DISCRETE AND CONTINUOUS MODELS AND APPLIED COMPUTATIONAL SCIENCE

Aim and Scope

Discrete and Continuous Models and Applied Computational Science arose
in 2019 as a continuation of RUDN Journal of Mathematics, Information
Sciences and Physics. RUDN Journal of Mathematics, Information Sciences
and Physics arose in 2006 as a merger and continuation of the series "Physics",
"Mathematics", "Applied Mathematics and Computer Science", "Applied Mathematics and Computer Mathematics".

Discussed issues affecting modern problems of physics, mathematics, queuing theory, the Teletraffic theory, computer science, software and databases
development.

It's an international journal regarding both the editorial board and contributing authors as well as research and topics of publications. Its authors
are leading researchers possessing PhD and PhDr degrees, and PhD and MA
students from Russia and abroad. Articles are indexed in the Russian and
foreign databases. Each paper is reviewed by at least two reviewers, the
composition of which includes PhDs, are well known in their circles. Author's
part of the magazine includes both young scientists, graduate students and
talented students, who publish their works, and famous giants of world science.

The Journal is published in accordance with the policies of COPE (Committee on Publication Ethics). The editors are open to thematic issue initiatives
with guest editors. Further information regarding notes for contributors, subscription, and back volumes is available at `http://journals.rudn.ru/miph`.

E-mail: `miphj@rudn.ru`, `dcm@sci.pfu.edu.ru`.

EDITORIAL BOARD

Contents

On the application of the Fourier method to solve the problem of correction of thermographic images

Obaida Baaj

Peoples' Friendship University of Russia (RUDN University),
6, Miklukho-Maklaya St., Moscow, 117198, Russian Federation

Abstract. The work is devoted to the construction of computational algorithms implementing the method of correction of thermographic images. The correction is carried out on the basis of solving some ill-posed mixed problem for the Laplace equation in a cylindrical region of rectangular cross-section. This problem corresponds to the problem of the analytical continuation of the stationary temperature distribution as a harmonic function from the surface of the object under study towards the heat sources. The cylindrical region is bounded by an arbitrary surface and plane. On an arbitrary surface, a temperature distribution is measured (and thus is known). It is called a thermogram and reproduces an image of the internal heat-generating structure. On this surface, which is the boundary of the object under study, convective heat exchange with the external environment of a given temperature takes place, which is described by Newton's law. This is the third boundary condition, which together with the first boundary condition corresponds to the Cauchy conditions — the boundary values of the desired function and its normal derivative. The problem is ill-posed. In this paper, using the Tikhonov regularization method, an approximate solution of the problem was obtained, stable with respect to the error in the Cauchy data, and which can be used to build effective computational algorithms. The paper considers algorithms that can significantly reduce the amount of calculations.

Key words and phrases: thermogram, ill-posed problem, Cauchy problem for the Laplace equation, integral equation of the first kind, Tikhonov regularization method

1. Introduction

Improving the quality and information content of images obtained by thermal imaging methods using a thermal imager that registers thermal electromagnetic radiation from the surface of the object under study in the infrared range by their mathematical (digital) processing is an urgent problem. In particular, in medicine, thermal imaging has become an effective diagnostic tool [1–4]. The image on the thermogram, which is a visualization of the temperature distribution on the surface of the patient's body, makes it possible

to assess functional anomalies in the state of his internal organs. At the same time, the image on the thermogram in some cases turns out to be somewhat distorted due to the processes of thermal conductivity and heat exchange. The paper proposes a method of image correction on a thermogram within a certain mathematical model. As an adjusted thermogram, the image of the temperature field on the plane near the density of heat sources is considered as more accurately transmitting the image of heat sources. It is proposed to obtain this field as a result of the continuation (similar to the continuation of gravitational fields in geophysics problems [5]) of the temperature distribution from the surface from which the initial thermogram is taken. The problem under consideration is ill-posed, since small errors in the initial data (the initial thermogram) may correspond to significant errors in solving the inverse problem. To construct its stable approximate solution, the Tikhonov regularization method [6] is used.

2.   Mathematical model and problem statement

Let's consider a physical and mathematical model, in which we set the task of continuing from the boundary of the stationary temperature distribution.

The physical model is a homogeneous heat-conducting body in the form of a rectangular cylinder, bounded by the surface S and containing heat sources with a time-independent density function that create a stationary temperature distribution in the body. We associate the density function of heat sources with the object under study. We assume that a given temperature distribution (equal to zero) is maintained on the lateral faces of the cylinder, and on the surface S there is convective heat exchange with the external environment of temperature U_0, described by Newton's law, according to which the density of the heat flux at the point of the surface S is directly proportional to the temperature difference inside and outside.

Let's move on to the mathematical model. In a rectangular cylinder

$$D^\infty = \{(x, y, z) : 0 < x < l_x, 0 < y < l_y, -\infty < z < \infty\} \subset \mathbb{R}^3 \qquad (1)$$

consider a cylindrical region

$$D(F, \infty) = \{(x, y, z) : 0 < x < l_x, 0 < y < l_y, F(x, y) < z < \infty\}, \qquad (2)$$

limited by the surface

$$S = \{(x, y, z) : 0 < x < l_x, 0 < y < l_y, z = F(x, y) < H\}. \qquad (3)$$

We'll assume that we also know that

$$a_1 < F(x, y) < a_2 < H, \quad (x, y) \in \Pi, \qquad (4)$$

$$\Pi = \{(x, y) : 0 < x < l_x, 0 < y < l_y\}. \qquad (5)$$

Let Γ be the set of side faces of the domain $D(F, \infty)$. In the domain $D(F, \infty)$ we consider the following mixed boundary value problem for the Laplace equation

$$\begin{cases} \Delta u(M) = \rho(M), \quad M \in D(F, \infty), \\ \dfrac{\partial u}{\partial n}\Big|_S = h(U_0 - u)\Big|_S, \\ u|_\Gamma = 0, \\ u \ \text{limited at} \ \ z \to \infty. \end{cases} \quad (6)$$

The problem (6) corresponds to the steady-state temperature distribution created by heat sources with the distribution density function ρ, on the surface S — a third boundary condition is set corresponding to convective heat exchange with a medium of temperature U_0 with a coefficient h, zero temperature is set at the boundary Γ.

We assume that the function ρ is such that the solution of the problem (6) exists in $C^2(D(F, \infty)) \bigcap C^1(\overline{D(F, \infty)})$. In particular, the solution of the problem (6) allows us to find $u|_S$, i.e. the temperature distribution of u on the surface S, which we will call a thermogram.

Now let the thermogram be obtained as a result of measurements. Let us now set the inverse problem. We set the problem of continuation of the temperature distribution from the surface towards the sources in order to obtain an adjusted thermogram as the temperature distribution $u|_{z=H}$ on the plane $z = H$, closer to the density carrier than the surface S.

We assume that the carrier of the function ρ is located in the domain $z > H$, then the solution of the problem (6) in the domain

$$D(F, H) = \{(x, y, z) : 0 < x < l_x, 0 < y < l_y, F(x, y) < z < H\} \quad (7)$$

satisfies the Laplace equation. The set of side faces of the domain $D(F, H)$ is denoted by Γ_H.

Inverse problem. Let the function be given within the framework of the model (6)

$$f = u\Big|_S. \quad (8)$$

It is required to find $u\Big|_{z=H}$. Since the value of H sufficiently arbitrarily defines the plane between the support of ρ and the surface S, then in fact the inverse problem consists in obtaining a solution u in the domain $D(F, H)$ of the boundary value problem

$$\begin{cases} \Delta u(M) = 0, \quad M \in D(F, H), \\ u\Big|_S = f, \\ \dfrac{\partial u}{\partial n}\Big|_S = h(U_0 - f)\Big|_S, \\ u\Big|_{\Gamma_H} = 0. \end{cases} \quad (9)$$

We assume that the function f in (8), (9) is taken from the set of solutions to the direct problem (6), so the solution to the inverse problem exists in $C^2(D(F, H)) \bigcap C^1(\overline{D(F, H)})$.

Note that in the problem (9) on the surface S of the form (3), Cauchy conditions are set, that is, the boundary values f of the desired function u and the values of its normal derivative are set, so the problem (9) has a unique solution. The boundary $z = H$ of the domain $D(F, H)$ is free and, thus, the problem (9) is unstable with respect to errors in the data, i.e. ill-posed.

The function $u|_{z=H}$ will be considered as an adjusted thermogram. Since the plane $z = H$ is located closer to the support of density ρ, it should be expected that the corrected thermogram more accurately conveys information about the distribution of heat sources than the original thermogram.

Further we give an explicit representation of the exact solution of the problem (9).

3.   Exact solution of the inverse problem

Based on the [7] scheme, an exact solution of the problem (9) is constructed in [8].

Let $\varphi(M, P)$ be the source function of the Dirichlet problem in the cylinder D^∞:

$$
\begin{aligned}
\Delta u(P) &= -\rho(P), \quad P \in D^\infty, \\
u\big|_{x=0,\, l_x} &= 0, \quad u\big|_{y=0,\, l_y} = 0, \\
u &\to 0 \quad \text{at} \quad |z| \to \infty.
\end{aligned}
\tag{10}
$$

In the domain $z_M < H$ in the cylinder (1), we introduce the notation

$$
\Phi(M) = \int\limits_S \left[h(U_0 - f(P))\varphi(M, P) - f(P)\frac{\partial \varphi}{\partial n_P}(M, P) \right] d\sigma_P.
\tag{11}
$$

In [8], the following representation of the solution of the problem is obtained (9)

$$
u(M) = v(M) + \Phi(M), \quad M \in D(F, H),
\tag{12}
$$

where the function Φ is calculated on the known functions f and f_1, and the function v has the form:

$$
\begin{aligned}
v(M) = -\sum_{n,m=1}^{\infty} \tilde{\Phi}_{nm}(a) \exp\{k_{nm}(z-a)\} \sin\frac{\pi n x}{l_x} \sin\frac{\pi m y}{l_y}, \\
M(x, y, z) \in D(-\infty, H),
\end{aligned}
\tag{13}
$$

where

$$
k_{nm} = \pi \left(\frac{n^2}{l_x^2} + \frac{m^2}{l_y^2} \right)^{1/2}
\tag{14}
$$

and $\tilde{\Phi}_{nm}(a)$ — Fourier coefficients of the function $\Phi(M)$

$$
\tilde{\Phi}_{nm}(a) = \frac{4}{l_x l_y} \int\limits_{\Pi(a)} \Phi(x, y, a) \sin\frac{\pi n x}{l_x} \sin\frac{\pi m y}{l_y} dx dy
\tag{15}
$$

on the auxiliary plane:

$$\Pi(a) = \{(x, y, z) : 0 < x < l_x, \, 0 < y < l_y, \, z = a\}, \quad a < a_1. \quad (16)$$

For a Φ function of the form (11) considering that $d\sigma_P = n_1(x_P, y_P)dx_P dy_P$, where the normal $\mathbf{n}_1$ to the surface S is calculated by the formula

$$\mathbf{n}_1 = \text{grad}\,(F(x, y) - z) = \nabla_{xy}F - \mathbf{k}, \quad n_1 = |\mathbf{n}_1|, \quad (17)$$

we will use the representation

$$\Phi(M) = \int\limits_{\Pi} \Big[h(U_0 - f(x_P, y_P))\varphi(M, P)\big|_{P \in S} n_1(x_P, y_P) - $$

$$- f(x_P, y_P)(\mathbf{n}_1, \nabla_P\varphi(M, P))\big|_{P \in S} \Big] dx_P dy_P. \quad (18)$$

When calculating the function $\Phi(M)\big|_{M \in \Pi(a)}$ on the rectangle $\Pi(a)$ for the source function $\varphi(M, P)$, you can use the formula

$$\varphi(M, P) = $$

$$= \frac{2}{l_x l_y} \sum_{n,m=1}^{\infty} \frac{e^{-k_{nm}|z_M - z_P|}}{k_{nm}} \sin\frac{\pi n x_M}{l_x} \sin\frac{\pi m y_M}{l_y} \sin\frac{\pi n x_P}{l_x} \sin\frac{\pi m y_P}{l_y}, \quad (19)$$

which for $z_M = a$ and $P \in S$ takes the form

$$\varphi(M, P) = \frac{2}{l_x l_y} \sum_{n,m=1}^{\infty} \frac{e^{-k_{nm}(F(x_P, y_P) - a)}}{k_{nm}} \times$$

$$\times \sin\frac{\pi n x_M}{l_x} \sin\frac{\pi m y_M}{l_y} \sin\frac{\pi n x_P}{l_x} \sin\frac{\pi m y_P}{l_y}. \quad (20)$$

The series converges uniformly, since the exponent is estimated by $\exp\{-k_{nm}(a_1 - a)\}$. When calculating the function Φ in (12), the source function at $a_2 < z_M < H$ and $P \in S$ takes the form

$$\varphi(M, P) = \frac{2}{l_x l_y} \sum_{n,m=1}^{\infty} \frac{e^{-k_{nm}(z_M - F(x_P, y_P))}}{k_{nm}} \times$$

$$\times \sin\frac{\pi n x_M}{l_x} \sin\frac{\pi m y_M}{l_y} \sin\frac{\pi n x_P}{l_x} \sin\frac{\pi m y_P}{l_y}. \quad (21)$$

The series converges uniformly on any fixed plane $z_M = \text{const}$, since the exponent is estimated by $\exp\{-k_{nm}(z_M - a_2)\}$, that is important for applications. At the points $z_M < a_2$, the source function can be calculated by the reflection method.

4.   Construction of an approximate solution to the problem

Let the function f in the problem (9) be given with an error, that is, instead of f, the function f^δ is given, so that

$$\|f^\delta - f\|_{L_2(\Pi)} \leqslant \delta. \tag{22}$$

In this case, the function (11) is calculated approximately

$$\Phi^\delta(M) = \int\limits_\Pi \Big[h(U_0 - f^\delta(x_P, y_P))\varphi(M, P)\big|_{P \in S} n_1(x_P, y_P) -$$
$$- f^\delta(x_P, y_P)(\mathbf{n}_1, \nabla_P \varphi(M, P))\big|_{P \in S} \Big] dx_P dy_P. \tag{23}$$

The approximate solution to the problem (9) is constructed using the Tikhonov regularization method [6] and in accordance with (12) has the form

$$u_\alpha^\delta(M) = v_\alpha^\delta(M) + \Phi^\delta(M), \quad M \in D(F, H), \tag{24}$$

where Φ^δ is a function of the form (23) and

$$v_\alpha^\delta(M) = - \sum_{n,m=1}^\infty \frac{\tilde{\Phi}_{nm}^\delta(a) \exp\{k_{nm}(z_M - a)\}}{1 + \alpha \exp\{2k_{nm}(H - a)\}} \sin \frac{\pi n x_M}{l_x} \sin \frac{\pi m y_M}{l_y}. \tag{25}$$

Note that the members of the series (25) differs from the members of the series (13) by the regularizing factor $(1 + \alpha \exp\{2k_{nm}(H - a)\})^{-1}$, ensuring the convergence of the series.

In the numerical solution, the bulk of the calculations is related to the calculation of the Fourier coefficients of the function Φ^δ by the formula (15). The next section is devoted to the calculation of Fourier coefficients with a significant reduction in the amount of calculations.

5.   Calculation of Fourier coefficients

As follows from the formulas (15), (23), (20), when calculating the Fourier coefficient for each pair of indices n and m, a superposition of the following calculations is required: summation of the series for φ, integration on the surface S, integration on the rectangle $\Pi(a)$. Thus, when discretizing [9] the problem (N_x points on the variable x, N_y points on the variable y) when calculating Fourier coefficients, about $O(N_x N_y)^4$ operations are required. This is the largest volume of operations when constructing a solution to the problem (9), during which, in addition to time, there is a loss of accuracy and an additional error is formed in calculating the Fourier coefficients and solving the problem as a whole.

It seems advisable to carry out some of these operations analytically, reducing the subsequent amount of calculations, namely. Let us carry out the

integration in the formula for calculating the Fourier coefficients (15) under the sign of the integral in (23) and under the sign of the sum in (20), and use the orthogonality of the complete system of functions

$$\left\{ \sin \frac{\pi n x}{l_x} \sin \frac{\pi m y}{l_y} \right\}_{n,m=1}^{\infty}. \tag{26}$$

Calculate the Fourier coefficient from the first term in (23)

$$\tilde{\Phi}_{1,nm}(a) = \frac{4}{l_x l_y} \int_{\Pi(a)} \Phi_1(x,y,a) \sin \frac{\pi n x}{l_x} \sin \frac{\pi m y}{l_y} dx dy =$$

$$= \frac{4}{l_x l_y} \int_{\Pi(a)} \sin \frac{\pi n x}{l_x} \sin \frac{\pi m y}{l_y} dx dy \times$$

$$\times \int_{\Pi} \left[h(U_0 - f^\delta(x_P, y_P)) \varphi(M,P) \big|_{P \in S} n_1(x_P, y_P) \right] dx_P dy_P. \tag{27}$$

By integrating on the rectangle $\Pi(a)$ under the sign of the integral on the rectangle Π, using the representation (20), we calculate the value

$$\frac{4}{l_x l_y} \int_{\Pi(a)} \sin \frac{\pi n x}{l_x} \sin \frac{\pi m y}{l_y} dx dy \varphi(M,P) \big|_{P \in S} =$$

$$= \frac{4}{l_x l_y} \int_{\Pi(a)} \sin \frac{\pi n x}{l_x} \sin \frac{\pi m y}{l_y} dx dy \frac{2}{l_x l_y} \sum_{n',m'=1}^{\infty} \frac{e^{-k_{n'm'}(F(x_P,y_P)-a)}}{k_{n'm'}} \times$$

$$\times \sin \frac{\pi n' x}{l_x} \sin \frac{\pi m' y}{l_y} \sin \frac{\pi n' x_P}{l_x} \sin \frac{\pi m' y_P}{l_y}. \tag{28}$$

By performing integration under the sign of the sum of uniformly convergent series and using the orthogonality of the system (26), we obtain

$$\frac{4}{l_x l_y} \int_{\Pi(a)} \sin \frac{\pi n x}{l_x} \sin \frac{\pi m y}{l_y} dx dy \varphi(M,P) \big|_{P \in S} =$$

$$= \frac{4}{l_x l_y} \frac{2}{l_x l_y} \sum_{n',m'=1}^{\infty} \frac{e^{-k_{n'm'}(F(x_P,y_P)-a)}}{k_{n'm'}} \frac{l_x l_y}{4} \delta_{nn'} \delta_{mm'} \sin \frac{\pi n' x_P}{l_x} \sin \frac{\pi m' y_P}{l_y} =$$

$$= \frac{2}{l_x l_y} \frac{e^{-k_{nm}(F(x_P,y_P)-a)}}{k_{nm}} \sin \frac{\pi n x_P}{l_x} \sin \frac{\pi m y_P}{l_y}. \tag{29}$$

Using (29), for the Fourier coefficients (27), replacing integration variables x_P and y_P with x and y, we get

$$\tilde{\Phi}_{1,nm}(a) = \frac{2}{l_x l_y k_{nm}} \times$$

$$\times \int\limits_{\Pi} \left[h(U_0 - f^\delta(x,y)) e^{-k_{nm}(F(x,y)-a)} n_1(x,y) \sin\frac{\pi n x}{l_x} \sin\frac{\pi m y}{l_y} \right] dx dy. \quad (30)$$

From the formula (30) it follows that to calculate the Fourier coefficient of the function Φ on the rectangle $\Pi(a)$ there is no need to calculate the function itself. You can use the formula (30), which formally coincides with the formula (15) for the Fourier coefficients on the system (26) of some function depending on the Fourier indices and including information about the surface S in the form of a function F and the normal n_1 calculated by the formula

$$n_1(x,y) = \sqrt{(F'_x(x,y))^2 + (F'_y(x,y))^2 + 1}.$$

In this case, the number of operations has the order of $O(N_x N_y)^2$, that is, the second order in terms of the number of points, which is two orders of magnitude less than the direct calculation of the Fourier coefficients by the formulas (15), (23), (20).

