Problematic Issues of Approximation and Interpolation in Signal Processing in Secure Information Systems

Olena Nehodenko[1], Svitlana Shevchenko[2], Nataliia Trintina[1], Volodymyr Astapenia[2], and Oleksandr Tereshchenko[3]

[1] *State University of Telecommunications, 7 Solomyanska str., 03110, Kyiv, Ukraine*

[2] *Borys Grinchenko Kyiv University, 18/2 Bulvarno-Kudriavska str., 04053, Kyiv, Ukraine*

[3] *National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute," 37 Peremohy ave., 03056, Kyiv, Ukraine*

Abstract

One of the basic indicators of information systems security is the availability and integrity of information. When transmitting a visual message, the signals are distorted due to the influence of various random factors. It is established that an ideal channel must have an absolutely rectangular frequency characteristic for the transmission of all components of the amplitude-frequency spectrum of a continuous random process without distortion and loss. Such frequency response is the result of a complex interaction of the parameters of the set of devices which are the part of the information transmission channel and perform the above procedures for conversion and signal processing. Some information and cybersecurity tasks require efficient interpolation, and in some situations a combination of interpolation and approximation techniques. The paper proposes a mathematical model of information signals based on fundamental trigonometric splines, which allow to take into account the differential properties of information signals. It is shown that it is feasible to use trigonometric splines as mathematical models of information signals, and it is more feasible to apply fundamental approximation trigonometric splines for recovery of signals as components of filters. The importance of this approach is explained by the fact that only fundamental functions are subject to processing when applying linear methods. This fact allows performing the necessary calculations for processing experimental data in two stages. In the first stage, calculations are performed related to the processing of fundamental functions (these calculations can be performed in advance). In the second stage, calculations are performed that take into account the values of the reproduced functions. It is shown that the method of phantom nodes should be used for interpolation of the useful signal in information networks, which allows to increase the accuracy of information processing.

Keywords

Information system, interpolation, approximation, fundamental trigonometric splines, signal processing.

1. Introduction

In information and communication systems, there are many problems where it is necessary to provide a high-quality approximation and interpolation of processes or objects with limited or inaccurate primary data. Thus, in the vast majority of information systems used by humans, the input and output signal is a

CPITS-II-2021: Cybersecurity Providing in Information and Telecommunication Systems, October 26, 2021, Kyiv, Ukraine
EMAIL: negodenkoav@i.ua (O. Nehodenko); s.shevchenko@kubg.edu.ua (S. Shevchenko); trintina2015@gmail.com (N. Trintina); v.astapenia@kubg.edu.ua (V. Astapenia); alexandr.tereschenko2014@gmail.com (O. Tereshchenko)
ORCID: 0000-0001-6645-1566 (O. Nehodenko); 0000-0002-9736-8623 (S. Shevchenko); 0000-0001-6827-4030 (N. Trintina); 0000-0003-0124-216X (V. Astapenia); 0000-0003-0536-2708 (O. Tereshchenko)

continuous function. Examples are voice messages, visual images, etc. (within the macroscopic model of the organs of formation and perception). It is important to note that these are random processes in all cases for the recipient. It is known that only such processes carry information [1, 2]. When communicating using telephone, our voice message (elastic vibrations of air) is converted into a continuous electrical signal in the microphone, which is a random process $A(t)$. Further transformations in modern systems include analog-to-digital conversion and a number of other procedures (error correction and cryptographic coding, modulation of carrier harmonic motion, amplification, radiation, and inverse transformations). The user (recipient) eventually needs the message in the "ideal" original form, but due to the influence of various random factors, they receive it in a distorted, somewhat different from the original form $\tilde{A}(t)$.

The same situation happens when transmitting a continuous visual message. This means that the channel between the source and the recipient of the continuous message is not ideal and needs to be improved to increase the availability and integrity of information as the basic indicators of the security of information systems. The ideal channel should have an absolutely rectangular frequency response to transmit all components of the amplitude-frequency spectrum of a continuous random process without distortion and loss. Such frequency response is the result of a complex interaction of the parameters of the set of devices which are the part of the information transmission channel and perform the above procedures for conversion and signal processing.

