

МАТЕМАТИЧЕСКОЕ И КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ ДИНАМИЧЕСКИХ СИСТЕМ И ПРОЦЕССОВ

УДК 517.927+681.5

*А. Керимбеков, akl7@rambler.ru**Кыргызско-Российский Славянский университет имени Б.Н.Ельцина*

О НЕКОТОРЫХ СВОЙСТВАХ УРАВНЕНИЯ РИККАТИ

В статье исследованы вопросы построения частных решений уравнения Риккати. Приведены ранее неизвестные свойства уравнения Риккати и разработан алгоритм построения частного решения.

Ключевые слова: уравнение Риккати, интегральное тождество, коэффициентно-сопряженные уравнения, дробно-линейное преобразование, определитель Якоби, алгоритм построения частного решения.

Введение

Уравнение Риккати с момента его появления (1724 г) привлекает внимание математиков тем, что, несмотря на обилие исследований (см. библиографию [1]), до настоящего времени не разработана методика построения его частного решения.

С появлением теории оптимального уравнения, т.е. начиная с 60-х годов 20-го столетия, интерес к уравнению Риккати еще более возрос, ибо вопросы разрешимости многих задач оптимизации, особенно при решении задачи синтеза оптимального уравнения, сводятся к разрешимости уравнения Риккати.

С другой стороны, уравнение Риккати тесно связано с линейными дифференциальными уравнениями 2-го порядка и линейными системами второго порядка дифференциальных уравнений с переменными коэффициентами, для которых также не разработаны методы построения их частных решений.

В данной статье приведены результаты исследований уравнения Риккати, в частности, обнаружен класс уравнений Риккати, коэффициенты которых обладают специфическими свойствами и разработан алгоритм построения их частных решений. Далее рассмотрены вопросы о приведении (преобразовании) произвольного уравнения Риккати к уравнению этого класса.

1. Коэффициентно-сопряженные (КС) уравнения Риккати

Рассмотрим уравнение Риккати вида

$$z'(x) = k_0(x) + k_1(x)z(x) + k_2(x)z^2(x), \quad (1)$$

где коэффициенты $k_0(x)$, $k_1(x)$ и $k_2(x)$ — непрерывные функции.

Теорема 1 (свойство 1). Пусть коэффициенты уравнения (1) удовлетворяют интегральному тождеству

$$\int k_0(x)e^{-\int k_0(x)dx}dx * \int k_2(x)e^{-\int k_1(x)dx}dx = 2, \quad (2)$$

и $y_0(x)$ является частным решением линейного неоднородного дифференциального уравнения

$$y'_0(x) - k_1(x)y_0(x) = k_0(x), \quad (3)$$

полученного при нулевом значении произвольной постоянной. Тогда функция

$$z(x) = -y_0(x) \quad (4)$$

является частным решением уравнения (1).

Доказательство утверждения теоремы несложное и может быть доказано несколькими способами, в частности по методике, изложенной в [2].

Утверждения теоремы мы проиллюстрируем на примере простейшего уравнения

$$z'(x) = e^{2x} + z(x) - 2e^{-2x}z^2(x), \quad (5)$$

где $k_0(x) = e^{2x}$, $k_1(x) = 1$, $k_2(x) = -2e^{-2x}$.

Поскольку

$$\int k_0(x)e^{-\int k_1(x)dx}dx = \int e^{2x} * e^{-x}dx = e^x$$

и

$$\int k_2(x)e^{\int k_1(x)dx}dx = -2 \int e^{-2x} * e^x dx = 2e^{-x},$$

то интегральное соотношение (2) имеет место. Теперь находим функцию $y_0(x)$ как решение уравнения

$$y'_0(x) - y_0(x) = e^{2x},$$

т.е. в виде

$$y_0(x) = e^x(c + y^x)|_{c=0} = e^{2x}.$$

Далее образуем функцию

$$z(x) = -y_0(x) = -e^{2x}. \quad (6)$$

и непосредственной проверкой убеждаемся, что она является частным решением уравнения Риккати

$$z'(x) = e^{2x} + z(x) - 2e^{-2x}z^2(x). \quad (7)$$

Заметим, что коэффициенты уравнения

$$z'(x) = k_2(x) - k_1(x)z(x) + k_0(x)z^2(x) \quad (8)$$

также удовлетворяют интегральному тождеству (2). Поэтому и для уравнения (8) справедлив аналогичный результат (свойство 2), т.е. если $y_0(x)$ является частным решением линейного неоднородного дифференциального уравнения

$$y'_0(x) + k_1(x)y_0(x) = k_2(x),$$

то функция

$$z(x) = -y_0(x)$$

является частным решением уравнения (8).

На примере уравнения (5) легко проверить, что функция $y_0(x) = 2e^{-2x}$, является частным решением линейного неоднородного дифференциального уравнения

$$y'_0(x) + y_0(x) = -2e^{-2x},$$

а функция

$$z(x) = -2e^{-2x}$$

частным решением уравнения Риккати вида

$$z'(x) = -2e^{-2x} - z(x) + e^{2x}z^2(x).$$

Определение 1. Уравнения (1) и (8), коэффициенты которых удовлетворяют интегральному тождеству (2), назовем коэффициентно-сопряженными (КС) уравнениями Риккати.

Множество коэффициентно-сопряженных уравнений образует класс уравнений Риккати, для каждого из которых удастся найти частное решение.

Класс коэффициентно-сопряженных (КС) уравнений Риккати является основным звеном при определении частного решения произвольного уравнения Риккати.

2. О преобразовании уравнения Риккати

Рассмотрим вопрос о преобразовании уравнения Риккати с произвольными коэффициентами в уравнение класса коэффициентно-сопряженных (КС) уравнений Риккати.

Рассмотрим уравнение Риккати

$$y'(x) = p(x) + q(x)y + z(x)y^2, \quad (9)$$

где коэффициенты $p(x)$, $q(x)$, $z(x)$ – произвольные непрерывные функции.

Общеизвестно, что уравнение (9) при дробно-линейной замене искомой функции $y(x)$ на новую искомую функцию $z(x)$ по формуле

$$y(x) = \frac{a(x)z(x) + b(x)}{(c(x)z(x) + h(x))}, \quad (10)$$

где $a(x)$, $b(x)$, $c(x)$, $h(x)$ – непрерывные и непрерывно-дифференцируемые функции, удовлетворяющие условию

$$Q(x) = a(x)h(x) - c(x)b(x) \neq 0, \quad (11)$$

может быть преобразовано в уравнение Риккати вида

$$z'(x) = k_0(x) + k_1(x)z(x) + k_2(x)z^2(x), \quad (12)$$

где

$$k_0(x) = \frac{1}{Q(x)} [p(x)h^2(x) + q(x)b(x)h(x) + r(x)b^2(x) - (b'(x)h(x) - b(x)h'(x))] =$$

$$\varphi_1[x, a(x), b(x), c(x), h(x), p(x), q(x), r(x)]; \quad (13)$$

$$\begin{aligned}
 k_1(x) &= \frac{1}{Q(x)} [2p(x)c(x)h(x) + q(x)c(x)b(x) + a(x)h(x) + 2r(x)a(x)b(x) \\
 &\quad - (a'(x)h(x) - a(x)h'(x) + b(x)c(x) - b(x)c'(x))] = \\
 &= \varphi_2[x, a(x), b(x), c(x), h(x), p(x), q(x), r(x),]; \quad (14)
 \end{aligned}$$

$$\begin{aligned}
 k_2(x) &= \frac{1}{Q(x)} [p(x)c^2(x) + q(x)a(x)c(x) + r(x)a^2(x) - (a'(x)c(x) - a(x)c'(x))] = \\
 &= \varphi_3[x, a(x), b(x), c(x), h(x), p(x), q(x), r(x)]. \quad (15)
 \end{aligned}$$

Лемма 1. Пусть (12) является уравнением класса (КС). Тогда при любых заданных коэффициентах $k_0(x), k_1(x), k_2(x)$ существует хотя бы одно уравнение Риккати с коэффициентами $p(x), q(x), r(x)$, которые при дробно-линейном преобразовании (10) переходят в уравнение (12).

Доказательство. При дробно-линейном преобразовании (10) коэффициенты уравнений (9) и (12) будут связаны соотношениями (13)-(15). Для доказательства утверждения теоремы достаточно проверить разрешимости системы (13)-(15) относительно коэффициентов $p(x), q(x), r(x)$ при заданных коэффициентах $k_0(x), k_1(x), k_2(x)$. Поскольку определитель Якоби для функции $\varphi_1(\cdot), \varphi_2(\cdot), \varphi_3(\cdot)$ удовлетворяет условию

$$\begin{aligned}
 D \left(\frac{\varphi_1(\cdot), \varphi_2(\cdot), \varphi_3(\cdot)}{p, q, r} \right) &= \frac{1}{[a(x)h(x) - c(x)b(x)]^3} \begin{bmatrix} h^2(x) & b(x)h(x) & b^2(x) \\ c^2(x) & a(x)c(x) & a^2(x) \\ 2c(x)h(x) & c(x)b(x) + q(x)h(x) & 2a(x)b(x) \end{bmatrix} \\
 &= -1,
 \end{aligned}$$

то, как общеизвестно, система (13)-(15) всегда разрешима относительно функции $p(x), q(x), r(x)$, то есть имеет место соотношения

$$\begin{aligned}
 p(x) &= \Psi_1(x, q(x), b(x), c(x), h(x)), \\
 q(x) &= \Psi_2(x, q(x), b(x), c(x), h(x)), \\
 r(x) &= \Psi_3(x, q(x), b(x), c(x), h(x)),
 \end{aligned}$$

где $a(x), b(x), c(x), h(x)$ – произвольные функции, удовлетворяющие ограничению (11).

Заметим, что условие выбора функций $a(x), b(x), c(x), h(x)$ позволяет их использовать и для определения коэффициентов $k_0(x), k_1(x), k_2(x)$ уравнения (12) при заданных коэффициентах $p(x), q(x), r(x)$ уравнения (9). В самом деле, согласно системе равенств (13)-(15), интегральное тождество (2) можно переписать в виде

$$\begin{aligned}
 &\left(\int \varphi_1[x, a(x), b(x), c(x), h(x)] e^{-\int \varphi_2[x, a(x), b(x), c(x), h(x)] dx} dx \right) * \\
 &\quad * \left(\int \varphi_3[x, a(x), b(x), c(x), h(x)] e^{\int \varphi_3[x, a(x), b(x), c(x), h(x)] dx} dx \right) \equiv 2,
 \end{aligned}$$

которое может быть выполнено за счет выборов функций $a(x), b(x), c(x)$ и $h(x)$, удовлетворяющих ограничению (11).

Теперь теоретические выводы проиллюстрируем на примере уравнения (5), где $k_0(x) = e^{2x}, k_1(x) = 1, k_2(x) = -2e^{-2x}$. В дробно-линейном преобразовании (10) положим

$$a(x) = e^x, b(x) = -e^{-x}, c(x) = e^x, h(x) = e^{-x}$$

и с учетом ограничения (11)

$$a(x)h(x) - c(x)b(x) = e^x e^{-x} + e^x e^{-x} = 2$$

выписываем систему равенств (13)-(15) относительно коэффициентов уравнения (9), т.е. получим систему

$$\begin{cases} p(x) - q(x) + r(x) = 2e^{4x}, \\ p(x) - r(x) = 3, \\ p(x) + q(x) + r(x) = -4e^{-4x}. \end{cases}$$

Решением этой системы являются функции

$$\begin{aligned} p(x) &= \frac{1}{2}(3 + e^{4x} - 2e^{-4x}), \\ q(x) &= -(e^{4x} + 2e^{-4x}), \\ r(x) &= -\frac{1}{2}(3 - e^{4x} + 2e^{-4x}), \end{aligned}$$

и уравнение (9) имеет вид

$$y'(x) = \frac{1}{2}(3 + e^{4x} - 2e^{-4x}) - (e^{4x} + 2e^{-4x})y(x) - \frac{1}{2}(3 - e^{4x} + 2e^{-4x})y^2(x). \quad (16)$$

Это уравнение при дробно-линейном преобразовании вида

$$y(x) = \frac{e^x z(x) - e^{-x}}{e^x z(x) + e^{-x}} = \frac{e^{2x} z(x) - 1}{e^{2x} z(x) + 1} \quad (17)$$

переходит в уравнение (5), коэффициенты которого удовлетворяют интегральному тождеству (2). Решение уравнения (5) $z(x) = -e^{2x}$ подставим в (17) и получим функцию

$$y(x) = \frac{e^{4x} + 1}{e^{4x} - 1},$$

которая является частным решением уравнения (16).

Заключение

В заключение отметим, что разработанная методика решения уравнения Риккати создает предпосылку для развития новых научных направлений и может быть полезной при разработке новых конструктивных методов решения прикладных задач естествознания, техники и оптимизации.

Литература

1. Егоров А.И. Уравнения Риккати. – М.: Изд-во «Солон-пресс», 2017. – 448 с.
2. Kerimbekov A. On a Class of Solutions of the Nonlinear Integral Fredholm Equation //Trends in Mathematics Research Perspectives (Analysis, Probability, Applications, and Computation. Proceedings of the 11th ISAAC Congress, Vaxjo (Sweden) 2017). Switzerland, Birkhauser. –2019. – P. 191–197.

УДК 517.97

О РАЗРЕШИМОСТИ ЗАДАЧИ ОПТИМИЗАЦИИ С МИНИМАЛЬНОЙ ЭНЕРГИЕЙ ПРИ ГРАНИЧНОМ УПРАВЛЕНИИ КОЛЕБАТЕЛЬНЫМ ПРОЦЕССОМ

Доулбекова Салтанат Байызбековна, к.ф.-м.н., доцент,

doulbekova25@mail.ru

Кыргызско-Российский Славянский университет

имени первого президента Российской Федерации Б.Н. Ельцина

Бишкек, Кыргызстан

Керимбеков Акылбек, д.ф.-м.н., профессор,

akl7@rambler.ru

Кыргызско-Российский Славянский университет

имени первого президента Российской Федерации Б.Н. Ельцина,

Бишкек, Кыргызстан

Баатов Авалкан Куканович к.ф.-м.н., доцент,

carterbek@mail.ru

Институт новых информационных технологий,

Кыргызского государственного университета имени И. Арабаева

Бишкек, Кыргызстан

В статье исследованы вопросы разрешимости задачи оптимизации колебательных процессов, описываемых интегро-дифференциальными уравнениями в частных производных с интегральным оператором Фредгольма, при минимизации интеграла энергии управляющей силы. Исследование проводилось с использованием понятия обобщенного решения краевой задачи управляемого колебательного процесса. В задаче оптимизации требуется найти управление, которое переводит колебательный процесс из одного состояния в другое заданное состояние. В процессе исследования установлено, что искомое оптимальное управление определяется как решение бесконечномерной системы интегральных уравнений Фредгольма первого рода и найдены достаточные условия существования решения этой некорректной задачи.

Ключевые слова: краевая задача, обобщенное решение, интеграл энергии, функционал, граничное управление, оптимальное управление.

Постановка задачи оптимизации и ее разрешимость. Рассмотрим задачу граничного управления колебательными процессами, где требуется минимизировать функционал

$$J[u(t)] = \int_0^T u^2(t) dt, \quad (1)$$

энергии управления (управляющих сил) при переводе управляемого процесса, описываемого краевой задачей

$$V_{tt} = V_{xx} + \lambda \int_0^T K(t, \tau) V(\tau, x) d\tau, \quad 0 < x < 1, \quad 0 < t \leq T \quad (2)$$

$$V(0, x) = \psi_1(x), \quad V_t(0, x) = \psi_2(x), \quad 0 < x < 1, \quad (3)$$

$$V_x(t, 0) = 0, \quad V_x(t, 1) + \alpha V(t, 1) = u(t) \quad (4)$$

из начального состояния $V(0, x) = \varphi_1(x)$ в заданное другое положение

$$V(T, x) = \xi(x) \quad (5)$$

за заданное время T . $K(t, \tau)$ – заданная функция, определенная в области $D = \{0 \leq t \leq T, 0 \leq \tau \leq T\}$, и удовлетворяет условию

$$\int_0^T \int_0^T K^2(t, \tau) d\tau dt = K_0 < \infty,$$

т.е. $K(t, \tau) \in H(D)$, $\xi(x)$, $\psi_1(x)$, $\psi_2(x)$ – заданные функции из пространства $H(0,1)$, причем функция $\psi_1(x)$ имеет обобщенную производную первого порядка; λ – параметр, T – фиксированный момент времени, постоянная $\alpha > 0$, $H(Y)$ – гильбертово пространство квадратично суммируемых функций, определенных на множестве Y .

Обобщенное решение краевой задачи. Известно [1,2], что краевая задача (2)-(4) при заданных условиях имеет единственное обобщенное решение

$$V(t, x) = \sum_{n=1}^{\infty} V_n(t) z_n(x), \quad (6)$$

где система функций $z_n(x) = \sqrt{\frac{2(\lambda_n^2 + \alpha^2)}{\lambda_n^2 + \alpha^2 + \alpha}} \cos \lambda_n x$ при каждом фиксированном $n = 1, 2, 3, \dots$, определяется как решение краевой задачи

$$z_n''(x) + \lambda_n^2 z_n(x) = 0, \quad z_n'(0) = 0, \quad z_n'(1) + \alpha z_n(1) = 0,$$

и образуют полную ортонормированную систему в гильбертовом пространстве $H(0,1)$, а соответствующие собственные значения λ_n определяются как решения трансцендентного уравнения $\lambda \operatorname{tg} \lambda = \alpha$ и удовлетворяют следующим условиям

$$\lambda_n \leq \lambda_{n+1}, \forall n = 1, 2, 3, \dots, \quad \lim_{n \rightarrow \infty} \lambda_n = \infty. \quad (n-1)\pi < \lambda_n < \frac{\pi}{2}(2n-1), \quad n = 1, 2, 3, \dots;$$

А коэффициенты Фурье $V_n(t)$ определяются соотношениями

$$V_n(t) = \psi_n(t, \lambda) + \frac{1}{\lambda_n} \int_0^T D_n(t, \tau, \lambda) z_n(1) u(\tau) d\tau, \quad (7)$$

$$\psi_n(t, \lambda) = \psi_n \left[\cos \lambda_n t + \lambda \int_0^t B_n(t, s, \lambda) \cos \lambda_n s ds \right] + \frac{1}{\lambda_n} \psi_{2n} \left[\sin \lambda_n t + \lambda \int_0^t B_n(t, s, \lambda) \sin \lambda_n s ds \right], \quad (8)$$

$$D_n(t, s, \lambda) = \begin{cases} \sin \lambda_n(t - \tau) + \lambda \int_0^t B_n(t, s, \lambda) \sin \lambda_n(s - \tau) ds, & 0 \leq \tau \leq t, \\ \int_0^t B_n(t, s, \lambda) \sin \lambda_n(s - \tau) ds, & t \leq \tau \leq T, \end{cases} \quad (9)$$

$$B_n(t, s, \lambda) = \sum_{i=1}^{\infty} \lambda^{i-1} K_{n,i}(t, s), \quad n = 1, 2, 3, \dots \quad (10)$$

– резольвента ядра

$$K_n(t, s) = \frac{1}{\lambda_n} \int_0^t \sin \lambda_n(t - \tau) K(\tau, s) d\tau, \quad K_n(0, s) = 0,$$

а повторные ядра $K_{n,i}(t, s)$ при каждом фиксированном $n = 1, 2, 3, \dots$, определяются по формулам

$$K_{n,i+1}(t, s) = \int_0^T K_n(t, \tau) K_{n,i}(\tau, s) d\tau, \quad i = 1, 2, 3, \dots, \quad (11)$$

$$K_{n,1}(t, s) \equiv K_n(t, s).$$

Согласно равенствам (10) - (11), имеет место следующая оценка

$$|B_n(t, s, \lambda)| \leq \frac{\sqrt{T}}{\lambda_n - |\lambda| \sqrt{K_0 T^2}} \cdot \sqrt{\int_0^T K^2(\tau, s) d\tau}, \quad (12)$$

которая выполняется для значений λ , удовлетворяющих следующее неравенство

$$|\lambda| \frac{T}{\lambda_n} \sqrt{K_0} < 1, \quad n = 1, 2, 3, \dots \quad (13)$$

Используя неравенства (12), получаем оценку:

$$\int_0^T |B_n(t, s, \lambda)|^2 ds \leq \frac{K_0 T}{\lambda_n - |\lambda| T \sqrt{K_0}}^2, \quad (14)$$

Ряд Неймана (10) абсолютно сходится при каждом фиксированном $n = 1, 2, 3, \dots$ для значений параметра λ , удовлетворяющего условию

$$\lambda < \frac{\lambda_n}{T \sqrt{K_0}}. \quad (15)$$

Заметим, что радиус сходимости $|\lambda| < \frac{\lambda_n}{T \sqrt{K_0}}$ ряда Неймана увеличивается с ростом n , и резольвента $B_n(t, s, \lambda)$ является непрерывной функцией как сумма абсолютно сходящегося ряда. Отметим, что при значениях параметра $|\lambda| < \frac{\lambda_1}{T \sqrt{K_0}}$ ряд Неймана абсолютно сходится к непрерывной функции для любого (!) $n = 1, 2, 3, \dots$

Доказано, что построенная функция $V(t, x)$ является элементом пространства $H(Q)$ и обобщенным решением краевой задачи (2) - (4).

Решение задачи оптимизации. Обобщенное решение (6) подставляем в (5), после несложных вычислений получим уравнение:

$$\int_0^T a_n(\eta) u(\eta) d\eta = h_n, \quad n = 1, 2, 3, \dots$$

где

$$a_n(\eta) = \frac{1}{\lambda_n} \left(\sin \lambda_n (T - \eta) + \lambda \int_{\eta}^T B_n(T, s, \lambda) \sin \lambda_n (s - \eta) ds \right) z_n(1); \quad (16)$$

$$h_n = \xi_n - \varphi_{1n} \left[\cos \lambda_n T + \lambda \int_0^T B_n(T, s, \lambda) \cos \lambda_n s ds \right] - \frac{\varphi_{2n}}{\lambda_n} \left[\sin \lambda_n T + \lambda \int_0^T B_n(T, s, \lambda) \sin \lambda_n s ds \right]. \quad (17)$$

Таким образом, искомое управление следует находить как решение системы интегральных уравнений Фредгольма первого рода. Эту систему, введя обозначения

$$a(t) = (a_1(t), \dots, a_n(t) \dots),$$

$$h = (h_1, \dots, h_n, \dots)$$

перепишем в матричной форме

$$\int_0^T a(t) u(t) dt = h. \quad (18)$$

Теперь исследуем разрешимость интегрального уравнения (18).

Решение ищем в виде:

$$u(t) = a^*(t) \cdot \alpha + \beta = \sum_{k=1}^{\infty} a_k(t) \alpha_k + \beta; \quad (19)$$

где $\alpha = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_2 \\ \dots \end{pmatrix}$ – неизвестный вектор, а β – произвольное постоянное, символ $*$ –

знак транспонирования.

Подставляя (19) в (18), получим бесконечномерную систему линейных неоднородных алгебраических уравнений вида :

$$\int_0^T \begin{pmatrix} a_1(t) \\ \vdots \\ a_n(t) \\ \dots \end{pmatrix} \cdot (a_1(t) \dots a_n(t) \dots) \cdot \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_2 \\ \dots \end{pmatrix} dt + \int_0^T \begin{pmatrix} a_1(t) \\ \vdots \\ a_n(t) \\ \dots \end{pmatrix} \cdot \beta dt = \begin{pmatrix} l_1 \\ \vdots \\ l_2 \\ \dots \end{pmatrix} \quad (20)$$

Введем бесконечномерную квадратную матрицу

$$A = \int_0^T a(t) \cdot a^*(t) dt = \int_0^T \begin{pmatrix} a_1(t) \\ \vdots \\ a_n(t) \\ \dots \end{pmatrix} \cdot (a_1(t) \dots a_n(t) \dots) dt =$$

$$= \begin{pmatrix} \int_0^T (a_1(t) \cdot a_1(t)) dt & \dots & \int_0^T (a_1(t) \cdot a_n(t)) dt & \dots \\ \dots & \dots & \dots & \dots \\ \int_0^T (a_n(t) \cdot a_1(t)) dt & \dots & \int_0^T (a_n(t) \cdot a_n(t)) dt & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix} \quad (21)$$

и

$$q = \int_0^T a(t) dt = \begin{pmatrix} q_1 \\ q_2 \\ \dots \\ q_n \\ \dots \end{pmatrix}.$$

Систему линейных неоднородных алгебраических уравнений (20) перепишем в матричной форме

$$A\alpha = H, \quad H(\beta) = h - \beta q, \quad (22)$$

Поскольку в системе (22) матрица A бесконечномерная, то при исследовании разрешимости системы (22) будем пользоваться вариационными методами, т.к. в этом случае методы решения конечномерных систем не пригодны. В этой связи ниже будем доказывать несколько утверждений:

Лемма 1. Вектор H является элементом пространства l_2 .

Лемма 2. При любом $\alpha \in l_2$ вектор $A\alpha$ является элементом пространства l_2 .

Доказательства лемм проводится непосредственно вычислением и не представляет труда.

Лемма 3. Бесконечномерная матрица A является положительно определенной.

Доказательство. Пусть $\gamma \in l_2$ – произвольный элемент. Тогда для скалярного произведения в l_2 имеет место неравенство

$$\langle Ay, \gamma \rangle = \gamma^* Ay = \gamma^* \int_0^T a(t) dt \cdot \gamma = \int_0^T \gamma^* a(t) a^*(t) \gamma dt = \int_0^T \|a^*(t) \gamma\|^2 dt \geq 0,$$

Откуда следует положительность матрицы A . Далее из равенства

$$\|a^*(t) \gamma\|^2 = 0, \Rightarrow \sum_{k=1}^{\infty} a_k(t) \gamma_k^2 = 0,$$

С учетом того, что функции $a_n(t)$, $n=1, 2, 3, \dots$ являются линейно независимыми на отрезке $[0, T]$, равенства $a_k(t) \gamma_k = 0$, выполняются тогда и только тогда, когда все $\gamma_k = 0$, $k=1, 2, 3, \dots$ Следовательно, матрица A является положительно определенной.

Теорема. Алгебраическая система (22) при каждом фиксированном β имеет единственное решение в пространстве l_2 .

Доказательство. В пространстве l_2 определим оператор $D[\alpha] = A\alpha$, $D: l_2 \rightarrow l_2$. Тогда оператор $D[\alpha]$ является положительно-определенным, что следует из Леммы 3, т.е.

$$\langle \alpha, D(\alpha) \rangle_{l_2} = \langle \alpha, A\alpha \rangle = \sum_{i,k=1}^{\infty} a_{ik} \alpha_i \alpha_k \geq 0.$$

Оператор $D[\cdot]$ является линейным, т.е. для любых произвольных постоянных C_1 и C_2 и произвольных векторов $\alpha^{(1)}$ и $\alpha^{(2)}$ пространства l_2 имеет равенства

$$D[C_1 \alpha^{(1)} + C_2 \alpha^{(2)}] = C_1 D[\alpha^{(1)}] + C_2 D[\alpha^{(2)}].$$

Из линейности следует, что оператор $D[\alpha]$ является взаимно однозначным оператором. Следовательно, существует обратный оператор $D^{-1}[\alpha]$, который, согласно теореме, из функционального анализа [3] является ограниченным оператором, т.е. оценка

$$\|D^{-1}[\alpha]\|_{l_2} \leq D_0 \|\alpha\|_{l_2}, \quad D_0 > 0 - \text{постоянная}, \quad (23)$$

имеет место для любого вектора $\alpha \in l_2$.

Таким образом, решение уравнения (22) определяется по формуле

$$\alpha = D^{-1}[H] = D^{-1}[h + \beta q].$$

Это решение подставляя в (19), получим решения интегрального уравнения Фредгольма первого рода (18) в виде

$$u(t) = a^*(t) \cdot \alpha + \beta = a^*(t) D^{-1}[h + \beta q] + \beta, \quad (24)$$

где β – произвольная постоянная.

Таким образом, установлено, что интегральное уравнение (18) имеет бесконечно много решений вида (24), среди которых может быть искомое управление $u^0(t)$, минимизирующее функционал (1).

Далее, для определения управления $u^0(t)$ рассмотрим функционал

$$J[u(t, \beta)] = \int_0^T u^2(t, \beta) dt = \int_0^T a^*(t) D^{-1}[h + \beta q] + \beta^2 dt, \quad (25)$$

И параметр β находим как решение экстремальной задачи вида

$$\Phi \beta = J[u(t, \beta)] \rightarrow \min, \quad \beta \in R, \quad (26)$$

Поскольку решается задача на безусловный экстремум, то для функции

$$\Phi(\beta) = \int_0^T u^2(t, \beta) dt = \int_0^T \left[a^*(t) D^{-1} h + \beta(1 + a^*(t) D^{-1} [q]) \right]^2 dt;$$

применяя классический метод решения, находим сначала критическую точку из условия

$$\begin{aligned} \Phi'(\beta) &= 2 \int_0^T \left[a^*(t) D^{-1} h + \beta(1 + a^*(t) D^{-1} [q]) \right] 1 + a^*(t) D^{-1} [q] dt = \\ &= 2 \int_0^T a^*(t) D^{-1} h 1 + a^*(t) D^{-1} [q] dt + 2\beta \int_0^T 1 + a^*(t) D^{-1} [q]^2 dt = 0; \end{aligned}$$

Т.е. критической точкой является

$$\beta^0 = - \frac{\int_0^T a^*(t) D^{-1} h 1 + a^*(t) D^{-1} [q] dt}{\int_0^T 1 + a^*(t) D^{-1} [q]^2 dt} \quad (27)$$

Далее, из неравенства

$$\begin{aligned} \Phi''(\beta) &= 2 \int_0^T \left[a^*(t) D^{-1} h + \beta(1 + a^*(t) D^{-1} [q]) \right] 1 + a^*(t) D^{-1} [q] dt = \\ &= 2 \int_0^T 1 + a^*(t) D^{-1} [q]^2 dt > 0; \end{aligned}$$

Следует, что значение β^0 является точкой минимума функции $\Phi(\beta)$.

Тогда для любого управления имеет место неравенство

$$J[u(t, \beta^0)] \leq J[u(t, \beta)].$$

Причем равенство имеет место лишь при $\beta = \beta^0$. Таким образом, искомое управление $u^0(t)$, на котором функционал (1) принимает наименьшее возможное значение, определяется по формуле

$$u^0(t) = a^*(t) D^{-1} h + \beta^0(1 + a^*(t) D^{-1} [q]). \quad (28)$$

Теперь проверим, что найденное оптимальное управление $u^0(t)$ является элементом гильбертового пространства $H(0, T)$ квадратично суммируемых функций, т.е. является допустимым управлением. Это следует из неравенства

$$\begin{aligned} \|u^0(t)\|_{H[0, T]}^2 &= \int_0^T u^0(t)^2 dt = \int_0^T \left[a^*(t) D^{-1} h + \beta^0(1 + a^*(t) D^{-1} [q]) \right]^2 dt \leq \\ &\leq 2 \int_0^T \left(\left| \langle a(t), D^{-1} h \rangle \right|^2 + 2\beta^{02} 1 + \left| \langle a(t), D^{-1} [q] \rangle \right|^2 \right) dt \leq \\ &\leq 2 \int_0^T \|a(t)\|_{l_2}^2, D_0^2 \|h\|_{l_2}^2 + 2\beta^{02} 1 + \|a(t)\|_{l_2}^2 D_0^2 \|q\|_{l_2}^2 dt < \infty, \end{aligned}$$

которое имеет место в силу следующих соотношений.

Заключение

В заключение отметим, что полученные результаты могут быть использованы на производстве, а также при разработке новых методов решения задач оптимального управления системами с распределенными параметрами. Рассматриваемая задача оптимизации является не корректной, что следует из системы интегральных уравнений Фредгольма первого рода, поэтому полученные результаты, в частности, разработанный

алгоритм построения решения системы интегральных уравнений Фредгольма первого рода, представляет практический и теоретический интерес.

Литература

1. Егоров А.И. Оптимальное управление тепловыми и диффузионными процессами. – М.: Наука, 1978. – 500 с.
2. Керимбеков А., Доулбекова С.Б. О разрешимости задачи нелинейной оптимизации колебательных процессов при появлении особых управлений //Вестник Евразийского национального университета имени Л.Н.Гумилева. Серия Математика. Компьютерные науки. Механика. – 2020. – Т.132, №3. – С. 6–16.
3. Люстерник Л.А., Соболев В.И. Элементы функционального анализа. – М.: Наука, 1965. – 520с.

УДК 519.85+539.424+538.9

*Г. Ч. Тукембаева, аспирант, e-mail: tukembaeva.g@gmail.com**Б. К. Темиров, д.ф.-м.н., профессор**КНУ им. Ж. Баласагына*

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ УСТОЙЧИВОСТИ СОЛИТОНА В МЕТАСТАБИЛЬНОМ СОСТОЯНИИ

В статье на базе операторно-разностных уравнений исследуется метастабильное состояние в различных приложениях науки и технологий. Объектом изучения являются двухфазные среды, свойства которых заданы произвольными параметрами без привязки к предмету исследования. Среда моделируется кубической функцией для реального газа по такой формуле, которая учитывает свойства и метастабильное состояние среды. Особое внимание уделено взаимодействию двух фаз, так как они генерируют колебания, которые, в свою очередь, солитон. Колебания среды отображаются в модели комплексными числами, поэтому формализация параметров среды нуждается в новых математических определениях для применения разнообразного математического аппарата. Тем самым нам пришлось рассматривать элемент объема в виде комплексных чисел, чтобы аналитически описать солитон и получить решение классическими методами математики. Для анализа применены метод Остроградского, исследование устойчивости полиномов по Гурвицу и интегрирование рациональных дробей. Таким образом, справедливость комплексных чисел для описания колебаний аппаратом математической физики позволяет изучать процессы под новым углом познания, чем способствуют созданию адекватных моделей в исследованиях.

Ключевые слова: двухфазная среда, осцилляция, метастабильное состояние, солитон, колебательная устойчивость, определители Гурвица.

Введение

Область двухфазных сред лежит под критической изотермой, где уравнение Ван-дер-Ваальса [1]–[10] расходится с опытом, особенно в метастабильном состоянии. Она описывается моделями двухфазного состояния [11] и вызывает интерес в связи с пузырьковым термоядерным синтезом [12]–[14]. В метастабильном состоянии протекают процессы, вызывающие научный интерес и порождающие странный аттрактор и динамический хаос, решение которых остается открытым [15]–[20]. В этой области кубическая функция должна качественно и количественно воссоздавать метастабильное состояние на возрастающем участке изотермы, который считается теоретическим. На нем находится точка перегиба, в которой участок пересекается с изобарой так, что кубическая функция разделена пополам, а площади под ней равны согласно правилу Максвелла. Точки минимума и максимума кубической функции по разные стороны от точки перегиба ограничивают возрастающий участок изотермы и не имеют физического обоснования, так как возрастающие участки изотермы неустойчивые. Для исследования таких процессов привлечена математическая теория Я.В. Быкова [21] для операторно-разностных уравнений 1-го порядка, развитая Б.К. Темировым для уравнений четного [22] и нечетного порядка [23]. Трудности возникают, когда значение элемента объема выражается комплексными корнями, а они появляются, когда решение – солитон в виде $\arctg$ -функции [24]. Это вызвано тем, что реальный газ содержит полином 2-й степени в знаменателе, корни которого комплексные и порождают колебания, где интегрирование такого полинома выражается через $\arctg$, см. с. 118 в [25] или с. 422 в [26]. Наличие колебаний указывает на взаимодействие двух фаз изучаемой среды.

Цель работы – моделирование двухфазной среды с учетом метастабильного состояния для процессов, описываемых кубической функцией.

Цель достигается тем, что результатом является солитон, так как он образуется на границе раздела фаз.

Описание модели

Рассмотрим график кубической функции (рис.) в системе координат VOP , где P – давление, V – молярный объем, T_k – критическая изотерма с критической точкой K , в которой три корня V_i сливаются в один кратный действительный V_k . Под ней расположена субкритическая изотерма T_M с возрастающим участком BMC так, что $T_M < T_k$.

Участок BMC имеет 3 критические точки: B , M и C . В точке минимума $B(V_{min}, P_{min})$

$$\frac{dP}{dV} = 0, \quad \frac{d^2P}{dV^2} > 0.$$

В точке максимума $C(V_{max}, P_{max})$

$$\frac{dP}{dV} = 0, \quad \frac{d^2P}{dV^2} < 0.$$

Точка $M(V_M, P_M)$ является точкой перегиба (метастабильного состояния), в которой

$$\frac{dP}{dV} = 0, \quad \frac{d^2P}{dV^2} = 0.$$

Она определяет новую, локальную систему координат $V'O'P'$, проходящую через точку M , где кубическая функция симметрична относительно начала координат O' , т.е. точки M .

Координатные оси разбивают плоскость на четыре квадранта: I, II, III и IV. Объем и давление в III квадранте принимают отрицательные значения. Такие значения определяют отрицательные элементы объема – $\Delta V < 0$ в символе Леви-Чивиты, что важно в тензорном анализе, но отрицательные давления известны. Тогда представление ΔV комплексными числами имеет смысл при изучении устойчивых колебаний, поскольку в этом случае удастся получить солитон, поэтому в [27] сформулирована лемма об отрицательном элементе объема. Колебания порождаются отрицательным дискриминантом, так как корень из него становится комплексным числом. Следовательно, особая роль отводится огибающей и дискриминантной кривой.

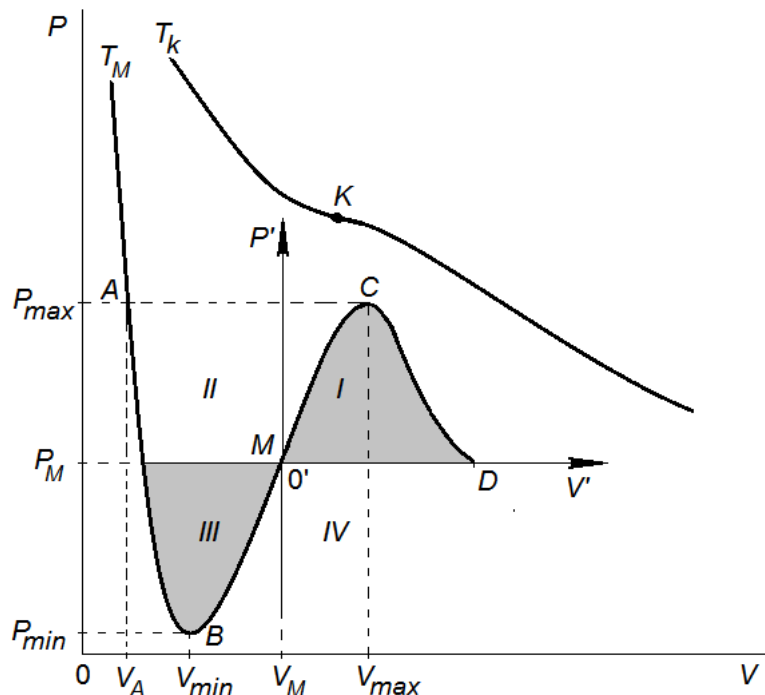


Рисунок – Возрастающий участок BMC под критической изотермой T_k

Локальная система координат $V'O'P'$ криволинейная и спрямляется по мере роста n в производной n -го порядка

$$\frac{d^n P}{dV^n} = 0.$$

Физически это означает: чем больше n , тем дальше от наблюдателя расположен горизонт прямой видимости. Например, противотуманные фары в туман, дождь, снегопад. Другой пример – радиоволны: с увеличением длины волны растет дальность их распространения. Математически это выражается операторно-разностными уравнениями n -го порядка. Для трехмерного пространства (x, y, z) вторая производная имеет вид матрицы Гессе, а точка перегиба переходит в седловую точку, т.е. в направлении x становится перевалом, ограниченным с обеих сторон склонами в направлении y .

Изучим поведение пузырька на субкритической изотерме T_M с внутренней энергией U в дифференциальной форме

$$dU = TdS - PdV, \quad (1)$$

где S – энтропия, давление

$$P = \frac{dF_n}{ds}$$

для расширяющегося газа F_n – нормальная составляющая силы приложена изнутри к элементу площади ds . Для газа, который описывается формулой Ван-дер-Ваальса

$$P = \frac{RT}{V-b} - \frac{a}{V^2}, \quad (2)$$

где V – молярный объем газа [$\text{м}^3/\text{моль}$], a [$\text{Па} \cdot \text{м}^6/\text{моль}^2$] и b [$\text{м}^3/\text{моль}$] – поправки Ван-дер-Ваальса; T – температура по Кельвину [K], P – давление [Па], $R=8,314$ Дж/моль·К – универсальная газовая постоянная; $a, b, R, T \in \mathbf{R}_+$, но $P \in \mathbf{R}$, $V \in \mathbf{C}$. Это необходимо для сравнения с моделью, которую предстоит разработать. Соответствующее формуле (2) уравнение Ван-дер-Ваальса имеет вид

$$PV^3 - (RT + Pb)V^2 + aV - ab = 0.$$

Тогда на изотерме $T=\text{const}$ интегрирование дает

$$\begin{aligned} U &= TS - \int \left(\frac{RT}{V-b} - \frac{a}{V^2} \right) dV, \\ U &= TS - RT \int \frac{dV}{V-b} + a \int V^{-2} dV, \\ U &= TS - RT \ln|V-b| - \frac{a}{V} + C. \end{aligned} \quad (3)$$

В случае, когда газ сжимается, то F_n меняет знак, поэтому

$$P = -\frac{dF_n}{ds}.$$

Следовательно, внутренняя энергия растет

$$U = T \int dS + \int PdV.$$

Подстановка по формуле (1) или (2) дает

$$\begin{aligned} U &= TS + \int \left(\frac{RT}{V-b} - \frac{a}{V^2} \right) dV, \\ U &= TS + RT \int \frac{dV}{V-b} - a \int \frac{dV}{V^2}, \\ U &= TS + RT \ln|V-b| + \frac{a}{V} + C. \end{aligned}$$

Значит, для сжатия имеем большую величину по сравнению с внутренней энергией, полученной в (3) для расширения. Тем самым в зависимости от направления нормальной составляющей силы принимаем объем положительным, если F_n отвечает расширению пузырька, или отрицательным объемом, если F_n соответствует сжатию пузырька.

