KOLMOGOROV-WIENER FILTER FOR CONTINUOUS TRAFFIC PREDICTION IN THE GFSD MODEL

Gorev V. N. – PhD, Associate Professor of the Department of Information Security and Telecommunications, Dnipro University of Technology, Dnipro, Ukraine.
Gusev A. Yu. – PhD, Associate Professor, Professor of the Department of Information Security and Telecommunications, Dnipro University of Technology, Dnipro, Ukraine.
Kornienko V. I. – Dr. Sc., Professor, Head of the Department of Information Security and Telecommunications, Dnipro University of Technology, Dnipro, Ukraine.

ABSTRACT

Context. We investigate the Kolmogorov-Wiener filter weight function for the prediction of continuous stationary telecommunication traffic in the GFSD (Gaussian fractional sum-difference) model.

Objective. The aim of the work is to obtain an approximate solution for the corresponding weight function and to illustrate the convergence of the truncated polynomial expansion method used in this paper.

Method. The truncated polynomial expansion method is used for the obtaining of an approximate solution for the Kolmogorov-Wiener weight function under consideration. In this paper we used the corresponding method on the basis of the Chebyshev polynomials of the first kind orthogonal on the time interval on which the filter input data are given. It is expected that the results based on other polynomial sets will be similar to the results obtained in this paper.

Results. The weight function is investigated in the approximations up to the eighteen-polynomial one. It is shown that approximations of rather large numbers of polynomials lead to a good coincidence of the left-hand side and the right-hand side of the Wiener-Hopf integral equation. The quality of the coincidence is illustrated by the calculation of the corresponding MAPE errors.

Conclusions. The paper is devoted to the theoretical construction of the Kolmogorov-Wiener filter for the prediction of continuous stationary telecommunication traffic in the GFSD model. The traffic correlation function in the framework of the GFSD model is a positively defined one, which guarantees the convergence of the truncated polynomial expansion method. The corresponding weight function is obtained in the approximations up to the eighteen-polynomial one. The convergence of the method is illustrated by the calculation of the MAPE errors of misalignment of the left-hand side and the right-hand side of the Wiener-Hopf integral equation under consideration. The results of the paper may be applied to practical traffic prediction in telecommunication systems with data packet transfer.

KEYWORDS: Kolmogorov-Wiener filter weight function, continuous telecommunication traffic, truncated polynomial expansion method, GFSD model, Chebyshev polynomials of the first kind.

ABBREVIATIONS

GFSD is the Gaussian fractional sum-difference;
ARMA is an autoregressive moving average;
ARIMA is an autoregressive integrated moving average;
FARIMA is a fractional autoregressive integrated moving average;
MAPE is a mean absolute percentage error.

NOMENCLATURE

$T$ is a time interval on which the input process data are observed;
$z$ is a time interval for which the forecast should be made;
$h(t)$ is the Kolmogorov-Wiener filter weight function;
$H$ is the Hurst exponent;
$R(t)$ is a traffic correlation function in the GFSD model;
$\Gamma(x)$ is the gamma function;
$\theta \in (0,1)$ is a constant which depends on the packet arrival rate;
$d$ is a fractional differencing parameter of the model;
$a,b$ are auxiliary constants;

$n$ is a number of polynomials in the corresponding approximations;
$g_s$ are the coefficients multiplying the polynomials;
$S_s(t)$ are the Chebyshev polynomials of the first kind orthogonal on $t \in [0,T]$;
$T_s(x)$ are the Chebyshev polynomials of the first kind orthogonal on $x \in [-1,1]$;
$\text{Left}(t)$ is the left-hand side of the Wiener-Hopf integral equation;
$\text{Right}(t)$ is the right-hand side of the Wiener-Hopf integral equation;
$G_s^b$ are the integral brackets;
$B_s^b$ are free terms in the linear set of algebraic equations in $g_s$;
$N$ is a number of points in the numerical integration.