This requires:

1. To know the essential components and the width of the amplitude-frequency spectrum of a continuous random process that is formed by the source (the width of this spectrum is formally infinite) with sufficient accuracy to efficiently recover the process.

2. to determine and minimize potential losses in the quality of recovery of the initial process after its sampling under conditions of artificial limitation of the spectrum when determining the sampling interval (sampling frequency) within the framework of Kotelnikov's theorem.

3. To identify and minimize potential losses in the quality of recovery of the initial process due to noise that occurs during quantization.

4. to ensure the formation of a rectangular frequency response of the channel in this band, taking into account the influence of all its components.

5. To identify and adequately take into account the impact on the frequency response of additive and multiplicative interferences which are inherent in the relevant channels (especially radio channels).

2. Some Examples which Demonstrate the Need of Identification of Processes and Images in Information Systems

All of the above cases are part of the problems of identification of processes and images.

This category of problems includes the approximation of the autocorrelation function (ACF) $R_a(\tau)$ according to the the results of its calculations based on the set of samples of the random process $A(t)$ in conditions when part of the samples is lost or distorted. Often this ACF is used to determine the amplitude-frequency spectrum of a random process $G_a(\omega)$ according to Khinchin's theorem [3].

$$G_a(\omega) = \int_{-\infty}^{\infty} R_a(\tau) \exp(-j\omega\tau) d\tau$$

where $\tau = t_2 - t_1$;; $\omega = 2\pi f$ is the angular frequency.

An example of the use of the ACF is also the prediction of a random process based on the parameters of this function. Such prediction (extrapolation of processes) is necessary for the formation of artificial intelligence algorithms in a broad sense and to solve relatively narrow problems of adaptation of individual devices ranging from industrial robots, drones, adaptive surveillance systems ending with smart home equipment.

When analyzing and processing visual information, the raster image can be considered as a random field, for which there is also the concept of spatial correlation function [4], and taking into account that such an image is a dynamic process in most cases, it is a spatial-temporal correlation function.

A number of tasks in the field of information and cyber security require effective interpolation, and in some situations a combination of methods of interpolation and approximation [5].

Such tasks include the following:

1) from the standpoint of ensuring the availability and integrity of information:

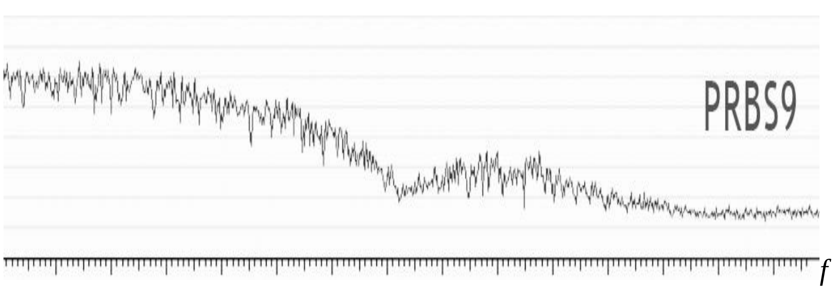
- restoration of the voice message in case of the loss of a certain part of its samples under the influence of interference or as a result of intentional distortion;
- restoration of the visual picture in case of the loss of a certain part of the fragments (pixels) under the influence of interference or as a result of intentional distortion (for example, during identification of a person, car numbers when monitoring the movement using surveillance cameras, etc.);
- as an alternative to error correction coding;
- recovery of information from the damaged drive.

In such cases (in digital representation) we are dealing with a random sequence of rectangular (in the first approximation) pulses, some of which are distorted or lost (when there are "windows" of damaged characters).

2) from the standpoint of information security, there are inverse tasks:

- determination of the boundaries of artificial distortions, beyond which the secrecy in the potential channel of information leakage can be considered reliable. That is, distortion of the message (voice, visual image, etc.) in such a channel, so that the opposing party does not recover the message. This may also be the case when you need to make the voice or appearance of an important witness, or certain details of the interior of a room involved in IoT technology unrecognizable;
- as a method of testing cryptographic procedures (ciphers) to assess their effectiveness;
- in cryptanalysis (decryption while the key is unknown).