Начнем с точки А, где за счет убыли внутренней энергии начинается рост пузырька. Размеры пузырька ограничены пределом сжатия электронного газа, слева от т. А. Поэтому в области отрицательных элементов объема, где F_n со знаком "-", соответственно системе координат $V'O'P'$ в III квадранте. Поправка Ван-дер-Ваальса b меняет знак. Вместо $(V-b)$ получаем $(V+b)$ за счет плотности электронного газа, т.е. его значительной внутренней энергии U . Уравнение Ван-дер-Ваальса выведено для разреженных газов, поэтому таких эффектов не учитывает. Более того, давление электронного газа зависит от электронов на внешних орбитах, поэтому сильнее проявляется в одном направлении по сравнению с другим направлением, так как эти электроны пытаются взаимодействовать с атомами по другую сторону оболочки пузырька. Давление в пузырьке ослабляется с P_{max} до P_{min} в точке В, так как P_{min} равно давлению окружающей среды. Затем объем пузырька увеличивается на возрастающем участке ВМС изотермы T_M .

Как только размеры пузырька достигнут значения V_M , выходим на горизонтальную площадку точки перегиба М. Горизонтальность обуславливает скапливание пузырьков на плоскости в метастабильном состоянии. Справа от V_M в системе координат $V'O'P'$ давление и объем принимают положительные значения. Дальнейший рост размеров пузырька благодаря внутренней энергии ограничен V_{max} и соответствующим P_{max} в точке С, так как P_{max} определяется коэффициентом поверхностного натяжения. Поэтому на оболочке пузырька возникает разрыв, и в среду впрыскивается электронный газ, а сам пузырек сжимается до размеров V_A , переходя по изобаре из т. С в т. А. Энергия среды повышается за счет электронов, отдаваемых атомами пузырька с внешних оболочек, которые тем самым превращаются в отрицательные ионы. Из т. А пузырек расширяется, так как при сжатии возросла его внутренняя энергия, а потому замыкается цикл АВМСА (см. рис.).

Исследование устойчивости

В работах [4], [8]-[10] приведены различные модели реального газа, в том числе по формуле Ван-дер-Ваальса (2): уравнение Редлиха – Квонга

$$P = -\frac{RT}{V-b} - \frac{a}{\sqrt{T}V(V+b)};$$

уравнение Бертло

$$P = -\frac{RT}{V-b} - \frac{a}{TV^2};$$

второе уравнение Дитеричи

$$P = -\frac{RT}{V-b} - \frac{a}{TV^{5/3}};$$

Соаве – Редлиха – Квонга

$$\left[P + \frac{a(\tilde{T})}{V(V+b)} \right] (V-b) = RT$$

Пенга – Робинсона

$$P = -\frac{RT}{V-b} - \frac{a\{T\}}{V(V+b) + b(V-b)};$$

Исикавы – Чанга – Лу

$$P = -\frac{RT(2V + b(T))}{V(2V - b(T))} - \frac{a(T)}{\sqrt{TV}(V + b(T))};$$

Фогельсона – Лихачева

$$\left[P + \frac{a}{(V + c)^{kTm}} \right] (V - b) = RT.$$

Общей чертой этих моделей является то, что содержат $(V - b)$ в знаменателе идеального газа, как в формуле Ван-дер-Ваальса (2).

Согласно вышеописанной модели, покажем, как в III квадранте меняется знак идеальной составляющей реального газа на примере формулы (2). Для этого надо переместить начало координат $V0P$ в точку перегиба M . Координаты т. M вычисляют с помощью первой и второй производных $dP/dV=0$ и $d^2P/dV^2=0$. Тогда имеем новую систему координат $V0'P'$, относительно которой в III квадранте элемент объема и давление в нем отрицательные. Значит, в $(V - b)$ надо поменять знак перед V на "–", поэтому в скобках получим $(-V - b)$. Теперь выносим знак "–" из скобок и имеем: $-(V+b)$. Следовательно, формула (2) обретет вид

$$P = -\frac{RT}{V + b} - \frac{a}{V^2},$$

как и в вышеперечисленных моделях.

Обобщение формулы (2) рассмотрим в следующем виде

$$P = \frac{RT}{V+b} + \frac{MV+N}{V^2+pV+q}, \quad (4)$$

где $b, p, q, T, M, N \in \mathbf{R}_+$. Квадратный трехчлен не разлагается на множители первой степени, так как дискриминант $D=(p/2)^2-q<0$, а поэтому $V \in \mathbf{C}$. Что касается давления, то оно может колебаться вместе с элементом объема, поэтому определимся с ним в ходе решения. Роль коэффициентов p и q следующая. В силу того, что $D<0$, подстановкой $z=V+p/2$ получим

$$P = \frac{RT}{V+b} + \frac{MV+N}{z^2+k^2},$$

где

$$V^2 + pV + q = \left(V + \frac{p}{2}\right)^2 + q - \left(\frac{p}{2}\right)^2, \\ k^2 = q - \left(\frac{p}{2}\right)^2,$$

поэтому

$$P = \frac{RT}{V+b} + \frac{MV+N}{(z+ik)(z-ik)}.$$

Тогда второе слагаемое формулы (4) интегрируется известным образом, см. с. 423-424 в [26]. Значит, простейшая дробь второго типа

$$\int \frac{MV+N}{V^2+pV+q} dV = \frac{M}{2} \ln(V^2 + pV + q) + \frac{2N-Mp}{\sqrt{4q-p^2}} \arctg \frac{2V+p}{\sqrt{4q-p^2}} + C \quad (5)$$

содержит солитон. Это указывает, что взаимодействие двух фаз между собой генерирует колебания, а они – солитон. Подстановка выражения (5) в формулу (1) позволяет найти внутреннюю энергию U . Однако сначала надо исследовать уравнение, определяемое формулой (4) на устойчивость. Например, интегрирование формулы Ван-дер-Ваальса (2) не дало солитон, чему свидетельствует выражение (3), так как только в случае $V \in \mathbf{C}$, подынтегральная функция содержит простейшие дроби второго типа.

Преобразуем формулу (4) в уравнение

$$P(V + b)(V^2 + pV + q) = RT(V^2 + pV + q) + (V + b)(MV + N).$$

Поскольку

$$P(V + b)(V^2 + pV + q) = PV^3 + P(p + b)V^2 + P(q + pb)V + Pqb,$$

то

$$PV^3 + P(p + b)V^2 + P(q + pb)V + Pqb = (RT + M)V^2 + (pRT + bM + N)V + qRT + bN.$$

Перенесем правую часть на левую сторону равенства и получим уравнение

$$PV^3 + [P(p + b) - (RT + M)]V^2 + [P(q + pb) - (pRT + bM + N)]V + (Pqb - qRT - bN) = 0,$$

где свободный член и коэффициенты при разных степенях взяты в скобки так, что

$$a_0V^3 + a_1V^2 + a_2V + a_3 = 0.$$

Необходимые условия устойчивости выполнены, если все коэффициенты

$$a_0 = P > 0, \quad a_1 = P(p + b) - (RT + M) > 0,$$

$$a_2 = P(q + pb) - (pRT + bM + N) > 0, \quad a_3 = Pqb - qRT - bN > 0.$$

С учетом определителей Гурвица имеем

$$\Delta_3 = \begin{vmatrix} P(p + b) - (RT + M) & Pqb - qRT - bN & 0 \\ P & P(q + pb) - (pRT + bM + N) & 0 \\ 0 & P(p + b) - (RT + M) & Pqb - qRT - bN \end{vmatrix} > 0,$$

$$\Delta_2 = \begin{vmatrix} P(p + b) - (RT + M) & Pqb - qRT - bN \\ P & P(q + pb) - (pRT + bM + N) \end{vmatrix} > 0,$$

$$Pqb > qRT + bN,$$

$$P(q + pb) > pRT + bM + N.$$

$$P(p + b) > RT + M,$$

$$P > 0.$$

Из условия $\Delta_3 > 0$ имеем

$$(Pqb - qRT - bN) \begin{vmatrix} P(p + b) - (RT + M) & Pqb - qRT - bN \\ P & P(q + pb) - (pRT + bM + N) \end{vmatrix} > 0.$$

На границе апериодической устойчивости $\Delta_3 = 0$, следовательно

$$Pqb - qRT - bN = 0,$$

$$\begin{vmatrix} P(p + b) - (RT + M) & Pqb - qRT - bN \\ P & P(q + pb) - (pRT + bM + N) \end{vmatrix} = 0.$$

Значит, давление на границе апериодической устойчивости

$$P = \frac{qRT + bN}{qb}.$$

На границе колебательной устойчивости

$$[P(p + b) - (RT + M)][P(q + pb) - (pRT + bM + N)] = P(Pqb - qRT - bN).$$

Раскрывая в неравенстве скобки, получим

$$(q + pb)(p + b)P^2 - (pRT + bM + N)(p + b)P - (RT + M)(q + pb)P + (RT + M)(pRT + bM + N) = P(Pqb - qRT - bN),$$

$$(q + pb)(p + b)P^2 - [(pRT + bM + N)(p + b) + (RT + M)(q + pb)]P + (RT + M)(pRT + bM + N) = qbP^2 - qRTp - bNP,$$

$$(p^2b + pb^2 + qp)P^2 - [(p + b)(pRT + bM + N) + (RT + M)(q + pb) + qRT + bN]P + (RT + M)(pRT + bM + N) = 0,$$

$$D = [(p + b)(pRT + bM + N) + (RT + M)(q + pb) + qRT + bN]^2 - 4(p^2b + pb^2 + qp)(RT + M)(pRT + bM + N),$$

$$D = [(p + b)(pRT + bM + N)]^2 + 2[(p + b)(pRT + bM + N)[(RT + M)(q + pb) + qRT + bN]] + [(RT + M)(q + pb) + qRT + bN]^2 - 4(p^2b + pb^2 + qp)(RT + M)(pRT + bM + N).$$

Обозначим

$$\alpha = p^2b + pb^2 + qp, \\ \beta = [(p + b)(pRT + bM + N) + (RT + M)(q + pb) + qRT + bN], \\ \gamma = (RT + M)(pRT + bM + N)$$

и получим квадратное уравнение

$$\alpha P^2 - \beta P + \gamma = 0.$$

Корни уравнения

$$P_{1,2} = \frac{\beta \pm \sqrt{\beta^2 - 4\alpha\gamma}}{2\alpha}$$

положительные, если дискриминант

$$D = \beta^2 - 4\alpha\gamma > 0.$$

С учетом коэффициентов

$$D = [(p + b)(pRT + bM + N)]^2 + 2[(p + b)(pRT + bM + N)[(RT + M)(q + pb) + qRT + bN]] + [(RT + M)(q + pb) + qRT + bN]^2 - 4(p^2b + pb^2 + qp)(RT + M)(pRT + bM + N).$$

Колебаниям, в данном случае неустойчивым, отвечает $D < 0$, т.е. неравенство

$$[(p + b)(pRT + bM + N)]^2 + 2[(p + b)(pRT + bM + N)[(RT + M)(q + pb) + qRT + bN]] + [(RT + M)(q + pb) + qRT + bN]^2 < 4(p^2b + pb^2 + qp)(RT + M)(pRT + bM + N).$$

Тогда давление будет комплексным числом $P \in \mathbb{C}$, так как колебания – это попеременное расширение – сжатие замкнутого элемента объема.

Отрицательное давление $P < 0$ указывает, что элемент объема внутри среды сжимается, поэтому формула (4) примет вид

$$-P = \frac{RT}{V + b} + \frac{MV + N}{V^2 + pV + q}.$$

Меняя знаки, чтобы a_0 удовлетворяло критерию Гурвица, получим уравнение

$$P(V + b)(V^2 + pV + q) = -RT(V^2 + pV + q) - (V + b)(MV + N), \\ P(V + b)(V^2 + pV + q) + RT(V^2 + pV + q) + (V + b)(MV + N) = 0, \\ PV^3 + [(p + b)P + RT + M]V^2 + [(q + pb)P + pRT + N + bM]V + (qbP + qRT + bN) = 0, \\ a_0V^3 + a_1V^2 + a_2V + a_3 = 0.$$

Теперь все коэффициенты положительные, в том числе $a_0 = P > 0$, поэтому по критерию Гурвица

$$P > 0, \quad (p + b)P + RT + M > 0, \quad (q + pb)P + pRT + N + bM > 0, \\ qbP + qRT + bN > 0,$$

$$\Delta_2 = \begin{bmatrix} (p + b)P + RT + M & qbP + qRT + bN \\ P & (q + pb)P + pRT + N + bM \end{bmatrix} > 0, \\ \Delta_3 = \begin{bmatrix} (p + b)P + RT + M & qbP + qRT + bN & 0 \\ P & (q + pb)P + pRT + N + bM & 0 \\ 0 & (p + b)P + RT + M & qbP + qRT + bN \end{bmatrix} > 0.$$

Проверяем на колебательную устойчивость $\Delta_2 > 0$ и имеем неравенство
 $[(p+b)P + RT + M][(q+pb)P + pRT + N + bM] - P(qbP + qRT + bN) > 0$.

Преобразуем его

$$[p^2b + pb^2 + qp]P^2 + [(p+b)pRT + (p+b)bM + pN + (q+pb)M + pbRT]P + [p(RT)^2 + NRT + bMRT + pMRT + NM + bM^2] > 0.$$

Обозначим

$$\begin{aligned}\alpha &= p^2b + pb^2 + qp > 0, \\ \delta &= (p+b)pRT + (p+b)bM + pN + (q+pb)M + pbRT > 0, \\ \gamma &= (RT+M)(pRT + bM + N) > 0,\end{aligned}$$

поскольку

$$(RT+M)(pRT + bM + N) = p(RT)^2 + NRT + bMRT + pMRT + NM + bM^2.$$

Следовательно, $\delta < \beta$, так как

$$\delta = (p+b)pRT + (p+b)bM + pN + (q+pb)M + pbRT,$$

но

$$\beta = (p+b)pRT + (p+b)bM + pN + (q+pb)M + pbRT + 2(qRT + bN).$$

Значит,

$$\beta = \delta + 2(qRT + bN).$$

Тогда имеем

$$\alpha P^2 + \delta P + \gamma > 0.$$

На границе колебательной устойчивости имеет место квадратное уравнение

$$\alpha P^2 + \delta P + \gamma = 0.$$

Найдем корни уравнения

$$P_{1,2} = \frac{-\delta \pm \sqrt{\delta^2 - 4\alpha\gamma}}{2\alpha}.$$

Поскольку дискриминант $D = \delta^2 - 4\alpha\gamma$, то паре действительных корней отвечает $D > 0$, где оба корня меньше нуля. Значит, в среде возникло отрицательное давление. Расширение сменилось сжатием, т.е. знак "+" поменялся на знак "-" в $dF_n = \pm P ds$.

Для устойчивых колебаний, если $4\alpha\gamma > \delta^2$, поскольку

$$\delta = \beta - 2(qRT + bN), \quad \delta^2 = [\beta - 2(qRT + bN)]^2,$$

то комплексные корни определяются неравенством

$$\begin{aligned}D &= \delta^2 - 4\alpha\gamma < 0, \\ D &= [\beta - 2(qRT + bN)]^2 - 4\alpha\gamma < 0.\end{aligned}$$

Тем самым колебания непрерывно переходят из области неустойчивых колебаний, где $P > 0$, в область устойчивых колебаний, отвечающих $P < 0$. В силу того, что

$$[\beta - 2(qRT + bN)]^2 - 4\alpha\gamma = \beta^2 - 4\beta(qRT + bN) + 4(qRT + bN)^2 - 4\alpha\gamma,$$

то при выполнении

$$[\beta - 2(qRT + bN)]^2 \geq 4\alpha\gamma$$

дискриминант $D \geq 0$, имеем границу действительных корней в случае равенства. Со стороны комплексных корней получаем неравенство

$$\begin{aligned}\beta^2 - 4\alpha\gamma &< 4\beta(qRT + bN) - 4(qRT + bN)^2, \\ D &< 4(qRT + bN)[\delta + qRT + bN],\end{aligned}$$

которое указывает, что давление стало комплексным числом, т.е. $P \in \mathbb{C}$.

Пусть идеальный газ в знаменателе представлен полиномом 1-й степени $(V - b)$

$$P = \frac{RT}{V - b} + \frac{MV + N}{V^2 + pV + q}.$$

Тогда уравнение примет вид

$$PV^3 + [P(p - b) - (RT + M)]V^2 + [P(q - pb) - (pRT - bM + N)]V - (Pqb + qRT + bN) = 0.$$

Поскольку

$$(Pqb + qRT + bN) < 0,$$

то система неустойчивая.

В случае отрицательного давления

$$PV^3 + [(p - b)P + RT + M]V^2 + [(q - pb)P + pRT + N - bM]V + (-qbP + qRT + bN) = 0.$$

Для устойчивости, во-первых, необходимо, чтобы выполнялось

$$\begin{aligned} P &< \frac{qRT + bN}{qb}, \\ P &< \frac{RT + M}{(b - p)}, \\ P &> \frac{bM - pRT - N}{(q - pb)}. \end{aligned}$$

Тем самым давление ограничено интервалом. Для колебательной устойчивости

$$\Delta_2 = \begin{bmatrix} (p - b)P + RT + M & -qbP + qRT + bN \\ P & (q - pb)P + pRT + N - bM \end{bmatrix} > 0,$$

т.е. требуется выполнение неравенства

$$\begin{aligned} &[(p - b)P + RT + M][(q - pb)P + pRT + N - bM] - P(-qbP + qRT + bN) > 0, \\ &(p - b)(q - pb)P^2 + [(p - b)pRT + (p - b)N - (p - b)bM + (RT + M)(q - pb)]P \\ &\quad + (RT + M)pRT + [(RT + M)N - (RT + M)bM] + qbP^2 - qRTP - bNP > 0, \\ &(pq - p^2b + pb^2)P^2 + [p^2RT - 2pbRT + b^2M + qM - 2pbM + pN - 2bN]P \\ &\quad + (RT + M)(pRT + N - bM) > 0. \end{aligned}$$

На границе колебательной устойчивости получим квадратное уравнение

$$(pq - p^2b + pb^2)P^2 + [p^2RT - 2pbRT + b^2M + qM - 2pbM + pN - 2bN]P + (RT + M)(pRT + N - bM) = 0.$$

Его дискриминант

$$\begin{aligned} D &= [(p^2RT + b^2M + qM + pN) - 2(pbRT + pbM + bN)]^2 \\ &\quad - 4(pq - p^2b + pb^2)(RT + M)(pRT + N - bM). \end{aligned}$$

Поскольку действительным значениям P отвечает $D \geq 0$, то

$$\begin{aligned} &(p^2RT + b^2M + qM + pN)^2 - 4(p^2RT + b^2M + qM + pN)(pbRT + pbM + bN) \\ &\quad + 4(pbRT + pbM + bN)^2 \geq 4(pq - p^2b + pb^2)(RT + M)(pRT + N - bM). \end{aligned}$$

С другой стороны, если $D < 0$, то P комплексное число. Тогда

$$\begin{aligned} &(p^2RT + b^2M + qM + pN)^2 + 4(pbRT + pbM + bN)^2 \\ &\quad < 4[(pq - p^2b + pb^2)(RT + M)(pRT + N - bM) \\ &\quad + (p^2RT + b^2M + qM + pN)(pbRT + pbM + bN)]. \end{aligned}$$

Выводы

Исследование колебаний в различных средах обуславливает применение соответствующего математического аппарата для формализации параметров среды, ее физических величин, так как колебания обнаруживаются опытом. Это касается тех физических величин, которые измеряются приборами. Например, скачущее атмосферное давление называют барической пилой, но, по сути, это колебания, значения которых

отражаются на приборе. Подобные вопросы подлежат математическому обоснованию и новому определению измеряемых величин, чтобы количественно и качественно описать измеряемые величины в виде формул. Из данного исследования следует, что зависимость давления от направления приложенной силы $dF_n = \pm P ds$ определяется знаком " $\pm$ ". Поэтому P может быть и отрицательным, и положительным, так как знак силы устанавливается по отклонению измерения в замкнутом элементе объема, внутри которого показания прибора равные по сравнению с соседним элементом объема. Иначе никакие методы конечных разностей или конечных элементов не применимы к исследованию изучаемых процессов. Более того, колебания давления в некоторых границах замкнутого элемента объема вынуждают принять P комплексным числом, как и элемент объема, ранее, особенно в метастабильном состоянии, когда колебания являются результатом взаимодействия двух разных фаз. Таким образом, справедливость комплексных чисел для описания колебаний аппаратом математической физики позволяет изучать процессы под новым углом познания, чем способствуют созданию адекватных моделей в исследованиях.

Литература

1. Савельев И.В. Курс общей физики. – Т. 1. – М.: Наука, 1970. – 511 с.
2. Белонучкин В.Е. Краткий курс термодинамики. – 2-е изд. – М.: МФТИ, 2010. – 164 с.
3. J. Kestin. A course in Thermodynamics. Vol. 1. New York: McGraw-Kill, 1979.
4. Фогельсон, Р.Л. Уравнение состояния реального газа. / Р.Л. Фогельсон, Е.Р. Лихачев // ЖТФ. – Т. 74, вып. 7. – 2004. – С. 129–130.
5. Колгатин С.Н. Критическая точка, бинадаль, спинодаль и отрицательные давления на примере уравнения Ван-дер-Ваальса // Научно-технические ведомости. СПбГПУ. Сер.: Наука и образование. – 2012. – № 3 (154), ч. 2. – С. 199–205.
6. Caupin F. and Anisimov M.A. Thermodynamics of supercooled and stretched water: Unifying two-structure description and liquid-vapor spinodal // J. Chem. Phys. 151, 034503 (2019); <https://doi.org/10.1063/1.5100228>
7. Виноградов В.Е. Исследование вскипания перегретых и растянутых жидкостей: спец. 01.04.14 «Теплофизика и теоретическая теплотехника»: автореф. дис. ... докт. физ.-мат. наук. – Екатеринбург, 2006. – 43 с.
8. Дей Е.А., Тюменков Г.Ю. Граничные параметры для состояния растянутой жидкости // Проблемы физики, математики и техники. – 2018. – № 4(37). – С. 18–20.
9. Дей Е.А., Тюменков Г.Ю. О приведенной форме термодинамических коэффициентов реальных газов // Проблемы физики, математики и техники. – 2022. – № 4(53). – С. 25–29.
10. Дей Е.А., Тюменков Г.Ю. Приведенные термодинамические коэффициенты в теории реальных газов // Проблемы физики, математики и техники. – 2023. – № 4(57). – С. 20–24.
11. Нигматулин Р.И. Динамика многофазных сред. Ч. 1. – М.: Наука. 1987. – 464 с.
12. Тукембаева Г.Ч., Темиров Б.К. Моделирование уравнений термодинамики динамическими системами // Проблемы автоматизации и управления. – 2023. – № 2(47). – С. 109–115.
13. Тукембаева Г.Ч., Темиров Б.К. Моделирование пузырькового термоядерного синтеза // Проблемы автоматизации и управления. – 2023. – № 3(48). – С. 26–34.
14. О термоядерных процессах в кавитирующих пузырьках / Нигматулин Р. И., Лэхи Р. Т. (мл.), Талейархан Р. П., Вест К. Д., Блок Р. С. // Успехи физич. наук. – 2014. – Т. 184, № 9. – С. 947–960.

15. Бойко В.Г., Могель Х.-Й., Сысоев В.М., Чалый А.В. Особенности метастабильных состояний при фазовых переходах жидкость – пар // Успехи физич. наук. – 1991. – Т. 161, № 2. – С. 77–111.
16. Ахромеева Т.С., Курдюмов С.П., Малинецкий Г.Г., Самарский А.А. Нестационарные структуры и диффузионный хаос. – М.: Наука, 1992. – 544 с.
17. Зелёный Л. М., Малова Х. В., Попова В. Ю. и др. Альберт Галеев: Проблемы метастабильности и взрывного пересоединения // Физика плазмы. – 2021. – Т. 47, 3 9. – С. 771–792.
18. Товбин Ю. К. Второе начало термодинамики, термодинамика Гиббса и времена релаксации термодинамических параметров // Журн. физической химии. – 2021. – Т. 95, № 4. – С. 483–507.
19. Канель Г. И. О наносекундной теплофизике (обзор) // Теплофизика высоких температур. –2020. – Т. 58, № 4. – С. 596–614.
20. Соболев Р.Н. Метастабильное состояние магматических систем // Бюлл. Моск. о-ва испытателей природы. Отдел геологический. – 2017. – Т. 92, № 2. – С. 83–89.
21. Быков Я.Б. Математические основы теории переходных процессов в импульсных системах.– Краснодар: Краснодар. политехн. ин-т, 1976. – 165 с.
22. Темиров Б.К. Осцилляция решений операторно-разностных уравнений с конечными разностями четного порядка : автореф. дис. ... канд.-физ. наук : 01.01.02 – Бишкек, 1995. – 15 с.
23. Темиров Б.К. Осцилляция решений операторно-разностных уравнений с конечными разностями произвольных нечетных порядков. – Бишкек: Maxprint, 2014. – 178 с.
24. Тукембаева Г.Ч., Темиров Б.К. Обобщение уравнения Ван-дер-Ваальса в пределах кубической функции // Проблемы автоматизации и управления. – 2023. – № 3(48). – С. 35–40.
25. Корн Г., Корн Т. Справочник по математике для научных работников и инженеров.– М.: Наука, 1984. – 832 с.
26. Выгодский М.Я. Справочник по высшей математике. – М.: Наука, 1976. – 872 с.
27. Темиров Б.К., Тукембаева Г.Ч. Элемент объема в малой окрестности особой точки // Тез. докладов V Борубаевских чтений./ Сб. докл. – Бишкек, 2024. – С. 70.

УДК 621.3.053.22:621.318.4

*К. Сатаркулов, Г.Ш. Эралиева, Г.Н. Ниязова,**А.К. Асанов, Н.А. Суюнтбекова**Институт машиноведения и автоматизации НАН КР,**Кыргызский государственный технический университет им. И. Раззакова*

КАТУШКА ИНДУКТИВНОСТИ С НОВЫМ СПОСОБОМ НАМОТКИ КАК УСТРОЙСТВО ДИАГНОСТИКИ

В работе рассмотрена катушка индуктивности (КИ) с новым способом намотки, которая использована в качестве устройства диагностики (УД), для оценки изменения спектрального состава тока нагрузки в проводе электропередачи от переходных процессов, возникающих в распределительной сети за счет подключения или отключения обычных нагрузок или нагрузок с нелинейными характеристиками. Конструктивно КИ представляет собой магнитопровод тороидальной формы, на которую намотана обмотка из пучка трех изолированных друг от друга эмалированных проводов. В результате получаются три независимые обмотки с равным количеством витков. Дано описание функционирования УД на конкретных примерах. Показано, как, соединяя концы двух параллельно намотанных проводов, можно изменить характер сопротивления с индуктивной на емкостное и обратно в отдельные моменты времени за счет емкости, образованной между этими проводами. Представлены компьютерные модели, описывающие работу УД. Полученные результаты моделирования показали работоспособность представленных устройств.

Ключевые слова: катушка индуктивности, переходной процесс; компьютерное моделирование; SimPowerSystems; воздушные линии; датчик тока; диагностирование.

Введение

Состояние энергетической эффективности электрических сетей рассмотрено в работе [1] выдающимися специалистами Российской Федерации по электроэнергетике В.Э. Воротническим, И.В. Жежеленко и крупным специалистом в области электроэнергетики и энергосбережения Г.Г. Трофимовым. Отмечено, что высокий уровень технических потерь в электрических сетях связан с такими причинами, как: низкий уровень компенсации реактивной мощности; 70-процентный физический и моральный износ сети при явно недостаточном использовании средств оптимизации режимов работы и регулирования напряжения; нерешенность проблем с качеством электрической энергии (КЭЭ). Даже в промышленно развитых странах Западной Европы несвоевременно внедряются решения, направленные на коррекцию КЭЭ, связанные с непрерывным ростом установленной мощности нелинейных, несимметричных и резкопеременных нагрузок. В [1] сказано, что нет необходимости доказывать значимость проблемы КЭЭ, как и проблемы энергоресурсосбережения, так как они относятся к числу важнейших проблем современной электроэнергетики и являются частью проблемы повышения энергоэффективности электрических сетей.

Практика эксплуатации электрических сетей показала, что универсальным средством улучшения энергоэффективности электрической сети, связанным с повышением его пропускной способности, снижением потерь мощности и электроэнергии в них, регулированием напряжения в контрольных точках сети, снижением уровня высоких гармоник и несимметрии напряжения, является внедрение устройств компенсации реактивной мощности различных типов фильтрокомпенсирующих и симметрирующих устройств. Но по ряду причин работа по компенсации реактивной мощности в электрических сетях стран СНГ ведется явно недостаточно [1].

Известно, что необходимость в компенсации реактивной мощности в электрических сетях вызвана тем, что нагрузка сети в основном имеет активно-индуктивный характер, и для повышения коэффициента мощности широко применяются разного рода

компенсирующие устройства (КУ). Например, для предприятий малой мощности представляют интерес компенсации реактивной мощности батареями конденсаторов (БК) [2].

Имеет значение, по мнению авторов, решение задачи изготовления катушек индуктивности с повышенным коэффициентом мощности путем использования специального способа его намотки. Решение такой задачи позволит улучшить энергоэффективность городских распределительных электрических сетей, так как в разных электрических устройствах сети функционирует множество катушек индуктивностей.

Кроме того, авторами ставится задача использования таких катушек, как диагностические устройства для оценки изменения гармонического состава тока нагрузки в проводах, питающие нагрузки, когда количественно соотношения линейных и нелинейных нагрузок изменяются случайным образом.

Вышесказанное связано с тем, что за последние годы возросла нагрузка жилых и административных зданий. С ростом нагрузки изменился и ее характер. Возросла доля электроприемников, таких, как компьютеры, телевизоры, СВЧ-печи, все активнее заменяются обычные лампы накаливания энергосберегающими, и внедряется частотно регулируемый электропривод. Несмотря на относительно небольшую мощность каждого отдельного электроприемника, их массовое применение приводит к существенному искажению синусоидальности кривых напряжения и тока в распределительных электрических сетях [1, 3 -11].

Несинусоидальность напряжений и токов оказывает влияние на работу электрооборудования, сокращая срок его службы, приводит к увеличению потерь напряжения и мощности в сети, уменьшению ее пропускной способности. Это создает условия для возникновения дополнительного риска отказа оборудования и роста электротехнического и технологического ущерба [1].

На сегодняшний день нормативно-правовая документация в области качества электрической энергии регламентирует уровни искажений не только по напряжению, но и по току. Отсутствуют также методики оценки воздействия несинусоидальности тока на работу электротехнического оборудования [3, 4].

В связи с этим представляется необходимым разработать способы оценки изменений спектрального состава тока нагрузки, а также влияния высших гармоник тока, создаваемых работой нелинейных электроприемников жилых и общественных зданий на оборудование распределительной электрической сети.

Постановка задачи

Представить способ и соответствующее устройство для оценки изменений спектрального состава тока нагрузки, т.е. высших гармоник тока, создаваемых работой нелинейных электроприемников.

Рассмотреть вопрос использования катушки с новым способом намотки для повышения ее коэффициента мощности.

Решение задачи

Аппаратная реализация для оценки изменения спектрального состава тока нагрузки от переходных процессов, которые возникают от всех возможных изменений нагрузок распределительной сети, реализована с помощью прибора специальной конструкции в виде датчика тока (СКДТ). Конструктивно СКДТ представляет катушку индуктивности (КИ), у которой магнитопровод имеет тороидальную форму и на которую намотана обмотка из пучка трех изолированных друг от друга эмалированных проводов (рис.1). В результате получаются три независимые обмотки с равным количеством витков. На рисунке 1 схематично показан ход намотки, где n_1 , n_2 , n_3 начало соответственно первой, второй и третьей обмотки, а k_1 , k_2 , k_3 соответствующие концы этих обмоток. Количество

витков W КИ должно быть достаточным, чтобы обеспечить соответствующую чувствительность датчика, который может быть уточнен в процессе эксперимента.

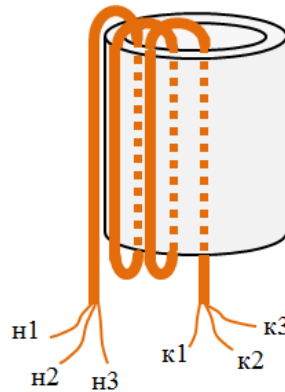


Рисунок 1 – СКДТ: $n1, n2, n3$ – начало соответственно первой, второй и третьей обмотки,
а $k1, k2, k3$ – соответствующие концы этих обмоток

Схема подключения СКДТ к соответствующим фазам физической модели исследуемой сети с устройством автоматического управления экспериментом (АУЭ) в данной работе не проводится.

На рисунке 2, *a, b, c* показана развернутая схема подключения СКДТ к конкретной фазе, исследуемой линий электропередачи в двух исполнениях.

Описание функционирования СКДТ на примере рисунка 2, *a*

На вход первой катушки СКДТ сигнал подается с выхода шунта 1, а на вход второй катушки СКДТ с выхода шунта 2 после предварительной фильтрации с помощью цифрового фильтра (ЦФ). ЦФ предназначен, чтобы на его выходе напряжение представляло чистую синусоидальную зависимость с частотой 50 Гц (первая гармоника). Заметим, что полярности напряжений u_1 и u_2 , подаваемые на входы первой и второй обмоток СКДТ, противоположны из-за особенности схемы подключений (рис.2, *a*), контакты выключателя Bk_2 разомкнуты.

Предположим, что в спектральном составе тока нагрузки $i(t)$ отсутствуют высшие гармоники и интергармоники из-за характера нагрузки, а шунты $\Pi 1, \Pi 2$ и фильтр спроектированы так, что действующие значения напряжений u_1 и u_2 равны. При таком предположении напряжения сигнала на выходе третьей обмотки СКДТ будут равны нулю, так как направления токов, протекающие по первой и второй обмоткам, из-за действия напряжений u_1, u_2 противоположны, а, следовательно, магнитное поле в сердечнике отсутствует.

Допустим, что характер нагрузки изменился из-за подключения нелинейной нагрузки Z_2 , т.е. контакты выключателя Bk_2 замкнуты. В этом случае на выходе третьей обмотки СКДТ возникнет сигнал с ненулевым напряжением. Это связано с тем, что в спектральном составе напряжения u_1 выхода первого шунта, кроме основной первой гармоники, появятся высшие не скомпенсированные гармоники, напряжения u_1 .

Магнитные поля, вызванные токами от действия напряжений, не скомпенсированной части гармоник, напряжения u_1 , наведут в третьей обмотке напряжения, сигнализируя о факте изменения характера переходного процесса в линии.

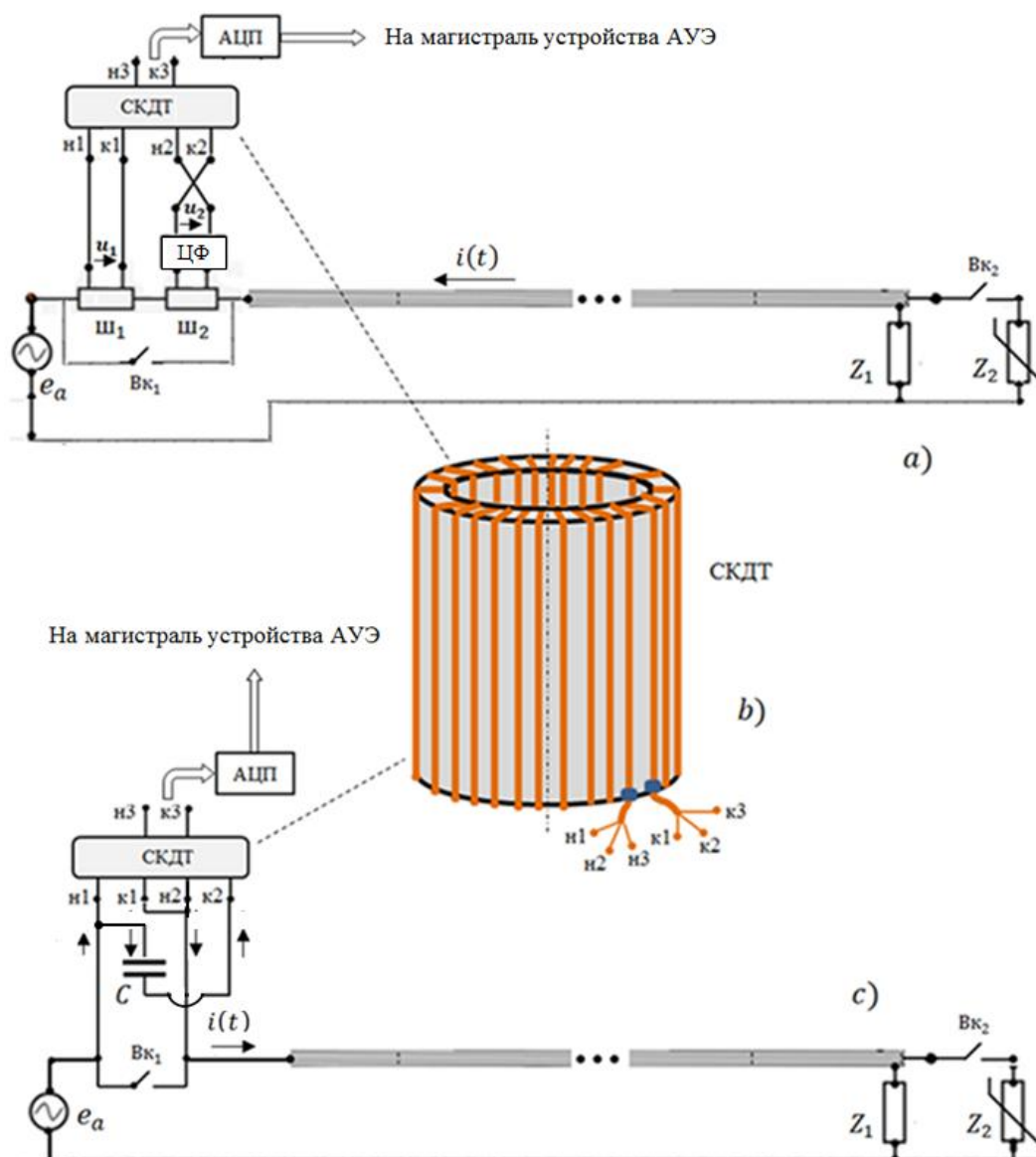


Рисунок 2 – *a, c* развернутые схемы подключения СКДТ к одной фазе, исследуемой провода электропередачи, в двух исполнениях; *b* – изображение СКДТ

Обоснования выше реализуем по результатам численного эксперимента, которые были получены на компьютерной модели СКДТ в среде SimPowerSystems.

В качестве компьютерного аналога СКДТ в среде SimPowerSystems можно предложить блок взаимной индуктивности Mutual Inductance (рис.3), предназначенный для моделирования катушек или проводников, имеющих магнитную связь [12].

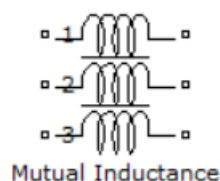


Рисунок 3 – Пиктограмма взаимной индуктивности Mutual Inductance

Параметры блока: собственные сопротивления и индуктивности первой, второй и третьей обмоток; взаимное сопротивление и индуктивность обмоток.