Similarly, the Fourier coefficient of the second term is calculated in the formula (23)

$$\tilde{\Phi}_{2,nm}(a) = \frac{4}{l_x l_y} \int\limits_{\Pi(a)} \Phi_2(x,y,a) \sin\frac{\pi n x}{l_x} \sin\frac{\pi m y}{l_y} dx dy =$$

$$= \frac{4}{l_x l_y} \int\limits_{\Pi(a)} \sin\frac{\pi n x}{l_x} \sin\frac{\pi m y}{l_y} dx dy \times$$

$$\times \int\limits_{\Pi} \left[f^\delta(x_P, y_P)(\mathbf{n}_1, \nabla_P \varphi(M,P))\big|_{P \in S} \right] dx_P dy_P. \quad (31)$$

Using the representation (20), we calculate the value

$$(\mathbf{n}_1, \nabla_P \varphi(M,P))\big|_{P \in S, M \in \Pi(a)} =$$

$$= \frac{2}{l_x l_y} \sum_{n',m'=1}^{\infty} \frac{e^{-k_{n'm'}(F(x_P,y_P)-a)}}{k_{n'm'}} \sin\frac{\pi n' x}{l_x} \times$$

$$\times \sin\frac{\pi m' y}{l_y} \cos\frac{\pi n' x_P}{l_x} \sin\frac{\pi m' y_P}{l_y} \frac{\pi n'}{l_x} F'_x(x_P, y_P) +$$

$$+ \frac{2}{l_x l_y} \sum_{n',m'=1}^{\infty} \frac{e^{-k_{n'm'}(F(x_P,y_P)-a)}}{k_{n'm'}} \sin\frac{\pi n' x}{l_x} \times$$

$$\times \sin\frac{\pi m' y}{l_y} \sin\frac{\pi n' x_P}{l_x} \cos\frac{\pi m' y_P}{l_y} \frac{\pi m'}{l_y} F'_y(x_P, y_P) +$$

$$+ \frac{2}{l_x l_y} \sum_{n',m'=1}^{\infty} e^{-k_{n'm'}(F(x_P,y_P)-a)} \sin \frac{\pi n' x}{l_x} \times$$

$$\times \sin \frac{\pi m' y}{l_y} \sin \frac{\pi n' x_P}{l_x} \sin \frac{\pi m' y_P}{l_y}. \quad (32)$$

By integrating on the rectangle $\Pi(a)$ under the sign of the integral on the rectangle Π, performing integration under the sign of the sum of a uniformly convergent series and using the orthogonality of the system (26), we obtain

$$\frac{4}{l_x l_y} \int_{\Pi(a)} \sin \frac{\pi n x}{l_x} \sin \frac{\pi m y}{l_y} dx dy \left(\mathbf{n}_1, \nabla_P \varphi(M,P)\right)\big|_{P \in S} =$$

$$= \frac{2}{l_x l_y} \frac{e^{-k_{nm}(F(x_P,y_P)-a)}}{k_{nm}} \left[\cos \frac{\pi n x_P}{l_x} \sin \frac{\pi m y_P}{l_y} \frac{\pi n}{l_x} F'_x(x_P, y_P) + \right.$$

$$\left. + \sin \frac{\pi n x_P}{l_x} \cos \frac{\pi m y_P}{l_y} \frac{\pi m}{l_y} F'_y(x_P, y_P) + k_{nm} \sin \frac{\pi n x_P}{l_x} \sin \frac{\pi m y_P}{l_y}\right]. \quad (33)$$

Hence and from (31) follows

$$\tilde{\Phi}_{2,nm}(a) = \frac{2\pi n}{l_x^2 l_y k_{nm}} \int_{\Pi} f^\delta(x,y) e^{-k_{nm}(F(x,y)-a)} \times$$

$$\times F'_x(x,y) \cos \frac{\pi n x}{l_x} \sin \frac{\pi m y}{l_y} dx dy +$$

$$+ \frac{2\pi m}{l_x l_y^2 k_{nm}} \int_{\Pi} f^\delta(x,y) e^{-k_{nm}(F(x,y)-a)} F'_y(x,y) \sin \frac{\pi n x}{l_x} \cos \frac{\pi m y}{l_y} dx dy +$$

$$+ \frac{2}{l_x l_y} \int_{\Pi} f^\delta(x,y) e^{-k_{nm}(F(x,y)-a)} \sin \frac{\pi n x}{l_x} \sin \frac{\pi m y}{l_y} dx dy. \quad (34)$$

Thus, the Fourier coefficient $\tilde{\Phi}_{2,nm}(a)$ is calculated as the sum of formally calculated Fourier coefficients over orthogonal systems

$$\left\{\sin \frac{\pi n x}{l_x} \sin \frac{\pi m y}{l_y}\right\}_{n,m=1}^{\infty}, \quad \left\{\cos \frac{\pi n x}{l_x} \sin \frac{\pi m y}{l_y}\right\}_{n,m=1}^{\infty},$$

$$\left\{\sin \frac{\pi n x}{l_x} \cos \frac{\pi m y}{l_y}\right\}_{n,m=1}^{\infty}. \quad (35)$$

of functions depending, among other things, on the indices of the Fourier coefficients. In this case, as well as when calculating the Fourier coefficient from the first term, the number of operations has the order of $O(N_x N_y)^2$, that is, the second order in terms of the number of points, which is two orders

of magnitude less than the direct calculation of the Fourier coefficients by the formulas (15), (23), (20).

Summing (30), (34), we get the Fourier coefficient

$$\tilde{\Phi}_{nm}(a) = \tilde{\Phi}_{1,nm}(a) + \tilde{\Phi}_{2,nm}(a).\qquad(36)$$

According to the remarks to the formulas (30), (34) in general, the number of operations when calculating the Fourier coefficients using these formulas relative to the number of $N_x N_y$ points on the thermogram has the order of $O(N_x N_y)^2$.

To calculate the Fourier coefficients using the formulas (30), (34), the Hamming method [10] is used.

6.　Conclusion and discussion

Stable solution of the inverse problem (9) can be used for mathematical processing of thermograms taken with a thermal imager, in particular, in medicine [4], in order to correct the image on the thermogram. Note that taking into account the blood flow leads to the need to use the metaharmonic equation [11, 12] in problem (9). As already mentioned, a thermogram, with one or another reliability, convey an image of the structure of heat sources inside the body. Refinement of the image on the thermogram can be carried out within the framework of the problem (9). In this case, the function f is associated with the original thermogram, and the function u_H is considered the result of processing the thermogram. Since the function u_H represents the temperature distribution on a plane closer to the studied heat sources than the original surface S, we can expect a more accurate reproduction of the image of the sources on the calculated thermogram u_H. The results of calculations carried out on a model example show the effectiveness of the proposed method and algorithm based on the formulas (24), (25), (23), (36), which can be used to process thermographic images.

References

[1]　E. F. J. Ring, "Progress in the measurement of human body temperature," *IEEE Engineering in Medicine and Biology Magazine*, vol. 17, no. 4, pp. 19–24, 1998. DOI: 10.1109/51.687959.

[2]　E. Y. K. Ng and N. M. Sudarshan, "Numerical computation as a tool to aid thermographic interpretation," *Journal of Medical Engineering and Technology*, vol. 25, no. 2, pp. 53–60, 2001. DOI: 10.1080/03091900110043621.

[3]　B. F. Jones and P. Plassmann, "Digital infrared thermal imaging of human skin," *IEEE Eng. in Med. Biol. Mag.*, vol. 21, no. 6, pp. 41–48, 2002. DOI: 10.1109/memb.2002.1175137.

[4]　G. R. Ivanitskii, "Thermovision in medicine [Teplovideniye v meditsine]," *Vestnik RAN*, vol. 76, no. 1, pp. 44–53, 2006, in Russian.

[5] A. N. Tikhonov, V. B. Glasko, O. K. Litvinenko, and V. R. Melihov, "On the continuation of the potential towards disturbing masses based on the regularization method [O prodolzhenii potentsiala v storonu vozmushchayushchih mass na osnove metoda regulyarizatsii]," *Izvestiya AN SSSR. Fizika Zemli*, no. 1, pp. 30–48, 1968, in Russian.

[6] A. N. Tikhonov and V. J. Arsenin, *Methods for solving ill-posed problems [Metody resheniya nekorrektnyh zadach]*. Moscow: Nauka, 1979, in Russian.

[7] E. B. Laneev, "Construction of a Carleman function based on the Tikhonov regularization method in an ill-posed problem for the Laplace equation," *Differential Equations*, vol. 54, no. 4, pp. 476–485, 2018. DOI: 10.1134/S0012266118040055.

[8] E. B. Laneev, N. Y. Chernikova, and O. Baaj, "Application of the minimum principle of a Tikhonov smoothing functional in the problem of processing thermographic data," *Advances in Systems Science and Applications*, vol. 1, pp. 139–149, 2021. DOI: 10.25728/assa.2021.21.1.1055.

[9] E. B. Laneev, M. N. Mouratov, and E. P. Zhidkov, "Discretization and its proof for numerical solution of a Cauchy problem for Laplace equation with inaccurately given Cauchy conditions on an inaccurately defined arbitrary surface," *Physics of Particles and Nuclei Letters*, vol. 5, no. 3, pp. 164–167, 2002. DOI: 10.1134/S1547477108030059.

[10] R. W. Hamming, *Numerical methods for scientists and engineers*. New York: McGraw-Hill Book Company, 1962.

[11] H. Pennes, "Analysis of tissue and arterial blood temperature in the resting human forearm," *J. Appl. Physiol.*, no. 1, pp. 93–122, 1948.

[12] J. P. Agnelli, A. A. Barrea, and C. V. Turner, "Tumor location and parameter estimation by thermography," *Mathematical and Computer Modelling*, vol. 53, no. 7–8, pp. 1527–1534, 2011. DOI: 10.1016/j.mcm.2010.04.003.

For citation:

Information about the authors:

Baaj, Obaida — postgraduate student of Nikolskiy Mathematical Institute of Peoples' Friendship University of Russia (RUDN University) (e-mail: 1042175025@rudn.ru, phone: +7(916)6890863, ORCID: https://orcid.org/0000-0003-4813-7981)

О применении метода Фурье для решения задачи коррекции термографических изображений

Обаида Бааж

*Российский университет дружбы народов,
ул. Миклухо-Маклая, д. 6, Москва, 117198, Россия*

Аннотация. Работа посвящена построению вычислительных алгоритмов, реализующих метод коррекции термографических изображений. Коррекция осуществляется на основе решения некоторой некорректно поставленной смешанной задачи для уравнения Лапласа в цилиндрической области прямоугольного сечения. Эта задача соответствует задаче аналитического продолжения стационарного распределения температуры как гармонической функции с поверхности исследуемого объекта в сторону источников тепла. Цилиндрическая область ограничена произвольной поверхностью и плоскостью. На произвольной поверхности измеряется (и таким образом, задано) распределение температуры, называемое термограммой и воспроизводящее изображение внутренней тепловыделяющей структуры. На этой поверхности — границе исследуемого объекта — имеет место конвективный теплообмен с внешней средой заданной температуры, который описывается законом Ньютона. Это третье краевое условие, которое в совокупности с первым краевым условием соответствует заданию условий Коши — граничным значениям искомой функции и ее нормальной производной. Задача некорректно поставлена. В статье применением метода регуляризации Тихонова получено приближённое решение поставленной задачи, устойчивое по отношению к погрешности к данным Коши, и которое может быть использовано для построения эффективных вычислительных алгоритмов. В работе рассматриваются алгоритмы, позволяющие существенно уменьшить объем вычислений.

Ключевые слова: термограмма, некорректная задача, задача Коши для уравнения Лапласа, интегральное уравнение первого рода, метод регуляризации Тихонова

Numerical simulation of cold emission in coaxial diode with magnetic isolation

Alexandr A. Belov[1,2]**, Oleg T. Loza**[2]**, Konstantin P. Lovetskiy**[2]**,
Sergey P. Karnilovich**[2]**, Leonid A. Sevastianov**[2]

[1] *Lomonosov Moscow State University,*
1, bld. 2, Leninskie Gory, Moscow, 119991, Russian Federation
[2] *Peoples' Friendship University of Russia (RUDN University),*
6, Miklukho-Maklaya St., Moscow, 117198, Russian Federation

(received: July 5, 2022; revised: July 18, 2022; accepted: August 8, 2022)

Abstract. Due to the emergence and active development of new areas of application of powerful and super-powerful microwave vacuum devices, interest in studying the behavior of ensembles of charged particles moving in the interaction space has increased. An example is an electron beam formed in a coaxial diode with magnetic isolation. Numerical simulation of emission in such a diode is traditionally carried out using particle-in-cell methods. They are based on the simultaneous calculation of the equations of motion of particles and the Maxwell's equations for the electromagnetic field. In the present work, a new computational approach called the point macroparticle method is proposed. In it, the motion of particles is described by the equations of relativistic mechanics, and explicit expressions are written out for fields in a quasi-static approximation. Calculations of the formation of a relativistic electron beam in a coaxial diode with magnetic isolation are performed and a comparison is made with the known theoretical relations for the electron velocity in the beam and for the beam current. Excellent agreement of calculation results with theoretical formulas is obtained.

Key words and phrases: coaxial diode with magnetic isolation, cold emission, point macroparticles

1. Introduction

Relativistic electron beams. The existing plasma relativistic microwave generators and amplifiers (plasma masers) are based on the interaction of tubular plasma with tubular high-current relativistic electron beam (REB) [1]. The explosive-emission cathode [2] forms a tubular REB with an internal radius of ∼2 cm and a thickness of ∼0.15 cm, which propagates in a magnetic field of 1 T created by a solenoid. The electron energy in such a beam is ∼10^6 eV, the electron current density is 10^3–10^4 A/cm^2. The power of

the REB, as a rule, exceeds 10^9 W, the current pulse lasts from several nanoseconds to several microseconds [3].

High-current relativistic electron beams are formed directly in the diode, which is supplied with a voltage pulse from the primary energy storage. Electrons receive energy only in the diode, no additional means of particle acceleration (similar to sections of linear inductive or resonant accelerators) are used. Installations for generating high-current REB are also, in some works, called direct-acting accelerators [4].

The creation of controlled beams (streams) of charged particles is carried out using a variety of devices, the main element of which is a source of charged particles. A fairly common element of such a system that provides an intense, well-focused electron beam is an electron gun. The most commonly used are thermionic guns, in which the primary element is a vacuum diode [5].

Calculation methods. To calculate the dynamics of electron beams, a gas-dynamic approximation is used (see, for example, [6]). As is known, the system of equations of gas dynamics is valid for thermodynamically equilibrium continuous media. Various types of equilibrium violation are taken into account using additional model assumptions. The success of this approach depends on how well the nonequilibrium model is chosen. Models that have proven themselves well in some applications (for example, nonequilibrium electronic processes of solid-state electronics) may not be applicable in other applications.

A more general approach is the kinetic Vlasov equation with respect to the distribution function [7], supplemented by a system of Maxwell's equations for electromagnetic fields. This model leads to a partial differential equation of the first order; it is a mathematical formulation of the well-known Liouville theorem on the conservation of phase volume [8]. The properties of the medium, such as particle concentration, charge density, average velocity, etc. are moments of the distribution function.

For the numerical solution of the kinetic equation, methods such as Particle-in-Cell (PiC) and Cloud-in-Cell (CiC) [9] are used. In these methods, the medium is replaced by a set of a finite number of particles possessing macrocharge that interact with each other. Each particle is attributed to the characteristics of the medium: charge, mass, momentum, energy, etc. The average values of these quantities are calculated as the sum of all model particles located in the considered region.

Macroparticles have a finite size, within which the spatial distribution of charge, mass, etc. is set. Most often, this distribution is chosen piecewise constant. In this case, the geometric dimensions of all particles are considered the same. In some works, more complex form-factors of the particles are considered.

The motion of macroparticles obeys the equations of Newtonian mechanics (or relativistic Lorentz equations). This leads to a system of ordinary differential equations (ODEs) for the coordinates and velocities of particles and a system of the Maxwell's equations for electromagnetic fields. For this system, the «leap-frog» scheme is traditionally used. First, electromagnetic fields are set and the change in the coordinates and velocities of the macroparticles is calculated in one time step. Then, according to the changed coordinates

and velocities, the electromagnetic fields are refined. After that, coordinates and velocities are calculated at the next time step, etc.

Based on this approach, Tarakanov developed the KARAT [10] code, which was widely applied to solving various problems of plasma physics. Among them are formation of a virtual cathode, formation of an electron beam in a coaxial diode with magnetic isolation, dynamics of a laser target and the initiation of deuterium-deuterium reactions, focusing of an electron beam and the development of hose instability, anisotropic Waibel instability and many others. We also note the works of Borodachev (see [11] and other works of this author). He proposed several improvements to this approach and performed calculations of a large number of tasks.

The main difficulty of the particle (cloud) method in a cell is the need to introduce space-time discretization separately for particles and separately for electromagnetic fields. This leads to a number of numerical artifacts. Among them are the stroboscopic effect (the onslaught of the phase of the electromagnetic field when its frequency does not match with the sampling frequency in time), non-conservativeness (either the momentum conservation law or the energy conservation law is fulfilled, but not simultaneously), the grid dispersion of the medium, the parasitic increase of shot and grid noise, and some others. This limits the accuracy of this method.

In the present work, a new method for calculating the emission problem of a coaxial diode with magnetic isolation is proposed. Instead of particles of finite size, it uses point macroparticles. Their motion is described by relativistic Lorentz equations. The electromagnetic field of the beam is calculated in a static approximation based on the instantaneous position of the particles: the electrostatic field is taken according to the Coulomb law and the magnetic field is according to the Biot–Savart–Laplace law. Edge effects at the cathode boundary are considered insignificant. Test emission calculations are performed and the beam velocity and current are compared with the well-known Fedosov's law. This comparison shows excellent accuracy of the proposed method: the discrepancy between the calculation and the specified theoretical law is no more than 1%. Such accuracy is obviously sufficient for applied calculations.

2.   Problem statement

Consider the problem of infinite electron emission in the model of a coaxial diode with magnetic isolation (CDMI) in a strong magnetic field [12]. A solid cylindrical conductive cathode with a radius of R_C with a negative potential of $-U$ is located in a cylindrical conductor-anode with a radius of R_A with zero potential in a strong longitudinal magnetic field (figure 1). The entire surface of the cathode, lateral cylindrical and end plane, has the property of infinite emission. As a result, a thin tubular electron beam with a radius of $R_B \approx R_C$ should be formed.

The parameters of the problem are: magnetic field $B = 1$ T $= 10^4$ Gs, diode length $L = 30$ cm, cathode radius $R_C = 1$ cm, anode radius $R_A = 2.72$ cm, cathode potential $U = 511$ kV $= 1.70 \cdot 10^3$ CGS.
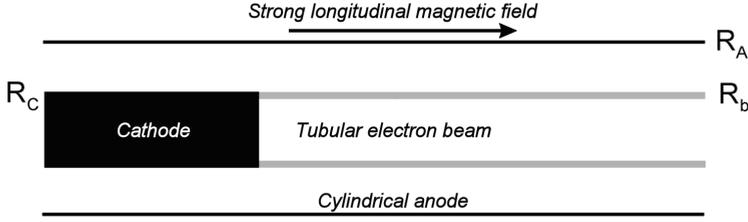
Figure 1. Scheme of the coaxial diode with magnetic isolation

For the current I of a steady beam, the Fedosov's empirical law is valid (see [13] and the cited literature)

$$I\,[\text{kA}] = \frac{17.02}{\ln(R_A/R_C)}\frac{\gamma - \gamma_b}{\gamma_b}\sqrt{\gamma_b^2 - 1},$$

$$\gamma = 1 + \frac{U\,[\text{kV}]}{511}, \quad \gamma_b = \sqrt{1/4 + 2\gamma} - 1/2. \tag{1}$$

A theoretical estimate of the z-components of the electron beam velocity is also known

$$v_z = c\sqrt{1 - 1/\gamma_b^2}. \tag{2}$$

Here c is the speed of light in vacuum. With the specified task parameters, we have $I \approx 2.7$ kA, $v_z \approx 2.3 \cdot 10^{10}$ cm/s.

3.   The method of point macroparticles

3.1.   Point macroparticles

Let the axis z of the coordinate system be directed along the magnetic field and the origin $x = y = z = 0$ correspond to the center of the cathode end face.

On the end surface of the cathode, we select J points $M^j = (x_0^j, y_0^j)$, evenly distributed over this surface. At the initial moment of time, macroparticles containing Z^j electrons fly out of all points. Their initial velocities are supposed to equal zero (i.e., the emission is cold). The emission at a given point of the cathode should be the stronger, the greater the magnitude of the cathode-anode field at this point. Therefore, we choose the value of Z^j proportional to the value of z-components of the cathode-anode field at the point M^j, i.e.,

$$Z^j = JZ^0 \frac{E_z(M^j)}{\sum\limits_j E_z(M^j)}. \tag{3}$$

Here Z^0 is the average charge of one particle. Since in practice $Z_0 \gg 1$, such a particle can be called a macroparticle.

Macroparticles are considered point-size (i.e., the interaction of electrons with each other inside one particle is not taken into account). Macroparticles interact with each other and with external electric and magnetic fields.

Further, at the moment of time $t = \tau$, the second portion of macroparticles flies out of all points M^j. For them, the electrostatic field of the first portion of macroparticles partially shields the cathode-anode field. If by the time $t = \tau$ the particles from the first portion do not have time to gain more speed, then they are located relatively close to the end of the cathode, and the shielding turns out to be stronger. Then, at the corresponding points M^j, the emission of particles of the second portion weakens. This can be taken into account by reducing the value of Z^j for the second portion. Conversely, if by the time of the second emission, the particles of the first portion managed to gain more speed, then the shielding turns out to be weaker. Then, at the corresponding points M^j, the highest charge is emitted, and the value of Z^j for these points must be increased.