In these situations, it is necessary to distinguish the true (which contains information) random sequence from a mixture of it with another interfering random sequence. The frequency spectra of such sequences have an envelope according to the shape of the spectrum of a rectangular pulse with chaotic filling. As an example, Fig. 1 presents the experimentally obtained spectrum of a pseudo-random sequence [6].



$1/\tau_i$

Figure 1: The spectrum of the pseudo-random sequence

Thus, the solution of these problems either in the time or in the frequency domain requires the involvement of a powerful mathematical apparatus.

Directional antenna systems play an important role in information retrieval and transmission technologies. They determine the angular coordinates of objects that reflect (radar, optoelectronic location, sonar) or emit waves (radio reconnaissance, radio-technical reconnaissance, ultrasonic reconnaissance, etc.), and in satellite, radio relay and tropospheric communication systems they ensure the availability and integrity of information for the consumer, which is in a certain angular direction. In such systems, the accuracy of the orientation of the maximum of the antenna directionality characteristic f (θ, φ), i.e. the angular spectrum that creates the antenna opening, is due to the amplitude-phase current distribution [7].

$$f(\theta,\varphi)=\sum_{i=0}^{N-1} A_i e^{j\left[\varphi_i+\frac{2\pi}{\lambda}\rho_i\cos\gamma_i\right]}$$

where: A_i is the normalized current amplitude in the antenna array element; φ_i is the phase of the current in the element; λ is the wavelength; ρ_i – the distance from the central to the i-th element of the array; γ_i – the angle between the direction to the point of space with the angular coordinates θ, φ and the direction ρ_i.

Under the influence of random factors, the amplitude-phase distribution in the opening may be distorted and the question arises of determining of the directional characteristic. These issues are considered using classical methods in the paper [8]. The results of measuring the angular distribution of the field strength, which created by the antenna in space, can be obtained with errors; here we can talk about identification of distortions on an aperture or detection of non-working vibrators [9,10].

At the present stage, these problems have not been fully resolved. One of the mathematical methods that can help to solve such problems is the use of spline approximation and spline interpolation [11,12].

3. Spline approximation and spline interpolation as a method of signal processing in secure information systems

Based on the analysis of the scientific literature [13–17], it was determined that there are properties of signals, without which the very statement of many tasks of signal processing does not make sense. Such properties of information signals are their smoothness properties, which characterize the behavior of the signal in some neighborhood of an arbitrary point belonging to the signal interval. These properties contain information about the existence of a certain number of continuous derivatives of the studied signal, as well as information about some analytical properties of these derivatives. On the basis of this theory the mathematical model of information signals was developed and explained using fundamental trigonometric splines which allow to take into account differential properties of information signals.

It is established [18] that the signal represented by the Fourier series can only be periodic. Signals of arbitrary shape can be represented by a Fourier series only approximately, since this provides for the periodic repetition of the signal interval outside its setting. At the junctions of the periods there may be breaks and fractures of the signal, as well as processing errors caused by the Gibbs phenomenon, to minimize which certain methods are used.

The paper proposes a method for attenuating the Gibbs phenomenon based on trigonometric Fourier series. This method allows to periodically extend the signal of arbitrary nature and at the same time get rid of breaks and fractures of the signal at the junctions of periods.

An improved discrete version of this method will be the basis for reproducing a useful signal in information transmission systems.

Consider the method of periodic continuation of non-periodic functions for the case when trigonometric splines are used as an approximating function.

Consider the function $f(t)$ on the interval $[0,2\pi]$. Set N interpolation nodes, $N=2n+1$, where $n=1,2,\ldots$, the uniform grid step is $h=2\pi\frac{i-1}{N}$, where $i=1,2,\ldots,N$.