INTRODUCTION

The problem of telecommunication traffic prediction is an urgent problem for telecommunications. For example, it is important for the optimization of network resources, for the detection of cyber-attacks and for network planning, see [1–3].
There are many approaches to telecommunication traffic prediction, which are used in different situations. In fact, the traffic may be treated as a time series. For example, the so-called gray model approach [4] may be used for the monotone, nonnegative and smooth time-series prediction, the ARMA model may predict stationary and some special non-stationary time series [1]. More sophisticated approaches, for example, such as ARIMA, FARIMA approaches and neural networks, may be used in more complex cases [2, 3].

However, another approach that may be applicable to the prediction of stationary and rather smooth telecommunication traffic is the approach based on the Kolmogorov-Wiener filter. The investigation of such an approach and its applicability may be of interest because of its simplicity in comparison with many approaches known in the literature. As far as we know, the investigation of the Kolmogorov-Wiener filter approach is not enough developed in the literature.

There are many mathematical models that may describe telecommunication traffic, see [5]. Our recent papers [6–8] were devoted to the theoretical construction of the Kolmogorov-Wiener filter for telecommunication traffic in the power-law structure function model and the fractional Gaussian noise model. In this paper we investigate the corresponding filter for the traffic prediction in the GFSD (Gaussian fractional sum-difference) model proposed in [9].

The object of study is the Kolmogorov-Wiener filter for the prediction of continuous stationary telecommunication traffic in the GFSD model.

The subject of study is the weight function of the corresponding filter.

The aim of the work is to obtain the weight function on the basis of the truncated polynomial expansion method.

1 PROBLEM STATEMENT

As is known [10], the Kolmogorov-Wiener weight function for the prediction of continuous time series obeys the Wiener-Hopf integral equation

\[ \int_0^t d\theta(t) R(t-\tau) = R(t+z). \] (1)

The problem statement is as follows: to obtain the unknown filter weight function as an approximate solution of the Wiener-Hopf integral equation (1) on the basis of the truncated polynomial expansion method.

2 REVIEW OF THE LITERATURE

The GFSD (Gaussian fractional sum-difference) model for mathematical traffic description was proposed in paper [9]. In [9] it is stressed that the corresponding model gives a good mathematical description of live packet traces for traffic in both directions of 3 Internet links: Auckland, Leipzig, and Bell.

Our previous papers [6–8] were devoted to the theoretical construction of the Kolmogorov-Wiener filter for telecommunication traffic in the power-law structure function model and the fractional Gaussian noise model, but we don’t know any papers devoted to the Kolmogorov-Wiener filter investigation for telecommunication traffic in the GFSD model. This fact justifies the scientific novelty of this paper.

In this paper we solve the integral equation (1) on the basis of the truncated polynomial expansion method. This method is a special case of the Galerkin method [11] and in this paper this method is based on the Chebyshev polynomials of the first kind orthogonal on the time interval \( t \in [0, T] \). Of course, another polynomial system may also be chosen. But, for example, in [8] three different polynomial systems (Chebyshev polynomials of the second kind, Chebyshev polynomials of the first kind, and polynomials orthogonal without weight) were investigated for two different traffic models, and it was shown that the results for all three polynomial systems are, in fact, similar. Therefore, we expect that the results for other polynomial systems will also be similar in the framework of the GFSD model.

3 MATERIALS AND METHODS

The traffic correlation function for the discrete GFSD model is as follows [9]:

\[ R(t) = (1-0) \frac{2(1-d)t^2 - (1-d)^2}{t^2 -(1-d)^2} \times \frac{\Gamma(1-d)}{\Gamma(d)} \frac{\Gamma(t+d)}{\Gamma(t-d+1)}, \quad t \geq 1. \] (2)

The value \( d = 0.31 \) is chosen [9]. First of all, we should propose an expression for \( R(t) \) for \( t \in (0,1) \). For \( t \geq 1 \) it is natural enough to require the correlation functions in the discrete and continuous cases to be the same. However, expression (2) is obviously not applicable for \( R(t) \) for \( t \in [0,1] \) because it is not convergent at \( t = 1-d \) and the inequality \([R(t)] \leq R(0)\) fails. In [9] the corresponding model is written for the traffic with a variance equal to 1, so we propose to define the correlation function for \( t \in [0,1] \) as

\[ R(t) = at^b + 1 \] (3)

where the constants \( a \) and \( b \) are calculated on the basis of the joining conditions