The particle movement obeys relativistic equations of motion which are written in terms of momentum, not velocity. Therefore, it is convenient to use the following rule for calculating new Z^j

$$Z^j = JZ^0 \frac{p_z^j}{\sum\limits_{j} p_z^j}. \tag{4}$$

3.2. System of equations

Let us write down the equations of motion of the j-th macroparticle. We choose momentum and radius vector as unknowns

$$\frac{d\mathbf{r}^j}{dt} = \frac{\mathbf{p}^j}{\sqrt{(p^j)^2/c^2 + (m^j)^2}}, \quad m^j = Z^j m, \tag{5}$$

$$\frac{d\mathbf{p}^j}{dt} = e^j \mathbf{E} + \frac{e^j}{c} \frac{[\mathbf{p}^j \times \mathbf{B}]}{\sqrt{(p^j)^2/c^2 + (m^j)^2}}, \quad e^j = Z^j e. \tag{6}$$

Here m, e are the mass and charge of the electron, $\mathbf{E}$, $\mathbf{B}$ are the total electric and magnetic fields in which the j-th particle moves.

Electric field $\mathbf{E} = \mathbf{E}^{\text{ext}} + \mathbf{E}^{el}$ is the sum of the external cathode-anode field $\mathbf{E}^{\text{ext}}$ and the electrostatic field $\mathbf{E}^{el}$ created by other electrons. According to the Coulomb law, we have

$$\mathbf{E}^{el} = \sum_{i \neq j, \, Z^i \neq 0} e^j \frac{\mathbf{r}^j - \mathbf{r}^i}{|\mathbf{r}^j - \mathbf{r}^i|^3}. \tag{7}$$

The sum is taken for all particles that have flown out of the cathode at a given time.

Magnetic field $\mathbf{B} = \mathbf{B}^{\text{ext}} + \mathbf{B}^{el}$ consists of an external solenoid field $\mathbf{B}^{\text{ext}}$ and a field $\mathbf{B}^{el}$, which is created by other electrons when moving. The field $\mathbf{B}^{el}$ is described by the Biot–Savart–Laplace law

$$\mathbf{B}^{el} = \sum_{i \neq j,\, Z^i \neq 0} \frac{e^j}{c} \frac{\left[\mathbf{v}^i, \mathbf{r}^j - \mathbf{r}^i\right]}{|\mathbf{r}^j - \mathbf{r}^i|^3}. \tag{8}$$

3.3.   Numerical method

The equations of motion are solved numerically according to the explicit Runge–Kutta scheme of the 4th order of accuracy. To increase the mathematical accuracy of the difference scheme, the time interval τ can be divided into several steps, during which the number of particles does not change.

The potential of the cathode-anode field is found by solving the boundary value problem for the Laplace equation. It is considered in a cylindrically symmetric formulation. On the walls of the anode, the potential is assumed to be zero, and on the cathode — equal to $-U$. The solution uses the package FreeFem++[14]. Figure 2 shows potential isolines and field strength vectors in coordinates $z - \rho$, where $\rho = \sqrt{x^2 + y^2}$ is the polar radius. For clarity, the range of variation of z is limited to the value of $z = 5$. Note that this cathode-anode field decreases rapidly as z increases. The maximum voltage is reached near the cathode rib.
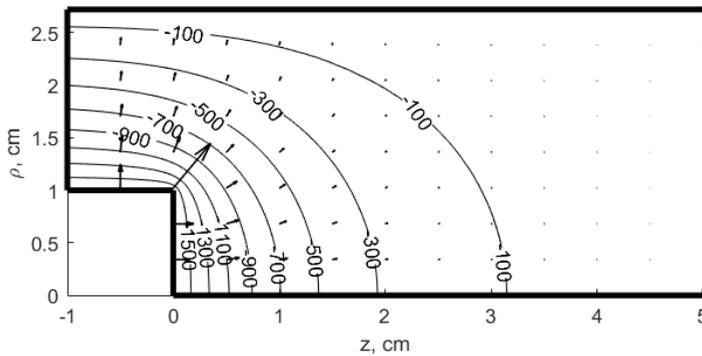
Figure 2. The potential and the cathode-anode field strength in units of CGS.
Solid line is the boundary of the computational domain

3.4.   Remarks

1) The fundamental difference between the proposed approach and the traditional PiC/CiC method is that macroparticles are considered point-like. This makes it possible to write out explicit expressions for the electric and magnetic fields created by particles. Therefore, during the calculation, it is not necessary to solve the system of the Maxwell's equations. This not only gives a gain in efficiency, but also eliminates the previously mentioned numerical artifacts that arise in the traditional particle method. This is an advantage of the point macroparticle method.

2) Quasi-static expressions for the fields (7) and (8) are approximate. They do not take into account the boundary effects at the anode boundaries.

The interaction of charges with walls can be taken into account by the method of electro- and magnetostatic images [15]. A separate issue is the accuracy with which the timebase of the fields (7) and (8) satisfy the Maxwell's equations. Related to the period of the REB formation, this issue requires additional research and is beyond the scope of this work. After the REB is formed and can be considered steady, temporal dependence of the field vanishes (at least, in some vicinity of the cathode). In this case, fulfillment of the Maxwell's equations is provided by employment of the Coulomb and the Biot–Savart–Laplace laws.

4.   Calculation results

Let us take $J = 3000$, $\tau = 10^{-11}$ s. Then $Z^0 = 5.956 \cdot 10^7$. The location of the points M^j at the end of the cathode corresponds to figure 3 (far left).

The calculation was carried out up to the time $t = 10^{-9}$ s, i.e., at the end of the calculation, the number of macroparticles reached $\sim 3 \cdot 10^5$. At the same time, the particles of the first portion reached the section $z \approx 28$, the steady beam corresponds to the segment $0 \leqslant z \leqslant 8$.
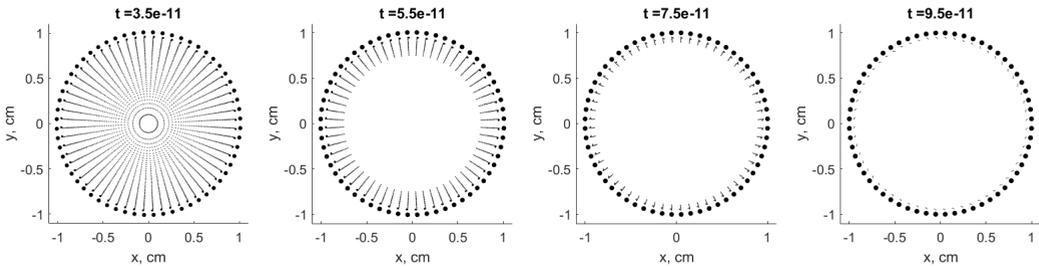
Figure 3. Cross sections of the beam with a plane $z = 0$ at time moments $t = 0.035, 0.055, 0.075, 0.095$ ns (from left to right)

Figure 3 shows the cross sections of the beam with the plane $z = 0$ at several consecutive moments of time. These moments are indicated above the graphs. The points are the intersection of the trajectories of the macroparticles with the specified plane. The size of the markers is proportional to the charge of the macroparticles. Particles with a charge less than $10^{-3} Z^0$ are not displayed.

It can be seen that at the initial moment of time, emission occurs from the entire surface of the cathode end face. Over time, the emission is suppressed first in the center of the cathode, and a tubular beam begins to form. Then the area in which the emission is suppressed expands, and the tube wall becomes thinner. Finally, a thin-walled beam is installed.

In figure 4, the dependence of the average z component of the electron velocity in a steady beam on the z coordinate is presented. It can be seen that with an increase in z, the value of v_z increases rapidly and at $z \approx 3$ cm goes to a constant value of $v_z = 2.30 \cdot 10^{10}$ cm/s. This value perfectly agrees with the theoretical value obtained from the Fedosov's law: the difference is only 0.01%.
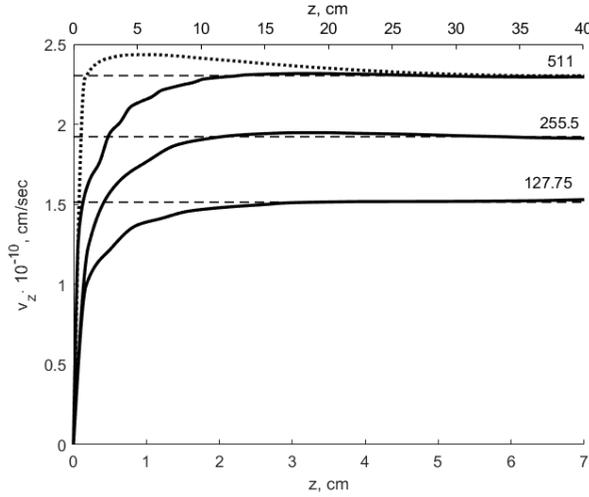
Figure 4. Average component v_z of the electron velocity. Solid line is calculation for $R_A/R_C = 2.72$, dotted line is calculation for $R_A/R_C = 5.44$ (upper axis of abscissa), dashed line is theoretical estimate from the Fedosov's law. The numbers near the lines are the value of the cathode-anode potential U

Figure 5 shows the dependence of the beam current I on the coordinate z. It is clearly seen that with the growth of z, the current value increases rapidly and reaches a constant value of $I = 2.86$ kA, which perfectly agrees with the Fedosov' law. The deviation of the right end of the curve from the horizontal is due to the fact that the beam is not fully established.
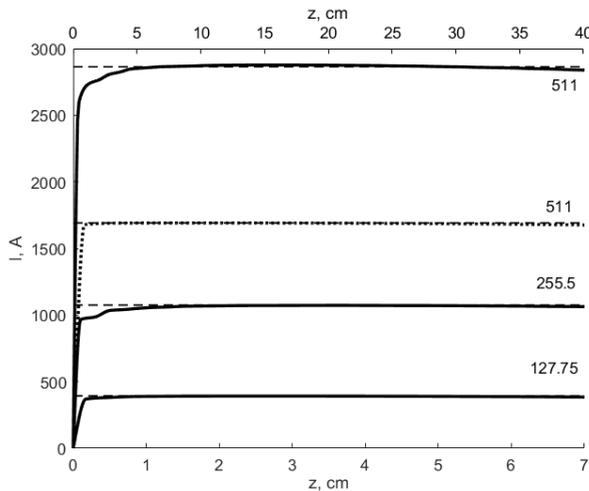
Figure 5. Beam current. Solid line is calculation for $R_A/R_C = 2.72$, dotted line is calculation for $R_A/R_C = 5.44$ (upper axis of abscissa), dashed line is theoretical estimate from the Fedosov' law. The numbers near the lines are the value of the cathode-anode potential U

5.   Parameter selection

In this section, we show that the velocity and current of the steady-state beam practically do not depend on the values of the user-defined parameters J and τ.

5.1.   Number of particles

Let us perform the calculation with $J = 120$, $\tau = 10^{-11}$ s. Then $Z^0 = 1.489 \cdot 10^9$. This value is 25 times different from the one specified in section 4. Figure 6 shows a comparison of the profiles of the average z-velocity components depending on z for calculations with $J = 3000$ and $J = 120$. It can be seen that the qualitative behavior of both curves coincides. The initial sections corresponding to a sharp increase in v_z are somewhat different. But at $z \geqslant 2$, both curves come out to the same constant value that coincides with the theoretical value (2). The calculated velocities of the steady-state beam are consistent with the theoretical estimate with an accuracy better than 0.5%.
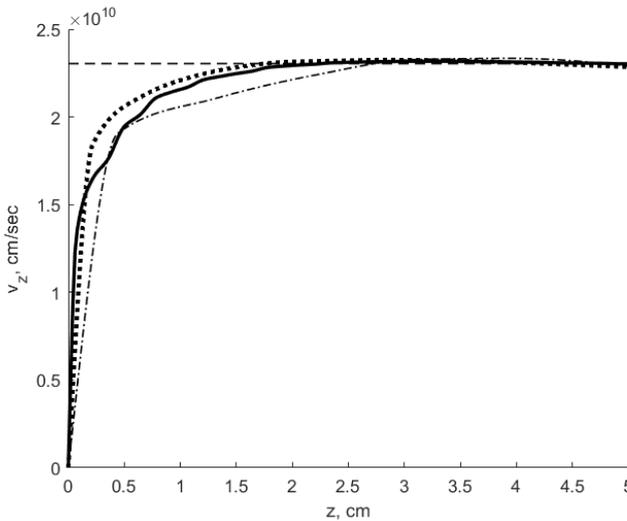
Figure 6. The average z speed component for $U = 511$ kV, $R_A/R_C = 2.72$. Solid line is calculation with $J = 3000$, $\tau = 10^{-11}$ s, dotted line is calculation with $J = 120$, $\tau = 10^{-11}$ s, dashed line is calculation with $J = 120$, $\tau = 2 \cdot 10^{-11}$ s, dashed line is theoretical estimate from the Fedosov's law

In figure 7, a similar comparison of beam current profiles for calculations with $J = 3000$ and $J = 120$ is presented. It can be seen that both curves practically coincide and quickly tend to the theoretical value determined by the Fedosov's law (1). The calculated values of the steady-state beam current are consistent with the theoretical estimate with an accuracy better than 0.3%.
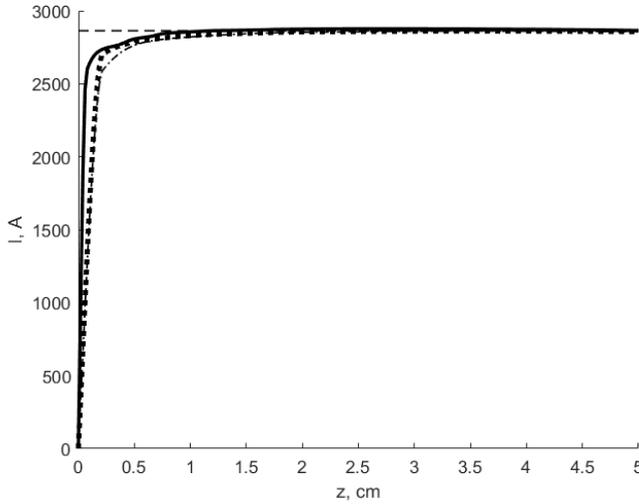
Figure 7. Beam current for $U_0 = 511$ kV, $R_A/R_C = 2.72$. The designations correspond to figure 6

5.2. Time interval between particle portions

Let us perform the calculation with $J = 120$, $\tau = 2 \cdot 10^{-11}$ s. To increase the mathematical accuracy, we divide the interval τ into 2 intermediate steps of the difference scheme. In this calculation, $Z^0 = 2.978 \cdot 10^9$. This value differs by 2 times from the one specified in 5.1. Figure 6 shows a comparison of v_z profiles depending on z for calculations with $\tau = 10^{-11}$ s and $\tau = 2 \cdot 10^{-11}$ s. It can be seen that both curves increase rapidly and tend to constant values that coincide with the theoretical estimate (2). The agreement with the theoretical estimate turns out to be better than 1%.

Figure 7 shows a similar comparison of beam current profiles for calculations with $\tau = 10^{-11}$ s and $\tau = 2 \cdot 10^{-11}$ s. It can be seen that both curves practically coincide and quickly reach the theoretical value of the current determined by the Fedosov's law. The accuracy with which the calculation is consistent with the theoretical estimate turns out to be better than 0.5%.

6. Verification of the model

6.1. Cathode-anode potential

To test the model, we will perform similar calculations for other values of the cathode-anode potential $U = 511/2 = 255.5$ kV and $U = 511/4 = 127.75$ kV. In both calculations, we put $J = 120$, $\tau = 10^{-11}$ s. The dependence of the z-components of the electron velocity on the z coordinate for both calculations is shown in figure 4. It can be seen that with an increase in z, the value of v_z increases from zero to $1.95 \cdot 10^{10}$ cm/s in the first calculation and to $1.52 \cdot 10^{10}$ cm/s in the second calculation. The obtained velocities are in

excellent agreement with the theoretical values: the difference does not exceed 0.3%.

Figure 5 shows the dependence of the beam current on the z coordinate in both calculations. It can be seen that with the growth of z, the current value increases rapidly and reaches a constant value of $I = 1.07$ kA for $U = 255.5$ kV and $I = 0.39$ kA for $U = 127.75$ kV. From figure 5, it can be seen that these values perfectly agree with the Fedosov current: in both calculations, the discrepancy does not exceed 0.5%.

6.2. Ratio of cathode and anode radii

Let us make calculations for a different ratio of the radii of the anode and cathode $R_A/R_C = 2.72 \cdot 2 = 5.44$. We take $J = 120$, $\tau = 10^{-11}$ s. The dependence of the average z-velocity component on the z coordinate is shown in figure 4. Since the beam is established at a significantly greater distance from the cathode than in the calculations of 4, a separate abscissa axis was used. It can be seen that with an increase in z, the value of v_z increases, passes through the maximum, then decreases and tends to a constant value. The latter perfectly agrees with the estimate from the Fedosov's law: the discrepancy does not exceed 0.5%.

Figure 5 shows the dependence of the beam current on the z coordinate (the upper axis of the abscissa). It can be seen that the current quickly tends to a constant value that coincides with the Fedosov current with an accuracy not worse than 0.6%.

7. Conclusion

The proposed algorithm provides self-consistent emission suppression in the center of the cathode and the formation of a tubular electron beam. The dependence of the average z component of the macroparticles velocity and the beam current on z reproduces the known quantitative regularities with good accuracy. This confirms the correctness of the calculation results. It is shown that the calculated particle velocity and current in a steady beam do not depend on the parameters set by the user.

Acknowledgments

References

[1] M. V. Kuzelev *et al.*, "Plasma relativistic microwave electronics," *Plasma Physics Reports*, vol. 27, no. 8, pp. 669–691, 2001. DOI: 10.1134/1.1390539.

[2] S. P. Bugaev, E. A. Litvinov, G. A. Mesyats, and D. I. Proskurovskii, "Explosive emission of electrons," *Physics Uspekhi*, vol. 18, no. 1, pp. 51–61, 1975. DOI: 10.3367/UFNr.0115.197501d.0101.

[3] O. T. Loza and I. E. Ivanov, "Measurements of the transverse electron
     velocities in high-current microsecond relativistic electron beams in
     a strong magnetic field," *Technical Physics*, vol. 48, no. 9, pp. 1180–1185,
     2003. DOI: `10.1134/1.1611905`.

[4] D. K. Ul'yanov *et al.*, "Controlling the radiation frequency of a plasma
     relativistic microwave oscillator during a nanosecond pulse," *Technical Physics*, vol. 58, no. 10, pp. 1503–1506, 2013. DOI: `10.1134/S1063784213100265`.

[5] S. Y. Belomyttsev, A. A. Grishkov, S. D. Korovin, and V. V. Ryzhov,
     "The current of an annular electron beam with virtual cathode in a drift
     tube," *Technical Physics Letters*, vol. 29, no. 7, pp. 666–668, 2003. DOI:
     `10.1134/1.1606783`.

[6] S. V. Polyakov, "Mathematical modeling using multiprocessor com-
     puting systems of electronic transport processes in vacuum and
     solid-state micro- and nanostructures [Matematicheskoye modelirovaniye
     s pomoshch'yu nogoprotsessornykh vychislitel'nykh sistem protsessov
     elektronnogo transporta v vakuumnykh i tverdotel'nykh mikro- i nanos-
     trukturakh]," in Russian, Diss. ... Doctor of Physical and Mathematical
     Sciences, M. V. Keldysh IAM, RAS, 2010.

[7] A. A. Vlasov, "The vibrational properties of an electron gas," *Physics
     Uspekhi*, vol. 10, no. 6, pp. 721–733, 1968. DOI: `10.3367/UFNr.0093.196711f.0444`.

[8] I. A. Kvasnikov, *Thermodynamics and statistical physics. Vol. 3. Theory
     of nonequilibrium systems [Termodinamika i statisticheskaya fizika, Tom 3,
     Teoriya ravnovesnykh sistem, Teoriya neravnovesnykh sistem]*. Moscow:
     URSS, 2003, in Russian.

[9] R. W. Hockney and J. W. Eastwood, *Computer simulation using particles*.
     McGraw-Hill Inc., 1981.

[10] V. P. Tarakanov, *User's Manual for Code KARAT*. Va, USA: BRA Inc.,
     1992.

[11] L. V. Borodachev, "Discrete modeling of low-frequency processes
     in plasma [Diskretnoye modelirovaniye nizkochastotnykh protsessov
     v plazme]," in Russian, Diss. ... Doctor of Physical and Mathematical
     Sciences, M. V. Lomonosov MSU, 2012.

[12] V. V. Andreev *et al.*, *Physical electronics and its modern applications
     [Fizicheskaya elektronika i yeye sovremennyye prilozheniya]*. Moscow:
     RUDN University, 2008, in Russian.

[13] S. E. Ernyleva, V. O. Litvin, O. T. Loza, and I. L. Bogdankevich,
     "Promising source of high-power broadband microwave pulses with
     radiation frequency variable up to two octaves," *Technical Physics*,
     vol. 59, no. 8, pp. 1228–1232, 2014. DOI: `10.1134/S1063784214080106`.

[14] F. Hecht, "New development in FreeFem++," *Journal of numerical
     mathematics*, vol. 20, no. 3–4, pp. 251–266, 2012. DOI: `10.1515/jnum-2012-0013`.

[15] V. I. Denisov, *Introduction to electrodynamics of material media [Vvedeniye v elektrodinamiku material'nykh sred]*. Moscow: M. V. Lomonosov MSU, 1989, in Russian.