The value of the function in the interpolation nodes is calculated. A sequence of values of the function $\{f(h(k\text{-}1))\}_{k=1}^{N}=\{f_k\}_{k=1}^{N}$ is obtained. Next, a trigonometric interpolation spline is constructed based on these nodes, which has the form:

$$S_{t_r}(f,\Delta_N,t)=\frac{a_0}{2}+\sum_{k=1}^{n}\alpha_k(r,N)\left[a_k^*\Phi_k^c(r,N,t)+b_k^*\Psi_k^s(r,N,t)\right], \tag{1}$$

where

$$\Phi_k^c(r,N,t)=\frac{\cos kt}{k^{r+1}}+\sum_{m=1}^{\infty}\left[\frac{\cos(mN+k)t}{(mN+k)^{r+1}}+\frac{\cos(mN-k)t}{(mN-k)^{r+1}}\right],$$

$$\Psi_k^s(r,N,t)=\frac{\sin kt}{k^{r+1}}+\sum_{m=1}^{\infty}\left[\frac{\sin(mN+k)t}{(mN+k)^{r+1}}-\frac{\sin(mN-k)t}{(mN-k)^{r+1}}\right],$$

$$\left[\alpha_k(r,N)\right]^{-1}=\frac{1}{k^{r+1}}+\sum_{m=1}^{\infty}\left[\frac{1}{(mN+k)^{r+1}}+\frac{1}{(mN-k)^{r+1}}\right],$$

$$a_0=\frac{2}{N}\sum_{i=1}^{N}f(t_i),\quad a_k^*=\frac{2}{N}\sum_{i=1}^{N}f(t_i)\cos kt_i,$$

$$b_k^*=\frac{2}{N}\sum_{i=1}^{N}f(t_i)\sin kt_i,\quad k=1,2,\dots,n.$$

The interpolation trigonometric spline S_{t_r} interpolates the function $f(t)$ at $N+1$ points, given on the segment 2π. Since the value of this trigonometric spline is definite at the point 2π and, due to the periodicity, it follows that $S_t(2\pi)=S_t(0)$. Therefore, we will consider the spline interpolation S_{t_r} only on the interpolation segment $2\pi-h$.

It is clear that the trigonometric interpolation spline is due to the fact that $f(0)\neq f(2\pi)$, in the neighborhood of the points 0 and 2π has the same defects as the Fourier series in the neighborhood of the breakpoints. Therefore, it is advisable to use the method of improving convergence, which is called the method of phantom nodes [13].

This method is as follows. An even number of phantom nodes is added to the sequence of interpolation nodes; the values in these nodes will be chosen taking into account the estimates of the derivatives, which we estimate using the divided differences in the neighborhoods of the points 0 and $2\pi-h.$. That is, we construct the function $\lambda(t), t\epsilon(2\pi-\alpha,2\pi)$ on the interval $(2\pi-\alpha,2\pi)$ based on the conditions

$$\lambda(2\pi-\alpha)=f(2\pi);\quad \lambda(2\pi)=f(0);$$
$$\lambda'^{(2\pi-\alpha)}=f'^{(2\pi)};\quad \lambda'^{(2\pi)}=f'^{(0)};$$
$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$
$$\lambda^{(k-1)}(2\pi-\alpha)=f^{k-1}(2\pi);\quad \lambda^{k-1}(2\pi)=f^{(k-1)}(0)$$

and find the values of the derivatives of the function at the corresponding points.

Addition of $2k$ $(k=1,2,\dots)$ phantom nodes increases the number of interpolation nodes on the segment $[0,2\pi]$, and reduces the step h of the interpolation grid, which now becomes equal to $h_k=2\pi\frac{i-1}{N+2k}$. Since the number of values of the interpolated function does not change, the decrease in the step of the interpolated grid leads to a decrease in the interpolation segment, which becomes equal to Nh_k.

A linear function $\varphi(t)$ is constructed on the segment $[2\pi-Nh_k,2\pi]$, that satisfies the necessary conditions

$$\varphi(t)=\begin{cases}f_N, & t=Nh_k;\\ f_1, & t=2\pi.\end{cases}$$

Calculating the values of this function in the phantom nodes $t_{N+1}, t_{N+2}, \dots, t_{N+2k-1}$, we can find the required values in these nodes.

An even number of phantom nodes is chosen in order the total number of interpolation nodes to be odd, because trigonometric interpolation of an odd number of points is more convenient. It is convenient to add a small number of phantom nodes, i.e. to set $k = 1,2$.