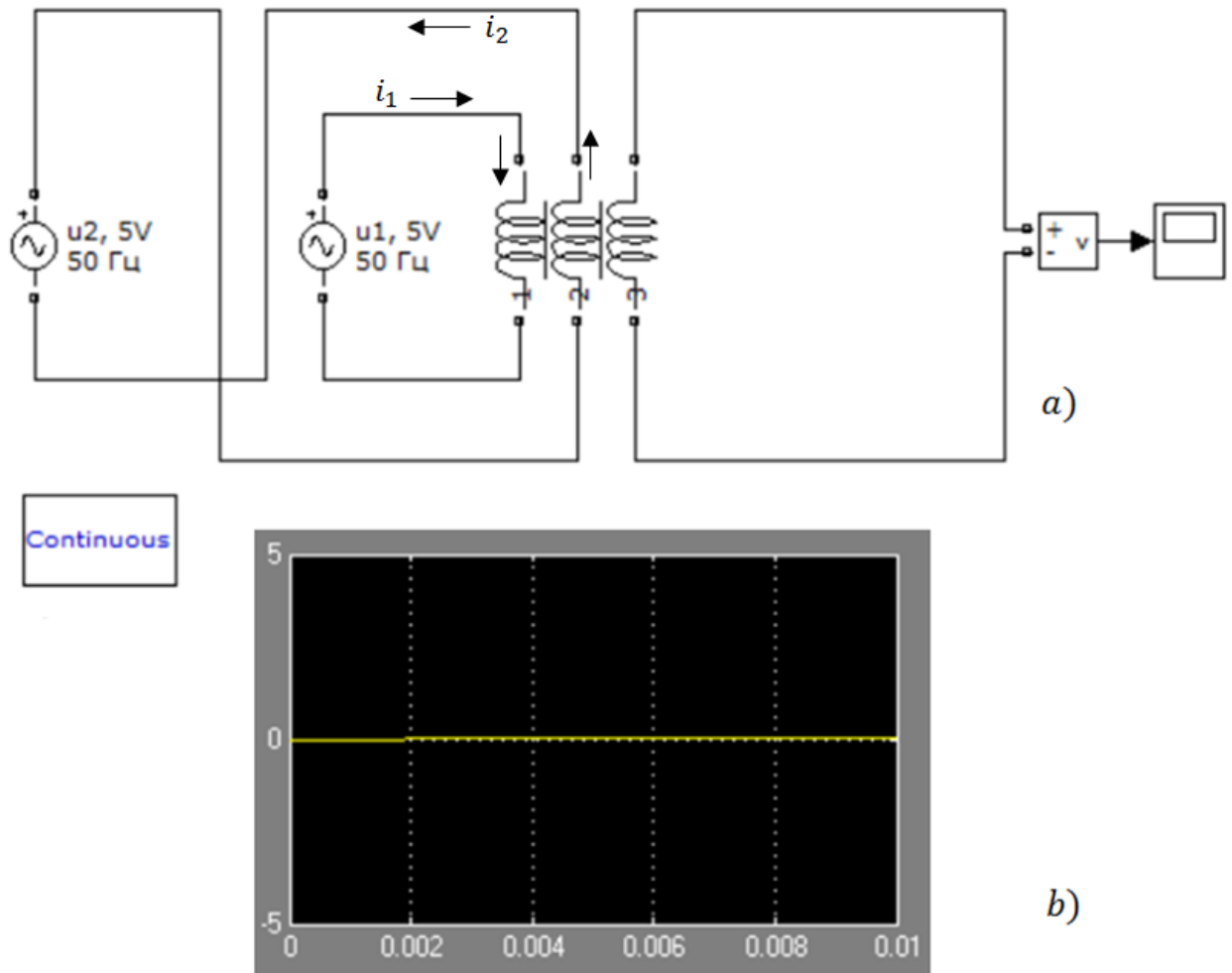


Рисунок 4 – Электрическая схема для исследования сигнала на выходе третьей обмотки СКДТ – *a*. Результат численного эксперимента – *b*

Электрическая схема для исследования сигнала на выходе третьей обмотки СКДТ при подаче сигналов на входы первой и второй обмоток показана на рисунке 4, *a*.

Предположим, что на вход первой обмотки подается сигнал напряжением 5 В, частотой 50 Гц с выхода первого шунта, а на вход второй обмотки подается сигнал напряжением также 5 В и частотой 50 Гц с выхода второго шунта после предварительной фильтрации. Так как по первой и второй обмоткам СКДТ протекают равные по амплитуде и частоте токи, то ток нагрузки не искажен, поэтому отсутствуют высшие гармоники и интергармоники. Из-за противоположности направления токов в первой и второй обмотках на третьей обмотке не наводится сигнал, который подтверждается численным экспериментом (рис. 4, *b*).

Рассмотрим случай, когда в спектральном составе тока нагрузки возникла, например, пятая гармоника (рис.5, *a*). В результате действия фильтра на выходе первого шунта возникнет сигнал, соответствующий первой гармонике пусть напряжением 5 В, тогда как на выходе второго шунта появится сигнал, соответствующий первой гармонике напряжением 5 В, и на него накладывается сигнал пятой гармоники с амплитудой например 1 В (из-за отсутствия фильтра). В результате на выходе третьей обмотки СКДТ появится сигнал от некомпенсированной части тока пятой гармоники. Это подтверждено

результатом численного эксперимента (рис.5, *b, c*).

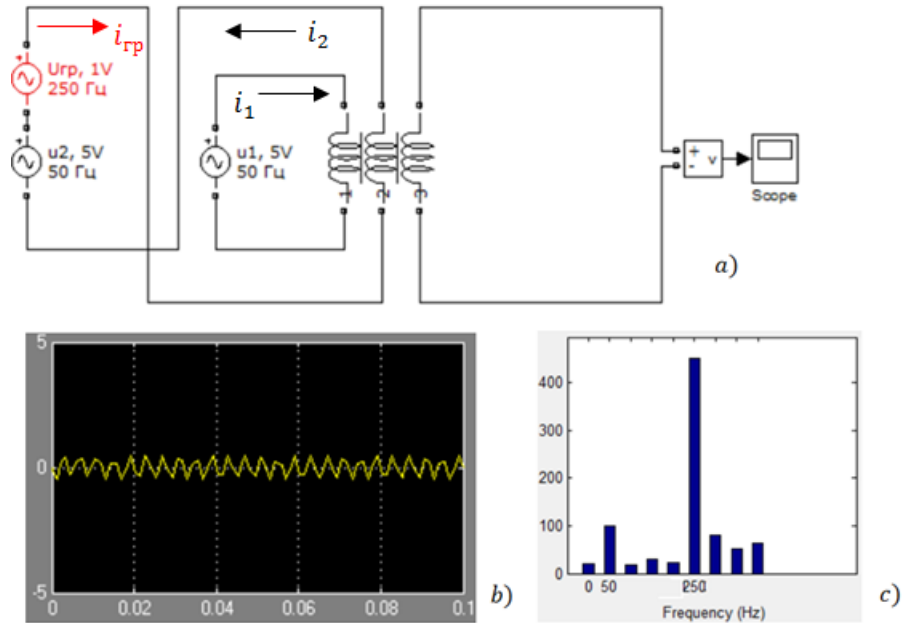
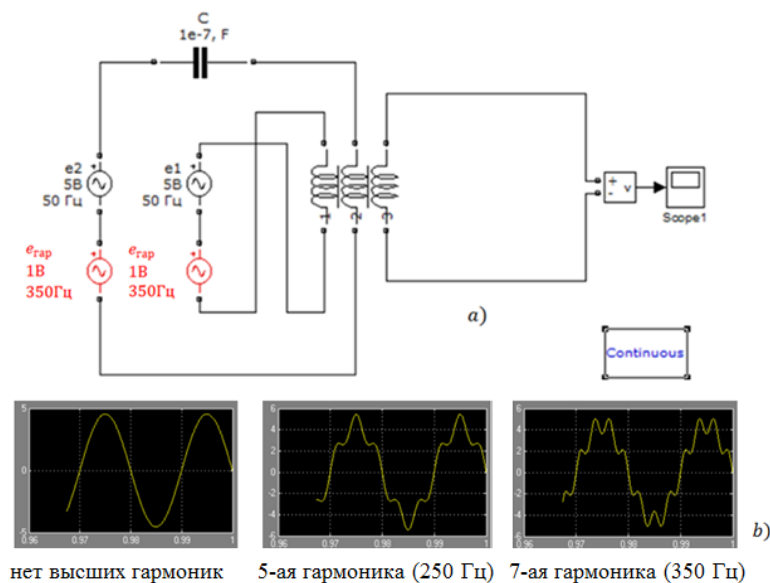


Рисунок 5 – Электрическая схема для исследования сигнала на выходе третьей обмотки СКДТ – *a*, при искаженном сигнале тока нагрузки. Результат численного эксперимента – *b, c*

Описание функционирования СКДТ на примере рисунка 2, *c*

Как видно из схемы подачи сигналов, на входы двух обмоток СКДТ, в отличие от выше рассмотренного случая (рис.4, *a*), ток нагрузки $i(t)$ протекает по виткам обмоток в следующей последовательности. Сначала ток $i(t)$ входит в начало ($n1$) первой обмотки, далее с выхода ($k1$) этой обмотки ток направляется по проводу в нагрузку, при этом конец первой обмотки $k1$ имеет соединение с началом второй обмотки. В то же время начало первой обмотки $n1$ соединено с концом второй обмотки $k2$ с помощью конденсатора C . В результате ток, проходящий через конденсатор C , входит в конец второй обмотки $k2$, пройдя вторую обмотку, с выхода $n2$ направляется по проводу в нагрузку. Такой способ соединения подходит, когда ток нагрузки соответствует сечению проводов обмотки.



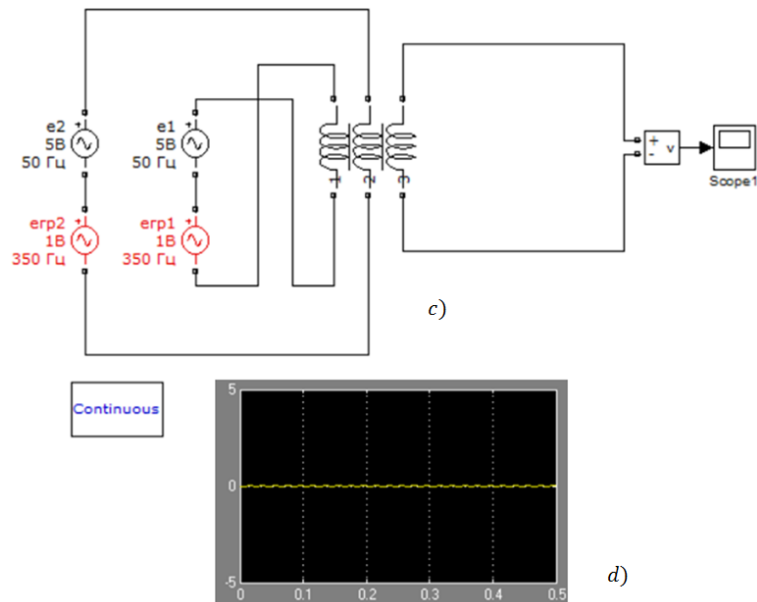


Рисунок 6 – Электрические схемы для исследования сигнала на выходе третьей обмотки СКДТ, с фильтром в виде конденсатора C – *a* и без фильтра – *c*

Компьютерная модель рассмотренного выше случая функционирования СКДТ представлена на рисунке 8, *a*. На рисунке 8, *b* представлены результаты численного эксперимента для трех случаев: первый соответствует, когда в спектральном составе тока нагрузки отсутствуют высшие гармоники; второй – в спектральном составе содержится 5-я гармоника и третий – содержится 7-я гармоника.

Как видно из рисунка 8, *d*, в случае отсутствия фильтра на рисунке 8, *c* в виде конденсатора датчик не реагирует на изменение спектрального состава тока нагрузки, что доказывает необходимость включения конденсатора.

Способ изменения характера сопротивления катушки индуктивности с индуктивного на емкостное и обратно в отдельные моменты времени

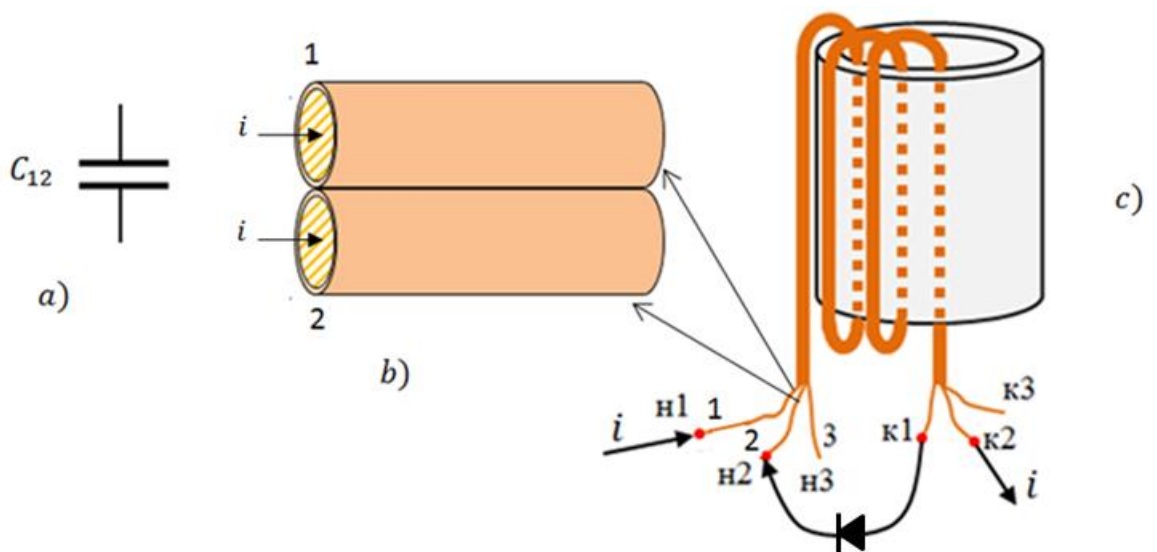


Рисунок 7 – Способ соединения концов двух проводов 1 и 2 обмотки КИ (*c*), когда конец провода 1 (к1) обмотки соединен с началом провода 2 (н2) через диод

На рисунке 7 показан способ соединения концов двух проводов 1 и 2 обмотки КИ (с) с помощью диода, приводящего к изменению характера нагрузки с индуктивного на емкостный и обратно. При этом конец первого провода $k1$ соединен с началом второго провода $n2$ обмотки с помощью диода, в результате по проводам 1, 2 обмотки протекает один и тот же ток i в одинаковом направлении (рис.7, б). Изменение характера нагрузки катушки с индуктивного на емкостный связано с электрической емкостью C_{12} между проводами 1 и 2 обмотки, который естественным образом образует эквивалентную емкость C_{12} в замене на индуктивности катушки в отдельные моменты времени. При этом емкость C_{12} вычисляется по формуле [13]:

$$C_{12} = \frac{\varepsilon l}{4 \log \left(\frac{d}{a} + \sqrt{\frac{d^2}{a^2} - 1} \right)}, \quad (1)$$

где ε – диэлектрическая проницаемость среды, геометрические параметры l, a, d – соответственно, радиус, длина, расстояние между проводами, входящие в выражение (1), представленное на рисунке 8.

В рассматриваемой задаче $2a$ – диаметр токопроводящей жилы эмалированного провода; l – общая длина провода обмотки КИ; $d = 2a + 2\Delta$ (Δ – толщина эмали – изоляции); $\varepsilon = \varepsilon_0 \cdot \varepsilon_r$, где $(\varepsilon_0, \varepsilon_r)$ соответственно электрическая постоянная и относительная диэлектрическая постоянная эмали-изоляции.

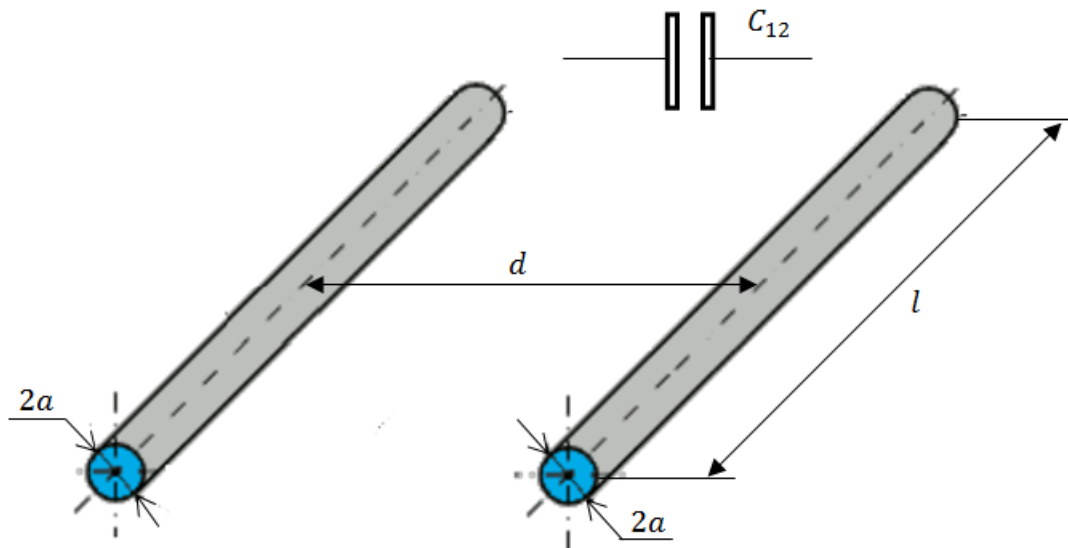


Рисунок 8 – Схема расположения двух параллельно расположенных проводов с указанием их геометрических параметров

Заключение

Катушку индуктивности (КИ) с новым способом намотки можно использовать в качестве устройства диагностики (УД) для оценки изменения спектрального состава тока нагрузки в проводе электропередачи от переходных процессов, возникающих в распределительной сети за счет подключений или отключений обычных нагрузок или нагрузок с нелинейными характеристиками. Показано, как, соединяя концы двух параллельно намотанных проводов с помощью диода, можно изменить характер сопротивления КИ с индуктивного на емкостный и обратно в отдельные моменты времени, за счет емкости образованные между этими проводами. Представлены

компьютерные модели, описывающие работу УД. Полученные результаты моделирования показали работоспособность представленных устройств.

Литература

1. Воротницкий В. Э., Жежеленко И.В., Трофимов Г.Г. Повышение энергетической эффективности электрических сетей / http://www.rtp3.ru/wp-content/uploads/2019/03/27-Povyshenie-energetich.-ef-ti-setej_na-sajt.pdf (дата обращения 20.09.2024).
2. Дрей, Н. М. Компенсация реактивной мощности на промышленных предприятиях с малой установленной мощностью электроприемников / Г.М. Михеев, Н.М. Дрей, А.Г. Зиганшин // Проблемы и перспективы развития энергетики, электротехники и энергоэффективности: IV Междунар. науч.-техн. конф. – Чебоксары: Изд-во Чуваш. ун-та, 2020. – С. 61– 67.
3. Тульский, В.Н. Управление качеством электроэнергии в электрических сетях [Текст] / В.Н. Тульский, И.И. Карташев, Н.М. Кузнецов, М.Г. Симуткин, Х.Б. Назиров // Горный журнал. – 2012. – №12. – С.52– 55.
4. Асанов А.К., Демидов А.Д., Симуткин М.Г., Тульский В.Н., Шведов Г.В. Электромагнитная совместимость систем электроснабжения и электроприемников жилых и общественных зданий // Энергосбережение – теория и практика: Труды Шестой международной школы-семинара молодых ученых и специалистов. – М.: Изд-ский дом МЭИ, 2012. – С. 279 – 283.
5. Боярская Н. П. Качество электроэнергии в городских распределительных сетях 0,4 кВ / Н. П. Боярская, С. А. Темербаев, В. П. Довгун, Я. С. Кунгс, А. Ф. Синяговский // Энергоэффективность систем жизнеобеспечения города 21-22 ноября 2012 г.: Материалы XIII Всероссийской научно-практической конференции. – Красноярск, 2012. – С. 113–121.
6. Дрей, Н. М. Гармонический анализ тока замыкания на землю в переходном режиме / М. Н. Атаманов, Н. М. Дрей, Г. М. Михеев, А. Н. Пулин // Проблемы и перспективы развития энергетики, электротехники и энергоэффективности: Материалы II Международной научно-технической конференции, Чебоксары, 18–19 октября 2018 года. – Чебоксары: Чувашский государственный университет имени И.Н. Ульянова, 2018. – С. 137–144.
7. Турукина, Т.Е. Алгоритм регулирования уровня несимметрии в распределительных сетях / Т.Е. Турукина, Е.А. Шутов, И.И. Елфимов, В.К. Сибирцев // Промышленные АСУ и контроллеры. – 2016. – № 10 . – С. 24–30.
8. Куско А. Качество энергии в электрических сетях / Пер. с англ. – М.: Додэка-XXI, 2008. – 336 с.
9. Akagi H. // Proceedings of the IEEE. Vol. 93, 2005, No. 12. P. 2128–2141.
10. Hu C-H., Wu C-J., Chen Y-W. // IEEE trans. on Power delivery. 1997. Vol. 12. No. 3. P. 1275–1281.
11. De Lima Tostes M., BezerraU., Silva R. // IEEE transactions on power delivery. 2005. Vol. 20. No. 1. Pp. 384–389.
12. Черных И.В. Моделирование электротехнических устройств в MATLAB, SimPowerSystems и Simulink. – М.: ДМК Пресс; СПб.: Питер, 2008. – 288 с.: ил.
13. Дж. Джексон. Классическая электродинамика, – М.: Мир, 1965. – 263 с.: ил.

УДК 624.131.437.311

*Н.К. Джусупбекова*¹, nazika11612@mail.ru*А.К. Асанов*¹, asanov_ak@mail.ru*Р.Б. Тентиев*¹, renattentiev@mail.ru¹*Кыргызский государственный технический университет им. И. Раззакова, Кыргызстан, 720044, г. Бишкек, пр. Мира, 66*

ЭКСПЕРИМЕНТАЛЬНО-РАСЧЕТНОЕ ОПРЕДЕЛЕНИЕ ЭЛЕКТРОМАГНИТНОЙ ОБСТАНОВКИ ПС 110 КВ «РАЗЗАКОВА»

Статья содержит экспериментально-расчетное определение электромагнитной обстановки ПС 110 кВ «Раззакова», расположенной в Баткенской области. На основе измеренных данных произведен расчет и моделирование параметров электромагнитной совместимости основного оборудования с системами вторичной коммутации. Произведены измерения удельных электрических сопротивлений грунта и сопротивления растеканию заземляющего устройства, на основе которых при коротких замыканиях и молниевом разряде определены опасности воздействия разностей потенциалов на вторичное оборудование. Произведен расчет среднего количества разряда молнии на территорию объекта. Произведены анализ и оценка состояния заземляющего устройства, сечений и коррозионного состояния элементов существующего заземляющего устройства, размещений вторичных цепей. Определены наличие и качество металlosвязей оборудования с заземляющим устройством объекта. Оценки результатов расчета сопоставлены с требованиями нормативно-технических документаций.

Ключевые слова: заземляющее устройство, удельное электрическое сопротивление грунта, сопротивление растеканию заземляющего устройства, вторичные цепи, короткое замыкание, импульсные помехи, молниезащита.

Введение. Система заземления и уравнивания потенциалов является важной частью современного энергообъекта. Характеристики ЗУ должны отвечать требованиям обеспечения электробезопасности обслуживающего персонала и надежной работы оборудования электроустановки в нормальных и аварийных условиях. ЗУ должно обеспечивать следующие эксплуатационные функции электроустановки:

- защиту людей от поражения напряжением шага и напряжением прикосновения при КЗ в электроустановках;
- действие релейных защит от замыкания на землю;
- действие защит от перенапряжений – разрядников, ОПН;
- отвод в грунт токов молнии с молниеприемников объекта и подходящих к нему тросовых молниеотводов;
- отвод рабочих токов (токов нулевой последовательности, тяговых токов и т.д.);
- снижение электромагнитных влияний на вторичные цепи и оборудование, в том числе импульсных перенапряжений при ударах молнии, протекании через ЗУ ВЧ-составляющей тока КЗ и т.п.;
- защиту проводящих коммуникаций (экранов и оболочек кабелей, труб воздухопроводов и т.п.) от токовых перегрузок;
- защиту от статического электричества;
- снижение разностей потенциалов, приложенных к изоляции цепей связи, уходящих с территории объекта – при КЗ на землю и молниевых разрядах.

В современных условиях состоянию заземляющего устройства и системы молниезащиты должно уделяться повышенное внимание, поскольку они оказывают существенное влияние на электромагнитную обстановку на объекте. Устройства РЗА, АСУ ТП, АСКУЭ и связи, основанные на микропроцессорных элементах, имеют широкие функциональные возможности и ряд других преимуществ перед электромеханическими устройствами. Однако, в отличие от них, современные устройства обладают сравнительно высокой чувствительностью к электромагнитным помехам. Энергообъекты (ПС, ТЭЦ, ГЭС и т.п.) являются мощными источниками электромагнитных полей и помех, поэтому

для нормального функционирования современных устройств защиты и автоматики необходимо обеспечивать их электромагнитную совместимость (ЭМС) с электромагнитной обстановкой (ЭМО) на энергообъектах. Система заземления играет важную роль в обеспечении ЭМС аппаратуры защиты, автоматики, управления, связи.

Таким образом, необходимо проводить обследование систем заземления и уравнивания потенциалов на действующих энергообъектах для обеспечения электробезопасности и ЭМС. По результатам должны разрабатываться мероприятия по модернизации заземляющих устройств.

Целью настоящей работы является обследование существующего контура заземляющего устройства ПС 110/35/10 кВ «Раззакова» с выявлением нарушения электромагнитной обстановки в части влияния на надежность работы первичного оборудования и систем вторичной коммутации.

Краткие сведения об объекте. Подстанция расположена рядом с с. Равват Лейлекского района Баткенской области. ПС 110/35/10 «Раззакова» представляет собой понижающую подстанцию, осуществляющую прием, преобразование и передачу электрической энергии. ПС эксплуатируется с постоянным дежурным персоналом на объекте.

Измерения производились на объекте в мае 2024 г. При выполнении работ использовались следующие документы [3]:

- нормальная схема электрических соединений ПС «Раззакова» на 2024 год;
- журнал уставок РЗА;
- документ «Токи короткого замыкания на шинах 6–500кВ подстанций Национальной электрической сети Кыргызстана (НЭСК) в максимальных и минимальных режимах по состоянию на 2023 г.».

В районе расположения ПС уровень грозовой активности составляет ~40–60 часов в год [2]. Среднестатистическая поражаемость грозой на территории составляет 35/100км² год. Рядом с территорией ПС находятся опоры ВЛ 500, 110 кВ и 10 кВ.

ПС имеет три распределительных устройства напряжением выше 1 кВ: ОРУ-110 кВ, выполнено по схеме: две рабочие системы шин и ремонтная перемычка; ОРУ – 35 кВ, РУ-10 кВ, выполнено в виде КРУН-10 кВ с двумя секциями шин. На ПС установлен один силовой трансформатор мощностью 16 МВА 110/35/10 кВ.

Размещение вторичных цепей

На открытой территории ПС (ОРУ-110 кВ, трансформатор, участок между ОПУ и ОРУ-35кВ) цепи проложены в железобетонных незаглубленных кабельных лотках (рис. 1). Трассы прокладки лотков указаны на рисунке 2. Часть цепей от ЭА до лотков выполнена в виде металлокоробов, в металлических трубах, а в некоторых отсутствует защитный покров цепей. Силовые кабели и контрольные проложены в лотках хаотично, без разнесения на существенное расстояние.



Рисунок 1– Фрагмент трассы кабельных лотков на ОРУ 110 кВ ПС «Раззакова»

Участки цепей от электроаппаратов до клеммных шкафов выполнены, как правило, в металлических трубах, металлокоробах или в ПВХ-кожухе. Клеммные шкафы электроаппаратов расположены непосредственно на стойках ЭА, к которым они относятся.

Вторичные цепи в здании ОПУ и ОРУ 35 кВ проложены в кабельных каналах и открыто по воздуху. В настоящее время на подстанции применяется МП-аппаратура АСКУЭ и связи.

Результаты измерений и расчетов

Определение состояния ЗУ

В связи с отсутствием проектной и исполнительной схемы ЗУ ОРУ-110 кВ и ОРУ-35кВ, с учетом измеренного значения сопротивления растеканию ЗУ, была построена адекватная модель ЗУ (рис. 2) с помощью ПО «Контур» [11]. Также были использованы исходные данные, которые показаны в таблице 1 [2].

Таблица 1

№	Измерения	Данные
1	Площадь ЗУ ПС	13500 м ²
2	Однослойный грунт с удельным сопротивлением	98 Ом*м
3	Длина пролета металлических опор напряжением 110 кВ	250 м
4	Стальной грозозащитный трос сечением	50 мм ²
5	Ток короткого замыкания в сети напряжением 110 кВ	$I_{кз}=6,8$ кА
6	Сопротивление заземления опор по ПУЭ 2.5.129	$RE1=10$ Ом
7	Вертикальный стержневой электрод длиной	$Lв=3$ м
8	Горизонтальный заземлитель	30 шт.
9	Вертикальный заземлитель	70 шт.
10	Сведения о грунте Протокол 1 от 10 мая 2024 года	

Растекание большей части тока КЗ и молниевое разряда происходит по ЗУ ПС и частично по системе «грозотрос» опоры ВЛ 110 кВ.

Оценка удельного сопротивления грунта

Измерения удельного сопротивления грунта проводились как на территории ПС, так и за ее пределами по методу вертикального электрического зондирования. Глубина промерзания грунта в зимний период составляет не более 2,0 м (согласно карте нормативных глубин промерзания грунтов) [1].

Согласно результатам измерений и аппроксимации, проведенной в соответствии с [7], данным [1, 6], принимаются следующие параметры эквивалентной модели грунта:

- летний период: двухслойный грунт с удельным сопротивлением 90 Ом*м (верхний слой) и 180 Ом*м (нижний слой). Глубина раздела слоев – 15 м;

- зимний период: двухслойный грунт с удельным сопротивлением 135 Ом*м (верхний слой) и 180 Ом*м (нижний слой). Глубина раздела слоев – 15 м.

Определение сопротивления растеканию ЗУ

Сопротивление растеканию является одной из главных характеристик ЗУ-электроустановок. Сопротивление растеканию измерялось с помощью классического метода «амперметра-вольтметра» интегрированным прибором MRU-120. Токовый и потенциальный зонды были вынесены за территорию ПС на расстояния порядка 1000 и 700 метров соответственно. Данные расстояния примерно в 3 раза превышают размер ЗУ-ПС по диагонали.

Измеренное сопротивление растеканию с учетом естественных заземлителей составляет 0,89 Ом.

На основании результатов измерений был проведен расчет распределения токов по отходящим коммуникациям. Распределение токов имеет следующий вид:

- 69% тока растекается с ЗУ ПС;
- 31% токов растекается по тросовым молниеотводам;

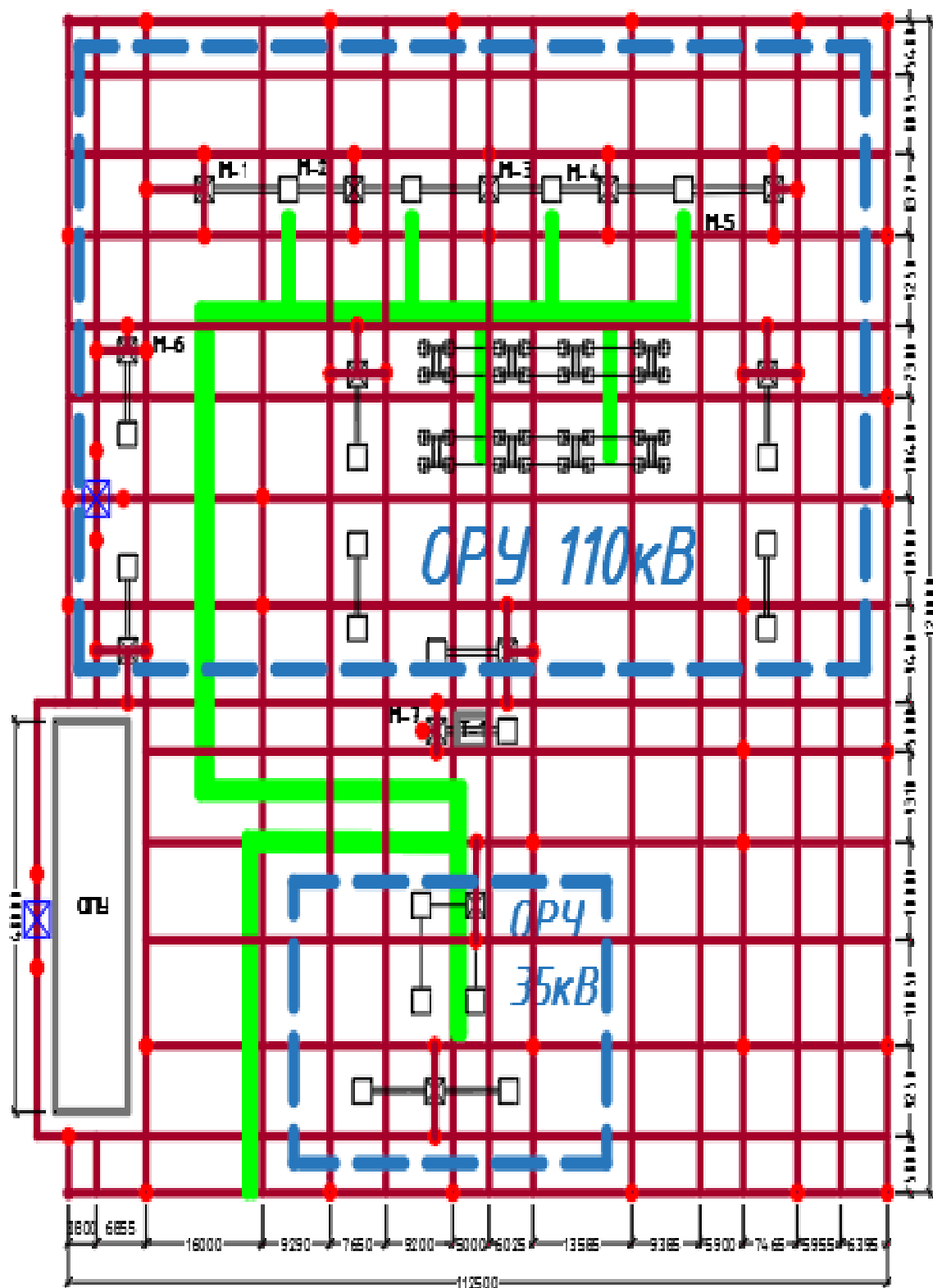


Рисунок 2 – Модель ЗУ ПС 110 кВ «Раззакова»

По результатам расчетов сопротивление растеканию ЗУ-объекта в зимний период составит около 1,1 Ом.

Согласно [1], сопротивление растеканию ПС не должно превосходить 0,5 Ом в любое время года с учетом естественных заземлителей [1, п.1.7.90]. Таким образом, сопротивление растеканию ЗУ ПС не соответствует требованиям НТД.

Расчет среднего количества разрядов молнии на территорию объекта за год

Поражение объекта молнией носит вероятностный характер и зависит от характеристик грозовой активности в регионе, геометрических параметров рассматриваемого объекта и характеристик окружающей местности. Поэтому выбор принимаемого в расчетах значения тока молнии будет определяться ожидаемым средним числом поражений объекта за год.

Расчет количества молниевых разрядов выполняется в соответствии с методикой [9].

Уровень грозовой активности (рис. 3, 4) в районе расположения объекта [2]: $T_d = 40 - 60$ часов в год.

Плотность ударов молнии на 1 кв. км в год рассчитывается по формуле:

$$N_g = 6.7 \cdot T_d / 100. \quad (1)$$

Если принять верхнее значение $T_d = 40$ часов в год, то $N_g = 4,02$.

Ожидаемое количество разрядов молнии в защищаемую территорию за год определяется по формуле:

$$ND = N_g \cdot Ad \cdot Cd \cdot 10^{-6}. \quad (2)$$

Здесь Ad – площадь сбора разрядов для рассматриваемой территории, Cd – коэффициент, учитывающий влияние относительного местонахождения защищаемого объекта [9]. Вокруг ПС отсутствуют объекты равной или большей высоты, поэтому $Cd = 1,0$.

Ожидаемое количество поражений молнией территории открытой части объекта определяется согласно описанной методике. Принимается: $Ad = 32200$ кв.м.

Для принятой грозовой активности $ND = 0,129$, что соответствует примерно 1 удару в ~8 лет.

По данным [2], в которых приведена эмпирическая формула для оценки вероятности молниевых разрядов с той или иной величиной тока молнии, молниевые разряды с амплитудой тока 15 кА и выше будут составлять 15% от общего числа молниевых разрядов.

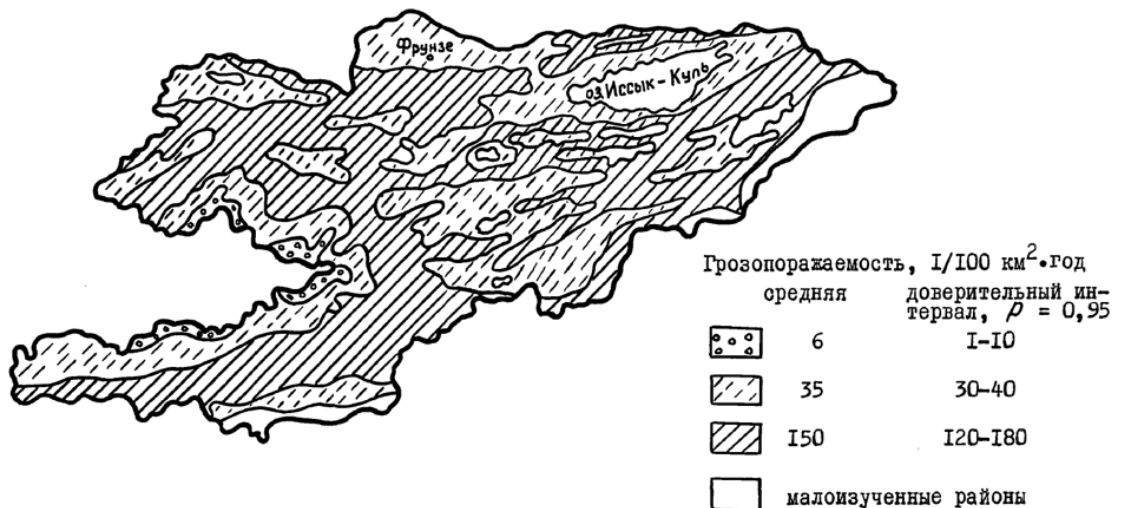


Рисунок 3 – Карта удельной грозопоражаемости территории Кыргызстана

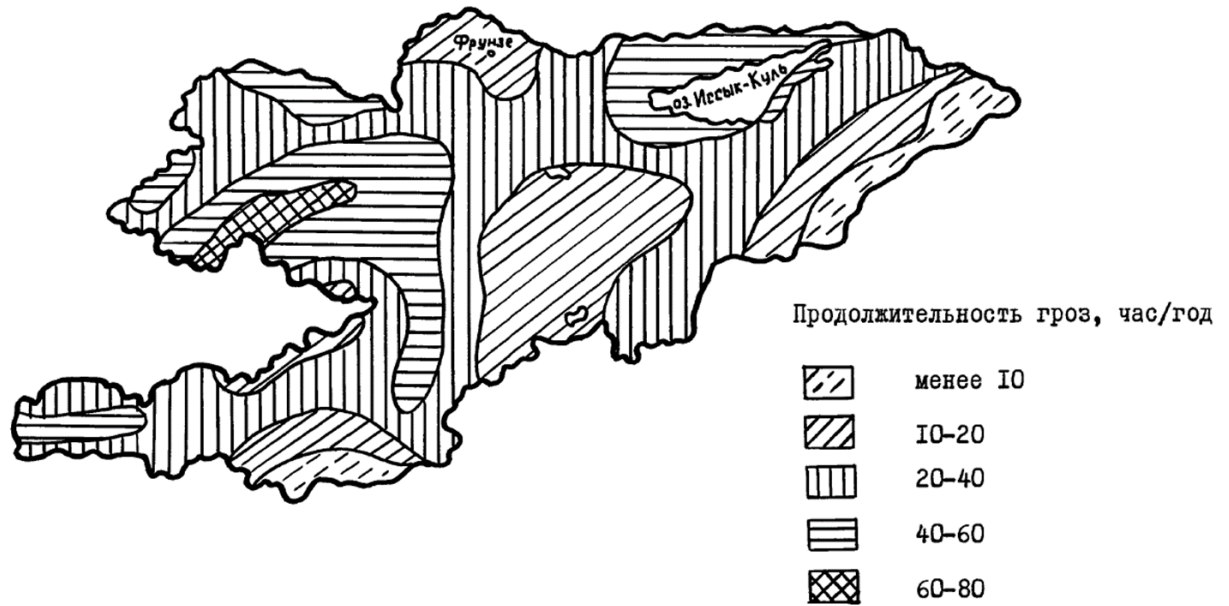


Рисунок 4 – Карта среднегодовой продолжительности гроз на территории Кыргызстана

Расчет импульсного сопротивления заземлителей

Рассчитываем импульсное сопротивление заземлителей $R_{\text{имп}}$. Оно зависит от импульсного коэффициента a_n

$$a_n = \sqrt{\frac{1500 \cdot \sqrt{S}}{(\rho + 320)(I_M + 45)}} = \sqrt{\frac{1500 \cdot \sqrt{13500}}{(98 + 320)(15 + 45)}} = 2,6,$$

где $I_M = 15 \text{ кА}$ – ток молнии, $\rho = 98 \text{ Ом/м}$ – эквивалентное удельное сопротивление грунта, $S = 13500 \text{ м}^2$ – площадь заземлителя.