Information about the authors:

Belov, Aleksandr A. — Candidate of Physical and Mathematical Sciences, Researcher of Faculty of Physics, M. V. Lomonosov Moscow State University; Assistant professor of Department of Applied Probability and Informatics of Peoples' Friendship University of Russia (RUDN University) (e-mail: `aa.belov@physics.msu.ru`, phone: +7(495)9393310, ORCID: https://orcid.org/0000-0002-0918-9263, ResearcherID: Q-5064-2016, Scopus Author ID: 57191950560)

Loza, Oleg T. — Doctor of Physical and Mathematical Sciences, Professor of Institute of Physical Research and Technology of Peoples' Friendship University of Russia (RUDN University) (e-mail: `loza-ot@rudn.ru`, phone: +7(495)9550822, ORCID: https://orcid.org/0000-0003-4676-6303)

Lovetskiy, Konstantin P. — Candidate of Physical and Mathematical Sciences, Associate professor of Department of Applied Probability and Informatics of Peoples' Friendship University of Russia (RUDN University) (e-mail: `lovetskiy-kp@rudn.ru`, phone: +7(495)9522572, ORCID: https://orcid.org/0000-0002-3645-1060)

Karnilovich, Sergey P. — Candidate of Physical and Mathematical Sciences, Assistant professor of Institute of Physical Research and Technology of Peoples' Friendship University of Russia (RUDN University) (e-mail: `karnilovich-sp@rudn.ru`, phone: +7(495)4344212, ORCID: https://orcid.org/0000-0001-5696-1546)

Sevastianov, Leonid A. — Doctor of Physical and Mathematical Sciences, Professor of Department of Applied Probability and Informatics of Peoples' Friendship University of Russia (RUDN University) (e-mail: `sevastianov-la@rudn.ru`, phone: +7(495)9522572, ORCID: https://orcid.org/0000-0002-1856-4643)

Численное моделирование холодной эмиссии в коаксиальном диоде с магнитной изоляцией

А. А. Белов[1,2], О. Т. Лоза[2], К. П. Ловецкий[2], С. П. Карнилович[2], Л. А. Севастьянов[2]

[1] *Московский государственный университет им. М. В. Ломоносова,*
Ленинские горы, д. 1, стр. 2, Москва, 119991, Россия
[2] *Российский университет дружбы народов,*
ул. Миклухо-Маклая, д. 6, Москва, 117198, Россия

Аннотация. В связи с появлением и активным развитием новых областей применения мощных и сверхмощных электровакуумных приборов СВЧ возрос интерес к изучению особенностей поведения ансамблей заряженных частиц, движущихся в пространстве взаимодействия. Примером является пучок электронов, формируемый в коаксиальном диоде с магнитной изоляцией. Численное моделирование эмиссии в таком диоде традиционно проводится с помощью методов типа «частица в ячейке». Они основаны на одновременном расчете уравнений движения частиц и уравнений Максвелла для электромагнитного поля. В данной работе предложен новый вычислительный подход, названный методом точечных макрочастиц. В нем движение частиц описывается уравнениями релятивистской механики, а для полей выписываются явные выражения в квазистатическом приближении. Выполнены расчеты формирования релятивистского электронного пучка в коаксиальном диоде с магнитной изоляцией и проведено сравнение с известными теоретическими соотношениями для скорости электронов в пучке и для тока пучка. Получено отличное согласование результатов расчета с теоретическими формулами.

Ключевые слова: коаксиальный диод с магнитной изоляцией, холодная эмиссия, точечные макрочастицы

Application of the method of continued boundary conditions to the solution of the problems of wave diffraction on various types of scatterers with complex structure

Dmitry V. Krysanov

Moscow Technical University of Communications and Informatics,
8a, Aviamotornaya St., Moscow, 111024, Russian Federation

Abstract. The article considers the application of the method of continued boundary conditions to the two-dimensional problem of diffraction of electromagnetic waves by a dielectric body with a cross section of complex geometry and to the problem of diffraction by a Janus sphere in the form of a permeable sphere partially covered by an absolutely soft or an absolutely rigid spherical screen. The results of calculating the scattering pattern for a large set of bodies of different geometry, including fractal-like scatterers, are obtained. It is illustrated that in the case of a smooth body boundary, the algorithm based on the Fredholm equations of the 1st kind makes it possible to obtain results with greater accuracy than for equations of the 2nd kind. The correctness of the method was confirmed by verifying the implementation of the optical theorem for various bodies and by comparing with the results of calculations obtained by other methods.

Key words and phrases: the method of continued boundary conditions, diffraction of waves on bodies of complex geometry, Janus sphere

1. Introduction

In the modern theory of diffraction, there is a growing need for the effective solution of increasingly complex problems, the construction of adequate mathematical models for a wide range of phenomena and processes. This, in turn, requires the development of increasingly universal methods for solving diffraction problems.

In this paper, the method of continued boundary conditions (MCBC) [1] is considered. In MCBC, the surface on which the observation point is chosen, denoted by S_δ, is located outside the scatterer at some sufficiently small distance δ from its boundary S, which is the carrier of the (auxiliary) current and over which integration is carried out. Due to the analyticity of the

wave field, the boundary condition will be approximately satisfied on the surface S_δ, and as a result, the diffraction problem is solved in an approximate formulation.

The advantages of MCBC are its versatility and simplicity. Moreover, the universality of the MCBC manifests itself, firstly, in the absence of restrictions on the geometry of the scatterer (including it is applicable both for scatterers with border breaks and for thin screens), and secondly, in the possibility of reducing the boundary value problem to Fredholm integral equations of the 1st, and of the 2nd kind [2]. In addition, when solving problems of diffraction on thin screens with the help of MCBC, it is easy to reduce the original boundary value problem to integral equations, both in the case of E- and in the case of H-polarization of the incident field (or problems of diffraction by bodies, on the boundary of which both the Dirichlet conditions and the Neumann conditions are satisfied). This is much more difficult to do, for example, when using the method of current integral equations. Another advantage of the MCBC is the ability to use various basis functions when solving the corresponding integral equations.

However, MCBC is an approximate approach, and computational algorithms based on MCBC have a lower convergence rate than, for example, algorithms based on discrete source method (in cases where the latter is applicable). At the same time, it is possible to improve the accuracy when using various basic functions (for example, splines) within the framework of the MCBC.

As an example of the application of MCBC, the problem of wave diffraction by a dielectric body of complex geometry is considered, which is very relevant and remains relatively poorly studied due to the complexity of its solution. The results of modeling the characteristics of wave scattering by dielectric bodies are of great interest in such areas as, for example, the optics of inhomogeneous media, laser flaw detection, the design of absorbing coatings, etc. [3].

The problem of diffraction on the Janus sphere in the form of a penetrable sphere partially covered by an absolutely soft or absolutely rigid spherical screen is also considered. Janus particles are of great interest in antenna engineering, medicine, and biology [4, 5]. Despite the practical significance of Janus particles, the scattering of waves by such structures has been studied rather poorly. There are a number of works in the literature devoted to both acoustic and electromagnetic problems of diffraction on the Janus sphere [6–9].

2. Solution of the problem of wave diffraction by a dielectric body of complex geometry

Let primary electromagnetic field $\mathbf{E}^0, \mathbf{H}^0$, be incident on an infinitely long magnetodielectric cylinder with a generator parallel to axis Oz and guide S. The geometry of the problem is shown in figure 1. Consider the case of E-polarization, when electric field intensity vector $\mathbf{E}$ has only one component E_z (below denoted by the letter U_- or U_+) parallel to the cylindrical body generator. The following coupling conditions will then take place at the boundary of the scatterer:

$$U_+\big|_S = U_-\big|_S, \qquad \frac{\partial U_+}{\partial n}\bigg|_S = \kappa \frac{\partial U_-}{\partial n}\bigg|_S, \tag{1}$$

where U_+ is the field inside the cylinder; $U_- = U^0 + U^1$ is the full field outside the body, where U^0 is falling and U^1 is scattered (secondary) fields; $\partial/\partial n$ is differentiation in the direction of the normal internal to S; and $\kappa = \mu_i/\mu_e$, where μ_i and μ_e are the relative magnetic permeabilities of the media inside and outside the body, respectively. The external medium ($D_e = \mathbb{R}^2\backslash\bar{D}$, $\bar{D} = D \cup S$, where D is the area bounded by curve S) and the medium inside the cylinder are assumed to be homogeneous, linear, and isotropic. At infinity, the standard radiation conditions for the scattered field are assumed to be met.
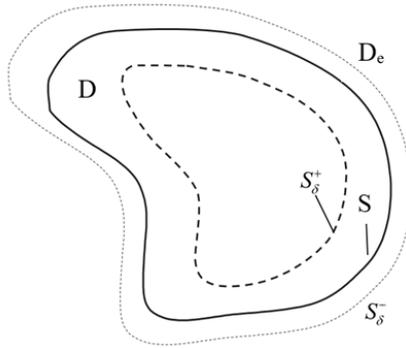
Figure 1. The geometry of the diffraction problem

Let us use the following representations to solve the Helmholtz equation in regions D and D_e, respectively [10]:

$$U_-(\mathbf{r}) = U^0(\mathbf{r}) + \int\limits_S \left\{ \frac{\partial U_-(\mathbf{r}')}{\partial n'} G_-(\mathbf{r};\mathbf{r}') - U_-(\mathbf{r}') \frac{\partial G_-(\mathbf{r};\mathbf{r}')}{\partial n'} \right\} ds',$$

$$U_+(\mathbf{r}) = -\int\limits_S \left\{ \frac{\partial U_+(\mathbf{r}')}{\partial n'} G_+(\mathbf{r};\mathbf{r}') - U_+(\mathbf{r}') \frac{\partial G_+(\mathbf{r};\mathbf{r}')}{\partial n'} \right\} ds', \tag{2}$$

in which $G_\pm(\mathbf{r};\mathbf{r}') = \dfrac{1}{4i} H_0^{(2)}\left(k_\pm|\mathbf{r} - \mathbf{r}'|\right)$ are the fundamental solutions of the scalar Helmholtz equation in $\mathbb{R}^2$ with material parameters of the media D_e and D, respectively, k_+ and k_- are the wavenumbers of the medium inside and outside the scatterer. Demanding, in accordance with MCBC, the fulfillment of conditions Eqs. (1) to be met on contour S_δ^- located in $\mathbb{R}^2\backslash\bar{D}$, and on contour S_δ^+ located in area D (see figure 1) using equations (2), we obtain the following systems of the Fredholm integral equations of the first or second kind, respectively:

$$\int\limits_{S} \left\{ \frac{\partial U(\mathbf{r}')}{\partial n'} \left(G_{-}(\mathbf{r}_{-};\mathbf{r}') + \kappa G_{+}(\mathbf{r}_{+};\mathbf{r}') \right) - \right.$$

$$\left. -U(\mathbf{r}') \left(\frac{\partial G_{-}(\mathbf{r}_{-};\mathbf{r}')}{\partial n'} + \frac{\partial G_{+}(\mathbf{r}_{+};\mathbf{r}')}{\partial n'} \right) \right\} ds' = -U^{0}(\mathbf{r}_{-}),$$

$$\int\limits_{S} \left\{ \frac{\partial U(\mathbf{r}')}{\partial n'} \left(\frac{\partial G_{-}(\mathbf{r}_{-};\mathbf{r}')}{\partial n} + \frac{\partial G_{+}(\mathbf{r}_{+};\mathbf{r}')}{\partial n} \right) - \right.$$

$$\left. -U(\mathbf{r}') \left(\frac{\partial^{2} G_{-}(\mathbf{r}_{-};\mathbf{r}')}{\partial n \partial n'} + \frac{1}{\kappa} \frac{\partial^{2} G_{+}(\mathbf{r}_{+};\mathbf{r}')}{\partial n \partial n'} \right) \right\} ds' = -\frac{\partial U^{0}(\mathbf{r}_{-})}{\partial n},$$

$$\tag{3}$$

$$U(\mathbf{r}) = \frac{1}{2}U^{0}(\mathbf{r}_{-}) + \frac{1}{2} \int\limits_{S} \left\{ \frac{\partial U(\mathbf{r}')}{\partial n'} \left(G_{-}(\mathbf{r}_{-};\mathbf{r}') - \kappa G_{+}(\mathbf{r}_{+};\mathbf{r}') \right) - \right.$$

$$\left. -U(\mathbf{r}') \left(\frac{\partial G_{-}(\mathbf{r}_{-};\mathbf{r}')}{\partial n'} - \frac{\partial G_{+}(\mathbf{r}_{+};\mathbf{r}')}{\partial n'} \right) \right\} ds',$$

$$\frac{\partial U(\mathbf{r})}{\partial n} = \frac{1}{1+\kappa} \frac{\partial U^{0}(\mathbf{r}_{-})}{\partial n} + \frac{1}{1+\kappa} \int\limits_{S} \left\{ \frac{\partial U(\mathbf{r}')}{\partial n'} \left(\frac{\partial G_{-}(\mathbf{r}_{-};\mathbf{r}')}{\partial n} - \right. \right.$$

$$\left. \left. -\kappa \frac{\partial G_{+}(\mathbf{r}_{+};\mathbf{r}')}{\partial n} \right) - U(\mathbf{r}') \left(\frac{\partial^{2} G_{-}(\mathbf{r}_{-};\mathbf{r}')}{\partial n \partial n'} - \frac{\partial^{2} G_{+}(\mathbf{r}_{+};\mathbf{r}')}{\partial n \partial n'} \right) \right\} ds',$$

$$\tag{4}$$

where observation points $M(\mathbf{r}_{\pm})$ belong to contours $S_{\delta}^{\pm}$ and point $M(\mathbf{r}) \in S$ and it is denoted that $U = U_{-}$. Note that the contours that are separated from S by a fairly small distance δ are most often chosen as $S_{\delta}^{\pm}$; i.e., equidistant contours are considered [1, 10]. Further, to solve system of equations (3), (4), the Krylov–Bogolyubov method is used. A generalization of the method to the problem of diffraction by a cylindrical body located in a homogeneous magnetodielectric half-space is given in [11].

Let us consider the results of numerical modeling. Thereafter, we will assume that the body is irradiated by a plane wave. As an example, let us first consider the diffraction problem on an elliptical cylinder, a cylinder with a quadrifolium cross section, and a cylinder with a rectangular cross-section. Figure 2 shows the angular dependences of the scattering pattern for the corresponding geometry obtained for the following values of the problem parameters: $k\delta = 10^{-4}$, $\varphi_{0} = 0$, $\mu_{i} = 1$, $\varepsilon_{i} = 4$ (the material parameters of the external medium are $\mu_{e} = 1$, $\varepsilon_{e} = 1$ everywhere). The dimensions of the bodies had the following values: the semiaxis or half the side lengths of the rectangle $ka = 5$, $kb = 1$ and the $ka = 5$, $\tau = 0.5$ parameters for the body with a cross section in the form of a quadrifolium. The results were compared with the patterns constructed using the modified discrete source method [10, 12]. Note that the modified discrete source method cannot be directly applied to the problem of the diffraction on bodies that have boundary breaks, and so the contour of the axial section of the body was approximated by a smooth contour to solve the problem using the modified discrete source method [12].

Note also that the modified discrete source method provides high accuracy of calculation for bodies with a smooth border, such as ellipses, multifoil, etc.

(a) Elliptical cylinder

(b) Body with a cross section in the form of a quadrifolium

(c) Body with a rectangular cross section
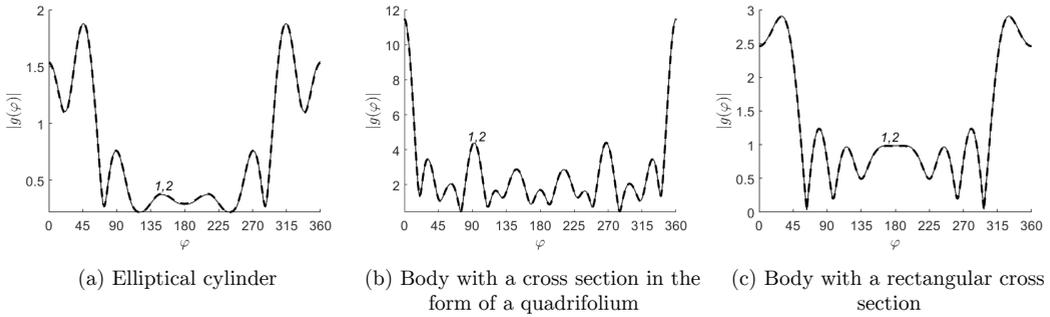
Figure 2. The angular dependence of the scattering pattern for different bodies: *(1)* the modified discrete source method and *(2)* the continued boundary conditions method

Figures 3 and 4 illustrate the angular dependences of the scattering pattern for the fractal-like cylinders with a cross section in the form of a Koch snowflake and Sierpinski curve (first iteration) [13] at the problem parameters of $k\delta = 10^{-4}$, $\mu_i = 1$, $\varepsilon_i = 4$. The maximum cross-sectional size of a body with a cross section in the form of the Koch snowflake and a body with a cross section in the form of the Sierpinski curve on the x axis was $kL = 10$. Two different angles of incidence $\varphi_0 = 0$ and $\varphi_0 = 45°$ were considered. As follows from the figures for the geometry under study, the maximum points of the angular dependences of the scattering pattern roughly coincide with the angles of incidence of the plane wave. It can also be seen that the dependences of the pattern for both a body with a section in the form of the Koch snowflake and a body with a section in the form of the Sierpinski curve have quite large side lobes.
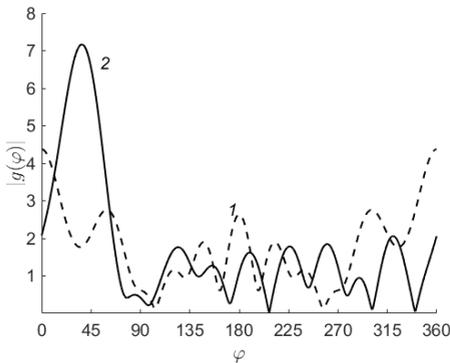
Figure 3. The angular dependence of the scattering pattern for a body with a cross section in the form of a Koch snowflake. The angle of incidence of the wave *(1)* $\varphi_0 = 0$ and *(2)* $\varphi_0 = 45°$

Figure 4. The angular dependence of the scattering pattern for a body with a cross section in the form of a Sierpinski curve. The angle of incidence of the wave *(1)* $\varphi_0 = 0$ and *(2)* $\varphi_0 = 45°$

Table 1 shows the differences in the scattering pattern modules of the specified geometry obtained by two methods: using the modified discrete source method and the continued boundary conditions method. As can be seen from the table 1, the difference in results 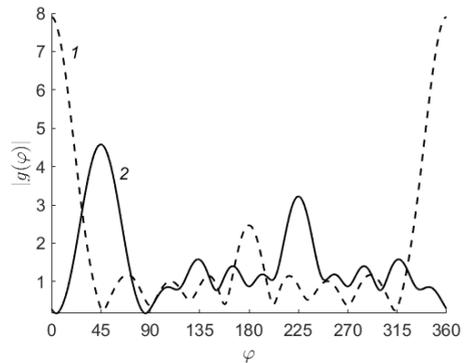decreases as the number of basic functions used increases. It also follows from the given data that for bodies with a smooth boundary, the use of the Fredholm equations of the 1st kind is more preferable, due to faster convergence. In the case of a body with a rectangular section, the use of Fredholm equations of the 2nd kind gives better results.