Phantom nodes can be added on both sides of the interpolation sequence. However, we have to construct two functions $\varphi_1(t)$ on the left and $\varphi_2(t)$ on the right. However, due to the periodicity of interpolation trigonometric splines, it is more convenient to do it on the one side.

When constructing the function $\varphi(t)$, we can require that its derivatives of a certain order also take certain values at the points Nh_k and 2π. Divided differences of the interpolated function can be used to find these values.

Also, in many cases, when constructing the function $\varphi(t)$, the exact values of the derivatives of the interpolated function $f(t)$ can be used. Figure 2 shows the algorithm for information processing.

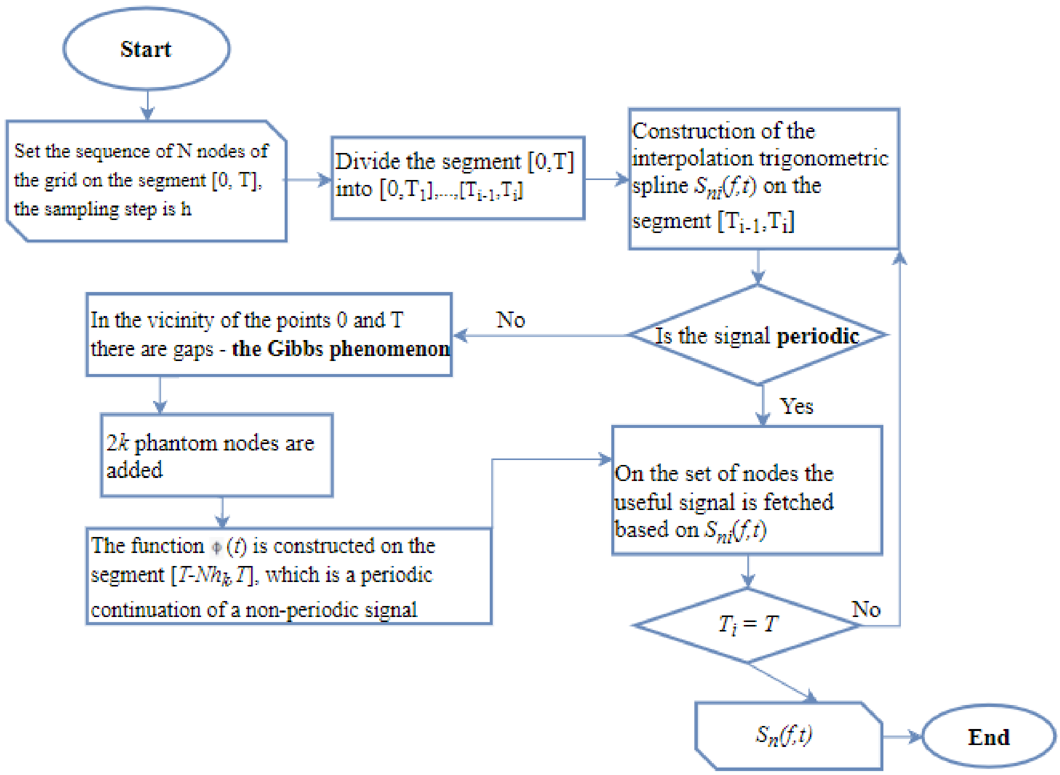


Figure 2: The scheme of the information processing algorithm

This algorithm is repeated on each of the specified segments for processing discrete information.

A generalized method for information processing during signal excretion on spline filters is proposed, which greatly simplifies the processing algorithm.

4. Conclusions

In the study of various errors of linear links in the theory of information systems for various purposes, such as linear amplifiers, filters and up to the frequency response of the information transmission channel in general, it is feasible to use periodic models of information signals. This feasibility is explained by the fact that the trigonometric functions used in the construction of periodic models are Eigen functions of linear operators, i.e. do not change within a constant when exposed to linear operators. Thus, it is proved that it is feasible to use trigonometric splines as mathematical models of information signals, and it is more feasible to use fundamental approximation trigonometric splines to recover signals as components of filters.

The importance of this approach is explained by the fact that when applying linear methods, only fundamental functions are subject to processing. This fact allows performing the necessary calculations for processing experimental data in two stages. In the first stage, calculations are performed related to the processing of fundamental functions (these calculations can be performed in advance). In the second stage, calculations are performed that take into account the values of the reproduced functions.