Тогда импульсное сопротивление заземлителей $R_{\text{имп}}$

$$R_{\text{имп}} = a_n \cdot R_{\text{зг}} = 2,6 \cdot 0,89 = 2,3 \text{ Ом}.$$

Анализ системы молниезащиты по условиям воздействия на вторичное оборудование

При ударах молнии в молниеотводы объекта возможно воздействие импульсных разностей потенциалов и наводок на вторичные цепи. Кроме того, опасность может представлять стекание импульсных токов с заземления высоковольтных ОПН. С целью выявления средств молниезащиты, протекание тока молнии через которые представляет опасность для вторичного оборудования, произведен предварительный визуальный осмотр системы молниезащиты. Список элементов системы молниезащиты, протекание тока молнии через которые представляет наибольшую опасность для первичного оборудования и систем вторичной коммутации, представлен ниже (см. табл. 2).

Таблица 2. Элементы системы молниезащиты, протекание тока молнии через которые представляет наибольшую опасность для вторичных цепей

Тип цепей	ЭСМЗ, представляющие наибольшую опасность	Обоснование
Цепи в ОПУ	М1	Расстояние элемента СЗМ до ОПУ не более 1,5 м

Заключение

Электромагнитную обстановку на ПС 110кВ «Раззакова» можно считать менее благоприятной. Наибольшую опасность представляют на промышленной частоте помехи при КЗ на ОРУ 110 кВ и импульсные помехи при молниевых разрядах.

1. Схема ЗУ на открытой части ПС (ОРУ-110, ОРУ-35, РУ-10 и трансформаторная группа) выполнена по требованию НТД.
2. Схема СУП в помещении ОПУ удовлетворяет требованиям НТД.
3. Площадь сечения заземлителей и заземляющих проводников удовлетворяет требованиям НТД.
4. Удельное сопротивление грунта в месте расположения объекта является вполне благоприятным по условиям растекания тока с заземлителей.
5. Сопротивление растеканию ЗУ ПС не удовлетворяет требованиям ПУЭ. Значительный вклад в растекание тока вносят искусственные заземлители, также растекание будет по тросовым молниеотводам.
6. На открытой части объекта (ОРУ-110, 35 кВ) не выявлена конструкция с нарушением локальной металлосвязи с ЗУ ПС.
7. Разности потенциалов на промышленной частоте при КЗ в сетях выше 1 кВ будут представлять опасность для вторичных цепей.
8. Помехи и перенапряжения при молниевом разряде представляют опасность для вторичных цепей.

Литература

1. Правила устройства электроустановок. –7-е изд. – М.: Изд-во НЦ ЭНАС, 2002.
2. Чичинский М.И. Дисс.... на соискание уч. степени к.т.н. «Особенности грозопоражаемости в горных районах и их учет при выборе грозозащиты воздушных линий электропередачи». – Фрунзе, 1984г.
3. Данные ОАО «Национальная электрическая сеть Кыргызстана» с 2012 по 2023гг.
4. ГОСТ Р 50571-4-44-2011 (МЭК 60364-4-44:2007) Электроустановки низковольтные. Часть 4-44. Требования по обеспечению безопасности. Защита от отклонений напряжения и электромагнитных помех. – М.: Стандартинформ, 2012.
5. ГОСТ Р 51317.6.5-2006 (МЭК 61000-6-5-2001) Совместимость технических средств электромагнитная. Устойчивость к электромагнитным помехам технических средств, применяемых на электрических станциях и подстанциях. Технические требования и методы испытаний.
6. Методические указания по контролю заземляющих устройств электроустановок. РД 153-34.0-20.525-00, М. СПО ОРГРЭС, 2000.
7. Коструба С.И. Измерение электрических параметров земли и заземляющих устройств. – М.: Энергоатомиздат, 1983.
8. Методические указания по определению электромагнитной обстановки и совместимости на электрических станциях и подстанциях. СО 34.35.311-2004. – М.: Изд-во МЭИ, 2004.
9. Методические указания по обеспечению электромагнитной совместимости на объектах электросетевого хозяйства. СТО-56947007-29.240.044-2010.
10. IEC 62305 Lightning Protection (МЭК 62305 Молниезащита).
11. Отчет НИР за 2023г. «Исследования электромагнитной обстановки объектов электроэнергетики и обеспечение электромагнитной совместимости устройств в электрических сетях» // НИИ Энергетики и связи при КГТУ им. И.Раззакова. – Бишкек, 2023. – 68 с.
12. Асанов А.К., Омокеева А.А. Экспериментально-расчетное определение электромагнитной обстановки ПС 110 кВ «Ананьево» // Проблемы автоматики и управления. – Бишкек, 2022. – №3 (45). – С. 42–51.
13. Тентиев Р.Б., Джусупбекова Н.К., Асанов А.К. Экспериментально-расчетное определение электромагнитной обстановки ПС 110 кВ «Парковая». // Проблемы автоматики и управления. –№3 (48). – Бишкек, 2023.– С. 117–123.

УДК 550.832.7

С.Н. Верзунов, verzunov@hotmail.com

К.А. Жолдошбаев

Кыргызско-Российский Славянский университет им. Б.Н. Ельцина

РАЗРАБОТКА ПРИНЦИПИАЛЬНЫХ ЭЛЕКТРИЧЕСКИХ СХЕМ ОСНОВНЫХ ФУНКЦИОНАЛЬНЫХ ЭЛЕМЕНТОВ ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ МОНИТОРИНГА СОСТОЯНИЯ ВОДЫ И ПОЧВ СЕЛЬСКОХОЗЯЙСТВЕННЫХ УГОДИЙ

В статье представлены разработки и принципы построения интеллектуальной системы мониторинга состояния воды и почв на сельскохозяйственных угодьях. Основное внимание уделено проектированию аппаратной части системы, включающей предусилители, режекторные и универсальные фильтры, аналого-цифровые и цифро-аналоговые преобразователи, а также микрокомпьютеры. Описаны схемотехнические решения и рекомендации по повышению точности измерений в условиях внешних помех. Приведены примеры интеграции оборудования, таких как микрокомпьютер Orange Pi Zero 3 и модуль ADS8688A, для обработки и анализа данных в режиме реального времени. Полученные результаты применимы для автоматизации управления водными ресурсами и улучшения экологического состояния сельскохозяйственных земель.

Ключевые слова: интеллектуальная система мониторинга, диэлектрическая проницаемость, сельскохозяйственные угодья, предусилитель ОРА627, режекторный фильтр, универсальный фильтр UAF42, аналого-цифровое преобразование, цифро-аналоговое преобразование, микрокомпьютер Orange Pi Zero 3, геотехнический мониторинг, метод свободного пространства, системы автоматизации, экология, управление водными ресурсами.

Введение

Эффективное управление состоянием сельскохозяйственных угодий, включая мониторинг воды и почв, является ключевым элементом устойчивого сельского хозяйства. В современных условиях возрастающей нагрузки на экосистемы разработка интеллектуальных систем мониторинга становится неотъемлемой частью решения данной задачи. Такие системы позволяют в реальном времени контролировать параметры окружающей среды, обрабатывать полученные данные и принимать обоснованные управленческие решения.

Создание надежной и точной системы мониторинга требует разработки аппаратной части, способной работать в условиях внешних помех и обеспечивать сбор высокоточных данных. Разработка таких устройств включает в себя проектирование предусилителей, фильтров, аналого-цифровых и цифро-аналоговых преобразователей, а также интеграцию вычислительных модулей, таких как микрокомпьютеры. Применение современных микропроцессорных решений в сочетании с алгоритмами машинного обучения открывает новые возможности для автоматизации управления водными и почвенными ресурсами.

Цель данной статьи – представить подходы к разработке аппаратных компонентов интеллектуальной системы мониторинга состояния воды и почв сельскохозяйственных угодий. В статье рассматриваются методы повышения точности измерений, решения для снижения уровня помех, а также примеры интеграции компонентов системы для реализации комплексного интеллектуального мониторинга в режиме реального времени.

Современные подходы к проектированию таких систем подробно описаны в исследованиях [1–4]. Кроме того, для успешного применения методов машинного обучения в интеллектуальной системе мониторинга состояния воды и почвы сельскохозяйственных угодий (ИСМСВП) необходимо не только корректно собрать и обработать данные, но и обеспечить их точное измерение в режиме реального времени.

Важную роль в этом играет разработка аппаратной части системы (рис. 1), которая включает в себя датчики и устройства обработки сигналов, обеспечивающие точность и достоверность данных для последующего анализа.

Именно на этом этапе особую важность приобретает использование аналоговых компонентов системы, обеспечивающих корректную работу сенсоров в условиях сильных внешних помех.

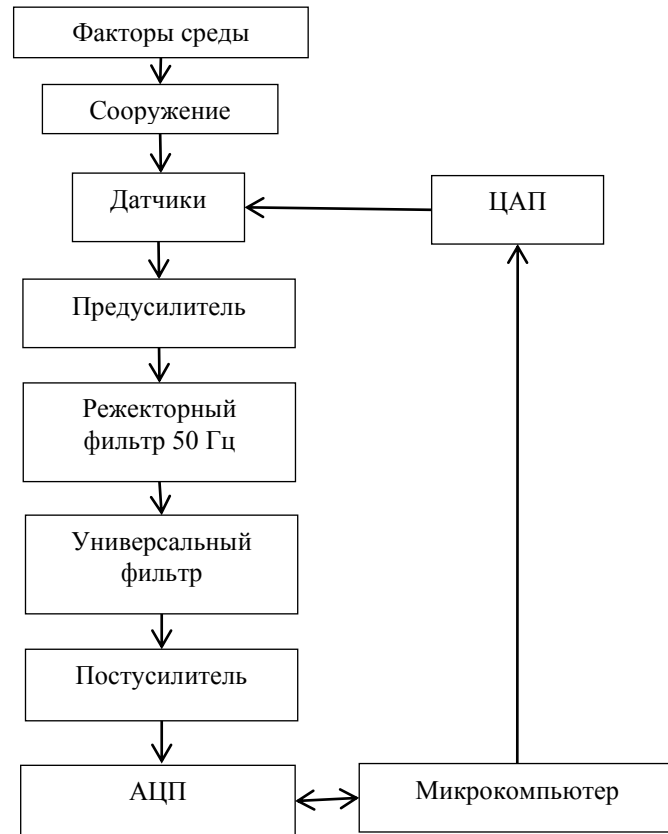


Рисунок 1 – Структурная схема аппаратной части ИСМСВП

Предусилитель, показанный на рисунке 2, является ключевым компонентом в системе сбора и обработки сигналов с датчиков, особенно когда необходимо усилить слабые аналоговые сигналы перед их дальнейшей обработкой. ОРА627 – это высокоточный операционный усилитель, который часто используется для усиления дифференциальных сигналов с датчиков при подавлении синфазных помех.

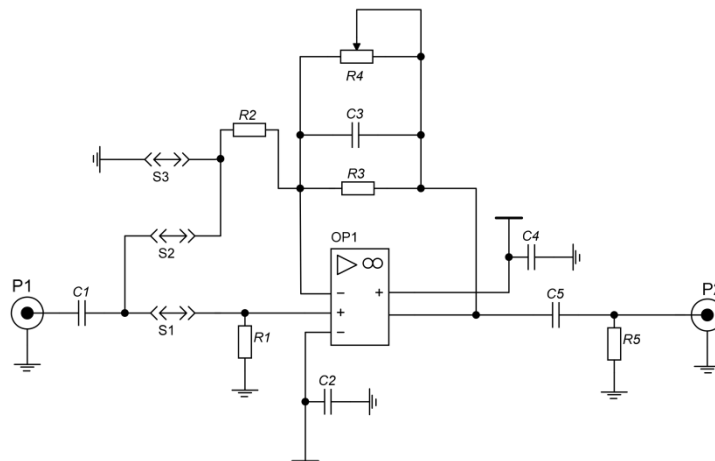


Рисунок 2 – Принципиальная электрическая схема предусилителя на базе ОРА627

Его преимущества включают низкий уровень шума, высокий коэффициент усиления и простоту использования [1]. Основные компоненты схемы, показанной на рисунке 2, включают: выводы питания $+V_s$ (вывод 7) и $-V_s$ (вывод 4) служат для подключения положительного и отрицательного питания соответственно. На вход V_{in+} (вывод 3) подключается сигнал с датчика. С выхода V_{out} (вывод 6) снимается усиленный сигнал. Резистор R_1 , подключенный между выводами 1 и 8, задает коэффициент усиления по формуле:

$$G = 1 + \frac{49.4 \text{ k}\Omega}{R_1}$$

Используется двухполярное питание ± 12 В или ± 15 В для обеспечения симметричного диапазона выходного сигнала. Сглаживающие конденсаторы, 0.1 мкФ – керамические и 10 мкФ – электролитические, между каждым выводом питания и землей служат для уменьшения шумов и пульсаций. Резисторы малого номинала (например, 50 Ω) функционируют последовательно с каждым входом для ограничения входного тока и защиты от перегрузок.

Последовательно входу может быть установлен конденсатор (например, 100 пФ) для формирования фильтра нижних частот и подавления высокочастотных помех. Кроме того, резистор и конденсатор могут быть подключены на выходе для формирования фильтра, ограничивающего полосу пропускания и уменьшающего шум.

Выходной сигнал подан на вход режекторного фильтра для дальнейшей обработки. Резистор R_2 на входе (50–100 Ω) защищает усилитель от токовых перегрузок. Конденсатор C_3 вместе с входными резисторами образуют фильтры нижних частот для подавления высокочастотных шумов. Резистор и конденсатор на выходе (R_5 и C_4) формируют фильтр нижних частот для сглаживания выходного сигнала.

Сглаживающие конденсаторы необходимо разместить как можно ближе к выводам питания ОРА627 для эффективного подавления пульсаций. Все земляные соединения должны быть выполнены с минимальным импедансом. Используется звездообразное заземление, чтобы предотвратить петли заземления. Печатная плата показана на рисунке 3.

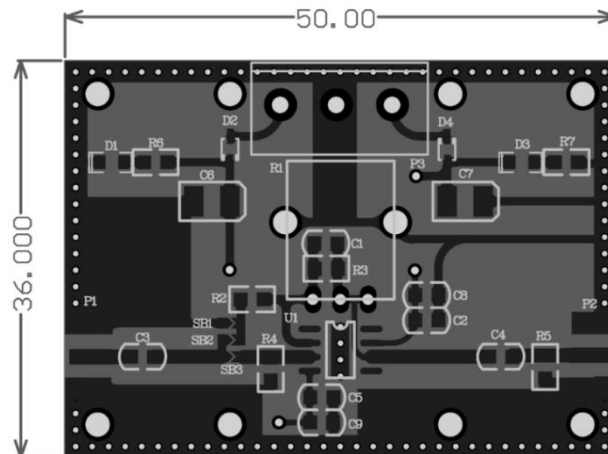


Рисунок 3 – Печатная плата предусилителя на базе ОРА627

Компоненты должны быть размещены максимально компактно, особенно конденсаторы фильтрации питания. Должна быть минимизирована длина проводников для входных сигналов, чтобы уменьшить наводки. Кроме того, необходимо использовать экранированные кабели для подключения датчиков. А также корпус предусилителя нужно поместить в металлический экран, соединенный с землей, для дополнительной защиты от электромагнитных помех. Схема предусилителя, который использует инструментальный усилитель ОРА627, предоставляет эффективное решение для усиления слабых сигналов с датчиков в системах геотехнического мониторинга.

Такая реализация схемы с учетом рекомендаций по питанию, фильтрации и разводке позволит получить высококачественный усиленный сигнал для дальнейшей обработки и анализа.

Режекторный фильтр предназначен для подавления сетевых помех в системах измерения и обработки сигналов. Он реализован на основе активной схемы, которая позволяет регулировать глубину подавления. Это дает возможность эффективно устранять помехи на частоте 50 Гц, не оказывая значительного влияния на полезный сигнал.

Основой конструкции фильтра является активный режекторный фильтр, построенный с использованием двойного Т-образного фильтра и операционного усилителя. Такая конфигурация обеспечивает глубокое подавление в амплитудно-частотной характеристике на частоте 50 Гц, а также высокую стабильность и гибкость в настройках фильтра. Т-образный фильтр состоит из набора резисторов и конденсаторов, рассчитанных на частоту 50 Гц. Операционный усилитель увеличивает сигнал и улучшает характеристики фильтра, обеспечивая возможность регулировки глубины и ширины провала в диапазоне от 6 дБ до 45 дБ.

Входной сигнал поступает на двойной Т-образный фильтр, который состоит из двух цепей: одна из них образована последовательным соединением резисторов и конденсаторов, а другая — параллельным. Эта комбинация формирует глубокий провал на частоте 50 Гц. Компоненты выбираются таким образом, чтобы фильтр находился в резонансе именно на этой частоте.

Операционный усилитель подключен к выходу Т-образного фильтра. Через обратную связь, реализованную с помощью потенциометра, можно регулировать усиление и, следовательно, глубину подавления. Усиленный и отфильтрованный сигнал подается на выход модуля. Эффект режекции достигается за счет резонанса в Т-образной цепи, которая обеспечивает максимальное подавление на частоте 50 Гц. Операционный усилитель компенсирует потери сигнала, увеличивая добротность фильтра и сужая полосу подавления. Глубина подавления регулируется изменением сопротивления потенциометра в цепи обратной связи операционного усилителя.

Для настройки фильтра на частоту 50 Гц необходимо правильно подобрать значения резисторов и конденсаторов согласно основной формуле для расчета центральной частоты:

$$f_0 = \frac{1}{2\pi RC},$$

где f_0 — центральная частота (50 Гц), R — сопротивление резисторов, C — емкость конденсаторов.

Схема подключения, учитывающая эти значения для создания требуемой АЧХ, показана на рисунке 4. На схеме — емкость конденсаторов: $C_1 = C_2 = C$; $C_3 = 2C$; сопротивление резисторов: $R_1 = R_2 = R$; $R_3 = \frac{R}{2}$. Если выбирать значение $C=0.1\text{мкФ}$, то сопротивление:

$$R = \frac{1}{2\pi f_0 C} = \frac{1}{2\pi \times 50 \times 0.1 \times 10^{-6}} \approx 31.8 \text{ кОм.}$$

где $f_0=50$ Гц.

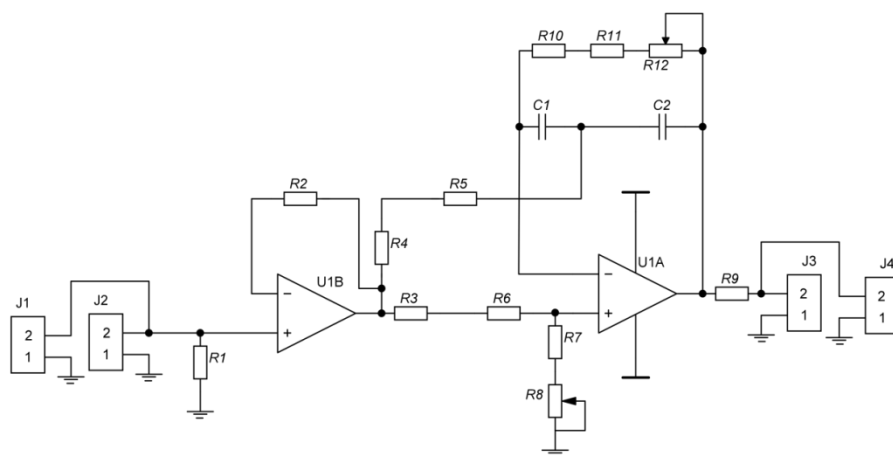


Рисунок 4 – Принципиальная схема режекторного фильтра

Если выбрать ближайшее стандартное значение резистора $R=33\text{ кОм}$, то $R1 = R2 = 33\text{ кОм}$, $R3 = 16,5\text{ кОм}$, $C1 = C2 = 0,1\text{ мкФ}$, $C3 = 0,2\text{ мкФ}$.

Инвертирующий вход операционного усилителя U1A (TL082) подключается к выходу Т-образного фильтра, в то время как неинвертирующий вход соединяется с выходом усилителя U1B (TL082) через потенциометр и заземляется. Питание операционных усилителей осуществляется от двухполярного источника напряжением $\pm 12\text{ В}$. Выход усилителя U1A служит основным выходом всей схемы. Для корректной работы рекомендуется использовать качественный линейный двуполярный источник питания с минимальными пульсациями, а также соблюдать полярность подключения, поскольку переполюсовка может повредить ОУ.

Регулировка глубины подавления осуществляется с помощью потенциометра RW2, при этом для стабильной работы рекомендуется не превышать глубину подавления более 40 дБ. Входной сигнал поступает на клеммы J3 и J4, а выходной сигнал снимается с клемм J1 и J2, при этом важно следить, чтобы уровни входного сигнала не превышали допустимые значения. Модуль необходимо устанавливать вдали от источников электромагнитных помех или экранировать, а также использовать экранированные кабели для подключения входов и выходов [2].

Таким образом, данная схема режекторного фильтра на частоту 50 Гц представляет собой эффективное решение для подавления сетевых помех в подсистемах сбора и обработки сигналов. Использование активного Т-образного фильтра с операционным усилителем обеспечивает высокую добротность и гибкость настройки, что позволяет адаптировать фильтр под конкретные требования. Правильная настройка и реализация этой схемы значительно улучшат качество измерений и повысят точность работы ИСМСВП.

Универсальный фильтр. Предлагаемый фильтр на базе микросхемы UAF42 представляет собой универсальный активный фильтр, способный реализовывать различные типы фильтрации, включая полосовую, низкочастотную и высокочастотную. В данной схеме используется полосовой фильтр, предназначенный для пропускания сигналов в заданном диапазоне частот и ослабления сигналов за его пределами. Центральная частота фильтра и его добротность (Q-фактор) настраиваются с помощью внешних резисторов и конденсаторов, подключенных к микросхеме UAF42.

Центральная частота фильтра f_0 рассчитывается по формуле:

$$f_0 = \frac{1}{2\pi\sqrt{R1 \cdot R2 \cdot C1 \cdot C2}}$$

где $R1$ и $R2$ – сопротивление резисторов в цепи фильтра (в Омах), а $C1$ и $C2$ – емкость конденсаторов (в Фарадах). Например, для фильтра с центральной частотой $f_0=1\text{ кГц}$ можно выбрать значения $R1=R2=10\text{ кОм}$ и $C1=C2=15,9\text{ нФ}$. Подставляя эти значения в формулу, получаем:

$$f_0 = \frac{1}{2\pi(10 \times 103 \, \Omega) \times (10 \times 103 \, \Omega) \times (15,9 \times 10 - 9 \, \Phi) \times (15,9 \times 10 - 9 \, \Phi)} \approx 1 \, \text{кГц}$$

Добротность Q , определяющая ширину полосы пропускания, рассчитывается по формуле:

$$Q = \frac{f_0}{BW}$$

где BW — ширина полосы пропускания. Если требуется, чтобы ширина полосы пропускания была $BW=100$ Гц, то добротность будет:

$$Q = \frac{1000 \, \text{Гц}}{100 \, \text{Гц}} = 10.$$

Основными компонентами предлагаемой схемы (рис. 5) являются микросхема UAF42, резисторы и конденсаторы, а также разъемы для подключения входных и выходных сигналов и питания.

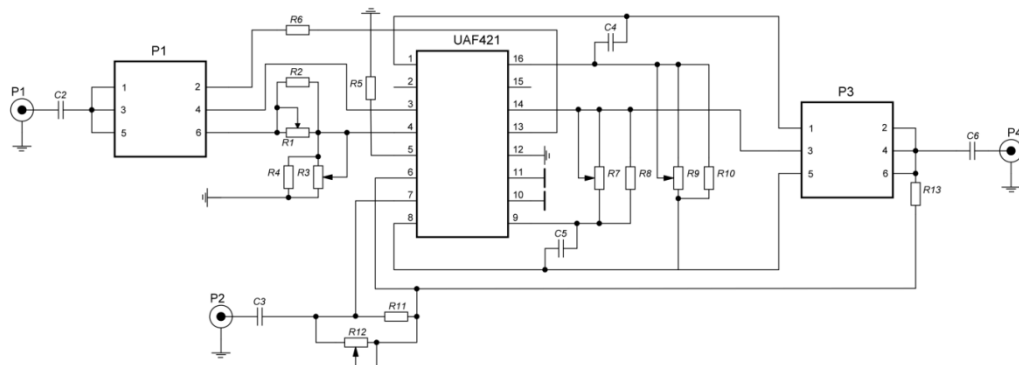


Рисунок 5— Универсальный фильтр

Резисторы $R1$ – $R15$ определяют параметры фильтра, такие как центральная частота и добротность, а подстроечные резисторы ($R4$ и $R6$) используются для точной настройки фильтра. Конденсаторы $C1$ – $C11$ влияют на реакцию фильтра на различные частоты сигнала. Разъемы $P1$ – $P6$ обеспечивают подключение входных и выходных сигналов, а также питания, которое осуществляется от источника ± 12 В.

Принцип работы схемы заключается в следующем: входной сигнал подается через разъемы $P2$ или $P3$ на входные контакты микросхемы UAF42. Фильтрация сигнала осуществляется внутри микросхемы, где внешние резисторы и конденсаторы задают необходимые параметры фильтра. Выходной сигнал снимается с вывода V_o микросхемы и передается через разъемы $P4$ (OUT) или $P5$ (OUT2) к последующим каскадам. Питание схемы поступает через разъем $P6$ (POWER), при этом дополнительная фильтрация осуществляется с помощью конденсаторов $C6$, $C7$, $C9$ и $C10$ для сглаживания пульсаций и подавления шумов.

Настройка и регулировка фильтра производится путем подбора значений резисторов и конденсаторов. Центральная частота f_0 регулируется изменением значений резисторов $R1$, $R2$ и конденсаторов $C1$, $C2$. Добротность Q настраивается с помощью подстроечных резисторов $R4$, $R6$, что позволяет изменять ширину полосы пропускания фильтра. Коэффициент усиления можно корректировать, изменяя номиналы резисторов $R3$, $R5$, $R7$, $R8$, $R9$, что позволяет адаптировать уровень выходного сигнала под требования конкретного применения.

Особенности конструкции включают фильтрацию питания, где конденсаторы $C6$, $C7$, $C9$, $C10$ обеспечивают стабильное питание микросхемы, уменьшая влияние внешних помех. Благодаря общей точке заземления обеспечивается стабильная работа схемы без возникновения нежелательных потенциалов. Наличие нескольких разъемов $P1$ – $P6$ облегчает интеграцию фильтра в систему и позволяет быстро изменять конфигурацию подключения.

Для обеспечения высоких характеристик фильтра рекомендуется использовать резисторы и конденсаторы с минимальным отклонением номинала, например 1% или ниже.

Использование качественных источников питания с минимальным уровнем пульсаций способствует предотвращению искажения сигнала. При работе с высокочастотными сигналами рекомендуется экранировать чувствительные участки схемы для защиты от электромагнитных помех [3].

Таким образом, схема полосового фильтра на базе UAF42 обеспечивает гибкую настройку и высокую точность фильтрации, что делает ее пригодной для широкого спектра датчиков в ИСМСВП. Тщательный подбор компонентов и внимательное проектирование обеспечивают стабильную работу и высокое качество обработки сигнала.

Цифро-аналоговый преобразователь. В качестве цифро-аналогового преобразователя предлагается использовать готовый модуль DAC8562/3, представляющий собой полностью интегрированное решение, предназначенное для приложений, требующих высокоточного аналогового выхода. В основе модуля лежит ЦАП DAC8562/3, который обеспечивает весь необходимый функционал для создания управляемого аналогового сигнала. Этот модуль включает в себя все необходимые компоненты для полноценного функционирования системы на базе DAC8562/3 и способен обеспечивать симметричное выходное напряжение ± 12 В при максимальной нагрузке до 50 мА. В составе модуля используется операционный усилитель OPA2277, отличающийся низким температурным дрейфом, что позволяет достичь высокой точности преобразования сигнала. Модуль поддерживает возможность настройки выходного сигнала как в однофазном, так и в дифференциальном режиме, что расширяет его функциональные возможности и область применения. Питание модуля осуществляется в диапазоне входного напряжения от 2,7 В до 5,5 В постоянного тока, при этом стандартная конфигурация обеспечивает коэффициент усиления дифференциального усилителя, равный 4,17. Номинальное сопротивление на выходе усилителя составляет 51 Ом, что обеспечивает его работу при подключении к различным нагрузкам. Модуль выполнен на четырехслойной печатной плате из материала FR4 и соответствует требованиям стандарта ROHS, что гарантирует его безопасность для окружающей среды и надежность при эксплуатации.

Внутренняя структура модуля представлена в виде блок-схемы, которая демонстрирует ключевые компоненты и их взаимодействие (рис. 6) [4].

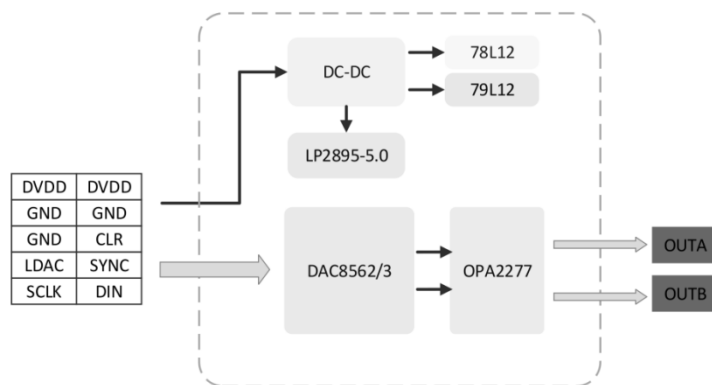


Рисунок 6 – Структурная схема модуля DAC8562/3

Конфигурация опорного напряжения дифференциального усилителя осуществляется путем выбора между резисторами R2 и R4. При пайке резистора R2 (0 Ом) опорное напряжение составляет 2,5 В, что приводит к биполярному выходу дифференциального усилителя. В случае пайки резистора R4 (0 Ом) опорное напряжение составляет 0 В, что обеспечивает формирование однополярного выходного сигнала. При этом резисторы R2 и R4 не могут использоваться одновременно, так как это может привести к некорректной работе усилителя (рис. 7).

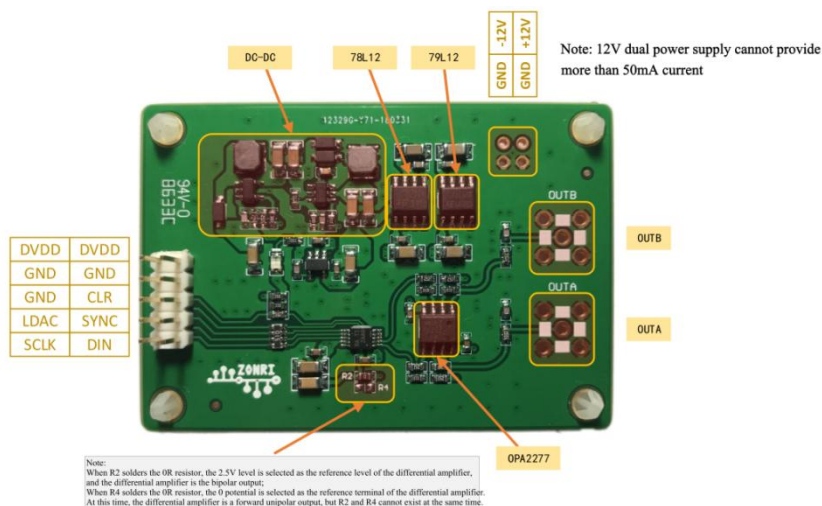


Рисунок 7 – Внешний вид DAC8562/3

Основными компонентами модуля являются высокоточный цифро-аналоговый преобразователь DAC8562/3, операционный усилитель OPA2277 с низким уровнем температурного дрейфа для обеспечения точного усиления сигнала, а также стабилизаторы напряжения LP2895-5.0 и линейные стабилизаторы 78L12 и 79L12, которые обеспечивают стабильное напряжение для поддержания точных характеристик выходного сигнала. Эти компоненты совместно обеспечивают высокую точность и надежность работы модуля, что делает его подходящим для использования в системах, требующих формирования прецизионных аналоговых сигналов.

Таким образом, модуль DAC8562/3 предназначен для использования в научных измерительных приборах, промышленных системах управления и других областях, где необходимо формирование высокоточных аналоговых сигналов. Он может применяться в задачах генерации сигналов, в устройствах для измерения и мониторинга, а также в различных промышленных установках, требующих точного контроля и управления процессами. Высокая точность, универсальность настройки и надежность модуля делают его оптимальным решением для многих приложений, связанных с обработкой аналоговых сигналов.

Аналогово-цифровой преобразователь. Для высокоточного аналого-цифрового преобразования предлагается использовать модуль ADS8688A (рис. 8), который представляет собой универсальное решение для преобразования аналоговых сигналов в цифровые с высокой точностью. Этот модуль основан на 16-разрядном АЦП ADS8688A и интегрирует все необходимые компоненты, обеспечивая высокие эксплуатационные характеристики в различных условиях. Он способен работать от внешнего однополярного источника питания с диапазоном входных напряжений от 2,7 В до 5,5 В постоянного тока, что делает его универсальным для широкого спектра приложений. Типичный рабочий ток составляет 33 мА при напряжении питания 3,3 В и 20 мА при 5,0 В, что способствует экономичному энергопотреблению и высокой эффективности использования. Для обеспечения стабильности преобразования модуль оснащен интегрированным опорным источником напряжения ADR444 с номиналом 4,096 В, позволяя выбирать между внутренней и внешней конфигурациями опорного напряжения в зависимости от специфических требований системы.

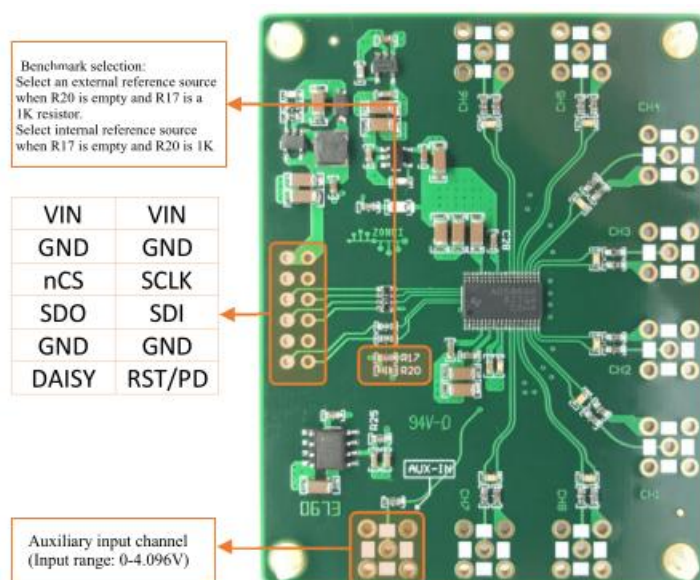


Рисунок 8 – Внешний вид ADS8688A

Дополнительный входной канал AUX может быть подключен к АЦП через встроенный мультиплексор, что позволяет передавать сигнал непосредственно на преобразователь, обходя остальные каналы. Это расширяет функциональные возможности модуля и обеспечивает дополнительную гибкость при конфигурации системы. Поддерживаемый диапазон входных напряжений включает значения $\pm 0,64$ В, $\pm 1,28$ В, $\pm 2,56$ В, $\pm 5,12$ В, $\pm 10,24$ В, а также $+5,12$ В и $+10,24$ В. Настройка диапазона входных напряжений и полярности может выполняться независимо для каждого из восьми каналов, что обеспечивает высокую адаптивность модуля к требованиям пользователя. Максимальная скорость выборки составляет 500 тыс. отсчетов в секунду, что делает модуль подходящим для высокоскоростных систем сбора и анализа данных.

Модуль ADS8688A реализован на четырехслойной печатной плате, что обеспечивает высокую устойчивость к электромагнитным помехам и минимизирует паразитные индуктивности и емкости.

Соответствие директиве RoHS гарантирует экологическую безопасность модуля, что особенно важно при использовании в медицинских и научных приборах. Для выбора опорного источника напряжения используются резисторы R17 и R20, обеспечивая гибкий выбор между внутренним и внешним опорным напряжением. Внешний источник активируется при отсутствии резистора R20 и установке резистора R17 номиналом 1 кОм; внутренний источник выбирается при отсутствии резистора R17 и наличии резистора R20 номиналом 1 кОм.

При использовании модуля без функции Daisy Chain соответствующий контакт необходимо подключить к земле, чтобы избежать некорректной работы системы. Daisy Chain (от англ. «гирлянда» или «цепочка») – это метод последовательного подключения нескольких устройств к одной шине коммуникации. Что касается модуля ADS8688A, функция Daisy Chain позволяет соединять несколько таких модулей последовательно по интерфейсу SPI. Это означает, что данные передаются от первого модуля ко второму и так далее, позволяя управлять всеми модулями через меньшее количество линий связи [5].

Модуль содержит восемь основных входных каналов и один вспомогательный канал AUX (рис. 9), что позволяет считывать данные с нескольких источников одновременно. Независимая настройка каждого канала обеспечивает возможность интеграции модуля в системы с различными требованиями к точности и диапазону измерений.

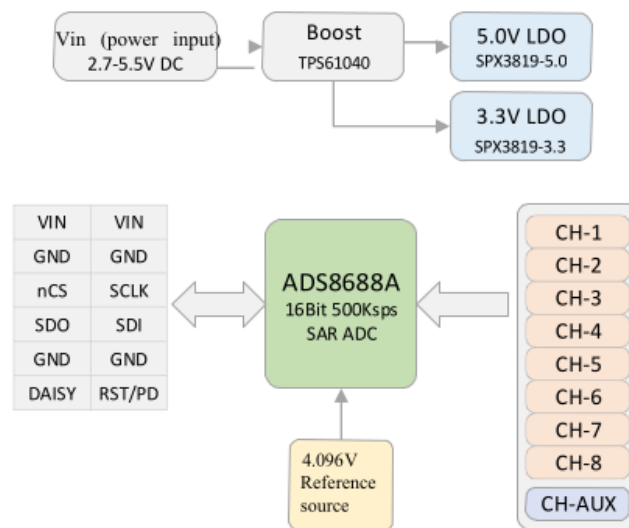


Рисунок 9 – Структурная схема модуля ADS8688A

Встроенные стабилизаторы напряжения, такие как TPS61040 и SPX3819, обеспечивают стабильное питание на уровнях 5,0 В и 3,3 В соответственно, что значительно улучшает характеристики преобразования и повышает надежность работы АЦП в условиях изменения внешних параметров.

Процесс первоначальной настройки и эксплуатации модуля ADS8688A подробно описан в руководстве пользователя [6]. Сначала необходимо подключить модуль к источнику питания и интерфейсу SPI. Питание подается через контакт Vin с диапазоном входных напряжений от 2,7 В до 5,5 В постоянного тока. Для обеспечения корректной работы рекомендуется подключить все контакты заземления модуля к общему заземлению управляющего микроконтроллера.

Таким образом, модуль ADS8688A предназначен для использования в системах, требующих прецизионного аналого-цифрового преобразования, включая научные приборы, системы автоматизированного управления и мониторинга, а также устройства контроля технологических процессов.

Высокая скорость выборки, гибкость настройки входных каналов, низкое энергопотребление и адаптивность к различным требованиям делают этот модуль хорошим выбором для широкого спектра задач.

Использование современных компонентов, таких как опорный источник ADR444 и стабилизаторы TPS61040, обеспечивает высокую стабильность и точность преобразования, что особенно важно для критически важных приложений, таких как медицинская аппаратура и промышленные системы управления. Модуль ADS8688A представляет собой оптимальное решение для промышленности, научных исследований и других областей, где требуется высокая точность и надежность аналого-цифрового преобразования сигналов.

Микрокомпьютер. В разрабатываемой ИСМЦВП предлагается использовать одноплатный микрокомпьютер Orange Pi Zero 3 в качестве основного локального вычислительного узла. Данный микрокомпьютер обладает достаточными вычислительными ресурсами и функциональными возможностями для обработки данных, поступающих от датчиков, а также для реализации алгоритмов потоковой предобработки в режиме реального времени.

Orange Pi Zero 3 основан на процессоре Allwinner H618, представляющем собой 64-битный четырехъядерный ARM Cortex-A53 с тактовой частотой до 1,5 ГГц. Это обеспечивает необходимую вычислительную мощность для выполнения сложных алгоритмов. Основные характеристики Orange Pi Zero 3 приведены в таблице 1 [7].

Таблица 1 – Основные характеристики Orange Pi Zero 3

Характеристика	Описание
Процессор	Allwinner H618 (4 ядра ARM Cortex-A53, до 1,5 ГГц)
Оперативная память	2 ГБ DDR3
Накопитель данных	Слот для карт MicroSD (до 128 ГБ)
Сетевые интерфейсы	- Ethernet 10/100 Мбит/с - Wi-Fi 802.11 b/g/n/ac - Bluetooth 5.0
Интерфейсы ввода-вывода	- 26-пиновый разъем GPIO - USB 2.0 - USB OTG
Поддерживаемые операционные системы	Различные дистрибутивы Linux (Debian, Ubuntu) и Android

Преимущества использования Orange Pi Zero 3 заключаются в следующем.

Мощный процессор и достаточный объем оперативной памяти позволяют обрабатывать большие объемы данных с датчиков и выполнять алгоритмы предварительной обработки в режиме реального времени.

Малые габариты и низкое энергопотребление делают микрокомпьютер идеальным для использования в условиях ограниченного пространства и ограниченных энергетических ресурсов, что особенно актуально для систем мониторинга в удаленных или труднодоступных местах.

Наличие разнообразных интерфейсов ввода-вывода облегчает подключение различных датчиков и модулей, а также интеграцию микрокомпьютера в существующую инфраструктуру системы.

Поддержка Wi-Fi и Ethernet обеспечивает стабильную передачу данных на сервер для дальнейшей обработки и анализа, а также позволяет осуществлять удаленное управление и обновление системы.