Table 1

Comparison of the results obtained using the modified discrete source method
and the continued boundary conditions method

N	System of integral equations of the first kind		System of integral equations of the second kind	
	absolute error	relative error	absolute error	relative error
Diffraction on an elliptical cylinder				
48	$1.295 \cdot 10^{-2}$	2.038%	$1.453 \cdot 10^{-1}$	24.297%
96	$1.904 \cdot 10^{-3}$	0.230%	$4.183 \cdot 10^{-2}$	7.238%
192	$6.096 \cdot 10^{-4}$	0.067%	$1.144 \cdot 10^{-2}$	2.003%
288	$5.834 \cdot 10^{-4}$	0.075%	$5.539 \cdot 10^{-3}$	0.977%
384	$5.607 \cdot 10^{-4}$	0.075%	$3.450 \cdot 10^{-3}$	0.612%
Diffraction on a body with a cross section in the form of a quadrifolium				
48	$1.643 \cdot 10^{-1}$	10.411%	$3.836 \cdot 10^{-1}$	21.169%
96	$2.499 \cdot 10^{-2}$	1.442%	$9.462 \cdot 10^{-2}$	5.125%
192	$5.802 \cdot 10^{-3}$	0.325%	$2.534 \cdot 10^{-2}$	1.370%
288	$2.984 \cdot 10^{-3}$	0.166%	$1.143 \cdot 10^{-2}$	0.619%
384	$2.176 \cdot 10^{-3}$	0.121%	$6.474 \cdot 10^{-3}$	0.351%
Diffraction on a body with a rectangular cross section				
48	$3.498 \cdot 10^{-2}$	4.781%	$5.035 \cdot 10^{-2}$	6.795%
96	$1.466 \cdot 10^{-2}$	1.956%	$1.417 \cdot 10^{-2}$	1.715%
192	$7.358 \cdot 10^{-3}$	0.879%	$4.773 \cdot 10^{-3}$	0.466%
288	$5.229 \cdot 10^{-3}$	0.561%	$3.122 \cdot 10^{-3}$	0.248%
384	$4.219 \cdot 10^{-3}$	0.429%	$2.641 \cdot 10^{-3}$	0.208%

3. Solution to the problem of wave diffraction on the Janus sphere

Let us consider the mathematical formulation of the problem. Let a homogenous sphere of radius a be covered with an infinitely thin spherical screen S with an opening angle $2\theta_J$. We introduce a spherical coordinate system in which the z is directed along the axis of the considered body of revolution (the Janus sphere). The geometry of the problem is depicted in figure 5.
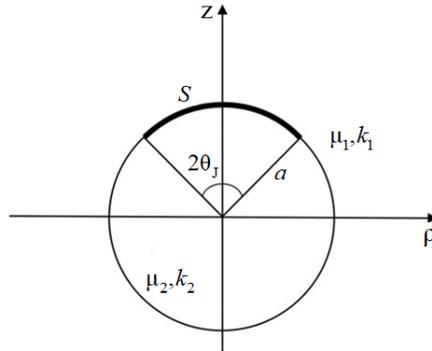
Figure 5. Axial section of a Janus sphere

We suppose that the wavenumbers and medium densities outside and inside the sphere are equal to k_1, μ_1 and k_2, μ_2, respectively. Thus, the wave field outside and inside the sphere satisfies the homogeneous Helmholtz equations

$$
\begin{aligned}
\Delta U + k_1^2 U = 0, \quad r > a, \\
\Delta U + k_2^2 U = 0, \quad 0 < r < a,
\end{aligned}
\tag{5}
$$

where r is the radial coordinate in the spherical coordinate system. For the sake of brevity, we consider only the case of an absolutely soft spherical screen. Then the boundary conditions on the surface of the screen have the form

$$
U\big|_{r=a,\,\theta<\theta_J} = 0.
\tag{6}
$$

At $\theta \in (\theta_J, \pi)$, the matching conditions are satisfied:

$$
[U] = 0, \quad \left[\frac{1}{\mu}\frac{\partial U}{\partial r}\right] = 0,
\tag{7}
$$

where μ is the density ($\mu = \mu_1$ for $r > a$, $\mu = \mu_2$ for $r < a$) and the square brackets indicate a jump of the corresponding quantity. We assume that the Janus sphere is irradiated by a plane wave, which has the form

$$
U_{\text{inc}} = \exp\left(-ik_1 r \left(\sin\theta \sin\theta_0 \cos\varphi + \cos\theta \cos\theta_0\right)\right).
\tag{8}
$$

Here, θ_0 is the angle of incidence of the plane wave. The scattered field U^1 satisfies the radiation condition at infinity. The total field also satisfies the Meixner condition at the edge of the spherical screen.

Let us apply the MCBC for solving the posed diffraction problem. For this purpose, we represent the field outside the permeable sphere in the form

$$U(\mathbf{r}) = U^0(\mathbf{r}) - \int\limits_S J(\mathbf{r}')G(\mathbf{r},\mathbf{r}')ds'. \tag{9}$$

Here, $U^0(\mathbf{r})$ is the primary field determined from the solution of the diffraction problem on the sphere in the absence of the screen,

$$J(\mathbf{r}') \equiv J(\theta',\varphi') = \left[\frac{\partial U}{\partial r'}\right]_{\substack{r=a \\ \theta<\theta_J}}.$$

In equations (9) $G(\mathbf{r},\mathbf{r}')$ is the Green function of the permeable sphere; for $r > a$, it has the form

$$G = G_0 + G_1, \tag{10}$$

where

$$G_0 = \frac{e^{-ik_1 R}}{4\pi R}, \quad R = |\mathbf{r} - \mathbf{r}'|, \tag{11}$$

$$G_1 = \frac{k_1}{4\pi i}\sum_{n=0}^{\infty}(2n+1)V_n H_n^{(2)}(k_1 r')H_n^{(2)}(k_1 r)P_n(\cos\gamma), \tag{12}$$

$$\cos\gamma = \sin\theta\sin\theta'\cos(\varphi-\varphi') + \cos\theta\cos\theta', \tag{13}$$

$$V_n = \frac{\mu_{12}J_n(k_1 a)\eta_n(k_2 a) - \eta_n(k_1 a)J_n(k_2 a)}{\xi_n(k_1 a)J_n(k_2 a) - \mu_{12}H_n^{(2)}(k_1 a)\eta_n(k_2 a)}, \tag{14}$$

$$\mu_{12} = \frac{\mu_1}{\mu_2}, \quad \eta_n(x) = xJ_n'(x), \quad \xi_n(x) = xH_n^{(2)\prime}(x), \tag{15}$$

and $J_n(x)$, $H_n^{(2)}(x)$ are the spherical Bessel and Hankel functions, respectively, $P_n(x)$ — Legendre polynomials. Note that the primary field outside the sphere has the form

$$U^0(\mathbf{r}) = U_{\text{inc}}(\mathbf{r}) + \sum_{n=-\infty}^{\infty} i^{-n}(2n+1)V_n H_n^{(2)}(k_1 r)P_n(\cos\gamma_0), \tag{16}$$

where $\cos\gamma_0(\theta,\varphi) = \sin\theta\sin\theta_0\cos\varphi + \cos\theta\cos\theta_0$.

According to the standard scheme of the MCBC, we then substitute formula (9) into boundary condition (6) imposed on the auxiliary surface S_δ shifted by a small distance δ from the surface S [1, 10, 14]. As a result, the problem will be reduced to solving a two-dimensional Fredholm integral equation of the first kind, which has the following form in spherical coordinates:

$$\int\limits_0^{2\pi}\int\limits_0^{\theta_J} K(\theta,\varphi,\theta',\varphi')J(\theta',\varphi')\sin\theta'd\theta'd\varphi' = B(\theta,\varphi), \tag{17}$$

$$\theta \in [0,\theta_J], \quad \varphi \in [0,2\pi],$$

where $K(\theta, \varphi, \theta', \varphi') = a^2\, G\big|_{r'=a, r=a+\delta}$, $B(\theta, \varphi) = U^0\big|_{r=a+\delta}$.

Equation (17) was solved using a piecewise-constant approximation of an unknown function with subsequent application of the Krylov–Bogolyubov method. The kernel of Eq. (17) was found using the acceleration of the convergence of series (12). In order to speed up the convergence of this series, the asymptotic behavior of the n-th term of the series as $n \to \infty$ was distinguished (this quantity can be called the singular part of the Green's function). The singular part of the Green's function was summarized analytically using the generating function of the Legendre polynomials. The remaining (regular) part of the Green's function was a fairly fast convergent series. A detailed derivation of the main relations, as well as the case of an absolutely rigid screen, was considered in [15].

Let us consider results of the numerical simulation. The results of calculating the intensity of the scattered field in the far zone for the Janus sphere were compared with the results obtained using the T-matrix method, which are given in [9]. The acoustic problem of diffraction was considered [9]. The wavenumbers and densities of the media inside and outside the sphere were equal to $k_1 = 1$, $\mu_1 = 1$ and $k_2 = 1.5$, $\mu_2 = 1.5$, respectively. Parameter δ in using the MCBC was taken to be 10^{-3} in all cases. The angle of incidence of the primary wave $\theta_0 = 0°$. Sphere radius $k_1 a = 6$. The half-opening angle was equal to $\theta_J = 90°$. The number of collocation points along both angular coordinates $N_1 = 25$, $N_2 = 100$. Figure 6 shows the angular dependences of the scattered field intensity obtained using the T-matrix method (curve *1*) and using the proposed approach based on the MCBC (curve *2*). It can be seen from Fig. 6 that the results coincide with graphic accuracy in the case of the Dirichlet condition. Due to the presence of the second normal derivative of the Green's function in the case of the Neumann condition on the screen, the accuracy of calculating the diagram using the MCBC is somewhat lower than in the case of the Dirichlet condition. For the considered problem, the accuracy of the optical theorem was verified. Calculations showed that the relative error in the fulfillment of the optical theorem did not exceed 2%.

(a) absolutely soft spherical screen   (b) absolutely rigid spherical screen
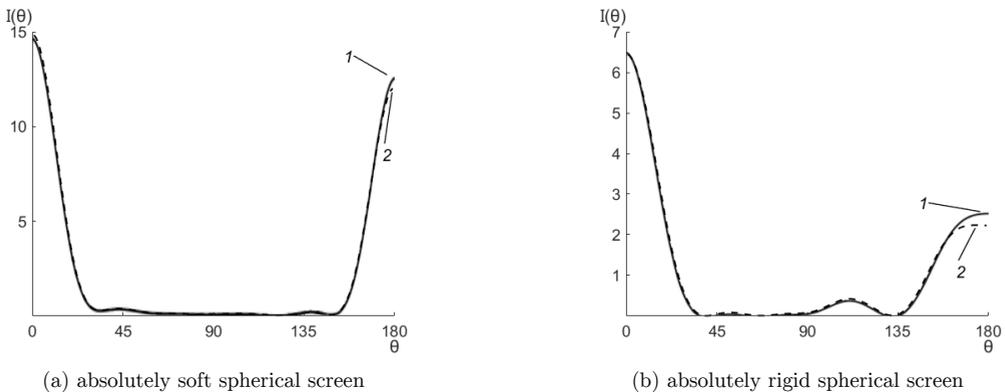
Figure 6. Angular dependences of the intensity of the scattered field of a Janus sphere in the form of a penetrable sphere partially covered with spherical screen, obtained using the T-matrix method and using an algorithm based on MCBC

Figure 7 depicts angular dependences of the scattered field intensity on the opening angle of the spherical screen. Curve *1* corresponds to diffraction by a permeable sphere not covered by a screen. Curve *6* in the figures shows the dependences obtained for the intensity upon diffraction of a plane wave by an absolutely soft (Fig. 7a) or absolutely rigid (Fig. 7b) sphere of the corresponding wave size using the modified discrete source method (MDSM) [16]. Curves *2–5* correspond to screen half-opening angles equal to 45°, 90°, 135°, and 179°. The wave size of the Janus sphere and the angle of incidence of the wave are $k_1 a = 6$, $\theta_0 = 0$. The material parameters of the media and wavenumbers are the same as for previous figure. It can be seen from the figure that in the case when the screen almost completely covers the sphere (curve *5*), the scattered field intensity graph coincides with the results for a perfectly reflecting sphere, which corresponds to the physical picture of the phenomenon under consideration. It is also seen that in all cases there is a sharp intensity maximum in the direction of the angle of incidence of the plane wave. In the case of a soft screen, the magnitude of the maximum has the greatest value for $\theta_J = 179°$ (that is, when the screen degenerates into a sphere). In the absence of a screen (that is, in the case of diffraction by a permeable sphere), the maximum in the direction of wave incidence is even greater (than for a covered sphere). In the backscattering direction (at $\theta = 180°$), there is also an intensity maximum, which takes on the largest values at $\theta_J = 45°$ and $\theta_J = 90°$. In the case of an absolutely rigid screen, the value of the intensity maximum in the direction of incidence of a plane wave is much greater for $\theta_J = 45°$ compared to other screen opening angles. The backscatter level is also maximum at $\theta_J = 45°$.

(a) Absolutely soft spherical screen   (b) Absolutely rigid spherical screen
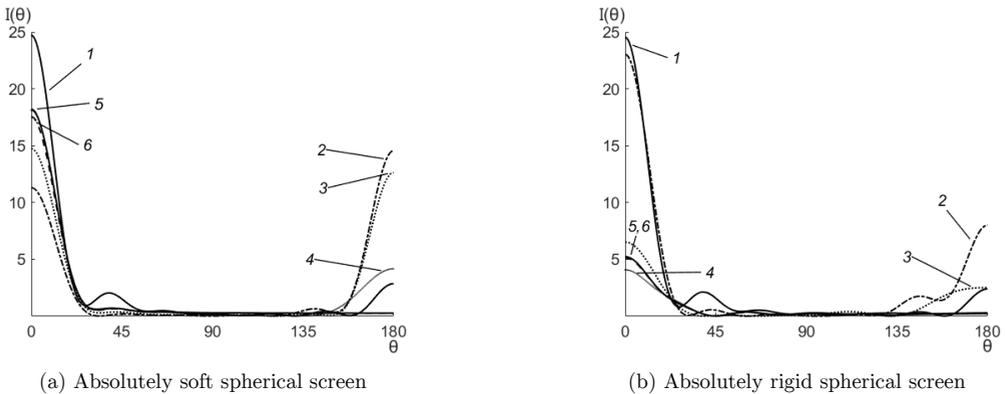
Figure 7. Angular dependences of the intensity of the scattered field of the Janus sphere for different opening angles of a spherical screen covering it

4. Conclusions

Based on the method of continued boundary conditions, an algorithm for solving the two-dimensional problem of plane wave diffraction by dielectric bodies with complex cross-sectional geometry is shown. A comparison is made with the results obtained using modified discrete source method. It is shown

that the MCBC makes it possible to obtain the results of scattering diagram calculations with a sufficiently high accuracy. The results of calculating the scattering diagram for a large set of bodies of different geometry, including fractal-like scatterers, are obtained. It is illustrated that in the case of a smooth body boundary, the algorithm based on the Fredholm equations of the 1st kind allows obtaining results with greater accuracy than for equations of the 2nd kind.

An algorithm for solving the scalar diffraction problem on the Janus sphere is shown on the basis of the MCBC. The results of calculating the intensity of the scattered field obtained using the proposed method are compared with the results found using the T-matrix method. It has been shown that the results coincide well. The angular dependences of the intensity of the scattered field for various opening angles of the reflecting screen are constructed and studied. A significant difference is shown between the behavior of the angular dependences of the intensity in the case of an absolutely soft and absolutely rigid screen.

References

[1] A. G. Kyurkchan and A. P. Anyutin, "The method of continued boundary conditions and wavelets," *Doklady Mathematics*, vol. 66, no. 1, pp. 132–135, 2002.

[2] A. G. Kyurkchan and A. P. Anyutin, "The well-posedness of the formulation of diffraction problems reduced to Fredholm integral equations of the first kind with a smooth kernel," *Journal of Communications Technology and Electronics*, vol. 51, no. 7, pp. 48–51, 2006. DOI: `10.1134/S1064226906010062`.

[3] M. I. Mishchenko, N. T. Zakharova, N. G. Khlebtsov, G. Videen, and T. Wriedt, "Comprehensive thematic T-matrix reference database: A 2015–2017 update," *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 202, pp. 240–246, 2017. DOI: `10.1016/j.jqsrt.2017.08.007`.

[4] J. Zhang, B. A. Grzybowski, and S. Granick, "Janus particle synthesis, assembly, and application," *Langmuir*, vol. 33, no. 28, pp. 6964–6977, 2017. DOI: `10.1021/acs.langmuir.7b01123`.

[5] M. Lattuada and T. A. Hatton, "Synthesis, properties and applications of Janus nanoparticles," *Nano Today*, vol. 6, no. 3, pp. 286–308, 2011. DOI: `10.1016/j.nantod.2011.04.008`.

[6] D. Kim, E. J. Avital, and T. Miloh, "Sound scattering and its reduction by a Janus sphere type," *Advances in Acoustics and Vibration*, vol. 2014, no. 392138, 2014. DOI: `10.1155/2014/392138`.

[7] A. Gillman, "An integral equation technique for scattering problems with mixed boundary conditions," *Advances in Computational Mathematics*, vol. 43, no. 2, pp. 351–364, 2017. DOI: `10.1007/s10444-016-9488-6`.

[8] S. C. Hawkins, T. Rother, and J. Wauer, "A numerical study of acoustic scattering by Janus spheres," *The Journal of the Acoustical Society of America*, vol. 147, no. 6, pp. 4097–4105, 2020. DOI: `10.1121/10.0001472`.

[9]   T. Rother, *Sound scattering on spherical objects*. Heidelberg: Springer, 2020.

[10] A. G. Kyurkchan and N. I. Smirnova, *Mathematical modeling in diffraction theory: based on a priori information on the analytical properties of the solution*. Amsterdam: Elsevier, 2015.

[11] D. V. Krysanov, A. G. Kyurkchan, and S. A. Manenkov, "Application of the method of continued boundary conditions to the solution of the problem of wave diffraction on scatterers of complex geometry located in homogeneous and heterogeneous media," *Optics and Spectroscopy*, vol. 128, no. 4, pp. 481–489, 2020. DOI: 10.1134/S0030400X20040141.

[12] A. G. Kyurkchan and S. A. Manenkov, "Application of different orthogonal coordinates using modified method of discrete sources for solving a problem of wave diffraction on a body of revolution," *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 113, no. 18, pp. 2368–2378, 2012. DOI: 10.1016/j.jqsrt.2012.05.010.

[13] R. M. Crownover, *Intoduction to fractals and chaos*. Boston: Jones and Bartlett Publishers, 1995.

[14] A. G. Kyurkchan and S. A. Manenkov, "Solution of the problem of diffraction by a plane screen in a plane layered medium with the help of the method of continued boundary conditions," *Journal of Communications Technology and Electronics*, vol. 65, no. 7, pp. 778–786, 2020. DOI: 10.1134/S1064226920060200.

[15] D. V. Krysanov, A. G. Kyurkchan, and S. A. Manenkov, "Two approaches to solving the problem of diffraction on a Janus sphere," *Acoustical Physics*, vol. 67, no. 2, pp. 108–119, 2021. DOI: 10.1134/S1063771021020020.

[16] S. A. Manenkov, "A new version of the modified method of discrete sources in application to the problem of diffraction by a body of revolution," *Acoustical Physics*, vol. 60, no. 2, pp. 127–133, 2014. DOI: 10.1134/S1063771014010102.

Information about the authors:

Krysanov, Dmitry V. — postgraduate student of Department of Probability Theory and Applied Mathematics of Moscow Technical University of Communications and Informatics (e-mail: d.v.krysanov@mtuci.ru, phone: +7(903)5418711, ORCID: https://orcid.org/0000-0001-5100-3007)

Применение метода продолженных граничных условий к решению задач дифракции на различных типах частиц сложной структуры

Д. В. Крысанов

Московский технический университет связи и информатики,
ул. Авиамоторная, д. 8а, Москва, 111024, Россия

Аннотация. В статье рассмотрено применение метода продолженных граничных условий к двумерной задаче дифракции электромагнитных волн на диэлектрическом теле с поперечным сечением сложной геометрии и к задаче дифракции на сфере Януса в виде проницаемого шара, частично покрытого абсолютно мягким или абсолютно жёстким сферическим экраном. Получены результаты расчёта диаграммы рассеяния для большого набора тел разной геометрии, в том числе фракталоподобных рассеивателей. Проиллюстрировано, что в случае гладкой границы тела алгоритм на основе уравнений Фредгольма 1-го рода позволяет получать результаты с большей точностью, чем для уравнений 2-го рода. Корректность метода подтверждена при помощи проверки выполнения оптической теоремы для различных тел и путём сравнения с результатами расчётов, полученных другими методами.

Ключевые слова: метод продолженных граничных условий, дифракция волн на телах сложной геометрии, сфера Януса

Development and analysis of models for service migration to the MEC server based on hysteresis approach

Dmitry S. Poluektov, Abdukodir A. Khakimov

Peoples' Friendship University of Russia (RUDN University),
6, Miklukho-Maklaya St., Moscow, 117198, Russian Federation

Abstract. Online video services are among the most popular ways of content consumption. Video hosting servers have a very high load every day, which we propose to reduce by migrating the application with the video content in demand to the local Multi-access Edge Computing (MEC) server of the target. This makes it possible to improve the quality of services (QoS) provided to users by reducing the transmission delay. Therefore, an architecture has been proposed that allows, at times of increased demand for the same video content, to migrate the video service application to the edge servers of the network operator. To evaluate the performance of this approach, a mathematical model was developed in the form of a queuing system. The results of the numerical experiment make it possible to optimize the time of using local MEC servers to provide video content.

Key words and phrases: queuing system, service migration, MEC, Markov process, truncated Markov process, video content

1. Introduction

In the modern world, demand for various multimedia services is increasing every year. For example, services for providing online video content are very popular and allow us to access information in a simple way anytime and anywhere from a device with Internet connection. However, with the growth in the amounts of video content and an increase in its demand, requirements for the quality of the services (QoS) are as well increased. Video service providers, in turn, are trying to reduce transmission delays and improve the quality of the video, which increases the size of video files and requires more channel bandwidth to be transmitted to the end user.

The idea of decentralized content placement by a video service provider is not new. In most countries, large cities, operators use the services of geographically distributed content delivery network architecture (CDN). It allows for the video data delivery optimization by using servers, located

much closer to the end user. Multi-access Edge Computing (MEC) servers utilization allows service providers to optimize the transmission process by placing the user-requested content on servers not only within a specific city, but also within a specific district or street. In the same way it solves the problem of high load on transport networks, which is beneficial for the video service provider, who gets the opportunity to provide high-quality content. Thus, the transport network operator can reduce network operation costs and receive additional profit for renting edge computing servers.

MEC introduces the cloud computing capabilities and IT service environment at the edge of the mobile network. The network edge includes base station infrastructure and data centers close to the radio network.