\[ R_{b=-1} = R_{b=-1,0} \frac{dR}{dt}_{b=-1,0} = \frac{dR}{dt}_{b=-1,0}. \] (4)
In such a case the inequality $|R(t)| \leq R(0)$ holds, the process variance $R(0) = 1$, the function $R(t)$ and its derivative are continuous on the time axis. It should also be stressed that if $T \gg 1$, then the leading order in the integrals calculated in the paper is given by the interval $t \in [1, T]$ rather than the interval $t \in [0,1]$, so the choice of $R(t)$ for $t \in [0,1]$ may not have any significant effect on the paper results. For example, for the value $\theta = 0.8$ the following values are obtained: $a = -0.845$, $b = -0.206$ (rounded off to 3 significant digits). It is also known that the correlation function of a stationary random process is even one. Therefore, finally, we propose the following form of the correlation function $R(t)$ in the continuous case:

$$R(t) = \frac{a|t|^2 + b|t|}{(1-0)^2 - (1-d)^2} \times \frac{n(1-d) - \Gamma|t| + d}{\Gamma|t| - \Gamma|t-1|}, \quad |t| \geq 1$$

the constants $a$ and $b$ are given by expressions (4).

In what follows, we solve the integral equation (1) with the correlation function (5) as follows. The unknown weight function is sought in the form of the truncated polynomial expansion:

$$h(t) = \sum_{k=0}^{n-1} g_k S_k(t),$$

where

$$S_k(t) = T_k\left(\frac{2t}{T} - 1\right),$$

see [8] and the definition of $T_k(x)$ in [12]. After substituting (6) into (1), multiplying by $S_k(t)$, and integrating over $t$, one can obtain the following expression for the coefficients $g_k$ in the $n$-polynomial approximation:

$$\begin{pmatrix} g_0 \\ g_1 \\ \vdots \\ g_{n-1} \end{pmatrix} = \begin{pmatrix} G_{00} & G_{01} & \cdots & G_{0,n-1} \\ G_{10} & G_{11} & \cdots & G_{1,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ G_{n-1,0} & G_{n-1,1} & \cdots & G_{n-1,n-1} \end{pmatrix}^{-1} \begin{pmatrix} B_0 \\ B_1 \\ \vdots \\ B_{n-1} \end{pmatrix},$$

(8)

where

$$G_{kk} = \int_0^T d\tau dS_k(t) S_k(\tau) R(\tau - \tau),$$

$$B_k = \int_0^T d\tau S_k(\tau) R(\tau + \tau).$$

(9)

On the basis of the fact that $R(t)$ is an even function and on the basis of the properties of the Chebyshev polynomials, one can derive the following properties of the integral brackets $G_{kk}$:

$$G_{kk} = G_{-k},$$

$$G_{kk} = 0 \text{ if } k \text{ and } s \text{ are of different parities},$$

(10)

see a similar derivation for another polynomial system in [6]. The properties (10) significantly reduce the computation time.

Let us discuss the calculation of the integral brackets (9) in detail. In this paper they are calculated in the Wolfram Mathematica package as follows. Unfortunately, the package fails to calculate the integral brackets on the basis of the explicit expression (9) with account for (5), so expression (9) is rewritten on the basis of the following change of the variables:

$$\begin{align*}
G_{kk} &= \int_0^T d\tau dS_k(t) S_k(\tau) R(\tau - \tau) \\
&= \{x = t - \tau, y = t + \tau\} = \\
&= \frac{1}{2} \int_0^T dx \left(2\left(1-d\right)x^2 - (1-d)^2\right) \frac{\Gamma(1-d)}{\Gamma(d)} \\
&\times \frac{\Gamma(-x+d)}{\Gamma(-x-d+1)} \int_{-x}^{x} dy S_k\left(\frac{x-y}{2}\right) S_k\left(\frac{x+y}{2}\right) \\
&+ \frac{1}{2} \int_0^T dx \left(2\left(1-d\right)x^2 - (1-d)^2\right) \frac{\Gamma(1-d)}{\Gamma(d)} \\
&\times \frac{\Gamma(x+d)}{\Gamma(x-d+1)} \int_{-x}^{x} dy S_k\left(\frac{x-y}{2}\right) S_k\left(\frac{x+y}{2}\right) \\
&+ \frac{1}{2} \int_0^T dx a(x^{k+1}) \\
&\times \int_{-x}^{x} dy S_k\left(\frac{x-y}{2}\right) S_k\left(\frac{x+y}{2}\right) \\
&+ \frac{1}{2} \int_0^T dx a(x^{k+1}) \\
&\times \int_{-x}^{x} dy S_k\left(\frac{x-y}{2}\right) S_k\left(\frac{x+y}{2}\right),
\end{align*}$$