Микрокомпьютер будет выполнять роль центрального узла в локальном модуле обработки данных. Его функции включают:

Прием данных от АЦП, таких как модуль ADS8688A, которые преобразуют аналоговые сигналы с датчиков в цифровой формат.

Предварительная обработка данных, включая фильтрацию сигналов, устранение шумов и выявление аномалий, что снизит нагрузку на серверную часть системы.

Передача обработанных данных на центральный сервер через Ethernet или Wi-Fi для хранения, дополнительного анализа и визуализации.

Реализация обратной связи: при необходимости микрокомпьютер может отправлять управляющие сигналы на исполнительные устройства через цифровые выходы GPIO или ЦАП DAC8562/3.

Аналоговые датчики различных физических параметров (деформации, вибрации, температуры и т.д.) могут подключаться к аналоговым входам ADC-модуля ADS8688A. ADC-модуль соединяется с Orange Pi Zero 3 через интерфейс SPI, обеспечивая высокоскоростную передачу данных. Цифровые датчики можно подключить к цифровым входам Orange Pi Zero 3. На микрокомпьютере выполняются алгоритмы предварительной обработки данных, а также могут быть запущены обученные нейросетевые модели для прогнозирования потенциальных рисков. Обработанные данные передаются на центральный сервер для дальнейшего анализа и принятия решений. При необходимости микрокомпьютер может управлять исполнительными механизмами (например, системами оповещения или аварийными клапанами) через выходные интерфейсы.

Для одноплатного микрокомпьютера Orange Pi Zero 3 используется операционная система Ubuntu Server, которая предоставляет стабильную и надежную платформу для разработки и исполнения приложений. Программное обеспечение может создаваться с использованием таких языков программирования, как Python, C++ и Rust, что позволяет эффективно интегрировать различные библиотеки и фреймворки. В частности, можно использовать следующие инструменты:

TensorFlow Lite или PyTorch Mobile для реализации и выполнения нейронных сетей непосредственно на устройстве;

Протокол MQTT для легковесной и эффективной передачи данных между устройствами и сервером;

NumPy и SciPy для численных вычислений и обработки сигналов;

GPIO-библиотеки для взаимодействия с аппаратными интерфейсами микрокомпьютера.

Для повышения надежности и безопасности системы рекомендуется использовать следующие средства:

Аппаратный сторожевой таймер (Watchdog Timer), который автоматически перезагружает систему в случае программных сбоев или зависания, что увеличивает надежность работы в автономном режиме;

Стабилизированное питание с качественными источниками и возможностью резервного питания от аккумуляторов для бесперебойной работы системы;

Шифрование данных с использованием протоколов TLS/SSL, а также настройка межсетевых экранов и регулярное обновление программного обеспечения для защиты от киберугроз.

В рамках эксплуатации и обслуживания могут быть предусмотрены следующие возможности:

удаленное управление и мониторинг, что позволяет обновлять программное обеспечение, настраивать параметры и проводить диагностику системы на расстоянии;

логирование и анализ системных журналов для отслеживания работы системы, анализа производительности и своевременного обнаружения проблем;

масштабирование системы, которое включает возможность объединения нескольких микрокомпьютеров в распределенную сеть для увеличения зоны мониторинга и повышения отказоустойчивости.

Таким образом, применение одноплатного микрокомпьютера Orange Pi Zero 3 в ИСМСВП оправдано с точки зрения его технических характеристик, экономической эффективности и гибкости. Это решение обеспечивает создание мощной и адаптивной платформы для сбора, обработки и анализа данных в реальном времени, что особенно важно для обеспечения состояния воды и почв сельскохозяйственных угодий в регионах с недостаточным количеством водных ресурсов.

Литература

1. Использование микрокомпьютеров в качестве управляющих элементов автоматизированных систем [Электронный ресурс]. – Режим доступа: <http://www.microsystems.ru/files/publ/492.htm> (дата обращения: 15.11.2024).
2. Аналого-цифровые преобразователи и их применение [Электронный ресурс]. – Режим доступа: <https://studfile.net/preview/3277492/page:5/> (дата обращения: 15.11.2024).
3. Ефимов С. Н. Микроэлектронные устройства в системах мониторинга [Электронный ресурс]. – Режим доступа: https://lib.ulstu.ru/venec/2002/1/Efimov_3.pdf (дата обращения: 15.11.2024).
4. Использование микрокомпьютеров в качестве управляющих элементов автоматизированных систем [Электронный ресурс] // CyberLeninka. – Режим доступа: <https://cyberleninka.ru/article/n/ispolzovanie-mikrokompyuterov-v-kachestve-upravlyayuschih-elementov-avtomatizirovannyh-sistem> (дата обращения: 15.11.2024).
5. ОРА627 [Электронный ресурс]. – Режим доступа: <https://www.ti.com/lit/ds/symlink/opa627.pdf?ts=1711067552629> (дата обращения: 15.06.2024).
6. TL082-N [Электронный ресурс]. – Режим доступа: <https://www.ti.com/lit/ds/symlink/tl082-n.pdf> (дата обращения: 15.06.2024).

7. UAF42 [Электронный ресурс]. – Режим доступа: <https://www.ti.com/lit/gpn/UAF42> (дата обращения: 15.06.2024).
8. DAC8562 [Электронный ресурс]. – Режим доступа: <https://www.ti.com/product/DAC8562> (дата обращения: 15.06.2024).
9. ADS8688A [Электронный ресурс]. – Режим доступа: <https://www.ti.com/lit/ds/sbas680/sbas680.pdf> (дата обращения: 15.06.2024).
10. Orange Pi Zero 3 [Электронный ресурс]. – Режим доступа: <http://www.orangepi.org/html/hardWare/computerAndMicrocontrollers/details/Orange-Pi-Zero-3.html> (дата обращения: 15.06.2024).

УДК 621.3.087.9

МЕХАНИЧЕСКОЕ МАСШТАБИРОВАНИЕ ДАТЧИКОВ ЛИНЕЙНЫХ РАЗМЕРОВ

*Дьяченко Е.И.**Институт машиноведения, автоматизации и геотехники НАН КР*

Работа посвящена рассмотрению устройств для масштабирования аналоговых сигналов. Рассматривается новый тип таких устройств – механический масштабизатор сигнала для датчиков линейных размеров. Данный масштабизатор представляет собой одну-единственную деталь, выполненную из какого-либо упругого материала. Особый профиль пустот в конструкции масштабизатора позволяет ему иметь большую жесткость в некоторых заданных направлениях и малую жесткость в других. Кинематические пары такого масштабизатора работают не за счет относительного движения его звеньев, а за счет собственной упругости материала. Это обстоятельство принципиально лишает такие масштабизаторы люфтов, износа, потребности в смазке и прочего, свойственного классическим кинематическим парам. Кроме того, благодаря современным аддитивным технологиям такое устройство зачастую дешевле аналогов и обладает рядом других положительных особенностей, о которых пойдет речь в статье. Также приводится расчет масштабизатора по перемещениям, с помощью метода конечных элементов в расчетной среде SolidWorks Simulation.

Ключевые слова: датчики, аналоговый сигнал, масштабизация, линейные перемещения, упругие механизмы, жесткость, АЦП.

Введение

Масштабирование аналоговых сигналов представляет собой важную техническую задачу, необходимую для обработки и анализа данных в различных областях, от электроники и телекоммуникаций до измерительной техники и медицинской диагностики. Этот процесс позволяет изменять амплитуду и динамический диапазон сигналов, обеспечивая оптимальные условия для их последующей обработки и интерпретации. В данной статье рассматриваются основные методы масштабирования аналоговых сигналов, их применение в различных областях и технические аспекты данного процесса.

Масштабирование аналоговых сигналов является неотъемлемой частью многих технических систем, играющей ключевую роль в обеспечении корректной обработки и интерпретации данных. Понимание основных методов и их применение в различных областях позволяют инженерам и ученым эффективно решать задачи по обработке и передаче аналоговых сигналов, способствуя прогрессу в различных технологических и научных дисциплинах.

Основные причины необходимости масштабирования аналоговых сигналов можно резюмировать так:

Совместимость с оборудованием. Разные устройства могут работать с различными уровнями напряжения или тока. Масштабирование позволяет адаптировать сигнал под требования конкретного оборудования.

Устойчивость к шумам. При повышении уровня сигнала (масштабировании) можно улучшить соотношение сигнал/шум, что делает обработку сигнала более надежной.

Упрощение обработки. Для обработки сигнала, например, с использованием аналоговых или цифровых схем, важно, чтобы уровень сигнала находился в определенных пределах. Масштабирование помогает привести сигнал к нужному диапазону.

Снижение искажений. Если сигнал слишком слабый, его сложно обработать без искажений. Масштабирование позволяет избежать потерь информации.

Масштабирование аналоговых сигналов

Масштабирование аналоговых сигналов включает в себя несколько основных методов [1]:

1. **Усиление.** Этот метод предназначен для увеличения амплитуды сигнала путем использования усилительных устройств. Усиление может быть линейным, когда усиливаются все частоты сигнала в одинаковой пропорции, или нелинейным, когда усиление зависит от частоты или амплитуды сигнала.

2. **Ослабление.** Используется для уменьшения амплитуды сигнала. Ослабление может быть полосовым или частотно-зависимым, чтобы обеспечить минимальное искажение сигнала при его изменении.

3. **Подавление постоянной составляющей.** Удаление постоянной составляющей сигнала позволяет корректно центрировать сигнал относительно нулевой точки для последующего анализа или передачи.

4. **Нормирование.** Процесс, при котором амплитуда сигнала масштабируется таким образом, чтобы его максимальная амплитуда соответствовала определенному предельному значению или стандарту.

5. **Компрессия динамического диапазона.** Этот метод позволяет уменьшить разницу между наибольшими и наименьшими значениями амплитуды сигнала, что полезно для обработки акустических сигналов или сигналов в режиме реального времени.

При реализации масштабирования аналоговых сигналов важно учитывать такие технические аспекты, как:

- **Линейность и искажения.** Усилители должны обеспечивать минимальные искажения сигнала в широком диапазоне частот и амплитуд.

- **Шум и динамический диапазон.** Необходимо контролировать уровень шума и обеспечивать достаточный динамический диапазон для сохранения качества сигнала.

- **Стабильность и надежность.** Устройства масштабирования должны быть стабильными и надежными в различных условиях эксплуатации и изменяющихся параметрах сигнала.

Для измерения аналоговых величин в цифровых системах используются аналогово-цифровые преобразователи (АЦП). АЦП характеризуются такими параметрами, как диапазон входных напряжений и разрядность, и ряд других. Диапазон входных напряжений зависит от опорного напряжения АЦП и определяет границы, в пределах которых аналоговый сигнал может быть оцифрован. Разрядность АЦП определяет шаг, с которым выполняется преобразование одной выборки аналогового сигнала.

Довольно часто возникает ситуация, когда диапазон входного измеряемого напряжения не соответствует входному диапазону АЦП. Если он больше, АЦП не сможет измерить его значения за пределами своей шкалы, если меньше, то АЦП, по сути, будет работать в узком диапазоне, и разрешение измеряемого напряжения будет низким.

Чтобы избежать этой ситуации, применяются масштабирующие схемы, которые «подгоняют» интересующий нас диапазон входных напряжений под диапазон АЦП. Эти схемы достаточно классические, обычно представляют из себя одну из следующих вариаций:

1. **Делитель напряжения с повторителем.** Его можно применять, когда входное измеряемое напряжение превышает диапазон АЦП и его нужно просто поделить (1). Операционный усилитель выполняет здесь роль повторителя с большим входным и маленьким выходным сопротивлением.

2. **Не инвертирующий усилитель на операционном усилителе.** Такая схема подойдет, если входной сигнал меньше диапазона АЦП и его нужно просто умножить.

3. **Операционный усилитель выполняет усиление сигнала и смещает его.** Схему следует применять, когда интересующий нас диапазон входного сигнала начинается не с нуля и его нужно «растянуть» на всю шкалу АЦП.

В общем случае усиление сигнала АЦП можно представить так (рис. 1):

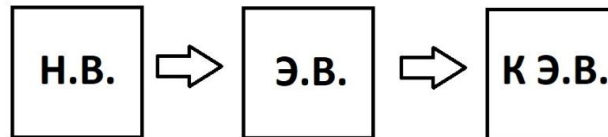


Рисунок 1 – Принципиальная схема усиления сигнала в АЦП (Н.В – неэлектрическая величина, Э.В. – электрическая величина, К – коэффициент масштабирования)

Предварительное масштабирование неэлектрической величины

В том случае, когда измеряется неэлектрическая величина, например, линейный размер, возможно применить схему, когда эта неэлектрическая величина предварительно масштабируется в такую же неэлектрическую величину, после чего уже масштабированная неэлектрическая величина конвертируется в электрическую. Такую схему в общем случае можно представить в следующем виде (рис.2):

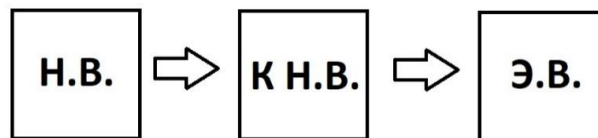


Рисунок 2 – Принципиальная схема усилителя с механическим масштабированием (Н.В – неэлектрическая величина, Э.В. – электрическая величина, К – коэффициент масштабирования)

Такой подход открывает возможность для более тонкой подстройки всех элементов датчика, то есть если диапазон неэлектрической величины, например, линейного размера заготовки изменяется в диапазоне БОльшим, чем рабочий диапазон, например индуктивного датчика, то можно предварительным механическим масштабированием подогнать диапазон этого изменения к диапазону измерения индуктивного датчика. И наоборот, например, если необходимо повысить чувствительность индуктивного датчика, а входная неэлектрическая величина достаточно мала, например, ее изменение от 0 до 0,1 а рабочий диапазон индуктивного датчика 5 мм, то возможно механически масштабировать входящую величину с диапазона 0 – 0,1 до диапазона 0 – 5 мм, то есть коэффициент К здесь будет 50. Благодаря этому достигаются широкие возможности для предварительного масштабирования входящей неэлектрической величины, что влечет за собой повышение надежности, уменьшение номенклатуры электронных компонентов, оптимизации конструкций датчиков, в ряде случаев улучшения массогабаритных характеристик и т.д.

В общем случае схемы механического масштабирования известны давно и активно применяются [2], например индикаторы часового типа – устройство для измерений и контроля отклонений от эталонной детали (рис.3). Прибор помогает определить биения, несоответствие геометрической формы и толщины деталей. Служит для проверки качества в рамках стандартов и допусков на различных предприятиях – от автомобилестроения до частных мастерских. Представляет собой, по сути, систему рычагов, где первый рычаг в виде штока, имевшего некоторое линейное перемещение, связан со стрелкой на циферблате, которая с некоторым коэффициентом редукции на заранее проградуированной шкале показывает перемещения штока.

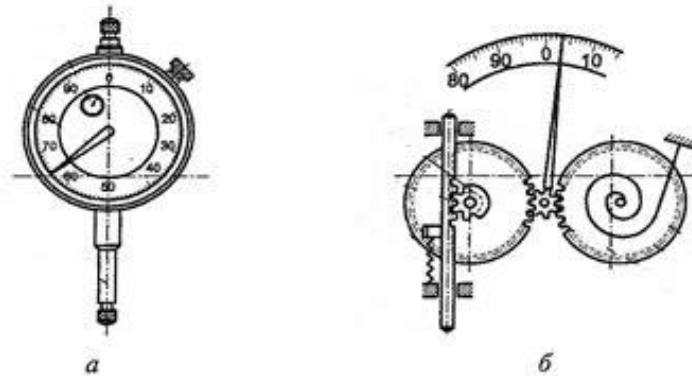


Рисунок 3 – а) Индикатор часового типа – общий вид; б) Устройство индикатора часового типа

Существуют индикаторы часового типа [3], которые уже имеют цифровую индикацию или цифровой вывод на компьютер (рис. 4).



Рисунок 4 – Индикатор часового типа с цифровым выводом

Такие устройства, по сути, являются датчиками размеров с механическим масштабированием неэлектрического сигнала с последующим преобразованием в электрический сигнал, но они имеют ряд минусов. Точность устройства строго зависит от качества исполнения механической части. Имеют достаточно ограниченный коэффициент масштабирования. Имеют стандартизированный ряд измеряемых размеров (обычно 0–10 мм). Массогабаритные характеристики достаточно однообразны. Имеют изнашивающиеся

части. Требуют периодического обслуживания ввиду наличия смазочного материала внутри. Достаточно хрупки. Достаточно требовательны к уровню вибраций.

Механический масштабизатор с гибкими связями

С повсеместным развитием компьютеризации производственных процессов появилась возможность создать достаточно сложные геометрические изделия. Компьютерное моделирование, а также комплексы программ, которые позволяют моделировать процессы, например напряжения и деформации, расширяют возможности современного производства, и становится возможным создать элементы с заранее заданной жесткостью ее узлов, что позволяет сформировать новый класс механизмов без кинематических пар. Кинематические пары в таких механизмах заменены их аналогами, но уже без подвижных частей, а подвижность элементов достигается распределением материала элемента, то есть собственной жесткостью каждого его узла. Такие механизмы имеют ряд принципиальных преимуществ:

1. Количество деталей. Зачастую номенклатура механизмов с гибкой связью в разы меньше их аналогов на классических кинематических парах.
2. Производственные процессы. Могут быть получены на широком спектре современного аддитивного и субтрактивного оборудования.
3. Цена. Зачастую цена таких механизмов в разы меньше их аналогов.
4. Точное движение. Ввиду отсутствия кинематического зазора в кинематических парах, а также замены внешнего трения, которое переменное от нагрузки и условий среды на внутреннее трение элементов гибкого механизма, возможно получать наноперемещения.
5. Габариты механизмов могут различаться от макромасштабного до наномасштабных.
6. Однозначность. Гибкие механизмы имеют низкую механическую флуктуацию. Измерения могут быть более повторяемы. Меньшая зависимость от условий окружающей среды.

В части приложения к измерительным устройствам нас больше всего интересует свойство таких механизмов, как высокая кинематическая точность ее элементов. Например, механический масштабизатор входной неэлектрической величины с последующим преобразованием уже масштабизированной в K раз величину может содержать чувствительный элемент с самым дешевым фоторезистором или датчиком Холла, даже с низкой чувствительностью. Ограничения чувствительности нивелируются механическим масштабизатором, который усиливает сигнал, в частности, что касается датчиков линейных размеров, возможно в несколько раз, а иногда и десятков раз увеличивать входное изменение размера до оптимальных для считывания датчиком.

На рисунке 5 показана конструкция такого механического масштабизатора. Он представляет собой одну-единственную деталь перемещения опорных планок, жестко связанных геометрическими размерами его звеньев и жесткостей элементов.

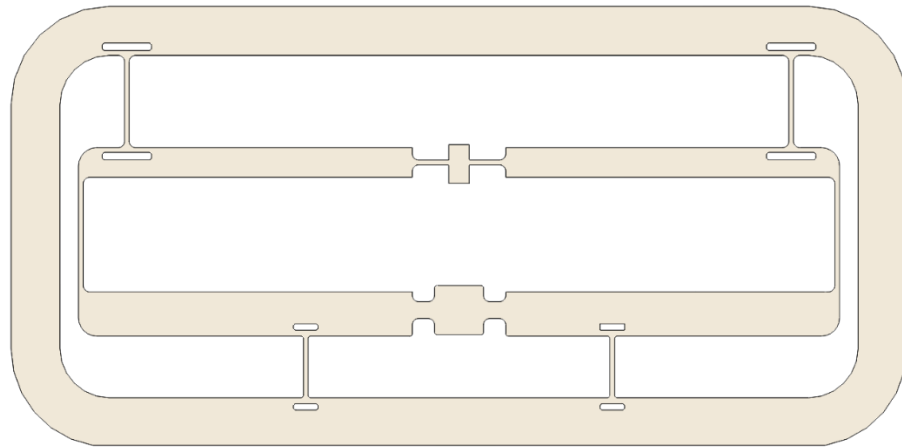


Рисунок 5 – Механический масштабизатор датчика линейных размеров – общий вид

Принципиальную схему работы устройства можно представить в виде рычажной системы – рисунок 6. Тут А – опорная площадка, которая воспринимает перемещение, В – опорная площадка с масштабированным перемещением, l , l_1 , l_2 – длины рычагов и балок, k_1 , k_2 , k_3 , k_4 , k_5 , k_6 – крутильные и изгибные жесткости элементов. Устройство работает следующим образом – опорная площадка А воспринимает усилие, возникающее от изменения размера элемента, который действует на нее, после чего по правилу рычага (правило рычага не соблюдается в чистом виде ввиду наличия крутильных и изгибных сопротивлений) усилие и перемещения преобразуются на другом конце балки, затем цикл повторяется уже на следующем участке, где находится опорная площадка В. В зависимости от геометрических характеристик устройства, а также жесткостей элементов происходит масштабизация входной неэлектрической величины в виде изменения линейного размера в такую же, только масштабизированную величину, уже эта величина считывается каким-либо способом с площадки В. Например, фоторезистором или датчиком Холла. Достигается эффект согласования чувствительности датчика линейного с изменением считываемой величины.

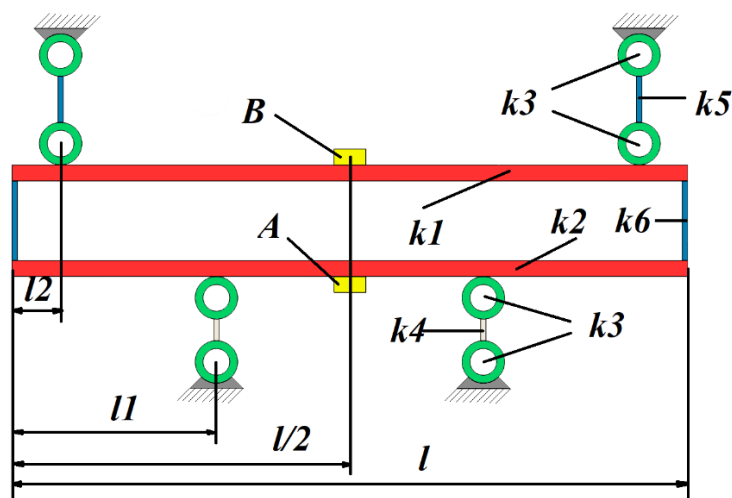


Рисунок 6 – Принципиальная схема масштабизатора датчика линейных размеров

В среде моделирования SolidWorks Simulation был произведен расчет перемещений (рис.7) опорного участка В в зависимости от перемещений участка А. Прилагаемое усилие 1 кГ. Материал масштабизатора в расчете – ABS-пластик. Перемещение участка А составило 0,066 мм, участка В – 0,327. Коэффициент масштабирования $K=5$.

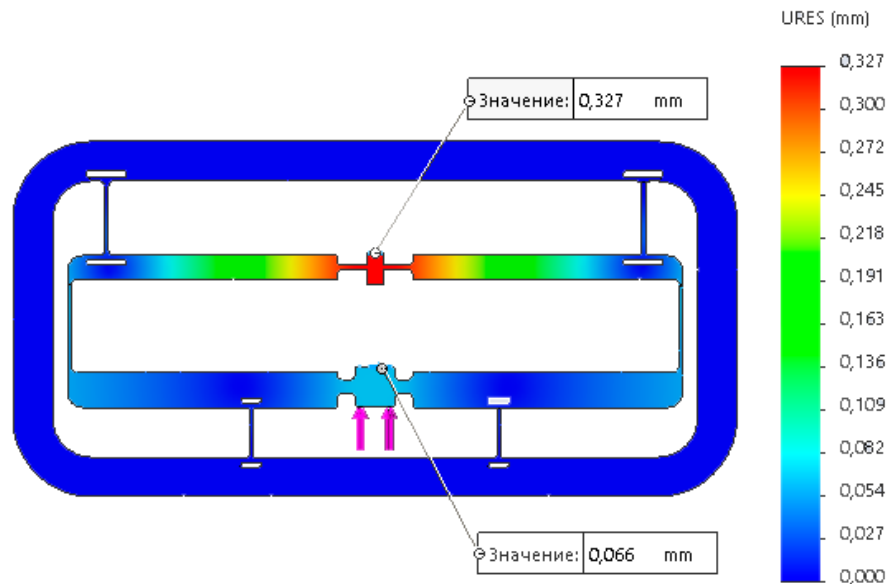


Рисунок 7 – Моделирование перемещения площадок А и В в SolidWorks Simulation

Синтез и анализ механизмов с гибкими связями

В настоящее время публикуется много работ касательно синтеза и анализа механизмов с гибкими связями, например [4, 5, 6]. По большей части работы на английском языке, можно привести в пример, где приводится достаточно серьезный математический аппарат [7, 8].

Проблема синтеза и анализа подобных механизмов достаточно нетривиальна, на русском практически не встречается публикаций, и достаточно приличный объем публикаций можно найти в зарубежной литературе. Конструирование описанных механизмов осложняется наличием не обычных кинематических пар, а гибких, также наличием некоего упругого предела, после которого начинается усталостное разрушение материала. Кроме того, наличие податливостей на многих участках механизма создает дополнительные трудности при их расчете. Основные аспекты проблематики создания механизмов данного типа:

Изучение деформаций. Оценка, как гибкие элементы реагируют на внешние нагрузки, и определение величины их деформаций.

Определение динамики. Исследование колебательных и динамических характеристик механизмов, включая резонансные явления.

Сила и нагрузка. Вычисление сил, действующих на гибкие связи, и их влияние на общую работу механизма.

Методы, используемые в анализе, можно разделить на следующие группы:

Аналитические методы. Использование уравнений движения и принципов статики для определения характеристик механизмов.

Численные методы. Применение методов конечных элементов (МКЭ) для моделирования сложных систем и их поведения под нагрузкой.

Экспериментальные методы. Проведение испытаний на физических образцах для верификации теоретических расчетов.

Выше был рассмотрен семеричный механизм с гибкими связями, аналог которого может быть построен с помощью простейших рычажных механизмов. Описанный выше способ, когда на основе известных классических механизмов строится гибкий, называется синтезом посредством замены твердого тела [9]. Говоря в общих чертах, используя этот способ для синтеза механизмов с гибкими связями, заменяются стандартные кинематические «жесткие» пары на упругие. Например, как на рисунке 8 показан переход от классического зажимного устройства к такому же зажимному, но уже являющемуся механизмом с гибкими связями.

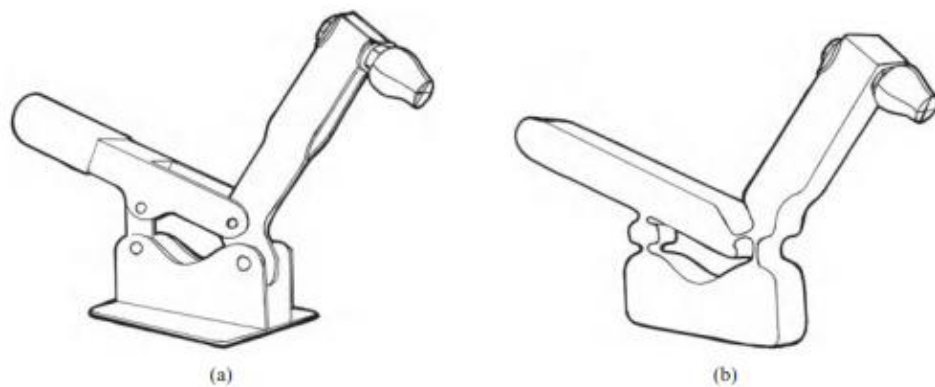


Рисунок 8 – Переход от классического механизма к механизму с гибкими связями

В [10] приводятся еще 3 способа синтеза подобных механизмов:

- Синтез посредством оптимизации топологии.
- Синтез посредством использования строительных блоков.
- Синтез через топологии свободы и ограничений.

В статье не будут рассматриваться эти способы, т. к. они выходят за рамки статьи, а для описанного в статье механизма масштабирующего перемещения использован способ синтеза посредством замены твердого тела.

Что касается анализа, то тут применяется стандартный аппарат теории упругости. Конечно же, МКЭ в различных САПР-программах значительно ускоряет разработку, т.к. дает возможность очень быстро прикладывать нагрузки, а после чего получать перемещения, напряжения и деформации на всех остальных участках механизма. На основе полученных после моделирования нагрузки данных изменяются длины, толщины и прочие параметры, влияющие на упругие взаимосвязи в механизме, с каждой последующей итерацией происходит приближение к требуемой по характеристикам работы механизма с упругими связями.

Заключение

Таким образом, предварительная механическая масштабизация входной неэлектрической величины в такую же неэлектрическую величину позволяет согласовать параметры датчика и параметры измеряемой величины, повысить его чувствительность, а также ряд других преимуществ.

Применение классических масштабизаторов, например, таких, какие используются в конструкциях измерительных головах часового типа, имеет ряд негативных моментов: жесткая привязка к номенклатурному ряду головок часового типа, ограниченный диапазон измерений, наличие механической флуктуации, связанной с неизбежным присутствием кинематического зазора, а также условий эксплуатации, например, деградация смазки или изменение ее свойств вязкости.

Применение механических масштабизаторов с гибкими связями позволяет нивелировать ряд негативных моментов, они имеют повышенную кинематическую точность, проще в изготовлении, зачастую дешевле, могут быть выполнены в любом габаритном исполнении, в том числе в наноразмерном и т.д.

Все это свидетельствует о перспективности применения таких масштабизаторов с гибкими связями. Теоретическая база для их расчета представляет собой непростую задачу ввиду наличия достаточно большого количества элементов с различной жесткостью, которая является функцией их геометрических параметров. МКЭ в целом позволяет оценить подвижность будущего элемента при симуляции, но сам синтез таких механизмов пока что затруднен.

Важное преимущество механических масштабизаторов с гибкими связями состоит в том, что имеется широкий спектр датчиков для считывания, например, линейных размеров после масштабизации, это во многом устраняет проблему подбора компонентов, а также позволяет использовать на менее дорогие из них. Снижение себестоимости конечного изделия, не снижая при этом его качественные параметры, одна из самых актуальных задач современной экономики.

Литература

1. Смит Стивен. Цифровая обработка сигналов. Практическое руководство для инженеров и научных работников. – ДМК ПРЕСС, 2018. – 58 стр.
2. Иванов А.И. Технические измерения – М., 1970. – Рис.105. – 78 стр.
3. https://aliexpress.ru/item/1005005116362275.html?sku_id=12000031738948893&spm=a2g2w.productlist.search_results.0.711269e5mBocQ8
4. Глазунов В.А., Хейло С.В. Механизмы перспективных робототехнических систем. – М.: Техносфера, 2021. – 296 с.
5. Скворцов П.А. Кинематический и прочностной анализ механизма параллельной структуры с упругими шарнирами/Известия высших учебных заведений. – Машиностроение. – №6. – 2024. – С. 40– 46.
6. Пожбелко В.И. Формализация структурного анализа и синтеза механизмов с кинематическими, гибкими и динамическими связями. Вестник ЮУрГУ, №11. – 2007. – С. 21 – 32.
7. Tian Y., Shirinzadeh B., Zhang D. Design and dynamics of a 3-DOF flexure-based parallel mechanism for micro/nano manipulation. Microelectron. Eng., 2010, vol. 87, no. 2, pp. 230– 241, doi: <https://doi.org/10.1016/j.mee.2009.08.001>
8. Lavanya S.B., Jayanth G.R., Mohanty A.K. Optimal design of bridge amplifiers for largesrange linear characteristics. IFToMM WC 2023. Springer, 2023, pp. 721–730, doi: https://doi.org/10.1007/978-3-031-45705-0_70
9. Larry L. Howell – Compliant Mechanisms. Wiley-Interscience 2001. – 112p.
10. Larry L. Howell – Handbook of Compliant Mechanisms. С 2013 John Wiley & Sons Ltd. – 78 p.

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ И ОБРАБОТКА ИНФОРМАЦИИ

УДК 004.75

*В. В. Гайдамако**Кыргызско-Российский (Славянский) университет имени Б.Ельцина, Бишкек
Институт машиноведения, автоматизации и геомеханики Национальной
академии наук, Бишкек***УЧЕТ ЭНЕРГОЗАТРАТ В SIMGRID-МОДЕЛИ ОБЛАЧНОЙ СИСТЕМЫ**

В статье рассматриваются способы учета энергозатрат для вычислительных и сетевых ресурсов при моделировании распределенных и облачных систем, в частности, для модели облачной системы, создаваемой с использованием фреймворка моделирования распределенных и облачных систем Simgrid. Описаны категории моделей энергопотребления, способы снижения энергозатрат, особенности учета энергозатрат в сетях передачи данных. Дано описание моделирования энергопотребления при отсутствии в модели сущностей, представляющих сетевые устройства.

Ключевые слова: облачные вычисления, центр обработки данных (ЦОД), энергозатраты, моделирование, Simgrid.

Введение

Облачные технологии представляют по требованию пользователя не только программы (Software-as-a-Service), но и платформу (Platform-as-a-Service), то есть аппаратное и соответствующее программное обеспечение (ПО), и полноценную инфраструктуру (Infrastructure-as-a-Service) – сетевое, вычислительное и другое оборудование, линии связи, системное и прикладное ПО, работающее на этом оборудовании, а также информационную безопасность и надежность. Организации могут перевести свою инфраструктуру или ее часть в облако, чтобы сократить затраты на создание и содержание собственных вычислительных мощностей. Некоторые задачи, требующие высокопроизводительных вычислений, могут быть непосильными для вычислительной мощности одной организации, но их можно решать, используя ресурсы облака. Однако прежде чем переводить часть или всю работу в облако, требуется оценить эффективность такого решения. Натурный эксперимент слишком затруднен или невозможен, поэтому для оценки различных показателей работы облачной инфраструктуры проводятся с помощью имитационного (компьютерного) моделирования.

Одним из важнейших показателей производительности облачной системы является потребляемая энергия и углеродный след. Углеродный след информационных и коммуникационных технологий (ИКТ) – это весь объем выбросов парниковых газов, вызванных производством, использованием и утилизацией продуктов и услуг ИКТ [1]. Количество энергии, потребляемой облачными центрами обработки данных (ЦОД), постоянно увеличивается, в 2022 году оно составляло 460TWh (тераватт-час), или 2% мирового потребления, и ожидается, что к 2026 году это будет более 1000TWh, если не будут приняты меры по энергосбережению [2], а углеродный след ИКТ, включающих ЦОД, сети, вычисления, компьютеры, смартфоны, серверы и маршрутизаторы, сопоставим с углеродным следом авиационного сектора и составляет 1,8–2,8% мирового углеродного следа [1]. Облачные ЦОД вносят наиболее значимый вклад и в энергопотребление, и в углеродный след, поэтому многие средства моделирования включают энергетические модели, позволяющие разрабатывать и тестировать стратегии оптимизации использования облачных ресурсов.

Средства моделирования, включающие учет энергозатрат

Первым известным симулятором с моделированием энергопотребления в сетях передачи данных является NS2 (Network Simulator 2) с моделированием передачи пакетов, для которого было также разработано расширение для моделирования энергопотребления сети на уровне потока. Он был развит до GreenCloud [3], который разработан для моделирования облаков с использованием различных архитектур: двухуровневой, трехуровневой и трехуровневой высокоскоростной. Он обеспечивает детальное моделирование и анализ энергии, потребляемой элементами сетевых серверов, маршрутизаторов и связей между ними. Кроме того, он анализирует распределение нагрузки по сети, а также коммуникации с высокой точностью (уровень пакетов TCP). С точки зрения энергии GreenCloud определяет три типа энергии: вычисление (ЦП), коммуникации и физический вычислительный центр (система охлаждения) и включает два метода снижения энергии: динамическое масштабирование напряжения для снижения напряжения коммутаторов и динамическое отключение сети, которое позволяет отключать коммутаторы, когда это возможно. Однако GreenCloud может использоваться только для моделирования облачных систем. Кроме того, поскольку он опирается на NS2, его масштабируемость ограничена с точки зрения количества моделируемых устройств и объемов обмена данными.

В качестве замены NS2 был предложен другой симулятор на уровне пакетов NS3. Он предоставляет несколько модулей, включая ECOFEN (Energy Consumption mOdel For End-to-end Networks) [4]. Но пакетная природа NS3 препятствует использованию ECOFEN для крупномасштабных сетей. Пакетные симуляторы дают точные прогнозы, но они не масштабируются. Частота событий симуляторов уровня пакетов очень быстро увеличивается на крупномасштабных платформах с реальными рабочими нагрузками.

Симуляторы уровня потока для создания эффективных моделей производительности работают на более высоком уровне абстракции и поэтому лучше подходят для моделирования крупномасштабных сетей. Они показывают хорошие способности к прогнозированию и могут использоваться для симуляций потребления энергии в сети.

CloudSim относится к симуляторам уровня потока и построен на основе GridSim [5]. Это известный симулятор для моделирования облачной среды с оценкой энергопотребления. Он реализует такие зеленые стратегии, как динамическое масштабирование частоты напряжения и энергосберегающие методы предоставления виртуальных машин. Но так как CloudSim разработан для облачной среды, он не может использоваться для моделирования других распределенных систем. Другой момент заключается в том, что модели CloudSim очень похожи на модели GridSim, которые, как известно, содержат отклонения модели совместного использования полосы пропускания [5].

SimGrid – также симулятор уровня потока, он очень универсален и может использоваться для моделирования высокопроизводительных вычислений, Grid, Peer-to-peer, Fog и т.д. Он хорошо масштабируем и расширения позволяют моделировать энергопотребление [6],

Модели энергопотребления

Понимание связи между использованием ресурсов и потреблением энергии необходимо и для пользователей, и для персонала, ответственного за работу ЦОД, но особенно для разработчиков алгоритмов эффективного распределения ресурсов, так как оптимизация энергопотребления в облаке должна проводиться автоматически. Например, энергоэффективность современного оборудования сильно снижается с понижением нагрузки, поэтому нужно стремиться к полной загрузке какого-то количества физических машин, в то время как другие могут быть отключены. Существует множество моделей энергопотребления, позволяющих проводить автоматическую оптимизацию, в [7] определены общие требования к этим моделям:

- Точность: модель должна быть достаточно точной, чтобы обеспечить желаемую экономию энергии, например, в среднем ошибка не должна превышать 10%, хотя в некоторых случаях (для некоторых алгоритмов оптимизации распределения ресурсов) может потребоваться и более высокая точность.
- Скорость: модель должна генерировать прогнозы достаточно быстро, чтобы ее можно было использовать в политиках энергосбережения; например, одна выборка в секунду достаточно быстра.
- Универсальность и переносимость: структура моделирования должна подходить для как можно большего числа систем, охватывая мобильные, настольные и серверные системы; различные семейства процессоров и технологии памяти и хранения; а также различные потребляемые мощности и динамические диапазоны. Она также должна допускать различное распределение потребляемой мощности между компонентами, а не предполагать доминирования каких-либо компонент в системе. Кроме того, создание моделей для новой системы не должно требовать тщательного исследования пространства проектирования или настройки.
- Низкая стоимость: создание и использование модели не должно требовать дорогостоящей или навязчивой инфраструктуры.
- Простота: при прочих равных условиях простые модели должны быть предпочтительнее сложных как с точки зрения количества входов, так и сложности взаимосвязи между входами и выходами.

Определены две категории моделей энергопотребления: комплексные (comprehensive) модели и высокоуровневые модели, или модели черного ящика. Комплексные модели центрального процессора (ЦП) были предложены в [8] и [9]. Эти модели достигают высокой точности с точки зрения энергопотребления ЦП и опираются на конкретные детали микроархитектуры ЦП. Высокоуровневые модели являются более общими и игнорируют конкретные низкоуровневые детали работы и характеристики моделируемой инфраструктуры. Но они повышают переносимость модели и скорость моделирования. Эти простые модели жертвуют точностью вычислений, но не требуют детального знания архитектуры используемых процессоров.

Высокоуровневый подход чаще применяется в моделировании, а именно линейная модель мощности, используемая при фиксированной частоте:

$$P_{TOT} = (1 - \alpha)P_{CPUIdle} + \alpha P_{CPUFull}$$

где α — загрузка ЦП, а $P_{CPUIdle}$ и $P_{CPUFull}$ — потребляемая ЦП мощность при 0% и 100% использовании соответственно.

Средства экономии энергии

Экономить энергию можно, минимизируя ее использование на каждой машине ЦОД, а также при передаче по сети. В [10] описаны три распространенных инструмента экономии потребляемой энергии, причем необходимо, чтобы все рассматриваемые хосты использовали технологию виртуализации.

Первое решение – метод включения-выключения(ON/OFF), отключает хосты, используемые недостаточно (по сравнению с пороговым значением) и включает их снова при необходимости. Все процессы, запущенные на таком хосте, переносятся на другие хосты, а затем хост отключается. И наоборот, когда все хосты интенсивно используются и поступают новые запросы на обслуживание, один или несколько хостов снова включаются. Этот процесс имеет только одну цель – сократить количество включенных хостов в заданное время.

Второе решение – миграция, перемещение процессов между [11] хостами. Собственно, переносятся не просто процессы, а виртуальные сущности – машины или контейнеры. Миграция позволяет переносить работающую виртуальную среду с одного хоста на другой. Миграция затратна и по времени, и по энергии, когда данные передаются между источником и местом назначения. Также важно учитывать общую стоимость таких действий. Эта техника позволяет освободить некоторые хосты и деактивировать их, в то же время оставшиеся работающие хосты могут использоваться на максимальном уровне (консолидация).