In work [1], the authors have presented a classification of application models and a study of the latest models of mobile cloud applications. In [2], a brief analysis of the requirements for mobile cloud computing (MCC) have been done, the main applications and upload technologies, the classification of contexts and context management methods. In [3], the authors have provided an overview of the definitions, architectures, and applications of MCC, as well as common problems and some existing solutions. In reference [4] a study of existing work on MCC platforms and intelligent access schemes can be found. Another group of scientists in [5] has investigated a detailed taxonomy of mobile cloud computing based on key issues and approaches to address them. Work [6] has introduced a comprehensive overview of the current MCC authentication mechanism and compared cloud computing. The authors in [7] have studied a taxonomy of MEC based on various aspects, including its characteristics, access technologies, applications, purposes, etc. Reference [8] has categorized deployed applications in MEC according to the technical metrics of MEC and the benefit brought by MEC to network stakeholders. A discussion of threats and security in boundary paradigms, as well as possible solution for each specific problem, can be found in [9]. In [10], representative applications and various aspects of the study of fog computing problems are highlighted. An overview on emerging security and privacy issues in fog computing, as well as cloud computing issues is closely discussed in [11]. A study of web caching and prefetching methods for improving network performance, as well as a classification of web caching policies, can be found in [12]. A description of the advantages and disadvantages of cache replacement strategies can be found in [13]. The model of interaction between the edge computing based on Software-defined networking (SDN) and Network functions virtualization (NFV) technology and the cloud computing in the next generation Internet of Things (IoT) is presented in [14]. In [15] authors consider the usage of a MEC server for processing home health monitoring data locally, making it possible to optimize the system-wide cost and the number of patients benefiting from MEC.

2. System description

As described above, the consumption of online video content is growing every year, since any information presented by a video sequence with sound accompaniment makes it easier to perceive or just spend leisure time. Requests for the provision of such services, especially of an entertainment nature, do

not have a constant intensity, but in most cases occur in avalanche bursts at different times or days. For example, in the evening, most people come home from work and watch their favorite series, talk shows, etc. This section describes the process of providing online video and possible scenarios for optimizing content delivery to users.
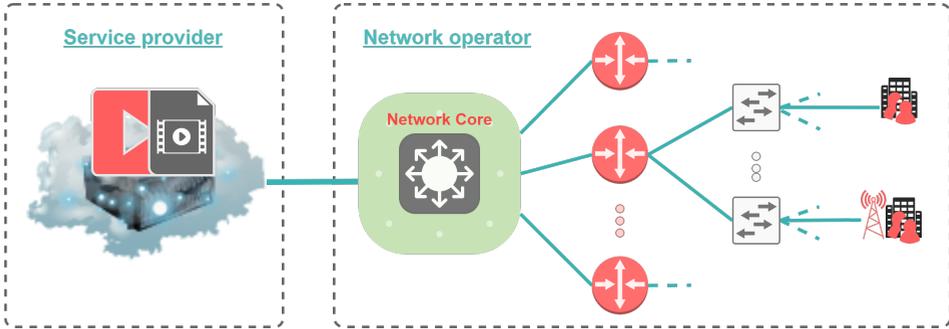

Figure 1. Network architecture for connecting users to a video content service

The figure 1 shows a diagram for providing online video services to users. On the left side the service provider's servers are shown. These servers host and process videos to provide the users on demand. On the right is the operator's last mile access network, which provides end users with access to the global network, and, in particular, connections to the service provider's servers. This network is presented in more detail and consists of

— 1st segment — it includes all elements of the operator's network core and, is responsible for routing traffic within the network and outside of it;
— 2nd segment — it consists of terminal switching devices for wired connection of users and/or base stations of a cellular network for a wireless connection of mobile users (this segment is variable and changes depending on the task).

Between the service provider and the network operator, there are backbone operators and traffic exchange points, which are shown as a direct connection, since they mainly just make data transfer delay, and are also not the main beneficiaries in optimizing the process of providing a service to users.

We consider the process of providing a service to users. The very process of establishing a connection for an online video service has been described and studied in detail in work [16]. Thus it is proposed to focus on the main points presented in the figure 2:

1) a user in the carrier's network sends a video viewing request to a service provider;
2) the service provider processes the request and sends a connection confirmation to the user;
3) a connection is established and the user starts watching the video.

The network operator serves several zones (districts) with N users who can potentially start watching video content. Most of the time, this service is not in great demand, which means that it does not process a large load on the operator's network and service provider's servers (the content delivery method shown in the figure 2). Nevertheless, at some point, a large number

of users start watching the same video content at the same time (for example, the release of a popular series or talk show), thereby a very high load on the servers and the network takes place.
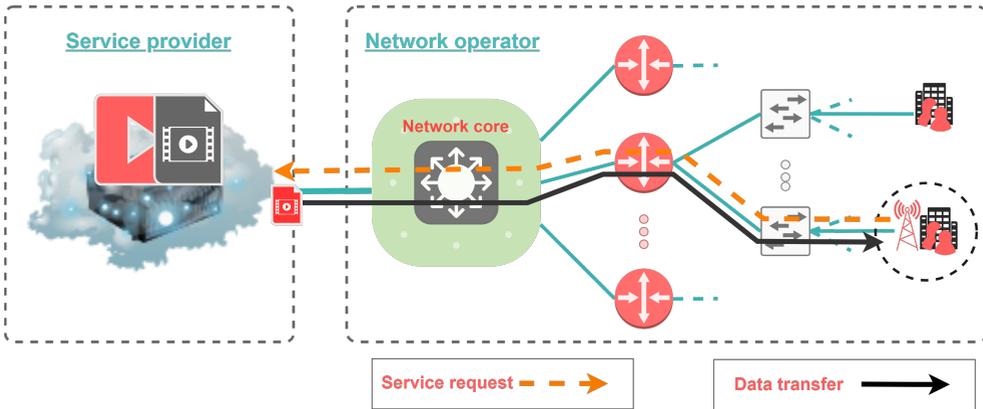

Figure 2. Data transmission for centralized online video providing

Such an avalanche surge of requests leads to an increase in the delay in the transmission of video content and a decrease in bandwidth available for each user, due to the limited bandwidth of the operator's network channel.

To reduce the load on the resources of the service provider and the operator's network, we propose placing MEC servers (figure 3) in each boundary switching zone. This allows us to temporarily migrate the application with access to the video to the facilities of the network operator, located close to the users. Then, the process of content delivery comes down to establishing a connection and transmitting video from MEC server for each high load zone.
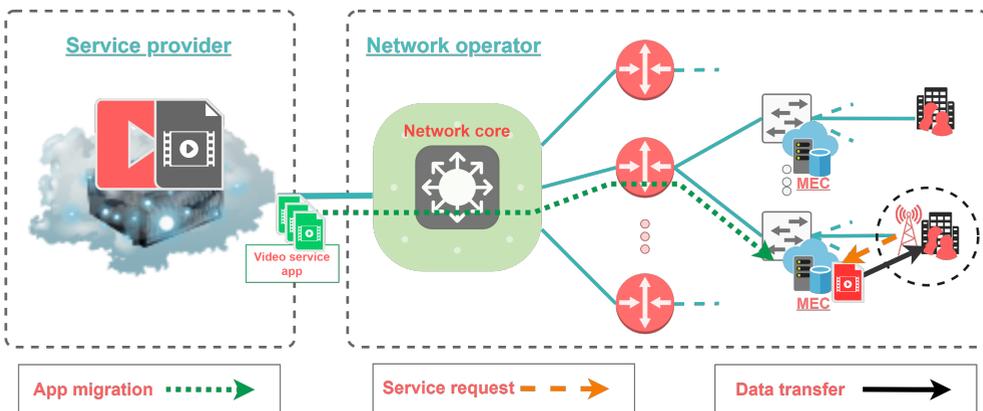
Figure 3. Data transmission for local online video providing with MEC

To formulate a system model, let's consider a single zone of the network operator in which there are N users who can potentially request the same service. $H < N$ users can watch video content directly from the service operator's servers. As soon as $n > H$ users request the service, the video service application is migrated to the MEC server, and all users are already

watching video from the local MEC server. Eventually, the demand for a large number of users in the service disappears, then when the number of active users $n \leqslant L$ (L — the threshold for the appropriate use of the MEC server) is reached, the process of deleting the service from the MEC server and switching the remaining users to the service provider's main server is initiated. The process of deleting and switching does not occur immediately if the number of users becomes $n \leqslant L$ during a certain period of time.

Figure 4 shows the sets of states of the system model and the transitions between them:

— $\mathcal{X}_0$ — serving users directly from the service provider's servers, without using MEC;
— $\mathcal{X}_1(L)$ — disable MEC, switch users to the main server;
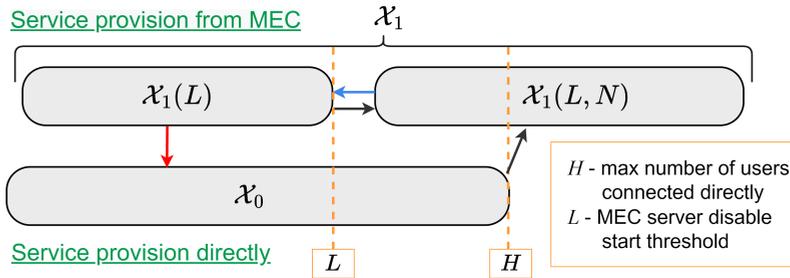— $\mathcal{X}_1(L, H)$ — services from the MEC server without the ability to switch to the main server.

Figure 4. The set of system states for migrating a video service application to MEC

3.   Queuing model service migration with hysteresis loop

To analyze the performance of the described system, we model it in the form of a queuing system (QS) model for migration of an application from a remote service provider server to the network operator's MEC server. The system receives a flow of user requests to provide streaming video services. This flow is assumed to be Poisson distributed with mean λ. When serving in the $\mathcal{X}_0$ group and reaching the number of users in the system $n > H$, users are served in the states $\mathcal{X}_1(L, H)$. Video content viewing duration by one user is exponentially distributed with average μ^{-1} minutes. When the system switches to the $\mathcal{X}_1(L)$ state group, the process of disabling the MEC server and switching user services to the service provider's server in the $\mathcal{X}_0$ state group is initialized. It takes an average of α^{-1} minutes to shut down the MEC server correctly, and this parameter is also exponentially distributed.

To analyze the queuing system, we introduce the Markov process $\mathcal{X}(t)$, which describes the behavior of the system at time t, with the state space:

$$\mathcal{X} = \mathcal{X}_0 + \mathcal{X}_1,$$

where

$$\mathcal{X}_0 = \{(s, n) \in \mathcal{X}_0 : \; s = 0, n = (0, H)\},$$

$$\mathcal{X}_1 = \{(s, n) \in (\mathcal{X}_1(L) \cup \mathcal{X}_1(L, H)) : \ s = 1, n = (1, N)\}.$$

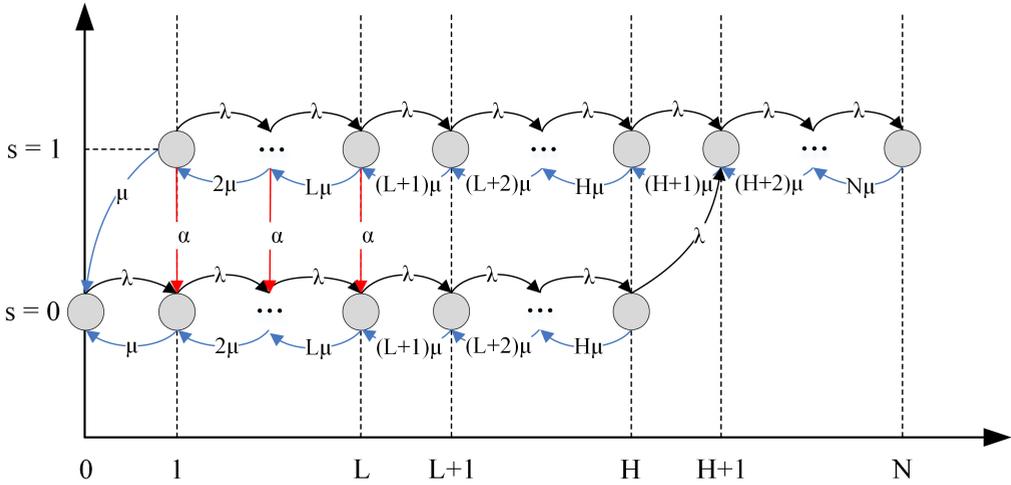The state transition diagram of the Markov process $\mathcal{X}(t)$ is shown in figure 5.

Figure 5. The state transition diagram of the Markov process $\mathcal{X}(t)$

Using the diagram, we write the infinitesimal generator $\mathbf{A}(a_{(s,n),(s',n')} : (s, n)(s', n') \in \mathcal{X})$ of Markov process $\mathcal{X}(t)$. Elements $a_{(s,n),(s',n')}$ are defined as follows:

$$a_{(s,n),(s',n')} = \begin{cases} \lambda, & s' = s, \ n' = n + 1 \text{ or } s' = s + 1, \ n' = n + 1 = H + 1, \\ n\mu, & s' = s, \ n' = n - 1 \text{ or } s' = s - 1, \ n' = n - 1 = 0, \\ \alpha, & s' = s - 1, \ n' = n \leqslant L, \\ *, & s' = s, \ n' = n, \\ 0, & \text{otherwise}, \end{cases}$$

where $* = -(\lambda \cdot \mathbf{1}\{n < N\} + n\mu + s\alpha \cdot \mathbf{1}\{n \leqslant L\})$.

The system probability distribution $p_{s,n}$ states $(s, n) \in \mathcal{X}$ is numerically calculated using the system of equilibrium equations

$$\begin{cases} \lambda p_{0,0} = \mu p_{0,1} + \mu p_{1,1}, \\ (\lambda + n\mu)p_{0,n} = n\mu p_{0,n+1} \cdot \mathbf{1}_{n<H} + \lambda p_{0,n-1} + \alpha p_{1,n} \cdot \mathbf{1}_{n\leqslant L}, & n = (1, H), \\ (\lambda + n\mu + \alpha)p_{1,n} = n\mu p_{1,n+1} \cdot \mathbf{1}_{n<N} + \\ \qquad\qquad\qquad + \lambda p_{1,n-1} \cdot \mathbf{1}_{n>1} + \lambda p_{0,H} \cdot \mathbf{1}_{n=H+1}, & n = (1, N), \\ \sum_{(s,n)\in X} p(s, n) = 1. \end{cases}$$

An important performance metric of the considered system is the time spent by users in the set of $\mathcal{X}_1$ states, which in the system model corresponds to the time that users watch video content from the MEC server until it is turned off. That is the lifetime of the service provider application after its migration to MEC. In the mathematical model, this is equal to the time interval from the moment when the Markov process $\mathcal{X}(t)$ reached the number of customers in the system H and passed into the set $\mathcal{X}_1$, i.e. into the state $(1, H + 1)$, until the moment when the process returned back to the set $\mathcal{X}_0$.

Let us denote τ_1 a random variable of the sojourn time of requests in the set $\mathcal{X}_1$. In order to find the cumulative distribution function (CDF) F_{τ_1} of the random variable τ_1, we can describe our systems by a truncated Markov process $\hat{\mathcal{X}}(t)$, which describes the behavior of the system at time $t > 0$ with the state space:

$$\hat{\mathcal{X}} = \mathcal{X}_0 + \hat{\mathcal{X}}_1^B,$$

where $\hat{\mathcal{X}}_1^B = \{(0, n) : n = 1, \ldots, L\}$.

The state transition diagram of the truncated Markov process $\hat{\mathcal{X}}(t)$ is shown in figure 6.
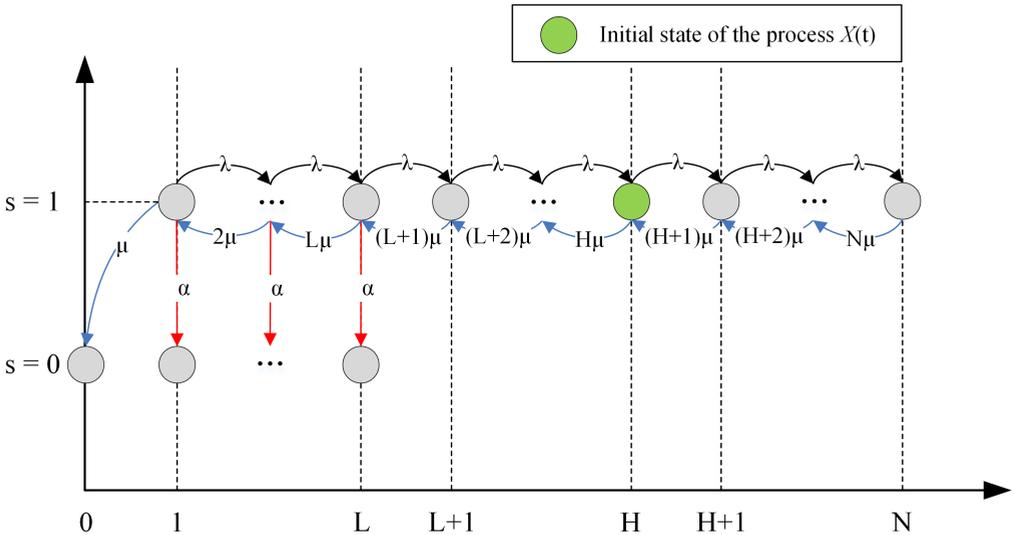
Figure 6. The state transition diagram of the truncated Markov process $\hat{\mathcal{X}}(t)$

The matrix $\hat{\mathbf{P}}(t)$ of transition probabilities can be written as follows:

$$\hat{\mathbf{P}}(t) = e^{\hat{\mathbf{A}}} = \sum_{n=0}^{\infty} \frac{\hat{\mathbf{A}} t^n}{n!}, \quad t \geqslant 0, \tag{1}$$

where $\hat{\mathbf{A}}$ is the infinitesimal operator of process $\hat{\mathcal{X}}(t)$. The distribution $\hat{\mathbf{p}}(t)$ of truncated process $\hat{\mathcal{X}}(t)$ satisfies the following equations:

$$\hat{\mathbf{p}}^T(t) = \hat{\mathbf{p}}^T(0)\hat{P}(t), \tag{2}$$

$$\frac{d}{dt}\hat{\mathbf{p}}^T(t) = \hat{\mathbf{p}}^T(0)\hat{\mathbf{A}}e^{\hat{\mathbf{A}}t}, \quad t \geqslant 0. \tag{3}$$

Initial probability vector $\hat{\mathbf{p}}^T(0)$:

$$\hat{\mathbf{P}}_{(s,n)}(0) = \begin{cases} 1, & (s,n) = (1, H+1), \\ 0, & \text{otherwise.} \end{cases} \tag{4}$$

Afterwards, we can find the cumulative distribution function $F_{\tau_1}(t)$ of the random variable τ_1 which equals

$$F_{\tau_1}(t) = \sum_{i=1}^{L} p_{(s,n)}(t), \quad t \geqslant 0. \tag{5}$$

The probability density function (PDF) of the random variable τ_1 is given by:

$$f_{\tau_1}(t) = \mu p_{(0,0)}(t) + \lambda \sum_{i=1}^{L} p_{(0,i)}(t). \tag{6}$$

We calculate the average time before the MEC shutdowns through the expectation of the random variable τ_1:

$$W_{\tau_1} = E(\tau_1) = \int_0^\infty t f_{\tau_1}(t)dt = \mu \int_0^\infty t p_{(0,0)}(t)dt + \lambda \sum_{i=1}^{L} \int_0^\infty t p_{(0,i)}(t)dt. \tag{7}$$

4. Numerical analysis

For numerical analysis, consider the network operator's service area, which is home to N users who are fans of the same series. New episodes of the series are released once a week and most users try to watch it as soon as it's possible, thereby creating a high load on the service provider's servers. In our scenario, every λ^{-1} minute there is a request to watch a video content. The duration of watching a video depends on the length and fascination of the episode, which takes μ^{-1} minute on average. The allowable load on the service provider's servers equals to H of user requests, when the limit is reached, the video content application is migrated to the nearest MEC server of the network operator. At some point, many users stop watching the series and the number of active users decreases. As soon as their number reaches the L threshold of active sessions, it becomes unreasonable to provide the service through the local MEC server and the disconnection process begins, which takes α^{-1} minutes on average.

For a more accurate estimate of the system, we introduce the parameter $\rho = \lambda/\mu$ describing the ratio of the rate of requests to watch a video to its average duration. This way, it is possible to determine the load created by users, which in our case correlates with the average number of active sessions.

We consider how the video viewing time through MEC changes at different values of the threshold for the start of its shutdown. The initial data is presented in the table 1.

Initial data

Notation	Value	Description
N	300	Maximum number of users watching videos
H	100	The max number of active user sessions on the service provider's server
L	85–100	MEC server disable threshold
μ^{-1}	45	Average video watch time, min
ρ	95, 100, 105, 110, 115	System load, number of active sessions
α^{-1}	15	average MEC disabling time, min

Figure 7 shows a plot of the average user service time through the MEC server depending on the threshold L value for the start of its disconnection at different loads (number of active sessions). It can be seen that as L increases, the MEC usage time decreases non-linearly, and the closer to the threshold H, the less significant the change. Periods of high load at which it is necessary to use MEC usually occur in the evening hours, therefore, it is proposed to choose the optimal time to use the MEC server in the range from 120 minutes (2 hours) to 300 minutes (5 hours), which is marked with a dotted line. With a value of 110 and 115 average number of active sessions, the disconnect threshold required to fall within the specified range is 97–100 users, which is quite close to threshold H. This shows the effectiveness of using MEC under such a load, but imposes increased costs on the service provider.