For the processing of information signals in practical calculations, an important place is occupied by differentiation operations to find derivative functions at the stage of interpolation, using only the values of these functions at individual points.

The improved method of signal processing in the information system on the basis of fundamental trigonometric splines allows to periodically extend the signal of arbitrary nature and at the same time to get rid of breaks and fractures of the signal at the junctions of periods. This method uses fundamental trigonometric splines, which allow real-time calculations. The method of phantom nodes should be used to interpolate the useful signal in information networks, which improves the accuracy of information processing.

Further improvement of the considered technique and application of spline approximation and spline interpolation in relation to the problems of signal processing listed above in the protected information systems is expected in the future.

5. References

[1] A. N. Kolmogorov, Three approaches to the definition of the "amount of information" // Communication problems. 1965. V. 1, 1st edition. pp. 46-49.

[2] R.L. Stratonovich, Information theory. Soviet radio, 1975.

[3] S. M. Rytov, Introduction to statistical radiophysics. Part 1. Random processes. Science, 1976. P. 496.

[4] S. M. Rytov, Y. A. Kravtsov, V. I. Tatarsky, Introduction to statistical radiophysics. Part 2. Random fields. _ M.: Science, 1978. P. 464.

[5] Cybernetics Dictionary. / Edited by V. S. Mikhalevich. 2nd ed. Ukrainian Soviet Encyclopedia named after M. P. Bazhana, 1989. P. 751.

[6] Astapenya, V., Sokolov, V. & Ageyev, D., 2020. Experimental Evaluation of an Accelerating Lens on Spatial Field Structure and Frequency Spectrum. 2020 IEEE Ukrainian Microwave Week (UkrMW). Available at: http://dx.doi.org/10.1109/ukrmw49653.2020.9252755.

[7] Y.S. Shifrin, Antennas. Kharkov: VIRTA, 1976.

[8] Y.S. Shifrin, Questions of the statistical theory of antennas. Soviet radio, 1970. P. 384.

[9] Vladymyrenko, M., et al. (2019). Analysis of Implementation Results of the Distributed Access Control System. 2019 IEEE International Scientific-Practical Conference Problems of Infocommunications, Science and Technology (PIC S&T). https://doi.org/10.1109/picst47496.2019.9061376

[10] Carlsson, A., et al. Sustainability Research of the Secure Wireless Communication System with Channel Reservation. 2020 IEEE 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET), 2020. https://doi.org/10.1109/tcset49122.2020.235583

[11] Encyclopedia of mathematics: Editor-in-chief I. M. Vinogradov. vol. 5. Soviet encyclopedia. 1984. P. 1248.

[12] V.P. Denisyuk, Fundamental functions and trigonometric splines: Monograph. PJSC Vipol, 2015. P. 296.

[13] A.P. Bondarchuk, The task of synthesizing algorithms for optimal processing of folding signals in broadcast communication systems / A.P. Bondarchuk, O. G. Varfolomeeva, N.V. Korshun, O.I.

Chumak, O.A. Kilmeninov // Scientific Notes of the Ukrainian Science and Research Institute of Communications. 2018. #1. pp. 12–22.
[14] R.M. Didkovsky, Comparative analysis of potential noise immunity of the main methods of noise signal modulation / R.M. Didkovsky // Scientific Notes of the Ukrainian Science and Research Institute of Communications. 2012. #2 (22). pp. 86–93.
[15] V.K. Dzyadyk, Introduction to the theory of uniform approximation of functions by polynomials. / V.K. Dzyadyk. Science, 1977. P. 512.
[16] I.F. Boyko, Transmission and receiving of digital signals in spline bases / I.F. Boyko, M.G. Gordeev, A.I. Kutin // Electronics and control systems: collection of scientific works. 2012. #3 (33). pp. 5–12
[17] Negodenko O. Mathematical models on the basis of fundamental trigonometric splines / V. P. Denysiuk, O. Negodenko // Science and Education a New Dimension. Natural and Technical Sciences. 2018. VI (18). Issue: 158. P.14–19