И третье решение – DVFS (Dynamic Voltage and Frequency Scaling, динамическая настройка напряжения и частоты) [12] может динамически изменять напряжение и частоту ЦП хоста в соответствии с его нагрузкой. В ядре Linux DVFS может быть активирован в пяти различных режимах: Performance, PowerSave, UserSpace, Conservative и OnDemand. Каждый режим имеет регулятор, который решает, следует ли изменять частоту (увеличивать или уменьшать) или нет. Три из этих пяти режимов используют фиксированную частоту: Performance использует ЦП на его самой высокой частоте, PowerSave использует самую низкую частоту, а режим UserSpace позволяет пользователю выбрать одну из доступных частот для ЦП.

Два последних режима, Conservative и OnDemand, являются динамическими. Это означает, что частота ЦП может меняться со временем в зависимости от нагрузки ЦП. Регуляторы этих двух режимов работают с пороговыми значениями (одним или двумя) и периодически проверяют, находится ли нагрузка ЦП ниже (или выше) этих пороговых значений, после чего принимается решение об изменении частоты ЦП.

Консервативный регулятор работает с верхним порогом `up_threshold` и нижним `down_threshold`. Когда нагрузка ЦП превышает `up_threshold`, частота увеличивается, а когда нагрузка ЦП ниже `down_threshold`, частота ЦП уменьшается. Этот режим является прогрессивным, и каждое изменение частоты ЦП выполняется с шагом один с использованием всех доступных частот.

Режим OnDemand (по запросу) использует только один порог и способствует производительности системы, напрямую устанавливая самую быструю частоту, когда нагрузка ЦП превышает порог. Уменьшение частоты ЦП, выполняемое с шагом один, как в консервативном режиме, выполняется, если нагрузка ЦП остается ниже порога в течение предопределенного периода времени. Более низкая частота, которая подразумевает более слабое напряжение, снижает энергопотребление ЦП, но также замедляет вычислительную мощность ЦП.

DVFS в Simgrid. Модуль учета энергопотребления Host Energy

Для учета потребления энергии в Simgrid используется подключаемый модуль (плагин) Host Energy. Этот плагин позволяет вести учет энергии и во время вычислений, и во время простоя, и даже рассеивания энергии. Плагин должен быть активирован перед загрузкой платформы (sg_host_energy_plugin_init()), для получения данных о потреблении данной ФМ используется sg_host_get_consumed_energy(). При включенной ФМ потребление энергии зависит от текущей загрузки ЦП и от профиля энергопотребления хоста. Simgrid считает, что потребление линейно зависит от количества ядер на полной скорости, с аномалией, когда все ядра простаивают (потребление не уменьшается до минимума) [6]. Энергетическая модель принимает 4 параметра:

- Idle – потребление при бездействии (т. е. мгновенное потребление в ваттах), когда хост запущен и работает, но ему нечего делать, у него нет нагрузки;
- Epsilon – потребление, когда все ядра находятся в состоянии 0 или Epsilon %, но не в состоянии бездействия;
- AllCores – потребление при полной загрузке всех ядер (100 %);
- Off – потребление, когда хост выключен.

Пример описания хоста:

```
<host id="HostA" speed="100.0Mf" core="4">
  <prop id="wattage_per_state" value="100.0:120.0:200.0" />
  <prop id="wattage_off" значение="10" />
</host>
```

Если заданы только два значения, для отсутствующего значения Idle используется Epsilon. В этом примере даны следующие параметры: Off = 10 Вт; холостой ход (Idle) 100 Вт; Эпсилон – 120 Вт, а AllCores – 200 Вт. Этого достаточно, чтобы рассчитать потребление в зависимости от количества загруженных ядер (таблица 1):

Таблица 1 – Потребление в зависимости от количества загруженных ядер

Загрузка ядер	Вт	Описание
0 (простой)	100	Значение Idle
0 (но нет простоя)	120	Значение Epsilon
1	140	Линейная экстраполяция между Epsilon и AllCores
2	160	Линейная экстраполяция между Epsilon и AllCores
3	180	Линейная экстраполяция между Epsilon и AllCores
4	200	Линейная экстраполяция между Epsilon и AllCores

Если у хоста несколько уровней потребления (pstates) DVFS (Dynamic voltage and frequency scaling) [12], например, у машин с графическими процессорами), указывается энергетический профиль каждого уровня потребления энергии (pstate):

```
<host id="HostC" speed="100.0Mf,50.0Mf,20.0Mf" core="4">
  <prop id="wattage_per_state"
    value="95.0:120.0:200.0, 93.0:115.0:170.0, 90.0:110.0:150.0" />
  <prop id="wattage_off" value="10" />
</host>
```

Получаемые значения приведены в таблице 2.

Таблица 2 – Значения потребления для различных режимов работы процессора

Уровень потребления (pstate)	Производительность	Idle	Epsilon	AllCores
0	100 Mflop/s	95 Вт	120 Вт	200 Вт
1	50 Mflop/s	93 Вт	115 Вт	170 Вт
2	20 Mflop/s	90 Вт	110 Вт	150 Вт

Энергопотребление каналов связи напрямую зависит от текущей загруженности трафиком. В файле платформы это потребление может быть указано следующим образом:

```
<link id="SWITCH1" bandwidth="125Mbps" latency="5us"
sharing_policy="SHARED" >
  <prop id="waitage_range" value="100.0:200.0" />
  <prop id="wattage_off" value="10" />
</link>
```

Первое свойство означает, что когда линия связи включена, но бездействует, она будет потреблять 100 Вт. При полной загрузке будет рассеиваться 200 Вт, при 50% нагрузки – 150 Вт. Второе свойство означает, при выключенной линии будет рассеиваться 10 Вт. Для запуска плагина используется функция `sg_link_energy_plugin_init()`, перед загрузкой платформы, для получения потребления линии связи – функция `sg_link_get_consumed_energy()` [6].

Особенности энергопотребления в сетях связи

Распределенные вычисления широко используются с момента появления Интернета, их роль и появление новых технологий (облачные системы, Интернет вещей. Fog-, Mist, Edge-системы) требуют повышения производительности сетей передачи данных, и их энергопотребление становится все более важным показателем для оптимизации. Именно в сетях связи наблюдается самый высокий рост потребления энергии – 10%, тогда как для пользовательских устройств он составляет 5%, для ЦОД – 4% [2]. Для сдерживания этой тенденции требуются энергоэффективные решения, для поиска которых и создаются модели. Модели энергопотребления в сетях, основанные на передаче пакетов, учитывающие каждый передаваемый пакет, точны, но с ростом количества устройств становятся малоприменимыми, слишком сложными для описания, и слишком сложными в интерпретации результатов.

Симуляторы уровня потока при работе не моделируют каждый пакет, тем самым экономя память и вычисления, практически представляя собой хороший компромисс между абстракцией сети и масштабируемостью с точки зрения использования памяти и времени вычислений. В [13] на примере симуляторов NS3 и Simgrid показано, что моделирование потребления энергии сети в симуляторе уровня потока приводит к прогнозам, которые соответствуют прогнозам симуляторов уровня пакета, а также, что простых линейных моделей достаточно для получения точных значений.

Классическая модель

Особенностью сетевых устройств является то, что даже если через них не проходит сетевой трафик, они должны оставаться включенными, чтобы в любой момент быть готовыми к принятию и передаче пакета. Эта часть энергопотребления в простое называется также статической. В некоторых случаях стоимость энергии простоя сетевого устройства может составлять 85% от его пиковой мощности [14]. Классически энергопотребление оценивается с использованием определения энергии

$$E(t) = \int_0^t P(t)dt$$

Потребление энергии для сетевого оборудования меняется со временем в соответствии с сетевым трафиком. Поскольку его мощность состоит из двух частей: статической и динамической, мы можем определить ее уравнением 1:

$$P_{netdev} = \underbrace{P_{idle}}_{\text{Static}} + \underbrace{\sum_{i \in S} N_{pi} \times P_{pi} + N_{bi} \times P_{bi}}_{\text{Dynamic}} \quad (1)$$

где S — набор портов на данном сетевом устройстве, N_{pi} и N_{bi} — соответственно количество пакетов и количество байтов, проходящих через порт i , а P_{pi} и P_{bi} — стоимость энергии на пакет и на байт для порта [13].

Существующие энергетические модели в литературе основаны на этом уравнении, и, следовательно, они несовместимы напрямую с симуляторами уровня потока. Действительно, симуляторы уровня потока не предоставляют никакого понятия порта в своих моделях или пакетах. Энергетические модели должны быть адаптированы к сетевому представлению потока.

Представление сетевого соединения в Simgrid

Симулятор SimGrid не содержит контроллеров сетевых интерфейсов и маршрутизаторов. Вместо этого пути между двумя хостами представлены маршрутами (Routes). Каждый маршрут состоит из нескольких сущностей, называемых связями (Link), которые представляют как саму линию связи, так и сетевой адаптер на каждом конце. Маршрут между двумя хостами может состоять из нескольких связей, но не обязательно из нескольких хостов. В симуляторах уровня потока нет понятия порта, но потребление энергии портами может быть включено в потребление энергии самими связями. Затем можно определить мощность всей сетевой платформы с помощью уравнения

$$P_{total} = \sum_{i \in L} LinkPower_i \times LinkUsage_i$$

где L представляет набор связей на моделируемой сетевой платформе.

Однако в каждой сети каждая связь подключена к двум разным портам, возможно, с разными схемами потребления энергии. Таким образом, абстракция симуляторов уровня потока создает трудности для транспонирования в них моделей энергии на уровне пакетов. Чтобы решить эту проблему, в [13] определяются две модели энергии, одна для однородной энергетической платформы (одинаковое потребление энергии для портов на данном соединении) и одну для неоднородной энергетической платформы. Это различие на самом деле можно упростить.

Однородная (гомогенная) модель для связи

Для линий связи (Link) нужно учесть потребление энергии каждым из портов и потребление при использовании полосы пропускания каждым из участников. Для достижения этой цели в [13] предлагается использовать одно сплит- или раздельно-дуплексное соединение (split-duplex) между конечными узлами, то есть полнодуплексная

связь в модели представляется двумя односторонними линиями, которые обозначаются как восходящая и нисходящая, или исходящая и входящая, или UP и DOWN. Так как в SimGrid нет портов, для учета двух портов (по одному на каждом конце соединения), динамическое потребление энергии каждым соединением удваивается. Поскольку раздельно-дуплексные каналы приводят к созданию двух соединений, потребление простоя уже будет удвоено, и нужно удвоить только динамическую часть потребления. Энергия, потребляемая портами, описывается линейной моделью с минимальным значением, равным потреблению энергии в режиме простоя, и максимальным значением, определяемым как:

$$max = idle + (BW \times ByteCons + PktCons/MTU) \times 2 \quad (2)$$

В этом уравнении BW представляет скорость порта в бит/с, ByteCons – потребление энергии на байт в джоулях, P ktCons – потребление энергии сетевым устройством в джоулях, необходимое для обработки пакета. Наконец, MTU, максимальная единица передачи, используется в качестве приближения размера пакета [13]. При рассмотрении односторонней связи канал потребляет энергию для передающего (Tx) и принимающего (Rx) портов одновременно и наоборот. Если есть несколько потоков, разделение полосы пропускания происходит на восходящих или нисходящих каналах в соответствии с направлениями потоков.

Эта энергетическая модель имеет несколько преимуществ. Во-первых, ее легко реализовать – на канал приходится только одно значение расхода энергии, что упрощает реализацию на симуляторах уровня потока. Эти симуляторы всегда используют каналы для связи (но не обязательно порты). Эта модель обеспечивает низкие вычислительные издержки, однако во многих реальных топологиях используются неоднородные сетевые устройства, и потребление энергии от одного сетевого устройства к другому варьируется в зависимости от его спецификации. Такая неоднородность не может быть явно представлена с помощью однородной модели.

Неоднородная (гетерогенная) модель для связи

В SimGrid маршруты от одного хоста к другому могут состоять из нескольких ссылок. Поэтому можно воспользоваться этой функцией для построения гетерогенной модели.

Для моделирования двух портов с разным потреблением энергии на одном канале в [13] вводятся маршруты из трех связей. Первая связь моделирует потребление энергии первого порта, вторая связь управляет совместным использованием полосы пропускания, и третья связь моделирует потребление энергии второго порта. Так как в модели используются разделенные дуплексные связи, SimGrid создаст две связи, и, следовательно, мощность простоя будет умножена на два. Чтобы решить эту проблему, мощность простоя каждого энергетического канала нужно разделить на два. Таким образом, эта энергетическая модель следует линейному поведению с минимальным значением, равным $idle/2$, и максимальным значением, определяемым как

$$max = idle/2 + (BW \times ByteCons + PktCons/MTU) \quad (3)$$

Итак, в Simgrid-модели возможно осуществлять учет потребления энергии и для вычислительных устройств, и для сетей передачи данных. Однако модели энергопотребления, перед тем как они будут включены в рабочую «большую» модель, должны пройти «обкатку» на небольших тестовых моделях, пройти валидацию [15].

Поэтому первым шагом в эксперименте будет создание «микромоделей», интерпретация наблюдений которых достаточно проста, результаты можно проверить теоретическими вычислениями или на доступных реальных системах.

Примеры определения количества потребленной энергии

Для учета энергопотребления ЦП Simgrid предоставляет плагины HostEnergy, HostDVFS, для связей – плагины LinkEnergy, WifiEnergy [6].

Плагин HostEnergy содержит функции, позволяющие получить количество потребленной хостом энергии (sg_host_get_consumed_energy), количество энергии, потребляемой хостом в состоянии простоя (sg_host_get_idle_consumption), минимальное и максимальное потребление в указанном состоянии (sg_host_get_wattmin_at, sg_host_get_wattmax_at).

Пример 1. Потребление хоста в разных состояниях. В файле платформы для хоста указывается потребление DVFS [6]:

```
<!-- Атрибут 'pstate' описывает начальное состояние, наименьшее значение pstate соответствует максимальной скорости процессора) -->
```

Первый хост, 4 ядра:

```
<host core="4" id="Host1" pstate="0" speed="100.0Mf,50.0Mf,20.0Mf">
<!--Список Idle:Epsilon:AllCores (Вт) соответствует потреблению во время простоя, при загрузке
epsilon у всех ядер, при полной загрузке всех ядер-->
<!--Список должен содержать энергетический профиль для каждого описанного состояния pstate,
здесь 100.0Mf,50.0Mf,20.0Mf -->
<prop id="wattage_per_state" value="100.0:93.33333333333333:200.0, 93.0:90.0:170.0,
90.0:90.0:150.0" />
```

При отключении питания

```
<prop id="wattage_off" value="10" />
</host>
```

Второй хост, одно ядро, работает при полной загрузке или простаивает (Epsilon=AllCores):

```
<host core="1" id="Host2" pstate="0" speed="100.0Mf,50.0Mf,20.0Mf">
<prop id="wattage_per_state" value="100.0:200.0:200.0, 93.0:170.0:170.0, 90.0:150.0:150.0" />
<prop id="wattage_off" value="10" />
</host>
```

Третий хост, одно ядро

```
<host core="1" id="Host3" pstate="0" speed="100.0Mf,50.0Mf,20.0Mf">
<prop id="wattage_per_state" value="100.0:200.0:200.0, 93.0:170.0:170.0, 90.0:150.0:150.0" />
<prop id="wattage_off" value="10" />
</host>
```

Описание коммуникации:

```
<link bandwidth="100kBps" id="bus" latency="0" sharing_policy="SHARED">
<!--диапазон потребления, близкий к реальному <prop id="wattage_range" value="10.3581:10.7479"
/> -->
<!-- для теста -->
<prop id="wattage_range" value="1:3" />
</link>
```

Маршруты:

```
<route dst="Host2" src="Host1">
<link_ctn id="bus" />
```

```

</route>
<route dst="Host3" src="Host1">
  <link_ctn id="bus" />
</route>
<route dst="Host3" src="Host2">
  <link_ctn id="bus" />
</route>
</zone>
</platform>

```

Эта простая платформа из трех хостов, все хосты соединены между собой одной линией связи, используется далее во всех примерах.

При запуске моделирования включаются все три хоста. Затем Host1 и Host2 выполняют действия, состоящие из периодов ожидания, процессор потребляет минимальное количество энергии, и вычислений $1E+08$ flops (операций с плавающей запятой).

В таблице 1 показаны затраты энергии на каждом этапе выполнения для работающего хоста Host1.

Таблица 1 – Результат выполнения для Host1, профиль 100.0:93.33333333333333:200.0, 93.0:90.0:170.0, 90.0:90.0:150.0

Действие	Pstate	Текущая скорость flop/s	пиковая	Время, с	Потребление Вт	Затраты энергии Дж
Sleeping	1	$1E+08$		10	0	1000
Вычисления	1	$1E+08$		1	93-200	1120
Вычисления	2	$2E+07$		5		1645
Sleeping	2	$2E+07$		10		2545

В таблице 2 Показаны затраты и энергии на каждом этапе выполнения для хоста Host2.

Третий хост был подключен к питанию, но не выполнял никаких действий.

Таблица 2 – Результат выполнения для Host2, профиль 100.0:200.0:200.0, 93.0:170.0:170.0, 90.0:150.0:150.0

Действие	Pstate	Текущая скорость flop/s	пиковая	Время, с	Потребление Вт	Затраты энергии Дж
Sleeping	1	$1E+08$		10		2000
Вычисления	1	$1E+08$		1	93-200	1200
Вычисления	2			5		1950
Sleeping	2	$2E+07$		4		2310

Общие затраты энергии, затраты на работающих хостах и на простаивающих хостах в Джоулях:

Total energy consumption: 9515.000000 Joules (used hosts: 5915.000000 Joules; unused/idle hosts: 3600.000000)

Затраты по хостам:

Energy consumption of host Host1: 3445.000000 Joules

Energy consumption of host Host2: 2470.000000 Joules

Energy consumption of host Host3: 3600.000000 Joules

Как видим, простаивающий хост потребляет очень много энергии, поэтому и нужны алгоритмы балансировки нагрузки и миграция ВМ, недогруженные машины потребляют много энергии, причем впустую.

Пример 2. Потребление при коммуникации (линия связи и сетевое оборудование источника и приемника). Используется та же платформа, с одной линией связи 100 Кб/с. Плагин активируется функцией `sg_link_energy_plugin_init()`; после этого выполняется передача сообщения через почтовый ящик, и в конце симуляции платформа выдает потребленную энергию. Передача осуществляется одним потоком (все сообщение сразу) или с помощью `put_async()` создается несколько потоков передачи, каждый передает сообщение, количество потоков и размер сообщения задаются аргументом командной строки.

Для сообщения 25000 байт одним потоком:

```
[10.250000] (0:maestro@) Total energy over all links: 10.750000
[10.250000] (0:maestro@) Energy consumption of link 'bus': 10.750000 Joules
```

Двумя потоками:

```
[10.541237] [link_energy/INFO] Total energy over all links: 11.623711
[10.541237] [link_energy/INFO] Energy consumption of link 'bus': 11.623711 Joules
```

Для 4 потоков:

```
[11.082474] [link_energy/INFO] Total energy over all links: 13.247423
[11.082474] [link_energy/INFO] Energy consumption of link 'bus': 13.247423 Joules
```

Для передачи 50000000 байт одним потоком:

```
[510.000000] (0:maestro@) Total energy over all links: 1510.000000
[510.000000] (0:maestro@) Energy consumption of link 'bus': 1510.000000 Joules
```

Двумя потоками:

```
[1092.474227] [link_energy/INFO] Total energy over all links: 3257.422680
[1092.474227] [link_energy/INFO] Energy consumption of link 'bus': 3257.422680 Joules
```

Для 4 потоков:

```
[2174.948454] [link_energy/INFO] Total energy over all links: 6504.845361
[2174.948454] [link_energy/INFO] Energy consumption of link 'bus': 6504.845361 Joules
```

Для относительно небольшого объема передачи увеличение количества потоков незначительно увеличивают и время передачи, и потребленную энергию. На больших объемах увеличение времени передачи и потребления очень значимо.

Пример 3. Затраты на виртуальных машинах. С хоста Host1 запускаются две ВМ на Host1 и Host2, используют по одному ядру:

```
auto* vm_host1 = host1->create_vm("vm_host1", 1);
vm_host1->start();
auto* vm_host2 = host2->create_vm("vm_host2", 1);
vm_host2->start();
```

На всех хостах запускаются по две активности:

- обе внутри ВМ на Host1;
- одна на ВМ, другая на хосте на Host2;
- обе на хосте Host3, без ВМ.

Все эти активности симулируют выполнение $1E+8$ операций с плавающей точкой.
Результат:

```
[ 10.000000] (0:maestro@) Energy consumption of host MyHost1: 1120.000000 Joules
[ 10.000000] (0:maestro@) Energy consumption of host MyHost2: 1600.000000 Joules
[ 10.000000] (0:maestro@) Energy consumption of host MyHost3: 1600.000000 Joules
```

Выполнение на двух ВМ привело к меньшему потреблению энергии, то есть использование виртуальных ресурсов приводит к заметной экономии энергии.

Заключение

Показатели потребления энергии как вычислительными ресурсами, так и сетями передачи данных, являются очень важными для оценки производительности распределенных систем, включая облачные, Fog- Mist- Edge-системы и Интернет вещей, прилагается множество усилий по оптимизации и, следовательно, снижению потребления энергии всеми элементами этих систем. Для разработки и тестирования алгоритмов оптимизации используются имитационные модели распределенных и облачных систем, в статье приведены примеры симуляторов NS2, NS3, GreenCloud, CloudSim. Для дальнейших экспериментов будет использоваться платформа моделирования Simgrid, для верификации моделей могут использоваться другие платформы. Для Simgrid приведены примеры описания энергопотребления хостов в состоянии простоя (Idle), загрузке Epsilon%, нагрузке 100% (AllCores), в выключенном состоянии (Off). Для хостов, имеющих несколько уровней потребления (DVFS) указывается энергетический профиль каждого уровня. Энергопотребление при передаче данных в Simgrid моделируется через сущность Link, представляющую как саму линию связи, так и сетевые устройства, которые она связывает, причем эти устройства могут иметь как одинаковое, так и разное энергопотребление. Таким образом, в Simgrid возможно создать модель энергопотребления для всей платформы.

Литература

1. Будущее углеродного следа сектора информации и связи (ИКТ). URL: https://www.databridgemarketresearch.com/ru/whitepaper/future-of-carbon-footprint-of-information-and-communication?srsId=AfmBOorN1KV53aAKCIHyakDtbNsZRnBPjZRYlYKISE_betTbMrPa1ZEz (дата обращения: 10.10.2024)
2. Global data center electricity use to double by 2026 - IEA report. URL: <https://www.datacenterdynamics.com/en/news/global-data-center-electricity-use-to-double-by-2026-report> (дата обращения: 10.10.2024)
3. Kliazovich, D., Bouvry, P., Audzevich, Y., Khan, S.U.: GreenCloud: A Packet-Level Simulator of Energy-Aware Cloud Computing Data Centers. In: IEEE GLOBECOM (2010)
4. Orgerie, A.C., et al.: Simulation Toolbox for Studying Energy Consumption in Wired Networks. In: Int. Conf. on Network and Service Management (2017)
5. CloudSim Energy-aware Simulations. URL: <https://cloudsimtutorials.online/cloudsim-energy-aware-simulations/> (дата обращения: 10.10.2024).

6. Simulation of Distributed Computer Systems // URL: <https://simgrid.org/> (дата обращения: 05.10.2024)
7. S. Rivoire, P. Ranganathan, C. Kozyrakis, A comparison of high-level full-system power models, in: Proceedings of the 2008 Conference on Power Aware Computing and Systems, HotPower'08, USENIX Association, Berkeley, CA, USA, 2008, p. 3.
8. R. Joseph, M. Martonosi, Run-time power estimation in high performance microprocessors, in: Proceedings of the 2001 International Symposium on Low Power Electronics and Design, ISLPED '01, ACM, New York, NY, USA, 2001, pp. 135–140
9. C. Isci, M. Martonosi, Runtime power monitoring in high-end processors: methodology and empirical data, in: Proceedings of the 36th Annual IEEE/ACM International Symposium on Microarchitecture, MICRO 36, IEEE Computer Society, Washington, DC, USA, 2003, p. 93.
10. Guérout, Tom and Monteil, Thierry and DaCosta, Georges and Neves Calheiros, Rodrigo and Buyya, Rajkumar and Alexandru, Mihai Energy-aware simulation with DVFS. (2013)Simulation Modelling Practice and Theory, vol. 39. pp. 76-91. ISSN1569-190X
11. C.C. Keir, C. Clark, K. Fraser, S. Hand, J.G. Hansen, E. Jul, C. Limpach, I. Pratt, A. Warfield, Live migration of virtual machines, in: Proceedings of the 2nd ACM/USENIX Symposium on Networked Systems Design and Implementation (NSDI), 2005, pp. 273–286.
12. T. Kolpe, A. Zhai, S. Sapatnekar, Enabling improved power management in multicore processors through clustered dvfs, in: Design, Automation Test in Europe Conference Exhibition (DATE), 2011, pp. 1–6
13. Loic Guegan, Betsegaw Lemma Amersho, Anne-Cécile Orgerie, Martin Quinson. A Large-Scale Wired Network Energy Model for Flow-Level Simulations. AINA 2019 - 33rd International Conference on Advanced Information Networking and Applications, Mar 2019, Matsue, Japan. pp.1047-1058, ff10.1007/978-3-030-15032-7_88ff. fffhal-02020045v2ff
14. Fiandrino, C., Kliazovich, D., Bouvry, P., Zomaya, A.Y.: Performance Metrics for Data Center Communication Systems. In: IEEE CLOUD (2015)
15. Velho, P., Schnorr, L.M., Casanova, H., Legrand, A.: On the validity of flow-level tcp network models for grid and cloud simulations. ACM Transactions on Modeling and Computer Simulation 23(4) (2013)

УДК 631.4:004.8

*М.С. Макаров**С.Н. Верзунов, verzunov@hotmail.com**Кыргызско-Российский Славянский университет им. Б.Н. Ельцина*

МОДЕЛИ АРХИТЕКТУРЫ И АЛГОРИТМЫ ФУНКЦИОНИРОВАНИЯ ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ МОНИТОРИНГА СОСТОЯНИЯ ВОДЫ И ПОЧВ СЕЛЬСКОХОЗЯЙСТВЕННЫХ УГОДИЙ

В статье представлены модели архитектуры и алгоритмы функционирования интеллектуальной системы мониторинга состояния воды и почв сельскохозяйственных угодий. Целью исследования является разработка адаптивной системы, обеспечивающей точный и оперативный мониторинг параметров почвы, таких как влажность, солевой состав и температура, с применением методов анализа больших данных и машинного обучения. Основное внимание уделено использованию сенсорной инфраструктуры для сбора данных, их предварительной обработке и интеграции с серверными модулями анализа. Рассматриваются подходы к прогнозированию возможных изменений в состоянии почвы с помощью нейронных сетей, что позволяет повысить точность агротехнических решений и минимизировать риски, связанные с неблагоприятными условиями. Предложенная архитектура системы отличается высокой масштабируемостью, гибкостью и возможностью интеграции с существующими информационно-управляющими системами в сельском хозяйстве. Результаты работы направлены на повышение эффективности использования ресурсов и устойчивость агроэкосистем в условиях изменяющегося климата.

Ключевые слова: интеллектуальная система мониторинга, сельскохозяйственные угодья, параметры почвы, машинное обучение, архитектура системы.

Введение

Интенсивное развитие сельского хозяйства в условиях изменяющегося климата требует внедрения современных технологий для повышения эффективности управления состоянием почвы и воды на угодьях. Ключевую роль в этом процессе играют интеллектуальные системы мониторинга, которые позволяют отслеживать изменения параметров среды в режиме реального времени, выявлять отклонения и прогнозировать возможные риски для урожайности [1, 2].

Диэлектрическая проницаемость почвы является важным индикатором, отражающим ее физико-химические свойства, такие как влажность, солевой состав и плотность [3]. Эти параметры существенно влияют на продуктивность сельскохозяйственных культур и эффективность использования ресурсов. Однако традиционные методы анализа почвы требуют значительных временных и финансовых затрат, что делает их малоэффективными для оперативного принятия решений [4].

В данной статье рассматриваются подходы к разработке интеллектуальных систем мониторинга состояния почвы и воды с использованием современной сенсорной инфраструктуры, методов анализа больших данных и алгоритмов машинного обучения [5]. Предложенные модели и алгоритмы направлены на создание адаптивных систем, которые обеспечивают сбор и обработку данных с сельскохозяйственных угодий, визуализацию результатов и формирование рекомендаций для агротехнических мероприятий [6].

Цель исследования – представить архитектурные решения и алгоритмы функционирования интеллектуальной системы мониторинга, обеспечивающей точное и своевременное управление состоянием почвы и воды, а также способствующей оптимизации сельскохозяйственного производства [7].

Модели архитектуры и алгоритмы функционирования ИСГМ

Для обеспечения точного и своевременного мониторинга состояния почвы интеллектуальной системой геоэкологического мониторинга (ИСГМ) используется набор

датчиков, собирающих информацию о ключевых параметрах окружающей среды и почвы. Данные, фиксируемые датчиками, передаются на локальные модули предварительной обработки, которые структурируют и передают метрики в центральную систему анализа. Обработка сигналов включает три основных этапа: усиление, фильтрацию и преобразование. Специализированные устройства выполняют эти процессы для повышения точности и устранения помех, обеспечивая надежность последующего анализа данных.

После первичной обработки локальные модули передают данные на серверный модуль посредством проводных или беспроводных технологий, таких как Wi-Fi, в режиме реального времени, обеспечивая непрерывный поток информации для интеграции и анализа.

На сервере реализованы алгоритмы для дальнейшей обработки данных, включая их сбор, первичную обработку, аналитическую оценку и интеграцию с другими системными компонентами. Серверы также поддерживают пользовательские интерфейсы, позволяющие пользователям взаимодействовать с системой, настраивать параметры мониторинга и получать доступ к аналитической информации.

Применяются статистические методы для определения тенденций и выявления аномалий в данных, включая расчет средних значений, стандартных отклонений и вероятностных распределений. Методы вероятностного моделирования, такие как метод Монте-Карло и байесовские сети, используются для оценки рисков и вероятностей их реализации.

Алгоритмы машинного и глубокого обучения в интеллектуальной подсистеме идентификации физико-химических свойств ИПИФХС анализируют данные для обнаружения сложных шаблонов и зависимостей, указывающих на потенциальные риски. Модели, включая нейронные сети, решающие деревья и ансамблевые методы, улучшают точность прогнозов [8], что помогает в планировании мер по минимизации последствий потенциальных аварий.

На основе анализа формируются рекомендации по улучшению почвы, планы мелиорации и другие меры реагирования на чрезвычайные ситуации. Систематическое обновление информации обеспечивает информирование всех заинтересованных сторон о возможных рисках и стратегиях их минимизации [9].

В качестве примера рассмотрим использование интеллектуальной системы мониторинга для наблюдения за состоянием сельскохозяйственных угодий. На поле устанавливаются датчики, такие как сенсоры влажности для измерения уровня воды в почве, датчики температуры для контроля микроклимата, а также сенсоры содержания солей и питательных веществ.

Эти устройства постоянно собирают данные и передают их на центральный сервер через беспроводные сети. Технология потоковой обработки данных, реализованная на сервере, позволяет мгновенно анализировать полученную информацию. Если анализ показывает, что влажность почвы снижается до критического уровня, система автоматически генерирует уведомление и отправляет его фермеру [10].

При дальнейшем ухудшении параметров система может инициировать автоматическое включение системы орошения или предложить план внесения удобрений для восстановления оптимального состояния почвы. Анализ собранных данных позволяет выявлять причины изменений, например, избыточное испарение или засоление, и разрабатывать долгосрочные стратегии по улучшению условий выращивания сельскохозяйственных культур. На рисунке 1 [11] представлена низкоуровневая модель работы интеллектуальной системы мониторинга состояния сооружений в формате Event-driven Process Chain (EPC), созданная в среде ARIS для моделирования рабочих процессов. EPC является методом моделирования бизнес-процессов, основанным на последовательности событий и функций, отражающих структуру и логику процессов.

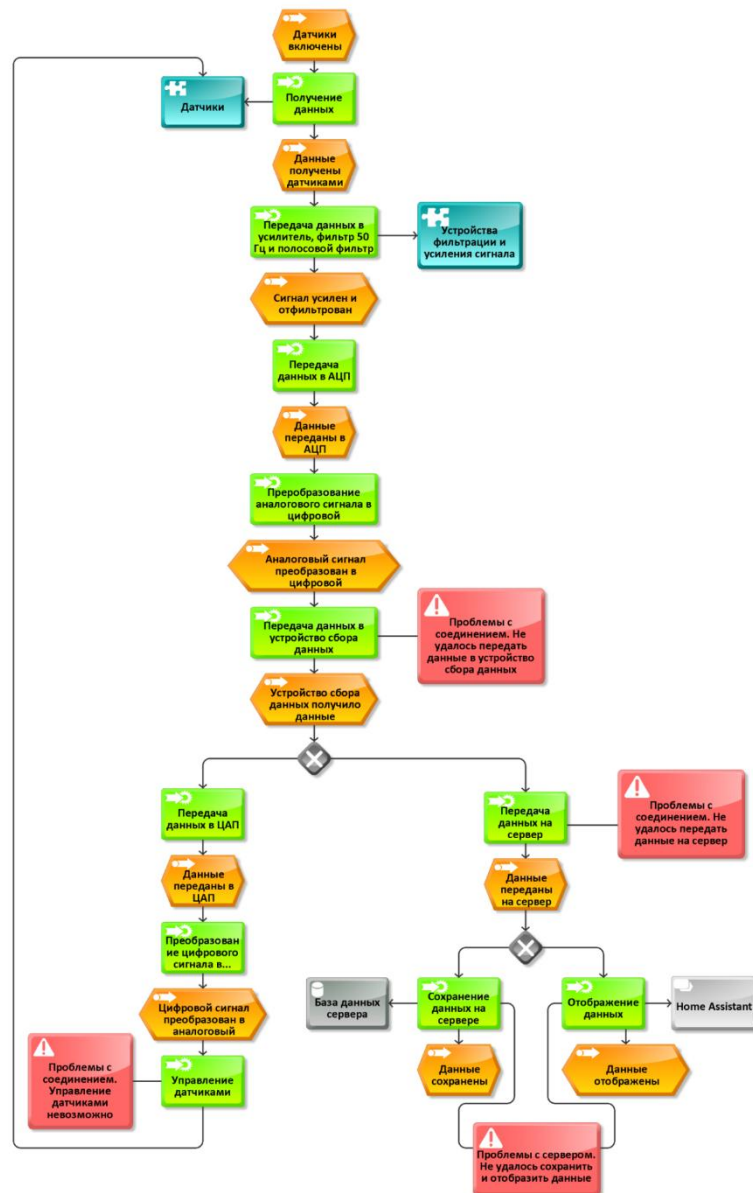


Рисунок 1 – Модель функционирования интеллектуальной системы геоэкологического мониторинга

События в EPC-диаграммах обозначают состояния, запускающие последующие действия, что делает EPC эффективным инструментом для анализа и оптимизации процессов.

Процесс мониторинга начинается с активации датчиков, осуществляющих сбор первичных данных о состоянии объектов. Эти данные усиливаются и фильтруются для улучшения качества перед цифровой обработкой. Аналоговые сигналы преобразуются в цифровую форму с помощью аналого-цифровых преобразователей (АЦП), после чего локальный модуль обрабатывает данные и подготавливает их к передаче на сервер.

В локальном модуле реализуется обратная связь через цифро-аналоговые преобразователи (ЦАП) для управления датчиками, позволяя корректировать их работу в реальном времени. Обработанные данные передаются на сервер через сетевые интерфейсы.

На сервере данные сохраняются в базе данных и отображаются в пользовательском интерфейсе системы управления.

Анализ модели выявляет потенциальные риски и недостатки реализации концепции:

1. Слабое или отсутствующее интернет-соединение может нарушить передачу данных на сервер, что приведет к сбоям в мониторинге в реальном времени.

2. Работа сервера может быть нарушена из-за ошибок в работе программного обеспечения, взаимодействующего с внешними устройствами и данными.
3. Необходимость постоянного обновления ядра и плагинов серверного ПО для поддержания совместимости и безопасности системы может стать проблемой. Стабилизация на одной версии программного обеспечения может повысить устойчивость системы, но со временем уменьшить ее функциональность из-за устаревания компонентов.

В представленной модели функционирования интеллектуальной системы геоэкологического мониторинга учтены эти ключевые аспекты. Так как основными задачами ИСГМ являются контроль параметров (влажность, количество солей, температура и т.д.), то важно обеспечить интеграцию с существующими системами управления объектами для прогнозирования потенциальных рисков. Это позволяет не только предотвращать аварии, но и оптимизировать эксплуатацию объектов.

Такие системы включают датчики, серверную инфраструктуру, программные решения для анализа данных и пользовательские интерфейсы, обеспечивая высокую адаптивность и масштабируемость, что особенно важно в засушливых регионах [7, 10].

Функциональные требования к ИСГМ

Функциональные требования к ИСГМ включают сбор, обработку и анализ данных, поступающих от различных сенсоров и измерительных устройств, установленных на объектах мониторинга. Система должна обеспечивать визуализацию обработанных данных, автоматизированное формирование отчетной документации и оповещение о выявленных критических изменениях в состоянии контролируемых объектов.

Ключевым компонентом функциональности является внедрение ИСГМ для прогнозирования возможных деформационных изменений зданий, сооружений или их фундаментов и оценки связанных с этим рисков. В таблице 1 представлена детализация функциональных требований системы.

Таблица 1 – Функциональные требования ИСГМ

Требование	Описание
Сбор данных	- Система должна собирать данные с различных типов датчиков в режиме реального времени. - Система должна поддерживать настройку режимов функционирования датчиков
Обработка и анализ данных	- Система должна выполнять первичную обработку и фильтрацию полученных данных. - Система должна проводить анализ данных для выявления отклонений от нормальных показателей
Прогнозирование	- Система должна осуществлять прогнозирование потенциальных проблем и других возможных геоэкологических изменений на основе исторических и текущих данных. - Система должна использовать методы машинного обучения для улучшения точности прогнозов с течением времени. - Система должна предоставлять визуализацию прогнозов в виде графиков, диаграмм и т.п. - Система должна генерировать предупреждения о потенциальных рисках на основе прогнозных данных. - Система должна позволять настраивать параметры прогнозирования для различных типов объектов и грунтов
Визуализация	- Система должна предоставлять упрощенные 2D или 3D-модели объектов мониторинга с отображением текущих показателей и прогнозных данных. - Система должна генерировать графики и диаграммы для отображения тенденций изменения параметров во времени
Оповещения и уведомления	- Система должна отправлять мгновенные уведомления при превышении критических порогов показателей. - Система должна поддерживать настройку пользовательских триггеров для оповещений, включая прогнозные триггеры

Управление объектами мониторинга	- Система должна позволять добавлять новые датчики и объекты мониторинга с указанием их характеристик. - Система должна обеспечивать возможность группировки объектов по различным критериям
Отчетность	- Система должна автоматически генерировать периодические отчеты о состоянии объектов, включая прогнозные данные. - Система должна позволять создавать простые отчеты по заданным параметрам, в том числе с учетом прогнозов
Интеграция	- Система должна предоставлять API для интеграции с внешними системами. - Система должна поддерживать экспорт исторических данных из собственных баз данных
Управление пользователями	- Система должна обеспечивать одноуровневый доступ без разграничения прав пользователей. - Система должна поддерживать аутентификацию пользователей с использованием современных протоколов безопасности.
Масштабирование и производительность	- Система должна обеспечивать возможность горизонтального масштабирования. - Система должна поддерживать обработку больших объемов данных без существенного снижения производительности

Нефункциональные требования к ИСГМ

Нефункциональные требования (табл. 2) охватывают вопросы производительности, надежности и безопасности ИСГМ. Требования к интерфейсам учитывают необходимость интеграции с существующими системами и оборудованием.

Таблица 2 – Нефункциональные требования ИСГМ

Требование	Описание
Производительность	- Система должна обрабатывать данные с минимум 1000 датчиков одновременно без заметной задержки. - Время отклика пользовательского интерфейса не должно превышать 2 секунды при нормальных условиях работы
Надежность	- Система должна обеспечивать доступность на уровне 99.9% в режиме 24/7. - Время восстановления после сбоя не должно превышать 15 минут
Безопасность	- Все передаваемые данные должны быть зашифрованы с использованием современных алгоритмов шифрования (например, AES-256). - Система должна соответствовать стандартам безопасности данных в строительной отрасли
Совместимость	- Система должна поддерживать интеграцию с основными ГИС-платформами через стандартные протоколы. - Клиентская часть должна работать в последних версиях популярных веб-браузеров (Chrome, Firefox, Safari, Edge)
Отказоустойчивость	- Система должна сохранять работоспособность при отказе до 20% датчиков на объекте. - При потере связи с сервером локальные компоненты должны продолжать сбор данных в автономном режиме не менее 24 часов
Соответствие нормативам	- Система должна соответствовать российским и международным стандартам в области геоэкологического мониторинга. - Все расчеты и прогнозы должны выполняться с точностью, соответствующей отраслевым нормативам

Функциональные возможности ИСГМ

После проведения всестороннего анализа функциональных требований ИСГМ становится возможным более детально рассмотреть функциональные возможности системы с использованием диаграммы деятельности, которая отражает основные процессы и взаимодействия компонентов системы. На диаграмме (рис. 2) представлен один ключевой актер, роль которого выполняет пользователь системы. Пользователь взаимодействует с широким спектром функций системы, каждая из которых отвечает за определенные этапы сбора, обработки и анализа данных, поступающих от различных сенсорных устройств,

установленных на контролируемых объектах. Взаимодействие с системой включает несколько ключевых этапов, начиная с получения данных от сенсоров и завершая формированием отчетов и прогнозированием потенциальных рисков.