It is also important to evaluate the impact of the average duration of video viewing and the duration of the MEC server shutdown on the time the service is provided through MEC. To do this, consider an optimally loaded system with a load $\rho = 100$ and several threshold values $L = 85, 90, 95, 100$, at which the average service time through MEC falls within the interval in figure 7.

The figure 8 depicts MEC server runtime under different values of threshold L. The average video viewing time is between 45 minutes and 2 hours, which corresponds to typical episodes of a TV series or a full-length film. It can be seen that while maintaining the average number of active sessions, an increase in the duration of each session can significantly affect the time of using MEC, especially at sufficiently low values of the threshold L. This is due to the fact that the total audience capture time becomes longer, and the possibility of disconnecting several users is lower. Although when using $L = H$, the average MEC usage time does not change significantly, which indicates frequent service switching between the service provider's servers and the local MEC servers.

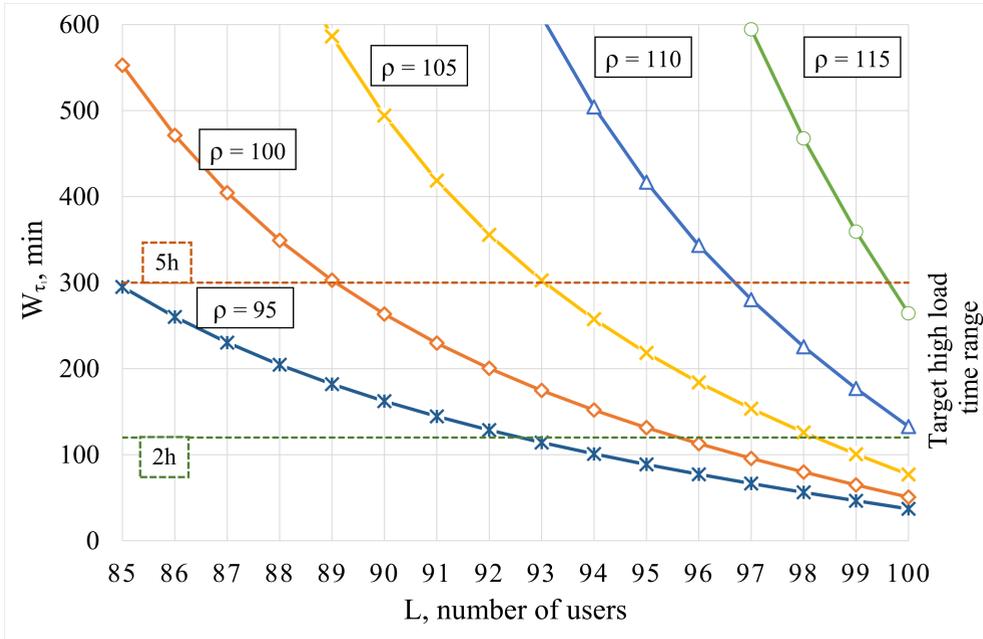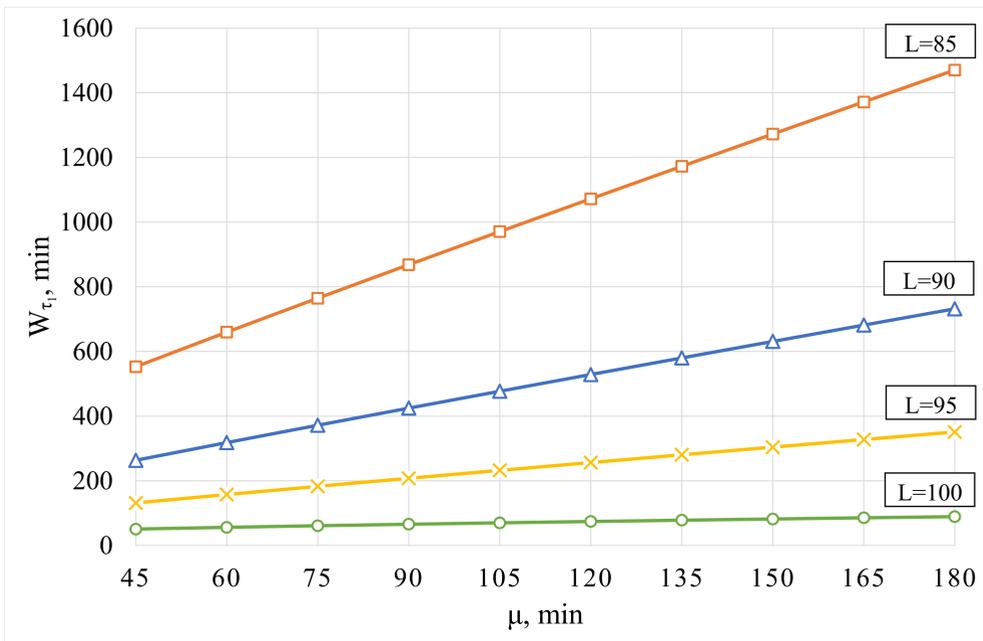Figure 7. MEC server runtime against the threshold disable value for different ρ

Figure 8. MEC server runtime against the average watch time for different L

Therefore, it is worth taking an assessment of the impact of the MEC disable delay on the duration of its use. In figure 9 this dependence is shown, with similar values of ρ and L for figure 8, average viewing time is 45 minutes

and average time required to disable MEC varies between 5–30 minutes. A similar behavior of the curves can be observed, which shows the expected increase in MEC usage time. It also allows, depending on the L threshold used, to select a more optimal time required to switch service back to the service provider's server.
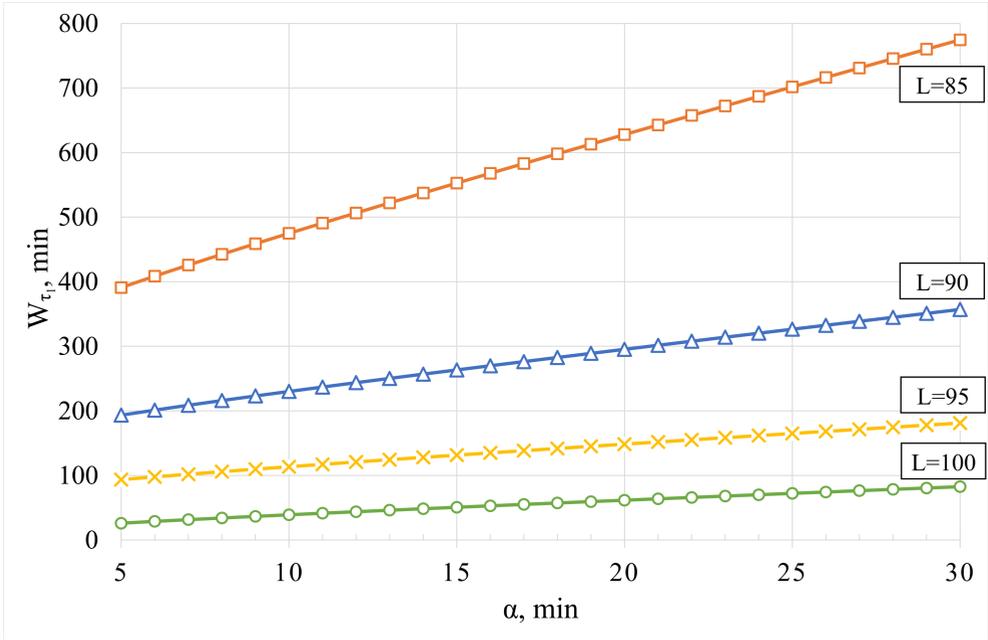

Figure 9. MEC server runtime against the MEC server disabling time for different L

5. Conclusions

The presented paper has investigated the scenario of providing services by a service provider of video content to users using the local MEC servers of the telecommunication operator under high service load. A mathematical model of the interaction for the case described in the scenario in the form of a Markov queuing system with hysteresis control using the MEC server has been developed. The resulting equation was derived for calculating the average time the MEC server is used to provide a video content depending on time. A numerical analysis of the scenario was carried out for one highly loaded zone of the telecommunication operator, in which users massively request to watch a video on an example of a popular TV series. It is shown how the changes in the MEC server disable initialization threshold, the duration of a video viewing and the duration of the MEC disabling have an effect on the average time of using local edge servers. This allows to set the optimization problem for various scenarios in the future.

Also, this work can be considered a continuation of [17], which allows, by placing the MEC server on the UAV, not only to reduce the load on the servers of the video content provider, but also to increase the QoS and QoE parameters for the mobile users.

Acknowledgments

References

[1]   A. Khan, M. Othman, S. A. Madani, and S. U. Khan, "A Survey of Mobile Cloud Computing Application Models," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 1, pp. 393–413, 2014. DOI: `10.1109/SURV.2013.062613.00160`.

[2]   L. Guan, X. Ke, M. Song, and J. Song, "A Survey of Research on Mobile Cloud Computing," in *2011 10th IEEE/ACIS International Conference on Computer and Information Science*, 2011, pp. 387–392. DOI: `10.1109/ICIS.2011.67`.

[3]   H. Dinh Thai, C. Lee, D. Niyato, and P. Wang, "A survey of mobile cloud computing: Architecture, applications, and approaches," *Wireless Communications and Mobile Computing*, vol. 13, Dec. 2013. DOI: `10.1002/wcm.1203`.

[4]   X. Fan, J. Cao, and H. Mao, *A survey of mobile cloud computing*, English, 2011.

[5]   N. Fernando, S. W. Loke, and W. Rahayu, "Mobile cloud computing: A survey," *Future Generation Computer Systems*, vol. 29, no. 1, pp. 84–106, 2013, Including Special section: AIRCC-NetCoM 2009 and Special section: Clouds and Service-Oriented Architectures. DOI: `10.1016/j.future.2012.05.023`.

[6]   M. Alizadeh, S. Abolfazli, M. Zamani, S. Baharun, and K. Sakurai, "Authentication in mobile cloud computing: A survey," *Journal of Network and Computer Applications*, vol. 61, pp. 59–80, 2016. DOI: `10.1016/j.jnca.2015.10.005`.

[7]   A. Ahmed and E. Ahmed, "A survey on mobile edge computing," in *10th International Conference on Intelligent Systems and Control (ISCO)*, 2016, pp. 1–8. DOI: `10.1109/ISCO.2016.7727082`.

[8]   M. Beck, M. Werner, S. Feld, and T. Schimper, "Mobile Edge Computing: A Taxonomy," in *The Sixth International Conference on Advances in Future Internet*, 2014, pp. 48–54.

[9]   R. Roman, J. Lopez, and M. Mambo, "Mobile edge computing, Fog et al.: A survey and analysis of security threats and challenges," *Future Generation Computer Systems*, vol. 78, part 2, pp. 680–698, 2018. DOI: `10.1016/j.future.2016.11.009`.

[10] S. Yi, C. Li, and Q. Li, "A Survey of Fog Computing: Concepts, Applications and Issues," in *Proceedings of the 2015 Workshop on Mobile Big Data*, ser. Mobidata '15, Hangzhou, China: Association for Computing Machinery, 2015, pp. 37–42. DOI: `10.1145/2757384.2757397`.

[11] S. Yi, Z. Qin, and Q. Li, "Security and Privacy Issues of Fog Computing: A Survey," Aug. 2015, pp. 685–695. DOI: `10.1007/978-3-319-21837-3_67`.

[12] W. Ali, S. M. Shamsuddin, and A. S. Ismail, "A Survey of Web Caching and Prefetching A Survey of Web Caching and Prefetching," *International Journal of Advances in Soft Computing and its Applications*, vol. 3, no. 1, 2011.

[13] S. Podlipnig and L. Böszörmenyi, "A Survey of Web Cache Replacement Strategies," *ACM Comput. Surv.*, vol. 35, no. 4, pp. 374–398, Dec. 2003. DOI: `10.1145/954339.954341`.

[14] Z. Lv and W. Xiu, "Interaction of Edge-Cloud Computing Based on SDN and NFV for Next Generation IoT," *IEEE Internet of Things Journal*, vol. 7, no. 7, pp. 5706–5712, 2020. DOI: `10.1109/JIOT.2019.2942719`.

[15] Z. Ning, P. Dong, X. Wang, X. Hu, L. Guo, B. Hu, Y. Guo, T. Qiu, and R. Y. K. Kwok, "Mobile Edge Computing Enabled 5G Health Monitoring for Internet of Medical Things: A Decentralized Game Theoretic Approach," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 2, pp. 463–478, 2021. DOI: `10.1109/JSAC.2020.3020645`.

[16] T. Stockhammer, "Dynamic Adaptive Streaming over HTTP: Standards and Design Principles," in *Proceedings of the Second Annual ACM Conference on Multimedia Systems*, ser. MMSys '11, San Jose, CA, USA: Association for Computing Machinery, 2011, pp. 133–144. DOI: `10.1145/1943552.1943572`.

[17] A. Khakimov, E. Mokrov, D. Poluektov, K. Samouylov, and A. Koucheryavy, "Evaluating the Quality of Experience Performance Metric for UAV-Based Networks," *Sensors*, vol. 21, no. 17, 2021. DOI: `10.3390/s21175689`.

Information about the authors:

Poluektov, Dmitry S. — postgraduate student of Department of Applied Probability and Informatics of Peoples' Friendship University of Russia (RUDN University) (e-mail: `poluektov-ds@rudn.ru`, phone: +7(495)9522823, ORCID: https://orcid.org/0000-0002-4246-8483)

Khakimov, Abdukodir A. — Researcher of Department of Applied Probability and Informatics of Peoples' Friendship University of Russia (RUDN University) (e-mail: `khakimov-aa@rudn.ru`, phone: +7(495)9522823, ORCID: https://orcid.org/0000-0003-2362-3270)

Разработка и анализ моделей гистерезисного управления миграцией сервисов на сервер граничных вычислений

Д. С. Полуэктов, А. А. Хакимов

*Российский университет дружбы народов,
ул. Миклухо-Маклая, д. 6, Москва, 117198, Россия*

Аннотация. Сервисы онлайн-видео являются одними из самых популярных способов потребления контента. На серверы видео хостинга приходится ежедневно колоссальная нагрузка, которую нами предложено снизить за счёт миграции приложения востребованным видеоконтентом на локальный сервер MEC целевой зоны. Это позволит повысить качество предоставляемых услуг пользователям за счёт сокращения задержки на передачу. Поэтому предложена архитектура, дающая возможность в моменты повышенного спроса на одинаковый видеоконтент производить миграцию приложения видеосервиса на граничные серверы оператора сети. Для оценки показателей эффективности такого подхода была построена математическая модель в виде системы массового обслуживания. Результаты численного эксперимента позволяют произвести оптимизацию времени использования локальных серверов MEC для предоставления видеоконтента.

Ключевые слова: миграция сервисов, граничные вычисления, марковский процесс, MEC, онлайн-видео, усечённый марковский процесс

Detection of cyber-attacks on the power smart grids using semi-supervised deep learning models

Eugeny Yu. Shchetinin[1], Tatyana R. Velieva[2]

[1] *Financial University under the Government of Russian Federation,*
49, Leningradsky Prospect, Moscow, 125993, Russian Federation
[2] *Peoples' Friendship University of Russia (RUDN University),*
6, Miklukho-Maklaya St., Moscow, 117198, Russian Federation

Abstract. Modern smart energy grids combine advanced information and communication technologies into traditional energy systems for a more efficient and sustainable supply of electricity, which creates vulnerabilities in their security systems that can be used by attackers to conduct cyber-attacks that cause serious consequences, such as massive power outages and infrastructure damage. Existing machine learning methods for detecting cyber-attacks in intelligent energy networks mainly use classical classification algorithms, which require data markup, which is sometimes difficult, if not impossible. This article presents a new method for detecting cyber-attacks in intelligent energy networks based on weak machine learning methods for detecting anomalies. Semi-supervised anomaly detection uses only instances of normal events to train detection models, which makes it suitable for searching for unknown attack events. A number of popular methods for detecting anomalies with semi-supervised algorithms were investigated in study using publicly available data sets on cyber-attacks on power systems to determine the most effective ones. A performance comparison with popular controlled algorithms shows that semi-controlled algorithms are more capable of detecting attack events than controlled algorithms. Our results also show that the performance of semi-supervised anomaly detection algorithms can be further improved by enhancing deep autoencoder model.

Key words and phrases: smart energy grids, cyber-attacks, semi-supervised anomaly detection, deep learning, autoencoder

1. Introduction

There are many problems in traditional power grids, such as the lack of automated analysis and situational awareness, poor visibility and slow response time, which makes them unable to meet the significantly increased demand and consumption of electricity in the 21st century [1]. With the help of modern information and communication technologies, intelligent

networks provide a bidirectional flow of electricity and information, which ensures a more efficient and stable supply of electricity and better demand management [2, 3]. The intelligent energy network consists of four main components: generation, transmission, distribution and consumption, which are connected through a three-level hierarchical structured communication network [4] (see figure 1). The first level of the communication network is the home network, which is responsible for communication at the consumption stage to connect smart devices in consumers' homes to the smart grid with smart meters for more efficient energy management and demand response. The second level of the communication network, the district network, is responsible for communication at the distribution stage, which collects data from smart meters and sends back control commands for advanced accounting applications.
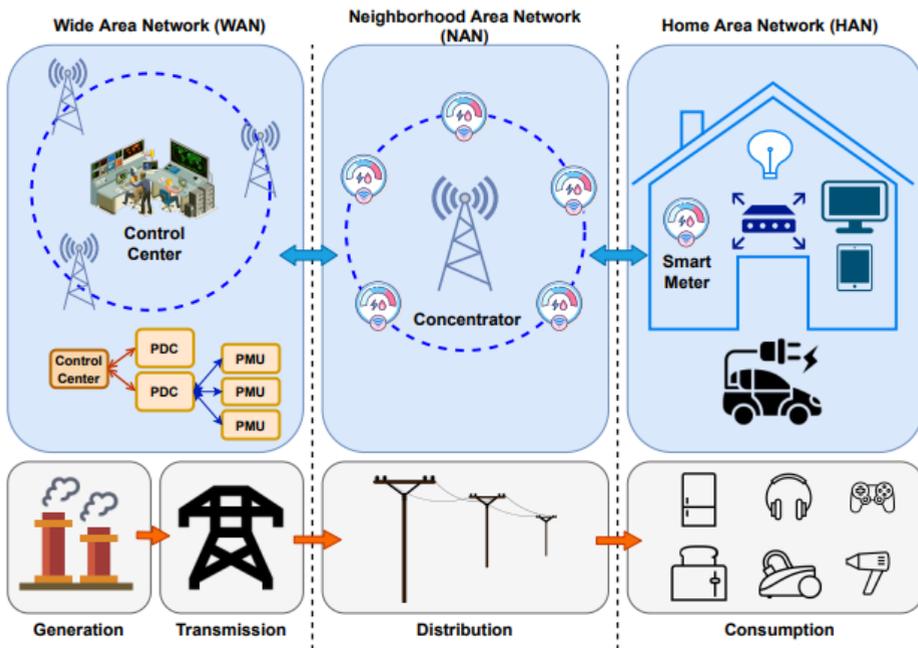

Figure 1. Diagram of smart grid energy consumption system

At the last level, the global network connects with utility management centers, forming the basis of an intelligent network for the communication needs of the stages of electricity production and transmission. Although the integration of advanced ICTs into traditional power grids brings significant benefits for the delivery and management of electricity, it also creates new vulnerabilities in security systems [5]. Cyber-attacks can target any of the four components of a smart grid — from smart home gateways in HAN to control rooms in the global network [6].

In this work, we used measurements from the Power Measurement Units (PMU) to detect cyber-attacks. PMU is a sensor device deployed at the global network level of the smart grid network, which provides real-time measurements of the state of the power system for a wide range of monitoring, protection and control. In the Global Monitoring System (GMS), several

PMUs are connected to a phasor data hub. The GMS central authority then collects information from the PDCS. PMU measurements combine both physical and cyber domains, making them a suitable choice for detecting cyber-attacks targeting the physical domain of an intelligent network, such as False Data Entry (FDE) attacks and malicious shutdown attacks.

In most widely used models have been built to detect cyber-attacks in intelligent networks using controlled learning algorithms. To train supervised algorithms, both normal and attacking data are required. However, collecting representative instances of various attack events is usually a difficult task, if not impossible, which can lead to poor model performance when detecting certain attacks, especially types of attacks not represented in the training data.

In this article, we proposed a method for detecting cyber-attacks in power smart grids with semi-supervised anomaly detection. Unlike supervised models, semi-supervised anomaly detection algorithms use only data from normal events to train a detection model that is capable of detecting unknown types of attacks. We have investigated a number of anomaly detection algorithms and identified the most effective ones for detecting cyber-attacks in smart energy grids. The performance of semi-supervised algorithms was compared with the characteristics of popular supervised algorithms to show their superiority in finding attack events. We have also supplemented semi-supervised anomaly detection with deep learning to extract features to further improve attack detection performance.

