W., Jia, W., Long, D., Chi, C. *Modeling autocorrelation functions of self-similar teletraffic in telecommunication networks based on optimal approximation in Hilbert space*. *Applied Mathematical Modelling* 27, 2003, 155-168
- [5] El-Abdouni-Khayari R., Sadre R., Haverkort, B., R., *The pseudo self-similar traffic model: application and validation*. *Performance Evaluation*, 56 (1-4), 2004, 3-22.
- [6] Doukhan, P., Oppenheim, G., Taqqu, M. *Theory and Applications of Long-Range Dependence*. Birkhauser, 2002.
- [7] E. Perrin, R. Harba, C. Berzin-Joseph, I. Iribarren, A. Bonami. "n-th-Order Fractional Brownian Motion and Fractional Gaussian Noise". *IEEE Transactions on Signal Processing*, V.2001 49/5.