Некоторые функции, такие как «Обработка и анализ данных», представлены на диаграмме как совокупность взаимосвязанных подфункций, обеспечивающих многоуровневую обработку информации. Например, «Первичная обработка данных» отвечает за сбор и фильтрацию данных на базовом уровне, тогда как «Вторичная обработка данных» включает более сложные аналитические процессы, такие как агрегация данных, проверка на отклонения от нормативных значений и выявление аномалий в поведении строительных объектов.

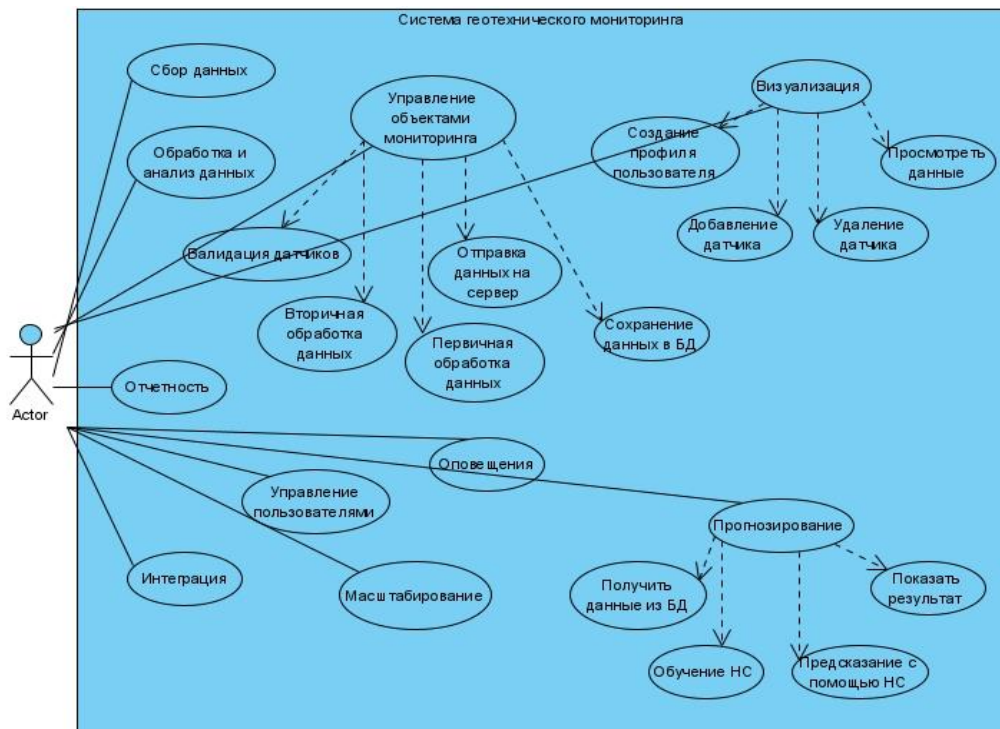


Рисунок 2 – Диаграмма вариантов использования ИСГМ

На диаграмме особое внимание уделено функции прогнозирования, которая является важнейшим элементом ИСГМ. Данный процесс включает последовательные этапы извлечения данных из базы, обучение искусственной нейронной сети (ИНС), адаптированной для решения задач прогнозирования, и генерацию предсказаний на основе текущих и исторических данных о состоянии объектов. Применение методов машинного обучения значительно повышает точность прогнозирования изменений физико-механических свойств, позволяя заранее выявлять потенциальные риски при эксплуатации сооружений.

Функции визуализации и отчетности играют важную роль в обеспечении понятной и удобной интерпретации данных для различных пользователей. Визуализация данных позволяет не только отслеживать текущие параметры состояния объектов в режиме реального времени, но и наглядно демонстрировать тенденции изменений и предсказания, основанные на анализе временных рядов данных.

Система должна предоставлять пользователю возможность легко настраивать интерфейсы и отображаемые параметры в зависимости от текущих задач мониторинга, будь то контроль за конкретным объектом или анализ совокупных данных нескольких объектов.

Функция отчетности, интегрированная в систему, обеспечивает автоматизированное формирование детализированных отчетов на основе собранных и обработанных данных, что значительно упрощает процессы документирования и планирования технического обслуживания. Генерация отчетов также включает предсказательные данные, что позволяет

пользователю принимать проактивные меры для предотвращения возможных аварийных ситуаций. Данные отчеты могут быть представлены в различных форматах в зависимости от требований конкретных пользователей, например, инженерных служб, управленческих структур или внешних аудиторов.

Таким образом, взаимодействие пользователя с системой включает в себя широкий спектр функций, направленных на обеспечение надежного и эффективного геоэкологического мониторинга, оптимизацию процессов управления объектами и минимизацию рисков при эксплуатации строительных сооружений.

Диаграмма классов ИСГМ

Проектирование архитектуры ИСГМ представляет собой сложный и многогранный процесс, требующий интеграции различных аспектов сбора, обработки, анализа и представления данных. Для создания эффективной и надежной системы мониторинга состояния сооружений необходимо учитывать не только специфику получаемой информации, но и особенности ее дальнейшего использования для прогнозирования и принятия решений. Диаграмма на рисунке 3, описывающая классы, представляет собой ключевую отправную точку для разработки специализированной платформы, предлагающей адаптируемые и расширяемые компоненты, способные эффективно решать задачи ГМ.

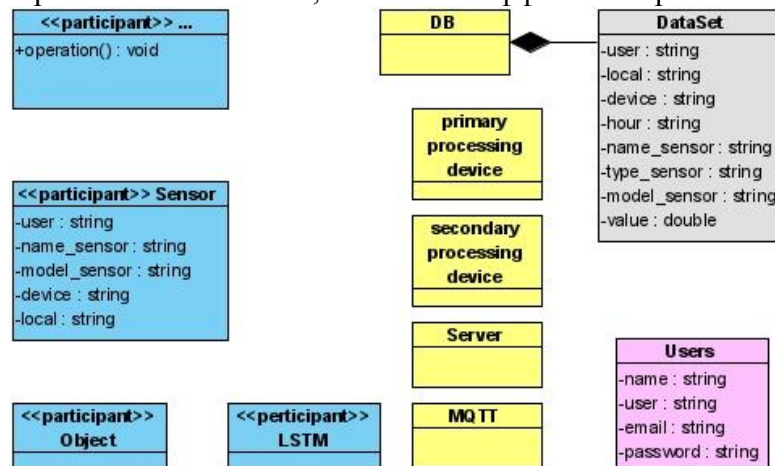


Рисунок 3 – Диаграмма классов ИСГМ

Диаграмма отражает высокоуровневое представление основных компонентов системы, включающее следующие элементы:

1. Классы-сущности Sensor, DataSet, Users.
2. Компоненты системы – база данных (DB), устройства первичной/вторичной обработки данных (primary/secondary processing devices), сервер (Server).
3. Интерфейсы взаимодействия – WebPage.
4. Специализированные модули – LSTM, Object.

Центральную роль в архитектуре системы занимают сенсоры (Sensors) — различные геоэкологические датчики, такие как датчики влажности, датчики солевого состава и датчики других характеристик почвы. Датчики влажности измеряют уровень воды в почве, обеспечивая данные для контроля ее состояния и оптимизации поливов. Датчики солевого состава фиксируют концентрацию солей, что позволяет выявлять проблемы засоления и корректировать агротехнические мероприятия. Другие датчики собирают информацию о

ключевых параметрах, таких как температура и плотность почвы, способствуя комплексному мониторингу и поддержанию ее плодородия.

Данные, собранные сенсорами, структурируются в формате DataSet и сохраняются в базе данных (DB). Обработкой данных и управлением запросами занимается серверная часть системы (Server), которая координирует взаимодействие между различными компонентами. Важным элементом является использование брокера сообщений MQTT, обеспечивающего эффективную и быструю передачу данных между сенсорами и сервером. Такая архитектура позволяет минимизировать задержки при передаче данных, что особенно важно для задач мониторинга в режиме реального времени.

Особую роль в системе играет модуль LSTM (Long Short-Term Memory), интегрированный в ИСИС, который отвечает за прогнозирование потенциальных изменений в состоянии объектов на основе анализа временных рядов данных. Применение архитектуры LSTM позволяет модели учиться на исторических данных, что значительно повышает точность прогнозов. Этот компонент играет ключевую роль в идентификации возможных деформаций и критических изменений на ранних этапах, что позволяет не только отслеживать текущее состояние объектов, но и предвидеть потенциальные риски, тем самым способствуя своевременным и превентивным мерам по обеспечению безопасности и оптимизации эксплуатационного обслуживания объектов.

Модуль LSTM благодаря своей способности обучаться на исторических данных и учитывать долгосрочные зависимости во временных рядах значительно повышает точность прогноза, что особенно важно для систем, ориентированных на геоэкологический мониторинг. В процессе прогнозирования сеть LSTM получает данные из базы данных через серверный интерфейс, обучается на этих данных и выполняет вычисление предсказательных значений, которые затем передаются пользователю через компонент WebPage или используются для формирования отчетов и предупреждений о возможных рисках.

Интерфейс взаимодействия с пользователями системы представлен компонентом WebPage, который обеспечивает доступ к результатам мониторинга и анализа для различных категорий пользователей. Пользователям интерфейс предоставляет возможность работы с данными в интерактивном режиме, включая отображение визуализаций, графиков и отчетов, что позволяет эффективно принимать решения на основе актуальной информации. Пользовательский интерфейс нацелен на интуитивное взаимодействие с системой и предоставляет гибкость в настройке отображаемых данных в зависимости от потребностей конкретного пользователя.

Таким образом, разработанная архитектура системы геоэкологического мониторинга учитывает множество аспектов работы с данными, начиная от их сбора сенсорами и

заканчивая их прогнозированием и визуализацией для конечного пользователя. Этот подход обеспечивает высокую адаптивность и расширяемость системы, что позволяет эффективно использовать ее в различных сценариях мониторинга строительных объектов.

Диаграмма компонентов ИСГМ

Архитектура сервисов была разработана на заключительном этапе процесса проектирования системы, основываясь на диаграмме классов, которая иллюстрирует взаимодействие между основными компонентами и интерфейсами системы, а также реализует связь по модели «издатель/пользователь». В этой архитектуре системные классы взаимодействуют посредством специальных портов, которые представлены в виде небольших прямоугольников и служат интерфейсами для передачи и приема данных между различными компонентами. На рисунке 4 подробно показано, как организован обмен сервисами между устройствами первичной и вторичной обработки данных, что позволяет наглядно отразить структуру взаимодействия компонентов системы.

Первичное устройство обработки данных выполняет функцию чтения данных с сенсоров, их предварительной фильтрации и вычисления средних значений параметров.

Эти обработанные данные затем передаются на вторичное устройство обработки через соответствующие интерфейсы, что позволяет вторичному устройству выполнять более сложный анализ, включая агрегацию данных, выявление аномалий и подготовку данных для последующего хранения или визуализации. Такая многоуровневая структура взаимодействия компонентов системы обеспечивает гибкость в обработке данных и позволяет динамически распределять вычислительные нагрузки между различными элементами архитектуры.

Особое место в системе занимает модуль на основе сети LSTM (Long Short-Term Memory), который отвечает за выполнение прогнозирования на основе временных рядов данных, хранящихся в базе данных. Диаграмма на рисунке 5 также иллюстрирует последовательность обмена сообщениями в контексте работы модуля LSTM, который взаимодействует с другими компонентами системы для обеспечения предсказания потенциальных деформаций или других критических изменений в состоянии объектов.

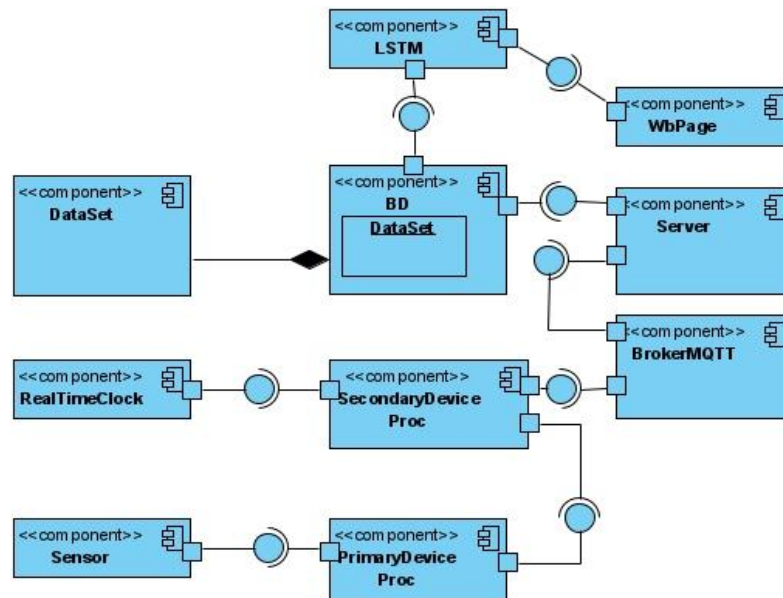


Рисунок 5 – Диаграмма компонентов ИСГМ

Таким образом, данная архитектура сервисов не только обеспечивает надежную обработку и анализ данных в режиме реального времени, но и интегрирует возможности прогнозирования на основе методов машинного обучения, что существенно повышает эффективность системы. Этот подход позволяет адаптировать систему к различным условиям эксплуатации, обеспечивая высокую степень масштабируемости и гибкости в зависимости от конкретных требований проекта.

Диаграмма последовательности работы ИПИФХС

Диаграмма представляет сервисную модель, описывающую процесс прогнозирования (use case «make prediction») в системе, использующей нейронную сеть типа LSTM (Long Short-Term Memory). LSTM — это тип рекуррентной нейронной сети (RNN), специально разработанный для обработки последовательных данных с сохранением долгосрочных зависимостей. В отличие от стандартных RNN, испытывающих трудности с удержанием информации на длительных временных интервалах, LSTM использует архитектуру с «ячейками памяти» и механизмами управления (вентильями), позволяющими выбирать, какие данные сохранять, обновлять или забывать. LSTM широко применяется в прогнозировании, обработке естественного языка, распознавании речи и анализе временных рядов.

Эта модель (рис. 6) демонстрирует ключевые этапы прогнозирования с применением методов машинного обучения на базе временных рядов данных в ИСГМ. Она служит базовой основой для разработки более детализированной архитектуры модуля прогнозирования, обеспечивая всесторонний подход к анализу и предсказанию критических изменений в строительных конструкциях.

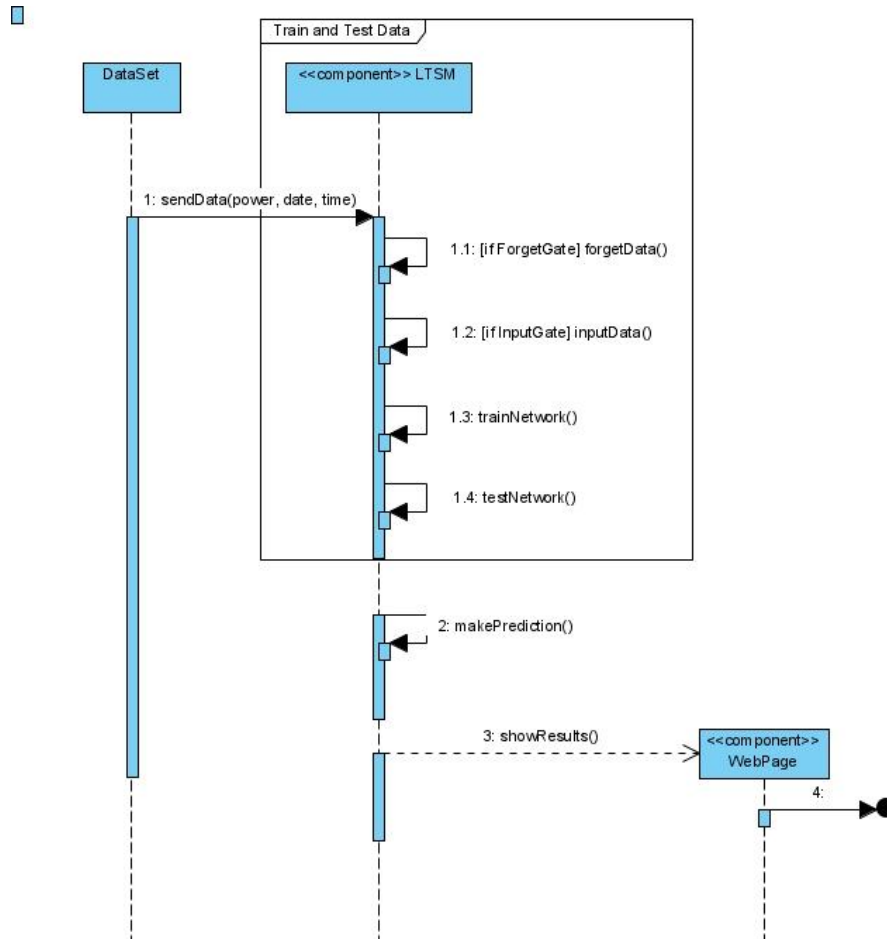


Рисунок 6 – Диаграмма последовательности работы ИПИФХС

Участники процесса

- **DataSet** – источник данных, который содержит информацию, собранную с различных сенсоров, установленных на объектах мониторинга.
- **LSTM** – модуль нейронной сети, выполняющий прогноз на основе анализа временных рядов данных.
- **WebPage** – пользовательский интерфейс, через который отображаются результаты прогнозирования.

Последовательность действий

1. **sendData(power, date, time)** – DataSet передает данные (включая временные метки) в модуль LSTM для дальнейшей обработки и прогнозирования.

Этапы внутри модуля LSTM (Train and Test Data)

1. **[ifForgetGate] forgetData()**
 - Данный шаг описывает работу «забывающего вентиля» (forget gate), являющегося важной частью LSTM-архитектуры.

- Forget gate отвечает за определение той информации, которая более не является полезной для текущего состояния сети, и ее следует исключить из состояния ячейки.
- Условие **ifForgetGate** проверяет необходимость активации данного механизма.
- Операция **forgetData()** выполняет фактическое удаление устаревших или ненужных данных, минимизируя накопление избыточной информации.

2. **[ifInputGate] inputData()**

- Следующий шаг связан с функционированием «входного вентиля» (input gate), который регулирует поступление новой информации в ячейку LSTM.
- Input gate определяет, какие новые данные должны быть добавлены в состояние ячейки на основании анализа текущего состояния и входных сигналов.
- Условие **ifInputGate** проверяет необходимость добавления новой информации.
- Операция **inputData()** добавляет новые данные в ячейку, тем самым обновляя состояние сети и обеспечивая адаптацию модели к новым данным.

3. **makePrediction()**

- На этом этапе LSTM выполняет предсказание на основе обработанных временных рядов данных. Важной особенностью LSTM является возможность учитывать длительные зависимости между данными, что делает модель особенно эффективной в контексте прогнозирования временных изменений в строительных конструкциях.

4. **showResults()**

- После выполнения прогноза результаты отправляются в пользовательский интерфейс (WebPage), где они визуализируются в удобном для пользователя формате, что позволяет осуществлять мониторинг состояния объектов и своевременно принимать меры по предотвращению аварийных ситуаций.

Обобщенные шаги работы системы:

1. Вначале нейронная сеть анализирует предыдущее состояние и определяет, какие данные необходимо «забыть» с помощью механизма **forget gate**.
2. Затем сеть решает, какие новые данные следует добавить с помощью механизма **input gate**.
3. После этого осуществляется процесс обучения и тестирования модели на основе актуальных данных, что обеспечивает ее адаптацию к изменяющимся условиям.

4. На заключительном этапе сеть, используя механизм **output gate**, выводит предсказательные данные, которые затем передаются на веб-интерфейс для отображения и дальнейшего анализа.

Таким образом, данная модель ИПИФХС обеспечивает не только точное прогнозирование на основе временных рядов данных, но и эффективное управление процессом обучения, что делает ее незаменимым компонентом для задач ГМ. Интеграция этой модели в ИСГМ способствует повышению оперативности и точности в управлении данными о состоянии объектов, минимизируя риски и повышая надежность эксплуатационных решений.

В рамках данной работы были рассмотрены основные аспекты проектирования ИСГМ, предназначенной для выполнения задач ГМ. Приведенные требования и функциональные возможности системы, такие как сбор, обработка и анализ данных, а также прогнозирование деформационных изменений с использованием нейронных сетей типа LSTM, отражают сложность и многоуровневость разрабатываемой системы.

Разработка базируется на использовании современных методологий проектирования систем Интернета вещей и сервисно-ориентированной архитектуры, что позволяет обеспечивать высокую степень масштабируемости и интеграции с широким спектром сенсорных устройств и внешних систем. Это гарантирует возможность мониторинга большого количества объектов в реальном времени и способствует оперативной диагностике потенциальных проблем, что крайне важно для повышения безопасности строительных объектов и предотвращения аварийных ситуаций.

Особое внимание уделено внедрению ИПИФХС, которая обеспечивает высокую точность прогнозирования возможных изменений на основе временных рядов данных. Это позволяет не только отслеживать текущее состояние объектов, но и заблаговременно выявлять потенциальные риски, связанные с эксплуатацией сооружений.

Интеграция ИПИФХС в ИСГМ способствует повышению эффективности принятия решений и планирования профилактических мероприятий. Разработанная архитектура платформы и предложенные методы анализа данных представляют собой надежное и гибкое решение для задач ГМ, обеспечивающее высокую производительность, надежность и безопасность при эксплуатации в условиях масштабных промышленных объектов.

Литература

1. Белозеров С. А., Козлов А. В. Системы мониторинга и управления в сельском хозяйстве: технологии и перспективы развития. — М.: Агропромиздат, 2018. — 256 с.
2. Goodfellow I., Bengio Y., Courville A. Deep Learning. — MIT Press, 2016. — 775 p. URL: <https://www.deeplearningbook.org/> (дата обращения: 14.11.2024).

3. Бородин О. Б., Гвоздева О. В., Сеница Ю. С., Колбнева Е. Ю. Цифровое сельское хозяйство: настоящее и будущее (обзор международной практики) // Московский экономический журнал. – 2021. – № 1. – С. 45–60. URL: <https://cyberleninka.ru/article/n/tsifrovoye-selskoye-hozyaystvo-nastoyascheye-i-budushchee-obzor-mezhdunarodnoy-praktiki> (дата обращения: 14.11.2024).
4. Советникова О. П. Основные направления развития сельского хозяйства в условиях цифровой трансформации // Вестник Витебского государственного технологического университета. – 2020. – № 2. – С. 112–118. URL: <https://rep.vstu.by/handle/123456789/15426> (дата обращения: 14.11.2024).
5. Добровлянин В. Д., Антинескул Е. А. Цифровизация сельского хозяйства: текущий уровень цифровизации в Российской Федерации и перспективы дальнейшего развития // Экономика сельского хозяйства России. – 2021. – № 3. – С. 29–34. URL: <https://cyberleninka.ru/article/n/tsifrovizatsiya-selskogo-hozyaystva-tekushiy-uroven-tsifrovizatsii-v-rossiyskoy-federatsii-i-perspektivy-dalneyshego-razvitiya> (дата обращения: 14.11.2024).
6. Белоусов П. В. Применение методов машинного обучения в аграрном секторе // Вестник Алтайской академии экономики и права. – 2022. – № 3. – С. 78–85. URL: <https://vaael.ru/ru/article/view?id=3680> (дата обращения: 14.11.2024).
7. Иванов Д. М. Спутниковый мониторинг и аналитика для устойчивого развития сельского хозяйства // Агроинформ. – 2023. – № 5. – С. 34–39. URL: <https://farmonaut.com/remote-sensing/спутниковый-мониторинг-и-аналитика-д/> (дата обращения: 14.11.2024).
8. Петров А. Н. Новые технологии в сельском хозяйстве и их применение // Агронаука. – 2022. – № 4. – С. 22–27. URL: <https://eos.com/ru/blog/novye-tehnologii-v-selskom-khozyajstve/> (дата обращения: 14.11.2024).
9. Сидоров В. Л. Цифровые технологии в сельском хозяйстве: новые перспективы // Платформы. – 2023. – № 1. – С. 15–20. URL: <https://platforms.su/articles/5180> (дата обращения: 14.11.2024).
10. Кузнецов М. Г. Будущее сельского хозяйства: как технологии могут обеспечить новый рост // Молодой ученый. – 2022. – № 26. – С. 117–119. URL: <https://moluch.ru/archive/473/104479/> (дата обращения: 14.11.2024).
11. М.С. Макаров, К.А. Жолдошбаев, С. Н. Верзунов. Анализ интеллектуальных систем геотехнического мониторинга // Проблемы автоматизации и управления. – 2024. – №2(50). – С. 27 – 36. URL: <https://pau.imash.kg/index.php/pau/article/view/478> (дата обращения: 14.11.2024).

УДК 004.056.5

С. В. Корякин^{1,2}

¹Старший преподаватель, Институт информационных технологий КГТУ им. И.Раззакова

² Научный сотрудник, Институт машиноведения автоматизации и геомеханики НАН КР

Салиев А.Б.

д.ф.-м.н., профессор, Институт информационных технологий КГТУ им.

И.Раззакова

Салиев С.М., salievsamat0@gmail.com

Студент гр. ИБ-2-22, Институт информационных технологий КГТУ им.

И.Раззакова

Верзунов С.Н.

Ведущий научный сотрудник, Институт машиноведения, автоматизации и геомеханики НАН КР, verzunov@hotmail.com

АНАЛИТИЧЕСКИЙ ОБЗОР ОСНОВНЫХ НАПРАВЛЕНИЙ, ПРИНЦИПОВ И МЕТОДОВ ОБЕСПЕЧЕНИЯ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ АИС

В статье рассматриваются основные направления, принципы и методы обеспечения информационной безопасности автоматизированных информационных систем (АИС). Основное внимание уделяется анализу возможных угроз и уязвимостей информационных систем, а также организационным, техническим, правовым и процедурным мерам по защите информации. Рассматриваются принципы конфиденциальности, целостности, доступности, аутентичности и неотказуемости как фундаментальные основы информационной безопасности. Производится обзор различных инструментов защиты, включая криптографические средства, системы управления доступом, антивирусные программы, системы обнаружения и предотвращения вторжений, а также резервное копирование и восстановление данных. Таким образом, статья направлена на систематизацию существующих знаний и методов для обеспечения информационной безопасности АИС, что позволит улучшить защиту информационных систем предприятий и организаций от различных угроз и атак.

Ключевые слова: информационная безопасность, аудит, меры, IDS/IPS, DLP, SIEM, автоматизированная информационная система, методы, инцидент, шифрование, GDPR, ISO/IEC 27001

В условиях стремительного роста объемов информации и увеличения сложности автоматизированных информационных систем (АИС) вопросы обеспечения их безопасности становятся в особенности важными для предприятий и организаций. Обеспечение информационной безопасности АИС направлено на защиту данных от несанкционированного доступа, изменения, уничтожения и других угроз, которые могут привести к значительным как финансовым, так и репутационным потерям.

Несмотря на значительные усилия и внедрение современных технологий защиты, сохраняются многочисленные уязвимости, которые могут быть использованы злоумышленниками для несанкционированного доступа, изменения или уничтожения данных. Проблема заключается в необходимости комплексного подхода к обеспечению информационной безопасности АИС, включающего организационные, технические, правовые и процедурные меры, которые могли бы эффективно противодействовать как существующим, так и новым угрозам.

Основные аспекты проблемы включают:

- Недостаточную осведомленность и ошибки пользователей, являющиеся одной из главных причин инцидентов безопасности.
- Использование устаревшего программного обеспечения и оборудования, не поддерживающего современные стандарты защиты.

- Неполное или некорректное внедрение политик управления доступом и мониторинга активности пользователей.
- Наличие сложных и разнообразных угроз, требующих применения многослойных и адаптивных методов защиты.

Для решения данной проблемы необходимо систематизировать знания о существующих угрозах и уязвимостях, а также проанализировать и оценить эффективность различных подходов и технологий обеспечения информационной безопасности.[1]

Считается, что под направлениями обеспечения информационной безопасности автоматизированных информационных систем (АИС) понимаются основные области деятельности и подходы, направленные на защиту информации и систем от различных угроз. Эти направления охватывают широкий спектр мер и стратегий, включающих как технические, так и организационные аспекты. Основные направления обеспечения информационной безопасности АИС включают:

1. *Организационные меры:*

- *Политики и процедуры.* Разработка и внедрение корпоративных политик и процедур по информационной безопасности, включая правила использования информационных систем, управления доступом и реагирования на инциденты.
- *Обучение и осведомленность.* Регулярное обучение сотрудников и повышение их осведомленности о правилах информационной безопасности.
- *Аудит и мониторинг.* Проведение регулярных аудитов и мониторинга для оценки текущего состояния безопасности и выявления потенциальных угроз.

2. *Технические меры*[2]:

- *Шифрование.* Использование криптографических методов для защиты данных при передаче и хранении.
- *Системы управления доступом.* Внедрение механизмов аутентификации и авторизации, таких как многофакторная аутентификация и ролевое управление доступом.
- *Защитное ПО.* Установка и обновление антивирусных программ, брандмауэров и систем обнаружения и предотвращения вторжений (IDS/IPS).

3. *Правовые меры* [3]:

- *Соответствие нормативным требованиям.* Обеспечение соответствия законодательным и нормативным требованиям в области информационной безопасности, таким как GDPR, ISO/IEC 27001 и другие.
- *Договоры и соглашения.* Разработка юридических документов, включая соглашения о конфиденциальности и контракты с поставщиками услуг.

4. *Процедурные меры:*

- *Реагирование на инциденты.* Разработка и внедрение процедур реагирования на инциденты информационной безопасности, включая план действий при утечках данных и кибератаках.
- *Управление рисками.* Оценка и управление рисками, связанными с информационной безопасностью, включая идентификацию, анализ и принятие мер по минимизации рисков.
- *Резервное копирование и восстановление.* Обеспечение регулярного резервного копирования данных и разработка планов восстановления после сбоев.

5. *Физические меры [4]:*

- *Контроль доступа к помещениям.* Обеспечение физической безопасности серверных помещений и рабочих мест.
- *Защита оборудования.* Использование сейфов, замков и других средств для защиты оборудования от кражи и несанкционированного доступа.

Эти направления взаимосвязаны и дополняют друг друга, создавая многослойную защиту, которая способна противостоять современным угрозам информационной безопасности (рис.1).



Рисунок 1 – Обобщенная блок-схема направлений обеспечения безопасности

Исходя из рисунка 1, предлагается выделить следующие принципы и методы, которые необходимо использовать в процессах обеспечения безопасности АИС, под которыми понимаются базовые правила и концепции, служащие основой для разработки и внедрения мер по защите информации и систем от различных киберугроз и направленных на создание комплексной и эффективной системы защиты. Основные принципы обеспечения информационной безопасности АИС включают [5]:

1. **Конфиденциальность:**

- Обеспечение доступа к информации только тем пользователям, которые имеют на это разрешение.
- Использование методов шифрования и контроля доступа для защиты данных от несанкционированного просмотра.[5]

2. **Целостность:**

- Защита данных от несанкционированного изменения или уничтожения.
- Применение методов контроля целостности данных, таких как контрольные суммы и цифровые подписи.[5]

3. **Доступность:**

- Обеспечение возможности доступа к информации и системам для уполномоченных пользователей в любое необходимое время.
- Защита от отказов в обслуживании (DoS/DDoS атаки) и обеспечение устойчивости системы к сбоям и отказам.[5]

4. Аутентичность:

- Гарантия подлинности и достоверности источника данных.
- Использование методов аутентификации, таких как пароли, биометрические данные и цифровые сертификаты.

5. Неотказуемость:

- Обеспечение невозможности отрицания факта отправки или получения информации.
- Использование цифровых подписей и журналов аудита для подтверждения действий пользователей.[5]

6. Разделение обязанностей:

- Разделение критических задач и функций между несколькими сотрудниками для уменьшения риска внутреннего мошенничества и ошибок.
- Введение различных уровней доступа и прав для выполнения различных операций.

7. Минимизация привилегий:

- Предоставление пользователям минимально необходимого уровня доступа для выполнения их задач.
- Регулярный пересмотр и ограничение прав доступа на основе текущих задач и ролей.

8. Мониторинг и аудит:

- Постоянное наблюдение за действиями пользователей и системными событиями.
- Ведение журналов аудита и проведение регулярных проверок для выявления и анализа подозрительной активности.[5,8]

9. Управление рисками:

- Оценка и управление рисками, связанными с информационной безопасностью.
- Идентификация, анализ и принятие мер по минимизации потенциальных угроз и уязвимостей.

10. Постоянное улучшение:

- Регулярное обновление и совершенствование мер безопасности в ответ на новые угрозы и технологии.
- Проведение обучения и повышения осведомленности сотрудников, а также тестирование и обновление защитных систем.

Эти принципы формируют основу для создания и поддержания надежной системы информационной безопасности, позволяя эффективно защищать данные и обеспечивать стабильную работу автоматизированных информационных систем (рис. 2).



Рисунок 2 – Принципы обеспечения информационной безопасности АИС

Под методами обеспечения информационной безопасности автоматизированных информационных систем (АИС) понимаются конкретные средства, технологии и процедуры, используемые для защиты информации и систем от различных угроз. Эти методы включают как технические, так и организационные меры, направленные на предотвращение, обнаружение и реагирование на инциденты безопасности.[5] Основные методы обеспечения информационной безопасности АИС включают:

1. Криптографические методы:

- **Шифрование данных.** Применение симметричных и асимметричных алгоритмов шифрования для защиты данных при передаче и хранении.
- **Цифровые подписи.** Использование цифровых подписей для обеспечения целостности и аутентичности данных.
- **Хеширование.** Применение хеш-функций для проверки целостности данных и хранения паролей.[6]

2. Методы аутентификации и авторизации:

- **Парольная аутентификация.** Использование сложных паролей для подтверждения личности пользователя.
- **Многофакторная аутентификация (MFA).** Применение нескольких факторов (пароль, смс-код, биометрия) для повышения надежности аутентификации.
- **Ролевая и атрибутивная модель управления доступом (RBAC/ABAC).** Назначение прав доступа на основе ролей и атрибутов пользователей.[6]

3. Средства обнаружения и предотвращения вторжений (IDS/IPS):

- **Системы обнаружения вторжений (IDS).** Мониторинг сетевого трафика и активности систем для выявления подозрительных действий.
- **Системы предотвращения вторжений (IPS).** Автоматическое принятие мер для блокировки или ограничения вредоносной активности.

4. Антивирусные и антишпионские программы:

- **Антивирусное ПО.** Использование программ для обнаружения и удаления вредоносного ПО (вирусы, трояны, черви).
 - **Антишпионское ПО.** Применение средств для обнаружения и удаления шпионского ПО.
5. **Методы защиты сетей:**
- **Брандмауэры (Firewall).** Установка межсетевых экранов для контроля и фильтрации сетевого трафика.
 - **Виртуальные частные сети (VPN).** Использование технологий для создания защищенных каналов связи через Интернет.
6. **Методы резервного копирования и восстановления данных:**
- **Резервное копирование.** Регулярное создание копий данных для предотвращения их потери в случае сбоя или атаки.
 - **Планы восстановления после сбоев.** Разработка и внедрение процедур для быстрого восстановления работы систем после инцидентов.
7. **Методы мониторинга и аудита:**
- **Журналирование.** Ведение журналов аудита для фиксации всех значимых действий в системе.
 - **Анализ логов.** Регулярный анализ журналов для выявления аномалий и инцидентов безопасности.[7]
8. **Методы управления уязвимостями:**
- **Патч-менеджмент.** Обновление программного обеспечения для устранения известных уязвимостей.
 - **Пентест (тестирование на проникновение).** Проведение тестов на проникновение для выявления и устранения уязвимостей.
9. **Обучение и повышение осведомленности:**
- **Тренинги по безопасности.** Регулярное обучение сотрудников правилам информационной безопасности.
 - **Фишинг-тесты.** Проведение тестов на осведомленность сотрудников о фишинговых атаках.[8]
10. **Методы реагирования на инциденты:**
- **Планы реагирования на инциденты.** Разработка и внедрение процедур для быстрого и эффективного реагирования на инциденты.
 - **Анализ инцидентов.** Проведение расследований инцидентов для выявления причин и разработки мер по предотвращению повторения.

Из перечисленного выше следует, что использование указанных методов при построении подсистем защиты АИС информации позволяет обеспечить многослойную защиту информации и систем, создавая комплексную подсистему безопасности, способную эффективно противостоять современным киберугрозам (рис.3).



Рисунок 3 – Методы обеспечения информационной безопасности

Обзор современных инструментов и технологий

Таким образом следует понимать, что в современном мире, где информация играет все более важную роль, обеспечение информационной безопасности (ИБ) автоматизированных информационных систем (АИС) становится критически важной задачей. С ростом использования информационных технологий и ценности данных киберугрозы становятся все более изощренными. Для защиты АИС от этих угроз необходимо использовать широкий спектр современных инструментов и технологий, таких как:

1. Системы контроля доступа (СКД):

- **Active Directory.** Централизованная система управления доступом, позволяющая управлять правами пользователей на доступ к ресурсам сети.
- **Kerberos.** Протокол аутентификации, обеспечивающий безопасный доступ к ресурсам сети.
- **SAML (Security Assertion Markup Language).** Протокол обмена аутентификационной информацией между различными системами.
- **SSO (Single Sign-On).** Единый вход в систему, позволяющий пользователям аутентифицироваться один раз и получать доступ к нескольким системам.
- **Biometric authentication.** Использование биометрических данных (отпечатки пальцев, распознавание лица) для аутентификации пользователей [9].

В таблице 1 приведен сравнительный анализ характеристик СКД.

Таблица 1– Сравнительный анализ характеристик СКД

Технология/Инструмент	Функциональность	Производительность	Масштабируемость	Надежность	Простота использования	Стоимость
Active Directory	Широкий набор функций, включая управление пользователями, группами, правами доступа и политиками безопасности	Высокая производительность	Хорошо масштабируется для больших организаций	Надежная система, используемая во многих крупных организациях	Может быть сложной для настройки и управления	Бесплатно для Windows Server
Kerberos	Простой и надежный протокол аутентификации	Высокая производительность	Хорошо масштабируется для больших организаций	Надежная система, используемая во многих крупных организациях	Может быть сложной для настройки и управления	Бесплатно
SAML	Стандартный протокол обмена аутентификационной информацией между различными системами	Позволяет использовать единый вход в систему для нескольких систем	Хорошо масштабируется для больших организаций	Надежный стандарт, поддерживаемый многими системами	Может быть сложной для настройки и управления	Стоимость зависит от конкретной реализации

2. Антивирусное программное обеспечение (АПО) [8]:

- **Kaspersky Internet Security.** Комплексное антивирусное решение, защищающее от вирусов, троянов, червей, шпионских программ и других угроз.
- **Eset NOD32 Antivirus.** Антивирусное решение, известное своей эффективностью и низким потреблением системных ресурсов.
- **Bitdefender Antivirus Plus.** Антивирусное решение, предлагающее широкий спектр функций.
- **Avast Antivirus.** Бесплатное антивирусное решение, обеспечивающее базовую защиту от вирусов и других угроз.
- **AVG Antivirus.** Еще одно бесплатное антивирусное решение, предлагающее различные платные функции.

В таблице 2 приведен сравнительный анализ характеристик АПО.

Таблица 2– Сравнительный анализ характеристик АПО

Технология/Инструмент	Функциональность	Производительность	Масштабируемость	Надежность	Простота использования	Стоимость
Kaspersky Internet Security	Широкий набор функций, включая защиту от вирусов, троянов,	Высокая производительность	Хорошо масштабируется для больших организаций	Надежная система, хорошо зарекомендовавшая себя	Легко использовать	Платная

	червей, шпионских программ и других угроз					
Eset NOD32 Antivirus	Эффективная защита от вирусов и других угроз	Низкое потребление системных ресурсов	Хорошо масштабируется для больших организаций	Надежная система, хорошо зарекомендовавшая себя	Легко использовать	Платная

3. Межсетевые экраны (МЭ) [16]:

- **Cisco Firepower Management Center.** Комплексная система межсетевой безопасности, включающая в себя межсетевой экран, систему обнаружения и предотвращения вторжений (IDS/IPS) и другие функции.
- **Palo Alto Networks PAN-OS.** Межсетевой экран нового поколения, предлагающий широкий спектр функций, включая защиту от APT-атак и URL-фильтрацию.
- **Fortinet FortiGate.** Межсетевой экран, известный своей надежностью и высокой производительностью.
- **SonicWall NSv.** Виртуальный межсетевой экран, удобный для развертывания в облачных средах.
- **Checkpoint SandBlast Network.** Межсетевой экран нового поколения, использующий технологии песочницы для защиты от неизвестных угроз.