2. Related work

Traditional approaches use PMU measurements to assess the state of the power system and compare the difference between the observed and estimated measurements with a threshold for detecting cyber-attacks. A lightweight scheme was proposed in the paper [4], which explores the spatial-temporal correlations between network state estimates and applies confidence voting to detect abnormal state estimates in intelligent networks caused by real-time FDI attacks.

Recently, machine learning has been widely used to detect cyber-attacks in smart grids, where most of the proposed approaches are based on supervised learning algorithms. In the paper [7] a number of supervised learning algorithms were investigated for recognizing violations in the power system and cyber-attacks. A One-Class SVM (OC-SVM) was used in [8] to create an intrusion detection module for detecting malicious attacks in a dispatch control system and data collection system using network traces. The paper [9] applied several popular supervised algorithms, including perceptron, k-nearest neighbor (KNN), support vector machines (SVM) and logistic regression (LR) with ensemble training and combining functions at the object level to predict FDI attacks. Their experimental results demonstrate that machine learning algorithms are superior to state-based algorithms. Singh [10] compared SVM, KNN models for detecting both direct and covert attacks in intelligent networks. Compared to number of methods based on supervised learning algorithms, only a limited number of studies have been conducted on the use of unsupervised anomaly detection algorithms to detect attacks in intelligent networks.

The data sets studied in our study were created based on the structure of the power system, consisting of intelligent electronic devices, dispatch control systems and network monitoring devices. The power system framework can simulate several operational scenarios to generate data corresponding to three types of events: absence of events, natural events, and attack events. The six types of events are described as follows:

— No event: Normal readings.
— Short circuit: There has been a single ground fault, which can be determined by reading the percentage range in the data.
— Line maintenance: Operators switch one or more IEDs to perform maintenance on certain parts of the power system and its components.
— Remote Shutdown Command Implementation Attack: Attackers can send commands that switch Improvised Explosive devices to switch switches when they can get into the system.
— Attack with changing relay settings: Attackers change the settings, for example, disable the main functions of the settings, as a result of which the IEDs do not switch the switches whenever an acceptable error or command occurs.
— Data Intrusion attack: Attackers modify PMU measurements such as voltage, current, and sequence components to simulate a real malfunction resulting in the disconnection of switches.

The system has four PMUs integrated with relays, where each PMI measures 29 features. A total of 116 functions were obtained from four PMUs. Depending on how to group the scenarios, three groups of datasets were created based on the generated data — binary class data and multi-class data from the framework. Since the purpose of our study was to distinguish attack events from other types of events, we adopted a binary group of datasets in which no events and normal events are treated as ordinary events.

3. Methodology

The proposed method contains two main components: deep representation learning and semi-supervised anomaly detection [11–15]. The first step of the proposed method is to prepare a training dataset that contains only examples of normal events. The dimension of the object space is then reduced by deep representation learning, when a low-dimensional hidden representation is extracted from the input data using a deep autoencoder. Finally, a semi-supervised anomaly detection algorithm using the representation of the studied features is used to train the detection model. At the detection stage, a hidden representation is first created from an unknown input instance by the deep autoencoder, which is then fed into the trained detection model to classify the instance as a normal event or an attack event.

In our study, deep autoencoder is used to extract features, which used to learn robust low-dimensional representations from multidimensional input data. PCA, a popular feature extraction method was used as a method for comparison. After training the autoencoder with the training dataset, the encoder and code layer are retained for feature extraction, while the decoder has been removed from the network. The hidden code-level representation will be used as input for a semi-supervised anomaly detection algorithm.

4.   Performance evaluation and results

The binary group of data sets on attacks on power systems adopted in our study contains 15 data sets covering 37 scenarios [16–18]. We used min–max normalization to normalize the data. The parameters used by the algorithms investigated in our study are listed in the table 1. The characteristics of semi-supervised anomaly detection algorithms using all 116 PMU functions were investigated first. Among the usual instances of the dataset, 50% were randomly selected to train the detection algorithm. The remaining 50% of normal instances and all attack instances were then used for testing. Examples of ROC curves obtained using eight algorithms on datasets 1 and 11 at one stage of the experiment are shown in the figure 2. The average AUC values of the algorithms calculated from the results of 10 runs for each of the 15 data sets are shown in the figure 2.

Table 1

Parameters used by the semi-supervised and supervised algorithms investigated in our study

Model	Parameters
OCSVM	RBF kernel, degree=3
LOF	K=25
KNN	K=10
IForest	iTrees=100
SVM	RBF kernel, degree=3, C=10
Deep Autoencoder	Batch_size=8

Figure 2 shows that three most effective algorithms in terms of average AUC are OCSVM, KNN and IForest. These algorithms show significantly better performance than the other four algorithms. Then we used the distance to the angle d to determine the detection threshold of the algorithm to obtain accuracy. The three best algorithms in terms of F1 average score are OCSVM, KNN and iForest.

In our study, two metrics were used to evaluate the effectiveness: the area under the ROC curve (AUC) and the F1 score. The ROC curve shows the relationship between the true positive rate (TPR) and the false positive rate (FPR) by changing the detection threshold. Equations (1) and (2) define TPR and FPR, where TP, TN, FP, and FN are true positive, true negative, false positive and false negative, respectively:

$$R = \frac{TP}{(TP + FN)}, \tag{1}$$
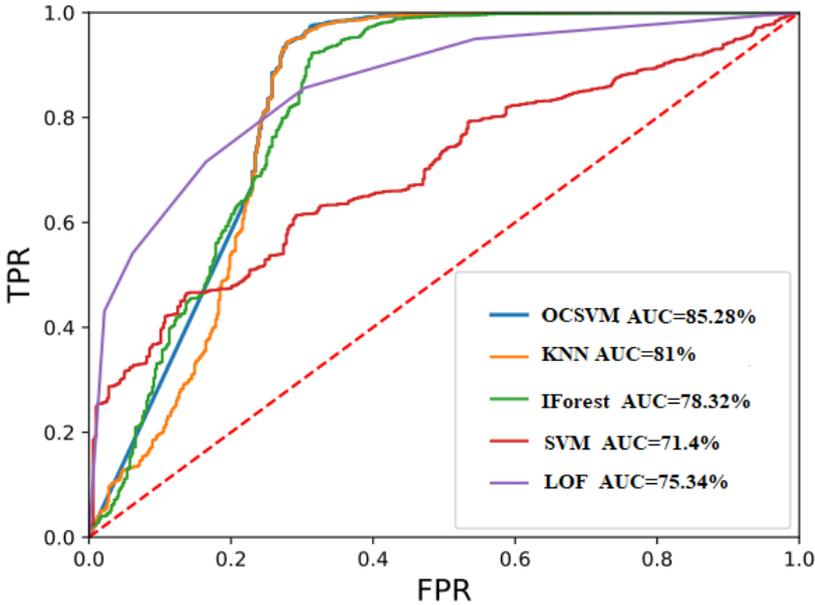
$$R = \frac{FP}{(FP + TN)}. \tag{2}$$

Figure 2. Examples of ROC curves obtained by semi-supervised and supervised algorithms

AUC measures the area under the ROC curve to indicate the performance of the model on distinguishing normal and attack events. A higher AUC value means that model has a better capability to distinguish normal and attack events. F1 score is defined as the harmonic mean of the precision and recall:

$$\text{Precision} = \frac{\text{TP}}{(\text{TP} + \text{FP})}, \tag{3}$$

$$\text{Recall} = \frac{\text{TP}}{(\text{TP} + \text{FN})}, \tag{4}$$

$$\text{F1} - \text{score} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{(\text{Precision} + \text{Recall})}. \tag{5}$$

We adopted the distance to corner (D) as the criterion, which determines the optimal threshold as the point on the ROC curve closest to the corner point $(0, 1)$:

$$D = \sqrt{(1 - \text{TPR})^2 + \text{FPR}^2}. \tag{6}$$

Then we compared three most effective semi-supervised algorithms in terms of AUC (OCSVM, LOF, IForest) with two popular controlled algorithms — KNN and SVM, which were used to detect cyber-attacks in power smart grids. For each of the 15 datasets, a training dataset for semi-supervised algorithms was generated by randomly selecting 50% of the regular instances. These regular instances were combined with the same number of randomly selected attack instances to form a training dataset for controlled algorithms. The

remaining 50% of normal instances and attack instances were used to form a test dataset for both semi-supervised and supervised algorithms.

Examples of ROC curves are shown in figures 2, which were obtained using five unsupervised and supervised algorithms. Examples of F1-scores are shown in figures 3, which were obtained using five unsupervised and supervised algorithms. Among all the algorithms, SVM had the worst performance, while the KNN algorithm has significantly better average AUC than other algorithms. The good performance of the controlled KNN algorithm in terms of AUC is due to its significantly better TPR compared to the three semi-supervised algorithms when the FPR is low. On the other hand, one of figure 2 also shows that as the FPR increases, the three semi-controlled algorithms can achieve a high TPR much faster than the controlled KNN algorithm.
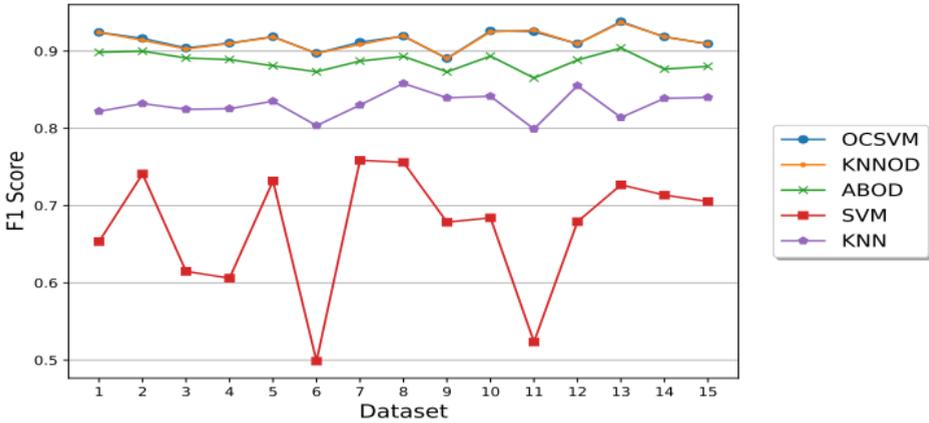

Figure 3. Performance comparison of the semi-supervised algorithms with supervised algorithms in terms of average precision, recall, and F1 score

5.   Performance improvement using a deep autoencoder

Finally, the impact of DAE-based deep representation training on the performance of semi-supervised anomaly detection algorithms for detecting cyber-attacks in power smart grids was investigated. The PCA method has been adopted as a reference method for comparison. We set the number of extracted objects to 30 for both DAE and PCA. The input and hidden encoder layers in the DAE have 116 and 60 nodes, respectively. The three most efficient semi-supervised algorithms OCSVM, KNN and IForest in terms of AUC, F1-score accuracy metrics were included in this study. The results in terms of the average AUC are shown in figure 4, which were obtained by averaging the results of all runs of all 15 datasets. Figures 4, 5 show that DAE can further improve the performance of three semi-controlled algorithms in terms of both performance indicators. Statistical tests (paired t-test with two samples, $\alpha = 0.05$) show that the AUC obtained using three semi-controlled algorithms with deep learning representation is significantly higher than when using all functions. The F1 scores obtained by OCSVM and KNN using deep representation learning are also significantly higher than when using all functions.
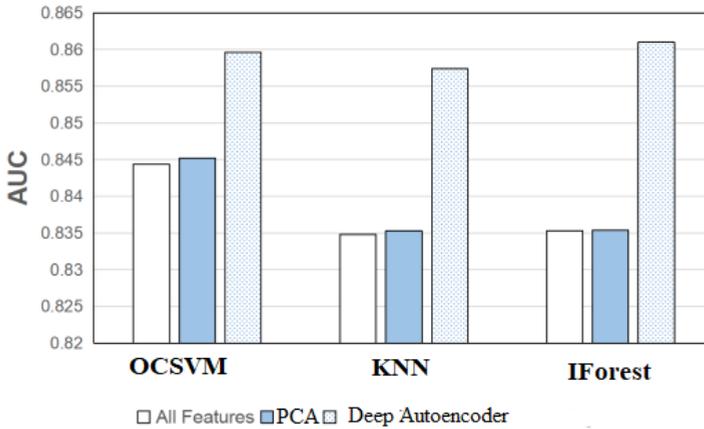
Figure 4. Performance of semi-supervised algorithms with and without feature extraction in terms of average AUC
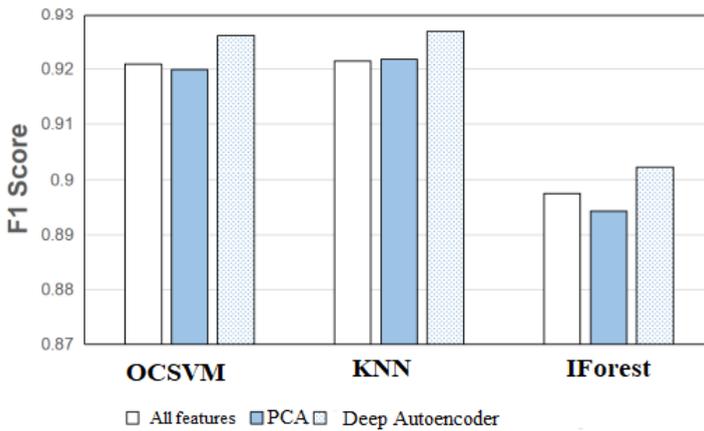

Figure 5. Performance of semi-supervised algorithms with and without feature extraction in terms of average F1 score

6.   Conclusion

Smart grids ensure efficient supply of electricity to various facilities and its management through the introduction of advanced digital technologies into traditional power grids. On the other hand, the vulnerabilities that have appeared in their security can be used to carry out cyber-attacks that lead to devastating damage. Using PMU measurements that connect the physical and cybernetic domains, the article develops a method based on semi-controlled anomaly detection and deep learning to detect cyber-attacks in smart energy grids. Unlike supervised algorithms, semi-supervised anomaly detection algorithms use only instances of normal events to train detection models, which makes them capable of detecting events of unknown types of attacks.

In our experiments, the most effective semi-supervised algorithms were identified using publicly available datasets on attacks on intelligent energy systems. A comparison of performance with popular controlled algorithms has shown that semi-supervised algorithms have a better ability to detect cyber-attacks. In addition, our results showed that the detection performance of semi-supervised algorithms can be further enhanced by deep representation training based on DAE.

References

[1] G. Dileep, "A survey on Smart Grid technologies and applications," *Renewable Energy*, vol. 146, pp. 2589–2625, 2020. DOI: 10.1016/j.renene.2019.08.092.

[2] V. C. Gungor, D. Sahin, T. Kocak, S. Ergut, C. Buccella, C. Cecati, and G. P. Hancke, "Smart Grid technologies: communication technologies and standards," *IEEE Transactions on Industrial Informatics*, vol. 7, no. 4, pp. 529–539, 2011. DOI: 10.1109/TII.2011.2166794.

[3] T. Flick and J. Morehouse, *Securing the Smart Grid: Next Generation Power Grid Security*. Syngress, 2010.

[4] S. Aftergood, "Cybersecurity: the cold war online," *Nature*, vol. 547, no. 7661, pp. 30–31, Jul. 2017. DOI: 10.1038/547030a.

[5] C. Chio and D. Freeman, *Machine learning and security: protecting systems with data and algorithms*. O'Reilly Media, 2018.

[6] D. S. Berman, A. L. Buczak, J. S. Chavis, and C. L. Corbett, "A survey of deep learning methods for cyber security," *Information*, vol. 10, no. 4, 2019. DOI: 10.3390/info10040122.

[7] D. Wang, X. Wang, Y. Zhang, and L. Jin, "Detection of power grid disturbances and cyber-attacks based on machine learning," *Journal of Information Security and Applications*, vol. 46, pp. 42–52, 2019. DOI: 10.1016/j.jisa.2019.02.008.

[8] S. Ahmed, Y.-D. Lee, S.-H. Hyun, and I. Koo, "Unsupervised machine learning-based detection of covert data integrity assault in Smart Grid networks utilizing isolation forest," *IEEE Transactions on Information Forensics and Security*, vol. 14, pp. 2765–2777, 2019.

[9] M. Ozay *et al.*, "Machine learning methods for attack detection in the Smart Grid," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 27, pp. 1773–1786, 2016.

[10] V. K. Singh and M. Govindarasu, "Decision tree based anomaly detection for remedial action scheme in Smart Grid using PMU data," in *IEEE Power & Energy Society General Meeting PESGM*, 2018, pp. 1–5. DOI: 10.1109/PESGM.2018.8586159.

[11] G. Pang, C. Shen, L. Cao, and A. V. D. Hengel, "Deep learning for anomaly detection: a review," *ACM Comput. Surv.*, vol. 54, no. 2, 2021. DOI: 10.1145/3439950.

[12] Z. E. Huma, S. Latif, J. Ahmad, Z. Idrees, A. Ibrar, Z. Zou, F. Alqahtani, and F. A. Baothman, "A hybrid deep random neural network for cyberattack detection in the Industrial Internet of Things," *IEEE Access*, vol. 9, pp. 55 595–55 605, 2021. DOI: `10.1109/ACCESS.2021.3071766`.

[13] M. S. Minhas and J. Zelek, "Semi-supervised anomaly detection using autoencoders," *Journal of Computational Vision and Imaging Systems*, vol. 5, no. 1, p. 3, 2019.

[14] M. Wieler. "Weakly supervised learning for industrial optical inspection." (2007), [Online]. Available: `https://hci.iwr.uni-heidelberg.de/node/3616`.

[15] R. Qi, C. Rasband, J. Zheng, and R. Longoria, "Semi-supervised outlier detection and deep feature extraction for detecting cyber-attacks in Smart Grids using PMU data," *Advances in Intelligent Systems and Computing*, vol. 1134, pp. 509–515, 2020. DOI: `10.1007/978-3-030-43020-7_67`.

[16] E. Y. Shchetinin, "On methods of quantitative analysis of the company's financial indicators under conditions of high risk of investments," *Discrete and Continuous Models and Applied Computational Science*, vol. 28, no. 4, pp. 346–360, 2020. DOI: `10.22363/2658-4670-2020-28-4-346-360`.

[17] E. Y. Shchetinin, "Modeling the energy consumption of smart buildings using artificial intelligence," in *CEUR Workshop Proceedings*, vol. 2407, 2019, pp. 130–140.

[18] E. Y. Shchetinin, "Development of Energy Saving Technologies for Smart Buildings by Using Computer Algebra," *Programming and Computer Software*, vol. 46, pp. 324–329, 2020. DOI: `10.1134/S0361768820050084`.

Information about the authors:

Shchetinin, Eugeny Yu. — Doctor of Physical and Mathematical Sciences, Lecturer of Department of Mathematics, Financial University under the Government of Russian Federation (e-mail: `riviera-molto@mail.ru`, ORCID: https://orcid.org/0000-0003-3651-7629)

Velieva, Tatyana R. — Candidate of Sciences in Physics and Mathematics, Senior lecturer of Department of Applied Probability and Informatics of Peoples' Friendship University of Russia (RUDN University) (e-mail: `velieva-tr@rudn.ru`, phone: +7(495)9520250, ORCID: https://orcid.org/0000-0003-4466-8531, ResearcherID: Q-6304-2016, Scopus Author ID: 56695390200)

Обнаружение кибератак на интеллектуальные энергосистемы с использованием неконтролируемых моделей глубокого обучения

Е. Ю. Щетинин[1], Т. Р. Велиева[2]

[1] *Финансовый университет при Правительстве Российской Федерации,*
Ленинградский проспект, д. 49, Москва, 125993, Россия
[2] *Российский университет дружбы народов,*
ул. Миклухо-Маклая, д. 6, Москва, 117198, Россия

Аннотация. Современные интеллектуальные энергосети объединяют передовые информационные и коммуникационные технологии в традиционные энергосистемы для более эффективного и устойчивого снабжения электроэнергией, что создаёт уязвимости в их системах безопасности, которые могут быть использованы злоумышленниками для проведения кибератак, вызывающих серьезные последствия, такие как массовые перебои в подаче электроэнергии и повреждение инфраструктуры. Существующие методы машинного обучения для обнаружения кибератак в интеллектуальных энергетических сетях в основном используют классические алгоритмы классификации, которые требуют разметки данных, что иногда сложно, а то и невозможно. В данной статье представлен новый метод обнаружения кибератак в интеллектуальных энергетических сетях, основанный на слабых методах машинного обучения для обнаружения аномалий. Полуконтролируемое обнаружение аномалий использует только экземпляры обычных событий для обучения моделей обнаружения, что делает его подходящим для поиска неизвестных событий атак. В ходе исследования был проанализирован ряд популярных методов обнаружения аномалий с полууправляемыми алгоритмами с использованием общедоступных наборов данных о кибератаках на энергосистемы для определения наиболее эффективных из них. Сравнение производительности с популярными управляемыми алгоритмами показывает, что полууправляемые алгоритмы лучше способны обнаруживать события атак, чем управляемые алгоритмы. Наши результаты также показывают, что производительность полуконтролируемых алгоритмов обнаружения аномалий может быть дополнительно улучшена за счёт усовершенствования модели глубокого автоэнкодера.

Ключевые слова: интеллектуальные энергетические сети, кибератаки, частично контролируемое обнаружение аномалий, глубокое обучение, автоэнкодер