(11)

the integral brackets are calculated in the Wolfram Mathematica package on the basis of the explicit expressions (11).

4 EXPERIMENTS

We investigate the results for the following numerical values of the parameters:

$$T = 100, \ z = 3, \ d = 0.31, \ \theta = 0.8.$$  

(12)
The integral brackets \( G_a \) are calculated on the basis of (11), the free terms \( B_a \) are calculated on the basis of (9), for each polynomial approximation the coefficients multiplying the polynomials are calculated on the basis of (8), the result for the corresponding weight function is given by (6). In order to justify the proposed algorithm, we compare the left-hand side and the right-hand side of the Wiener-Hopf integral equation (1) by calculation of the MAPE errors:

\[
\text{Left} (t) = \int_{\tau}^{T} dh(t) R(t - \tau),
\]
\[
\text{Right} (t) = R(t + z),
\]
\[
\text{MAPE} = \frac{1}{T} \int_{0}^{T} \left( \frac{\text{Left} (t) - \text{Right} (t)}{\text{Right} (t)} \right) dt \cdot 100\%.
\]

In this paper the MAPE error is roughly estimated as

\[
\text{MAPE} = \frac{1}{N} \sum_{j=0}^{N} \left( \frac{\text{Left} \left( \frac{jT}{N} \right) - \text{Right} \left( \frac{jT}{N} \right)}{\text{Right} \left( \frac{jT}{N} \right)} \right) \cdot 100\%,
\]

\[N = 10^2.\]

It should be stressed that \( \text{Left} (t) \) is calculated on the basis of the Wolfram Mathematica package using the following explicit expressions:

\[
\text{Left} (t) = \begin{cases} 
L_r (t), & t \in (1, T-1) \\
L_2 (t), & t \geq T - 1 \\
L_1 (t), & t \leq 1
\end{cases}
\]

where

\[
L_r (t) = \int_{i-T}^{i} dh(t - x) \frac{2(1-d)x^2 - (1-d)^2}{x^2 - (1-d)^2} \times \\
\times (1-0) \frac{\Gamma (1-d)}{\Gamma (d)} \frac{\Gamma (-x+d)}{\Gamma (-x-d+1)} + \\
\int_{0}^{1} h(t - x) (a(-x)^6 + 1) dx + \\
\int_{0}^{1} h(t - x) (ax^6 + 1) dx + \\
\int_{0}^{1} dh(t - x) \frac{2(1-d)x^2 - (1-d)^2}{x^2 - (1-d)^2} \times \\
\times (1-0) \frac{\Gamma (1-d)}{\Gamma (d)} \frac{\Gamma (x+d)}{\Gamma (x-d+1)} .
\]

The calculation of all the integrals in this paper is made on the basis of the NIntegrate function built in the Wolfram Mathematica package.

**5 RESULTS**

The obtained results for the MAPE are given in Table 1.

Table 1 – MAPE for the approximations of \( n \) polynomials rounded off to two decimal places

<table>
<thead>
<tr>
<th>( n )</th>
<th>MAPE, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>26.59</td>
</tr>
<tr>
<td>2</td>
<td>17.95</td>
</tr>
<tr>
<td>3</td>
<td>11.44</td>
</tr>
<tr>
<td>4</td>
<td>8.43</td>
</tr>
<tr>
<td>5</td>
<td>5.92</td>
</tr>
<tr>
<td>6</td>
<td>4.65</td>
</tr>
<tr>
<td>7</td>
<td>3.51</td>
</tr>
<tr>
<td>8</td>
<td>2.93</td>
</tr>
<tr>
<td>9</td>
<td>2.35</td>
</tr>
<tr>
<td>10</td>
<td>2.11</td>
</tr>
<tr>
<td>11</td>
<td>1.79</td>
</tr>
<tr>
<td>12</td>
<td>1.67</td>
</tr>
<tr>
<td>13</td>
<td>1.47</td>
</tr>
<tr>
<td>14</td>
<td>1.41</td>
</tr>
<tr>
<td>15</td>
<td>1.27</td>
</tr>
<tr>
<td>16</td>
<td>1.24</td>
</tr>
<tr>
<td>17</td>
<td>1.14</td>
</tr>
<tr>
<td>18</td>
<td>1.14</td>
</tr>
</tbody>
</table>

In order to illustrate the coincidence of the left-hand side and the right-hand side, we build the corresponding graphs for the 18-polynomial approximation, see Fig. 1. Therefore, it may be concluded that the truncated polynomial expansion method is convergent for the problem under consideration, and approximations of rather large numbers of polynomials are rather accurate.
The traffic in the GFSD model. Results of the truncated polynomial expansion method will be almost the same for different polynomial systems for similar in the fractional Gaussian noise model and in the structures of the correlation function for 1.

The Kolmogorov-Wiener filter weight function for the prediction of continuous stationary telecommunication traffic in the GFSD (Gaussian fractional sum-difference) model is investigated. The truncated polynomial expansion method based on the Chebyshev polynomials of the first kind is used in a search for an approximate solution of the Wiener-Hopf integral equation. The method is realized on the basis of the Chebyshev polynomials of the first kind orthogonal on the required time interval. The traffic correlation function in the GFSD (Gaussian fractional sum-difference) model is a non-negative one, which justifies the convergence of the proposed method. Approximations up to the eighteen-polynomial one are investigated, and the method convergence is illustrated by the calculation of the corresponding MAPE errors of misalignment of the left-hand side and the right-hand side of the Wiener-Hopf integral equation under consideration. It is shown that approximations of a large number of polynomials are rather accurate.

In [8], the truncated polynomial expansion method is investigated for three polynomial systems (Chebyshev polynomials of the second kind, Chebyshev polynomials of the first kind, and polynomials orthogonal without weight) for two different traffic models (powel-law structure function model and fractional Gaussian noise model), and it is shown that the results for all the three polynomial sets are, in fact, the same. It should also be stressed that the structures of the correlation function for $t \gg 1$ are similar in the fractional Gaussian noise model and in the GFSD model (see [9]). Therefore, it is expected that the results of the truncated polynomial expansion method will be almost the same for different polynomial systems for the traffic in the GFSD model.

CONCLUSIONS

The Kolmogorov-Wiener filter weight function for the prediction of continuous stationary telecommunication traffic in the GFSD (Gaussian fractional sum-difference) model is calculated on the basis of the truncated polynomial expansion method. Approximations up to the 18-polynomial one are investigated. It is shown that approximations of rather large numbers of polynomials lead to a good coincidence between the left-hand side and the right-hand side of the Wiener-Hopf integral equation.

The results of this paper may be useful for the practical prediction of telecommunication traffic in systems with data packet transfer.

The scientific novelty of the paper is the fact that for the first time the Kolmogorov-Wiener filter weight function is calculated for the telecommunication traffic prediction in the GFSD (Gaussian fractional sum-difference) model.

The practical significance is that the obtained results may be applied to the practical prediction of telecommunication traffic in systems with data packet transfer.

Prospects for further research are to obtain a practical prediction on the basis of the obtained results and to investigate the solutions for the weight function on the basis of a non-polynomial orthogonal function system.

ACKNOWLEDGEMENTS

The work is made in the framework of the Project “Research of methods of increase of efficiency of the automated control of thermal work of units of high power of industrial and household function” (State Registration No. 0122U002601) of the Dnipro University of Technology.

REFERENCES


© Gorev V. N., Gusev A. Yu., Kornienko V. I., 2022
DOI 10.15588/1607-3274-2022-3-3
ФІЛЬТР КОЛМОГОРОВА-ВІНЕРА ДЛЯ ПРОГНОЗУВАННЯ НЕПЕРЕРВНОГО ТРАФІКУ У GFSD МОДЕЛІ

Горев В. М. – канд. фіз.-мат. наук, доцент кафедри безпеки інформації та телекомунікацій, Національний технічний університет «Дніпровська Політехніка», Дніпро, Україна.

Гусєв О. Ю. – канд. фіз.-мат. наук, доцент, професор кафедри безпеки інформації та телекомунікацій, Національний технічний університет «Дніпровська Політехніка», Дніпро, Україна.

Корнієнко В. І. – д-р техн. наук, професор, завідувач кафедри безпеки інформації та телекомунікацій, Національний технічний університет «Дніпровська Політехніка», Дніпро, Україна.

АКТУАЛЬНІСТЬ. Досліджено вагову функцію фільтра Колмогорова-Вінера для прогнозування неперервного стаціонарного телекомунікаційного трафіку у GFSD (Gaussian fractional sum-difference) моделі.

МЕТА РОБОТИ. Метою роботи є отримати наближене розв’язок для відповідної функції та проілюструвати збіжність методу обірваних розв’язків за поліномами.

МЕТОД. Метод обірваних розв’язків за поліномами використано для отримання наближеного розв’язку для досліджуваної функції трансформації Колмогорова-Вінера. В цій статті нами використано відповідний метод на основі поліномів Чебишева першого роду як експоненції положення на часовому відрізку, на якому задані вхідні дані фільтра. Одержано функцію, що відповідає відповідному розв’язку, що збігається з наближеннями до вісімнадцяти поліномів включно.

РЕЗУЛЬТАТИ. Вагова функція фільтра Колмогорова-Вінера для прогнозування неперервного стаціонарного телекомунікаційного трафіку у GFSD моделі. Переглянуто вагову функцію фільтра Колмогорова-Вінера для прогнозування неперервного стаціонарного телекомунікаційного трафіку у GFSD моделі.

Висновки. Статтю присвячено теоретичній побудові фільтра Колмогорова-Вінера для прогнозування неперервного стаціонарного телекомунікаційного трафіку у GFSD моделі. Кореляційна функція трафіку в рамках GFSD моделі є позитивно визначеною, що гарантує збіжність методу обірваних розв’язків за поліномами. Відповідна вагова функція отримана у наближеннях до вісімнадцяти поліномів включно. Збіжність методу проілюстрована обчисленням середніх або абсолютних помилок нев’язки.

КЛЮЧОВІ СЛОВА: вагова функція фільтра Колмогорова-Вінера, неперервний телекомунікаційний трафік, метод обірваних розв’язків за поліномами, GFSD модель, поліноми Чебишева першого роду.

© Gorev V. N., Gusev A. Yu., Korniienko V. I., 2022
DOI 10.15588/1607-3274-2022-3-3
данные фильтра. Ожидается, что результаты, которые будут базироваться на других полиномиальных системах, будут ана- логичны результатам, полученным в данной статье.

Результаты. Весовая функция исследована в приближениях до восемнадцати полиномов включительно. Показано, что приближения достаточно большого числа полиномов приводят к хорошему совпадению левой и правой частей интегрального уравнения Винера-Хопфа. Качество совпадения проиллюстрировано вычислением соответствующих средних абсолютных ошибок невязки.

Выводы. Статья посвящена теоретическому построению фильтра Колмогорова-Винера для прогнозирования непрерывного стационарного телекоммуникационного трафика в GFSD модели. Корреляционная функция трафика в рамках GFSD модели положительно определена, что гарантирует сходимость метода оборванных разложений по полиномам. Соответствующая весовая функция получена в приближениях до восемнадцати полиномов включительно. Сходимость метода проиллюстрирована вычислением средних абсолютных ошибок невязки лево й и правой частей исследуемого интегрального уравнения Винера-Хопфа. Результаты работы могут быть применены к практическому прогнозированию трафика в телекоммуникационных сетях с пакетной передачей данных.

КЛЮЧЕВЫЕ СЛОВА: весовая функция фильтра Колмогорова-Винера, непрерывный телекоммуникационный трафик, метод оборванных разложений по полиномам, GFSD модель, полиномы Чебышева первого рода.

ЛИТЕРАТУРА