В таблице 3 приведен сравнительный анализ МЭ.

Таблица 3– Сравнительный анализ МЭ

Технология/Инструмент	Функциональность	Производительность	Масштабируемость	Надежность	Простота использования	Стоимость
Cisco Firepower Management Center	Комплексная система межсетевой безопасности, включающая в себя межсетевой экран, систему обнаружения и предотвращения вторжений (IDS/IPS) и другие функции	Высокая производительность	Хорошо масштабируется для больших организаций	Надежная система, используемая во многих крупных организациях	Может быть сложной для настройки и управления	Платная
Palo Alto Networks PAN-OS	Межсетевой экран нового поколения, предлагающий широкий спектр функций, включая защиту от APT-атак и URL-фильтрацию	Высокая производительность	Хорошо масштабируется для больших организаций	Надежная система, хорошо зарекомендовавшая себя	Может быть сложной для настройки и управления	Платная
Fortinet FortiGate	Межсетевой экран,	Высокая производительность	Хорошо масштабируется	Надежная система,	Легко использовать	Платная

	известный своей надежностью и высокой производительностью	ость	я для больших организаций	хорошо зарекомендовавшая себя	ь	
--	---	------	---------------------------	-------------------------------	---	--

4. Системы обнаружения вторжений (COB) [16]:

- **Snort.** Open-source система обнаружения вторжений, использующая сигнатурный анализ и методы статистического анализа трафика.
- **Suricata.** Еще одна open-source система обнаружения вторжений, основанная на Snort.
- **Deepwatch Network Security.** Коммерческая система обнаружения вторжений, использующая машинное обучение для выявления неизвестных угроз.
- **McAfee Network Security Platform.** Комплексная система сетевой безопасности, включающая в себя IDS/IPS, межсетевой экран и другие функции.
- **Cisco Security Center.** Комплексная система безопасности, включающая в себя IDS/IPS, межсетевой экран, SIEM и другие функции.

В таблице 4 приведен сравнительный анализ COB.

Таблица 4 – Сравнительный анализ COB

Технология/Инструмент	Функциональность	Производительность	Масштабируемость	Надежность	Простота использования	Стоимость
Snort	Open-source система обнаружения вторжений, использующая сигнатурный анализ и методы статистического анализа трафика	Средняя производительность	Хорошо масштабируется для больших организаций	Бесплатная система, но может требовать значительных ресурсов для работы	Может быть сложной для настройки и управления	Бесплатно
Cisco Security Center	Комплексная система безопасности, включающая в себя IDS/IPS, межсетевой экран, SIEM и другие функции	Высокая производительность	Хорошо масштабируется для больших организаций	Надежная система, хорошо зарекомендовавшая себя	Может быть сложной для настройки и управления	Платная

5. Системы предотвращения вторжений (СПВ) [16]:

- **Cisco Firepower NGIPS.** Система предотвращения вторжений нового поколения, основанная на Cisco Firepower Management Center.
- **Palo Alto Networks PAN-OS.** Система предотвращения вторжений, интегрированная с межсетевым экраном PAN-OS.
- **Fortinet FortiGate.** Система предотвращения вторжений, интегрированная с межсетевым экраном FortiGate.
- **SonicWall NSv.** Виртуальная система предотвращения вторжений, удобная для развертывания в облачных средах.

- **Checkpoint SandBlast Network.** Система предотвращения вторжений нового поколения, использующая технологии песочницы для защиты от неизвестных угроз.

В таблице 5 приведен сравнительный анализ СПВ.

Таблица 5 – Сравнительный анализ СПВ

Технология/Инструмент	Функциональность	Производительность	Масштабируемость	Надежность	Простота использования	Стоимость
Cisco Firepower NGIPS	Система предотвращения вторжений нового поколения, основанная на Cisco Firepower Management Center	Высокая производительность	Хорошо масштабируется для больших организаций	Надежная система, используемая во многих крупных организациях	Может быть сложной для настройки и управления	Платная
Fortinet FortiGate	Система предотвращения вторжений, интегрированная с межсетевым экраном FortiGate	Высокая производительность	Хорошо масштабируется для больших организаций	Надежная система, хорошо зарекомендовавшая себя	Легко использовать	Платная

6. Системы шифрования [9]:

- **PGP (Pretty Good Privacy).** Open-source система шифрования, использующая криптографию с открытым ключом.
- **GPG (GNU Privacy Guard).** Свободная реализация PGP.
- **VeraCrypt.** Бесплатная программа шифрования, позволяющая шифровать файлы, диски и разделы диска.
- **BitLocker.** Система шифрования дисков, встроенная в Windows.
- **FileVault 2.** Система шифрования дисков, встроенная в macOS.

В таблице 6 приведен сравнительный анализ PGP.

Таблица 6– Сравнительный анализ PGP

Технология/Инструмент	Функциональность	Производительность	Масштабируемость	Надежность	Простота использования	Стоимость
PGP	Open-source система шифрования, использующая криптографию с открытым ключом	Средняя производительность	Хорошо масштабируется для больших организаций	Надежная система, используемая во многих крупных организациях	Может быть сложной для использования для неискушенных пользователей	Бесплатно
GPG	Свободная реализация PGP	Средняя производительность	Хорошо масштабируется	Надежная система,	Может быть сложной	Бесплатно

		сть	я для больших организаций	используемая во многих крупных организациях	для использования для неискушенных пользователей	
VeraCrypt	Бесплатная программа шифрования, позволяющая шифровать файлы, диски и разделы диска	Средняя производительность	Хорошо масштабируется для больших организаций.	Надежная система, хорошо зарекомендовавшая себя	Легко использовать	Бесплатно

7. Системы резервного копирования [8]:

- **Veeam Backup & Replication.** Коммерческая система резервного копирования, предлагающая широкий спектр функций.
- **Acronis Backup & Recovery.** Еще одна коммерческая система резервного копирования, известная своей простотой использования.
- **CrashPlan.** Облачная система резервного копирования, позволяющая хранить резервные копии данных в защищенном дата-центре.
- **UrBackup.** Open-source система резервного копирования, предлагающая широкий спектр функций.
- **Backupper.** Бесплатная программа резервного копирования, позволяющая создавать резервные копии файлов, дисков и разделов диска.
- **Duplicati.** Бесплатная программа резервного копирования с открытым исходным кодом, позволяющая создавать резервные копии данных в локальном хранилище или в облаке.

В таблице 7 приведен сравнительный анализ Системы резервного копирования.

Таблица 7 – Сравнительный анализ систем резервного копирования

Технология/Инструмент	Функциональность	Производительность	Масштабируемость	Надежность	Простота использования	Стоимость
Veeam Backup & Replication	Коммерческая система резервного копирования, предлагающая широкий спектр функций	Высокая производительность	Хорошо масштабируется для больших организаций	Надежная система, хорошо зарекомендовавшая себя	Может быть сложной для настройки и управления	Платная
Acronis Backup & Recovery	Еще одна коммерческая система резервного копирования, известная своей простотой использования	Средняя производительность	Хорошо масштабируется для средних и крупных организаций	Надежная система, хорошо зарекомендовавшая себя	Легко использовать	Платная

8. Системы управления информационной безопасностью (СУИБ):

- **IBM Security QRadar.** Коммерческая СУИБ, предлагающая широкий спектр функций, включая мониторинг событий безопасности, анализ журналов и реагирование на инциденты.
- **McAfee ePolicy Orchestrator.** Коммерческая СУИБ, известная своей масштабируемостью и надежностью.
- **LogRhythm SIEM.** Коммерческая СУИБ, предлагающая функции мониторинга событий безопасности, анализа журналов и реагирования на инциденты.
- **AlienVault USM Anywhere.** Облачная СУИБ, удобная для использования малыми и средними предприятиями.
- **OSSEC.** Open-source СУИБ, предлагающая базовые функции мониторинга событий безопасности и анализа журналов.

В таблице 8 приведен сравнительный анализ СУИБ.

Таблица 8 — Сравнительный анализ СУИБ

Технология/Инструмент	Функциональность	Производительность	Масштабируемость	Надежность	Простота использования	Стоимость
IBM Security QRadar	Коммерческая СУИБ, предлагающая широкий спектр функций, включая мониторинг событий безопасности, анализ журналов и реагирование на инциденты	Высокая производительность	Хорошо масштабируется для больших организаций	Надежная система, используемая во многих крупных организациях	Может быть сложной для настройки и управления	Платная
McAfee ePolicy Orchestrator	Коммерческая СУИБ, известная своей масштабируемостью и надежностью	Высокая производительность	Хорошо масштабируется для больших организаций	Надежная система, хорошо зарекомендовавшая себя	Может быть сложной для настройки и управления	Платная
OSSEC	Open-source СУИБ, предлагающая базовые функции мониторинга событий безопасности и анализа журналов	Средняя производительность	Хорошо масштабируется для организаций всех размеров	Надежная система, хорошо зарекомендовавшая себя	Может быть сложной для настройки и управления	Бесплатно

9. SIEM-системы [9]:

- **Splunk.** Коммерческая SIEM-система, предлагающая широкий спектр функций, включая сбор и анализ журналов, мониторинг событий безопасности и реагирование на инциденты.
- **ArcSight Investigator.** Коммерческая SIEM-система, известная своей мощностью и масштабируемостью.
- **LogRhythm SIEM.** Коммерческая SIEM-система, предлагающая функции мониторинга событий безопасности, анализа журналов и реагирования на инциденты.
- **Elastic Stack (ELK).** Open-source SIEM-система, основанная на Elasticsearch, Logstash и Kibana.
- **Graylog.** Open-source SIEM-система, предлагающая базовые функции сбора и анализа журналов.

В таблице 9 приведен сравнительный анализ SIEM-систем.

Таблица 9 — Сравнительный анализ SIEM-систем

Технология/Инструмент	Функциональность	Производительность	Масштабируемость	Надежность	Простота использования	Стоимость
LogRhythm SIEM	Коммерческая SIEM-система, предлагающая функции сбора и анализа журналов, мониторинга событий безопасности и реагирования на инциденты	Высокая производительность	Хорошо масштабируется для больших организаций	Надежная система, хорошо зарекомендовавшая себя	Может быть сложной для настройки и управления	Платная
Graylog	Open-source SIEM-система, предлагающая базовые функции сбора и анализа журналов	Средняя производительность	Хорошо масштабируется для организаций всех размеров	Надежная система, хорошо зарекомендовавшая себя	Легко использовать	Бесплатно
Splunk	Коммерческая SIEM-система, предлагающая широкий спектр функций, включая сбор и анализ журналов, мониторинг событий	Высокая производительность	Хорошо масштабируется для больших организаций	Надежная система, используемая во многих крупных организациях	Может быть сложной для настройки и управления	Платная

	безопасности и реагирование на инциденты					
--	--	--	--	--	--	--

10. DLP-системы [9]:

- **McAfee DLP Suite.** Коммерческая DLP-система, предлагающая широкий спектр функций, включая контроль конфиденциальной информации, шифрование данных и предотвращение утечек.
- **Symantec DLP Suite.** Коммерческая DLP-система, известная своей эффективностью и точностью.
- **Cisco Secure Content Manager.** Коммерческая DLP-система, интегрированная с другими решениями Cisco по безопасности.
- **Digital Guardian DLP.** Коммерческая DLP-система, предлагающая функции контроля конфиденциальной информации, шифрования данных и предотвращения утечек.
- **Irius Risk Manager.** Open-source DLP-система, предлагающая базовые функции контроля конфиденциальной информации.

В таблице 10 приведен сравнительный анализ DLP-системы.

Таблица 10 — Сравнительный анализ DLP-системы

Технология/Инструмент	Функциональность	Производительность	Масштабируемость	Надежность	Простота использования	Стоимость
McAfee DLP Suite	Коммерческая DLP-система, предлагающая широкий спектр функций, включая контроль конфиденциальной информации, шифрование данных и предотвращение утечек	Высокая производительность	Хорошо масштабируется для больших организаций	Надежная система, хорошо зарекомендовавшая себя	Может быть сложной для настройки и управления	Платная
Symantec DLP Suite	Коммерческая DLP-система, известная своей эффективностью и точностью	Высокая производительность	Хорошо масштабируется для больших организаций	Надежная система, хорошо зарекомендовавшая себя	Может быть сложной для настройки и управления	Платная
Cisco Secure Content Manager	Коммерческая DLP-система,	Высокая производительность	Хорошо масштабируется для больших	Надежная система, хорошо	Может быть сложной	Платная

	интегрированная с другими решениями Cisco по безопасности		организаций	зарекомендовавшая себя	для настройки и управления	
--	---	--	-------------	------------------------	----------------------------	--

Помимо перечисленных технологий и инструментов, существует множество других решений, которые могут быть использованы для обеспечения информационной безопасности АИС, при этом выбор конкретных технологий и инструментов должен зависеть от: потребностей организации, уровня ее рисков, бюджета, кадрового обеспечения. Также важно, чтобы все используемые технологии и инструменты были интегрированы друг с другом и работали в единой системе, поскольку обеспечение информационной безопасности АИС – это непрерывный процесс, который требует постоянного внимания и усилий.

Именно поэтому предлагается регулярно проводить организационные и технические мероприятия такие, как оценка рисков, обновление технологий и инструментов, обучение персонала, тестирование системы безопасности, не забывая о том, что только комплексный подход к обеспечению информационной безопасности позволит защитить АИС от киберугроз.

Таким образом, на основе проведенного анализа следует отметить, что при построении системы защиты информации для АИС важно учитывать множество факторов, включая специфику бизнеса, объем и чувствительность данных, а также существующие угрозы и уязвимости. Именно поэтому предлагаются следующие рекомендации по выбору наиболее подходящих инструментов для различных аспектов информационной безопасности:

1. Управление доступом и аутентификация

Рекомендации:

Microsoft Active Directory (AD). Для централизованного управления учетными записями пользователей и их правами доступа.

Okta или Azure AD. Для реализации многофакторной аутентификации (MFA) и единого входа (SSO).

Преимущества:

Централизованное управление пользователями и их правами доступа; повышенная безопасность за счет многофакторной аутентификации.

2. Защита сети

Рекомендации:

Cisco ASA или Palo Alto Networks. Для защиты периметра сети с помощью межсетевых экранов и систем предотвращения вторжений (IDS/IPS);

Snort. Для глубокого анализа трафика и обнаружения угроз.

Преимущества:

Эффективная защита от внешних атак и несанкционированного доступа; возможность обнаружения и предотвращения подозрительной активности в реальном времени.

3. Шифрование данных

Рекомендации:

BitLocker. Для шифрования данных на локальных дисках и съемных носителях.

OpenSSL. Для обеспечения безопасности передачи данных через Интернет с использованием SSL/TLS.

Преимущества:

Защита данных на устройствах и в процессе их передачи; поддержка высоких стандартов безопасности и совместимость с различными системами[10].

4. Антивирусное ПО

Рекомендации:

Symantec Endpoint Protection или **McAfee Total Protection**. Для защиты от вирусов, троянов и других вредоносных программ.

Malwarebytes. Для дополнительной защиты и удаления вредоносного ПО.

Преимущества:

Эффективное обнаружение и удаление вредоносных программ; постоянное обновление баз данных угроз для защиты от новых вирусов.

5. Мониторинг и управление инцидентами

Рекомендации:

Splunk или **IBM QRadar**. Для централизованного сбора, анализа и корреляции событий безопасности (SIEM).

Archer от **RSA**. Для управления рисками и инцидентами, а также обеспечения соответствия нормативным требованиям.

Преимущества:

Обеспечение видимости и контроля над событиями безопасности в реальном времени; автоматизация процессов обнаружения, анализа и реагирования на инциденты.

6. Резервное копирование и восстановление данных

Рекомендации:

Veeam Backup & Replication. Для регулярного создания резервных копий данных и их восстановления в случае утраты.

Acronis Backup. Для обеспечения непрерывности бизнеса и защиты данных от сбоев и катастроф.

Преимущества:

Надежное резервное копирование и быстрое восстановление данных; защита от потерь данных из-за аппаратных сбоев или кибератак.

7. Защита от утечек данных (DLP)

Рекомендации:

Symantec Data Loss Prevention. Для мониторинга и предотвращения утечек конфиденциальной информации.

Forcepoint DLP. Для контроля и защиты данных как на уровне конечных точек, так и при передаче через сеть.

Преимущества:

Предотвращение несанкционированного распространения конфиденциальной информации, поддержка различных политик и правил для защиты данных[11].

Таким образом, в рамках проведенного аналитического обзора были выявлены ключевые направления, принципы и методы защиты информации, а также проведена оценка современных инструментов и технологий, используемых для формирования надежной системы безопасности. Информационная безопасность в современных условиях требует применения комплексного подхода, включающего взаимосвязанные организационные и технические меры, которые в совокупности обеспечивают минимизацию рисков и предотвращение инцидентов, связанных с несанкционированным доступом или нарушением целостности данных.

Организационные методы, такие как разработка и внедрение политик безопасности, управление правами доступа, а также обучение сотрудников основам информационной безопасности, создают базовые условия для обеспечения защищенности на уровне всей организации. Эти меры способствуют формированию устойчивой системы управления рисками, направленной на минимизацию вероятности человеческих ошибок, а также обеспечивают соответствие установленным нормативам и стандартам в сфере информационной безопасности.

Технические средства представляют собой основу защиты IT-инфраструктуры. К ним относятся межсетевые экраны, системы обнаружения и предотвращения вторжений (IDS/IPS), антивирусные программы, технологии шифрования данных, а также решения для резервного копирования информации. Указанные инструменты обеспечивают многоуровневую защиту, способную нейтрализовать широкий спектр угроз и уменьшить потенциальные последствия инцидентов информационной безопасности.

Таким образом, формирование эффективной системы защиты информации возможно лишь при интеграции организационных и технических мер, что позволяет создать устойчивую и адаптивную IT-инфраструктуру, способную противостоять современным киберугрозам. Данный подход обеспечивает комплексное управление рисками, повышая надежность системы безопасности и снижая уязвимость информационных ресурсов.

Литература:

1. Smart-Soft. Защита информации в автоматизированных системах. URL: https://www.smart-soft.ru/blog/zaschita_informatsii_v_avtomatizirovannyh_sistemah (дата обращения: 11.06.2024).
2. Методы и средства разработки автоматизированных информационных систем на основе онтологии. URL: <https://cyberleninka.ru/article/n/metody-i-sredstva-razrabotki-avtomatizirovannyh-informatsionnyh-sistem-na-osnove-ontologii-upravlenie-kachestvom-programmno/viewer> (дата обращения: 15.06.2024).
3. Автоматизированные информационные системы и их возможности. URL: <https://cyberleninka.ru/article/n/avtomatizirovannye-informatsionnye-sistemy-ais-i-ih-vozmozhnosti-ispolzovanie-programmy-ais-kachestvo-produktsii-na-predpriyatii/viewer> (дата обращения: 21.06.2024).
4. Функциональная структура автоматизированной информационной системы профилактики безнадзорности и правонарушений. URL: <https://cyberleninka.ru/article/n/funktsionalnaya-struktura-avtomatizirovannoy-informatsionnoy-sistemy-profilaktiki-beznadzornosti-i-pravonarusheniy/viewer> (дата обращения: 11.07.2024).
5. Методы обеспечения информационной защиты. URL: <https://domgadalki.ru/foto/metodiy-obespecheniya-informatsionnoy-zashitiy> (дата обращения: 21.08.2024).
6. Концепция обеспечения безопасности информации. URL: <https://res.sm-center.ru/sibir/File/6b3b75e5-03f9-4ae1-a11e-9bd850a7c513:Концепция.pdf> (дата обращения: 11.09.2024).
7. Solodjannikov. Информационная безопасность. URL: <https://infosec.spb.ru/wp-content/uploads/2020/08/solodjannikov.pdf> (дата обращения: 15.09.2024).

8. Защита информации в автоматизированных системах. URL: <https://searchinform.ru/services/outsource-ib/zaschita-informatsii/v-informatsionnykh-sistemakh/v-avtomatizirovannykh-informatsionnykh-sistemakh/> (дата обращения: 18.09.2024).
9. 12 топовых IDS/IPS инструментов. URL: <https://www.csoonline.com/article/569085/12-top-idsips-tools.html> (дата обращения: 11.10.2024).
10. Selectel. Мониторинг информационной безопасности. URL: <https://habr.com/ru/articles/790002/> (дата обращения: 11.10.2024).
11. Инструменты для управления идентификацией. URL: https://livebusiness.org/tools/identity_management/ (дата обращения: 11.10.2024)

УДК 004.056.5

С.В. Корякин, srgkoryakin1@gmail.com¹Институт информационных технологий КГТУ им. И.Раззакова²Институт машиноведения автоматизации и геомеханики НАН КР

С.Н. Верзунов,

Институт машиноведения автоматизации и геомеханики НАН КР

П.Л. Вейс, veis_p@iuca.kg

Международный университет в Центральной Азии

А.К. Оруналиева,

Институт информационных технологий КГТУ им. И. Раззакова

АНАЛИТИЧЕСКИЙ ОБЗОР МЕТОДОВ И СПОСОБОВ РЕАЛИЗАЦИИ СИСТЕМ ЗАЩИТЫ ПЕРСОНАЛЬНЫХ ДАННЫХ НА ПРЕДПРИЯТИИ

Аннотация. В статье приводится анализ методов и технологий, применяемых для построения систем защиты персональных данных (ТЗПД) на предприятиях. Рассматриваются и предлагаются различные подходы к их реализации. Особое внимание уделено как правовому регулированию в Кыргызской Республике, так и международным стандартам. Приведены примеры эффективных решений, применяемых для защиты информации, а также обсуждаются стратегии управления рисками, связанные с утечкой и несанкционированным доступом к данным.

Ключевые слова: защита данных, персональные данные, информационная безопасность, инструменты безопасности, технологии защиты, шифрование, системы IDS/IPS, управление доступом (IAM), антивирусное ПО, файрволлы, многофакторная аутентификация (MFA), резервное копирование данных, DLP-системы, SIEM-системы, управление инцидентами безопасности, аудит доступа, обучение сотрудников, предотвращение утечек данных, киберугрозы, защита конфиденциальной информации.

Защита персональных данных, обрабатываемых организациями, в настоящее время является одной из наиболее актуальных проблем в сфере информационной безопасности и кибербезопасности. Увеличение объемов данных, а также частота инцидентов, связанных с утечками, создают необходимость разработки комплексных систем защиты, которые соответствуют не только требованиям бизнеса, но и законодательным нормам. В Кыргызской Республике данный вопрос регулируется Законом «О персональных данных» № 58 от 14 апреля 2008 года [1], который устанавливает обязательства организаций по защите персональных данных, а также штрафные санкции за несоблюдение этих обязательств. Следует отметить, что в глобальном контексте важную роль играет GDPR — нормативно-правовой акт Европейского союза, касающийся обработки персональных данных [2].

Целью проводимого в данной работе исследования является глубокий анализ современных подходов к созданию систем защиты персональных данных в корпоративной среде, а также оценка их эффективности с точки зрения минимизации возможных рисков. Именно по этой причине при построении ТЗПД предлагается выделить следующие основные аспекты, которые наиболее полно способствуют реализации функциональных возможностей и эффективности разрабатываемых ТЗПД: правовые основы и регуляторные требования, классификация подходов к защите данных, технические решения для защиты данных, управление рисками и предотвращение утечек данных, адаптивные меры защиты в условиях изменяющихся угроз.

1. Правовые основы и регуляторные требования

Первостепенным аспектом в вопросе защиты данных на предприятиях является выполнение законодательных и нормативных требований. В Кыргызской Республике, как уже было упомянуто, нормативной базой для защиты персональных данных выступает Закон № 58 «О персональных данных» [1], который определяет обязательства субъектов данных и операторов. Этот закон устанавливает принципы, по которым компании обязаны обрабатывать и хранить персональные данные, включая необходимость обеспечения их безопасности.

Следует подчеркнуть, что данные нормы предполагают внедрение комплексных мер защиты, направленных на предотвращение как внешних, так и внутренних угроз. Эти меры охватывают как технические аспекты защиты данных, так и организационные мероприятия, обеспечивающие их соблюдение [3].

2. Классификация подходов к защите данных

Существующие подходы к защите персональных данных можно условно разделить на три большие категории: организационные, технические и физические меры. Каждая из них нацелена на устранение или минимизацию различных типов угроз.

Организационные меры включают разработку внутренних политик и процедур, регламентирующих порядок обработки данных и поведение сотрудников при работе с ними. Обучение сотрудников и контроль за выполнением установленных правил играют ключевую роль в снижении рисков, связанных с человеческим фактором [4].

Технические меры реализуются через использование специализированного программного обеспечения и аппаратных решений. Например, одним из основных методов защиты данных является шифрование, которое делает информацию недоступной для третьих лиц в случае утечки. Используются как симметричные (AES), так и асимметричные алгоритмы шифрования (RSA) [5]. Также активно применяются технологии многофакторной аутентификации (2FA), что существенно снижает риск несанкционированного доступа [6].

Физические меры направлены на защиту инфраструктуры и оборудования. К ним можно отнести контроль доступа к серверным комнатам, использование биометрических систем и видеонаблюдение. Эти методы обеспечивают защиту от физических угроз, таких как кража оборудования или саботаж [7].

3. Технические решения для защиты данных

Одним из ключевых элементов любой системы защиты является шифрование данных, которое гарантирует, что даже при утечке информации злоумышленники не смогут ее расшифровать. В современных условиях симметричное шифрование (например, алгоритм AES) широко используется для защиты больших объемов данных, тогда как асимметричное шифрование (RSA, ECC) применяется для более специфичных задач, таких как защита ключей и аутентификация [5].

Кроме шифрования, важную роль в защите персональных данных играет использование систем мониторинга и анализа событий информационной и кибербезопасности так называемых SIEM-систем. Такие системы позволяют в реальном времени отслеживать инциденты, проводить анализ сетевого трафика и оперативно реагировать на возможные угрозы [8]. SIEM-инструменты дают возможность собирать и анализировать данные о событиях, происходящих в информационной системе, что позволяет предотвратить потенциальные атаки до их реализации.

Особое внимание следует уделить многофакторной аутентификации (MFA), которая позволяет повысить уровень безопасности, требуя от пользователей предоставления нескольких доказательств их личности. Это может быть сочетание пароля, физического токена и биометрических данных, что значительно усложняет задачу злоумышленникам [9].

4. Управление рисками и предотвращение утечек данных

Для обеспечения должного уровня безопасности важно не только применять технические и организационные меры, но и грамотно управлять рисками. Оценка рисков позволяет выявить наиболее уязвимые точки в системе защиты и разработать адекватные меры для их минимизации [10]. В частности, применение методик управления рисками,

основанных на стандартах ISO/IEC 27001, позволяет структурировать процессы и создать эффективную систему управления информационной безопасностью (СУИБ) [11].

Кроме того, одним из важнейших инструментов для предотвращения утечек данных являются системы DLP (Data Loss Prevention). Эти решения позволяют контролировать перемещение данных внутри организации, предотвращать их утечку за ее пределы и блокировать несанкционированные действия пользователей. DLP-системы обеспечивают фильтрацию трафика, контроль за копированием данных и анализ подозрительных активностей [12].

5. Адаптивные меры защиты в условиях изменяющихся угроз

Современная киберсреда постоянно эволюционирует, что требует от предприятий адаптации к новым типам угроз. С каждым годом растет количество кибератак, направленных на получение персональных данных (ПД). В этой связи важным становится внедрение технологий, основанных на искусственном интеллекте и машинном обучении, которые позволяют предсказывать и эффективно реагировать на новые угрозы [13].

Применение ИИ и машинного обучения в области информационной безопасности и кибербезопасности открывает новые горизонты для построения гибких и устойчивых систем защиты. Эти технологии способны анализировать огромные объемы данных в режиме реального времени, выявлять аномалии и самостоятельно принимать решения о блокировке подозрительной активности [14].

Следует отметить, что на сегодняшний день в условиях стремительного развития информационных технологий, а также с учетом всеобщей цифровизации объектов критической инфраструктуры организаций наблюдается увеличение объема обрабатываемых данных внутри киберфизических систем. В связи с этим защита персональных данных становится одной из ключевых задач для предприятий. Угроза утечек и несанкционированного доступа к конфиденциальной информации требует внедрения комплексных мер безопасности. В Кыргызской Республике защита персональных данных регулируется рядом законодательных актов, таких как Закон «О персональных данных», «Закон о кибербезопасности» и Уголовный кодекс Кыргызской Республики, предусматривающий ответственность за нарушение правил обработки и защиты персональных данных.

Считается, что для создания систем защиты ПД используются следующие современные инструменты и технологии, которые предлагается разделить на основные – базовые и дополнительные – опциональные.

К основным – базовым системам, предлагается отнести следующие: системы предотвращения вторжений, шифрование данных, системы управления идентификацией и доступом (IAM), антивирусное ПО и антималяварные решения, файрволлы (межсетевые экраны), системы резервного копирования и восстановления данных.

Дополнительными или опциональными предлагается называть: многофакторная аутентификация (MFA), DLP-системы (Data Loss Prevention), SIEM-системы (Security Information and Event Management), учет и аудит доступа к данным, управление инцидентами безопасности.

1. Системы предотвращения вторжений (IDS/IPS)

Эти системы отслеживают сетевой трафик и активность на предмет подозрительной деятельности и попыток вторжений. IDS (Intrusion Detection Systems) обнаруживают и сообщают о подозрительных действиях, в то время как IPS (Intrusion Prevention Systems) не только обнаруживают, но и предотвращают такие действия (рис. 1).

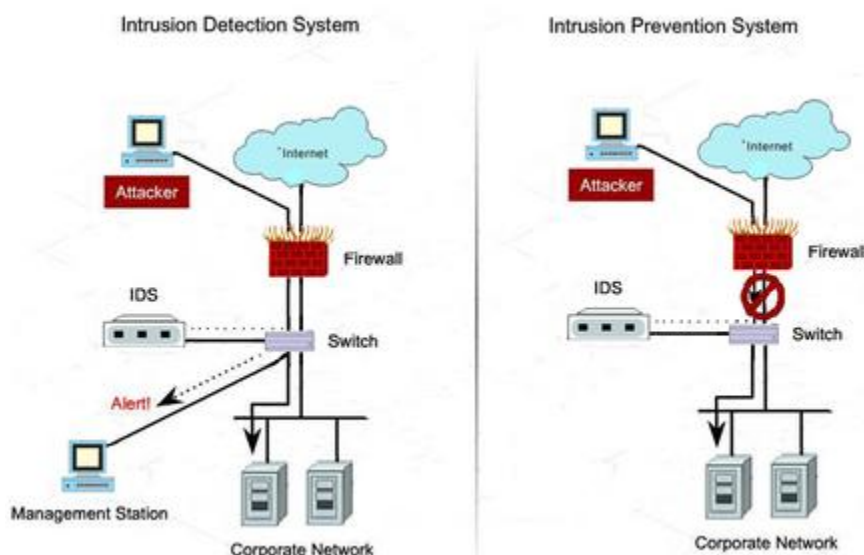


Рисунок 1– Обобщенная структурная схема IDS/IPS

В качестве самых распространенных в использовании из существующих IDS/IPS систем в настоящее время считаются: Snort – популярная IDS с открытым исходным кодом, Cisco Firepower – комплексное решение, включающее IDS и IPS[15]. Кроме того, следует отметить, что существует очень большое количество IDS и IPS-решений, которые широко используются специалистами по кибербезопасности, при этом выбор того или иного решения зависит от личных предпочтений специалиста, а также требований стандартов и нормативных документов, используемых при реализации систем защиты персональных данных.

2. Шифрование данных

Шифрование данных обеспечивает защиту информации путем преобразования ее в зашифрованный текст, который может быть расшифрован только при наличии соответствующего ключа. Существует несколько типов шифрования, включая симметричное и асимметричное (рис. 2).

Примеры решений для шифрования:

- AES (Advanced Encryption Standard) – стандарт симметричного шифрования.
- RSA – алгоритм асимметричного шифрования, широко используемый для защиты данных[16].

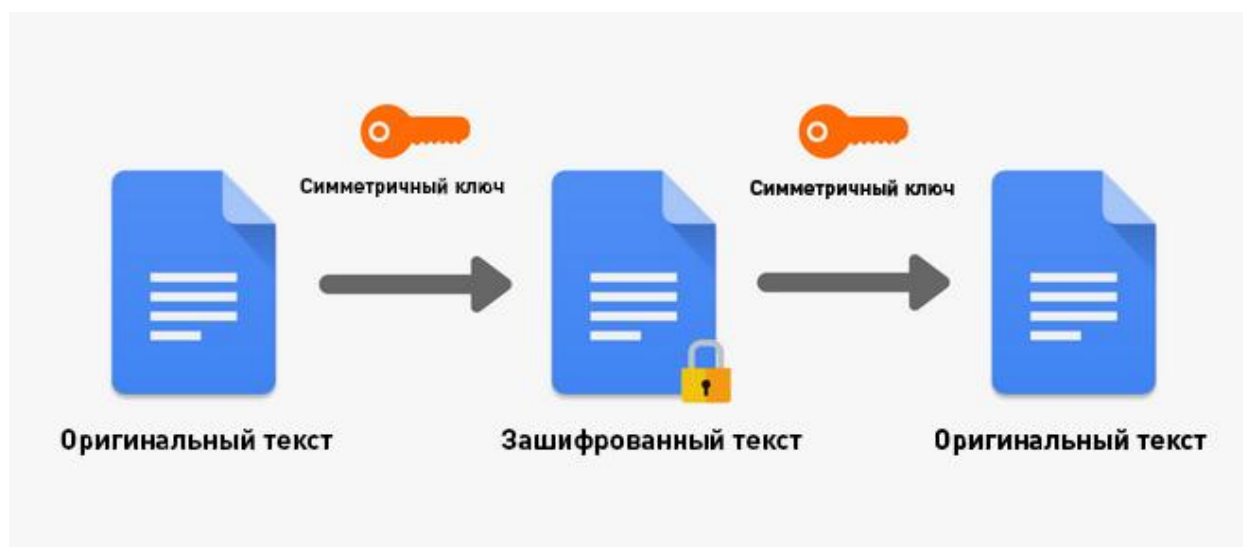


Рисунок 2 – Обобщенный вариант процесса шифрования данных

3. Системы управления идентификацией и доступом (IAM)

IAM (Identity and Access Management) позволяет контролировать доступ сотрудников к информационным ресурсам на основе их ролей и полномочий. Эти системы включают в себя управление учетными записями пользователей, аутентификацию и авторизацию (рис.3).

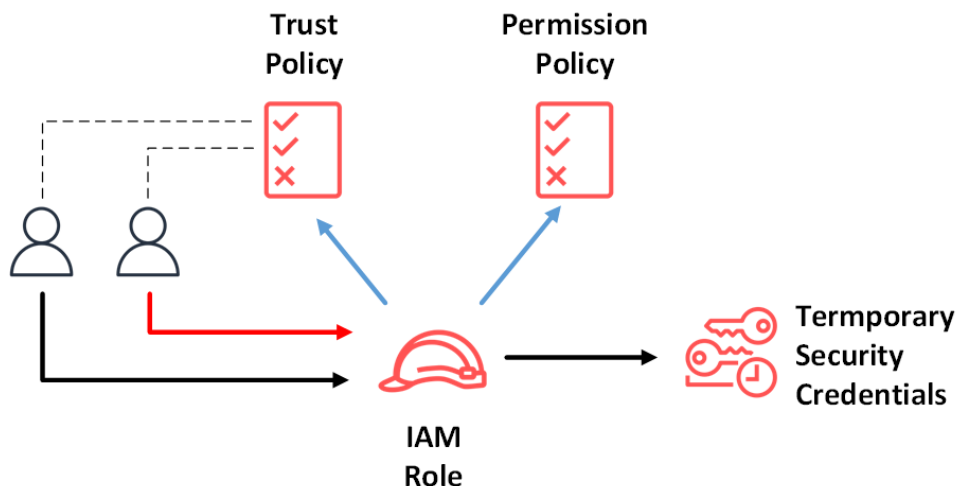


Рисунок 3 – Обобщенная схема работы IAM

Из наиболее используемых IAM-систем предлагается выделить: Microsoft Azure Active Directory – облачная служба управления идентификацией, Okta – платформа управления идентификацией и доступом[17].

4. Антивирусное ПО и антималварные решения

Антивирусные программы и антималварные решения предназначены для обнаружения, предотвращения и удаления вредоносных программ. Они работают на основе сигнатур и методов эвристического анализа, применяемого в СЗПД для защиты от известных и новых киберугроз. Наиболее известными примерами антивирусного ПО являются: Kaspersky Antivirus – антивирусное ПО, известное своей высокой эффективностью, Norton Security – комплексное решение для защиты от вредоносного ПО.

5. Файрволлы (межсетевые экраны)

Файрволлы ограничивают несанкционированный доступ к сетям предприятия, фильтруя входящий и исходящий трафик на основе заданных правил безопасности. Они могут быть аппаратными или программными.

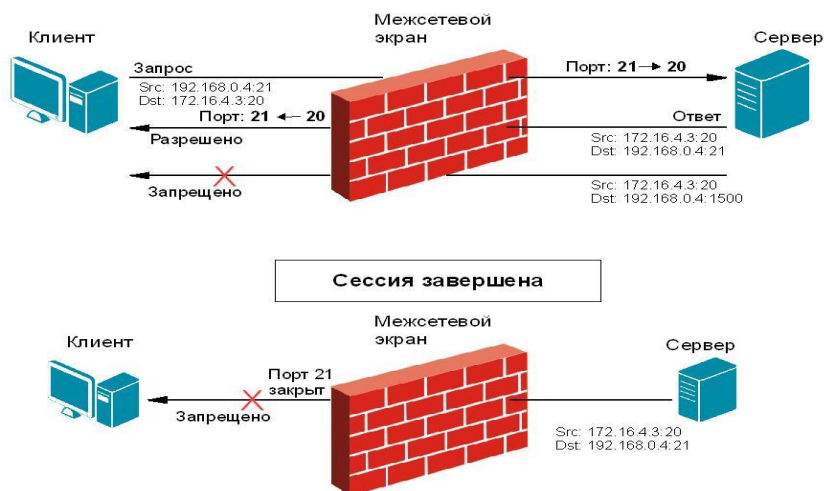


Рисунок 4 – Принцип работы межсетевого экрана

Следует выделить такие решения, как Cisco ASA – аппаратный фаерволл, pfSense – программный фаерволл с открытым исходным кодом, которые получили очень широкое применение в СЗПД[15.18].

6. Системы резервного копирования и восстановления данных

Считается, что резервное копирование данных обеспечивает их сохранность в случае сбоев или атак. Эти системы позволяют восстанавливать данные из резервных копий, хранящихся на удаленных серверах или внешних носителях (рис. 5).

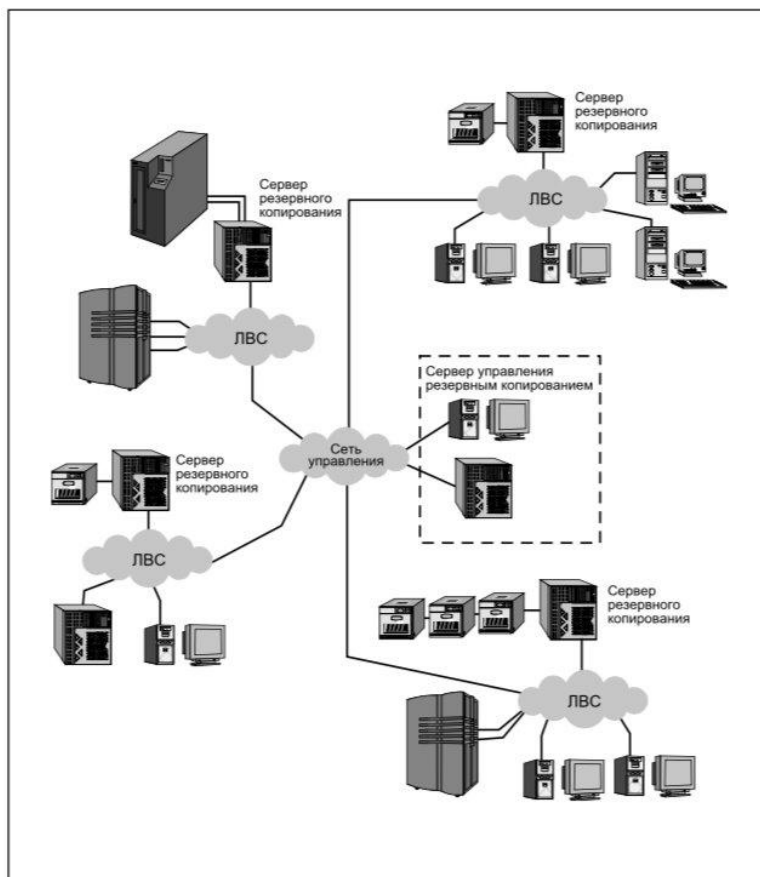


Рисунок 5 – Обобщенная структурная схема системы резервного копирования

Из наиболее распространенных вариантов систем для резервного копирования данных выделяют: Veeam Backup & Replication – решение для резервного копирования и восстановления данных и Acronis Backup – программное обеспечение для создания резервных копий и восстановления данных.

Дополнительные технологии и инструменты информационной безопасности

1. Многофакторная аутентификация (MFA)

Многофакторная аутентификация добавляет дополнительный уровень безопасности, требуя от пользователей подтверждать свою личность несколькими способами, такими как пароли, отпечатки пальцев или SMS-коды.

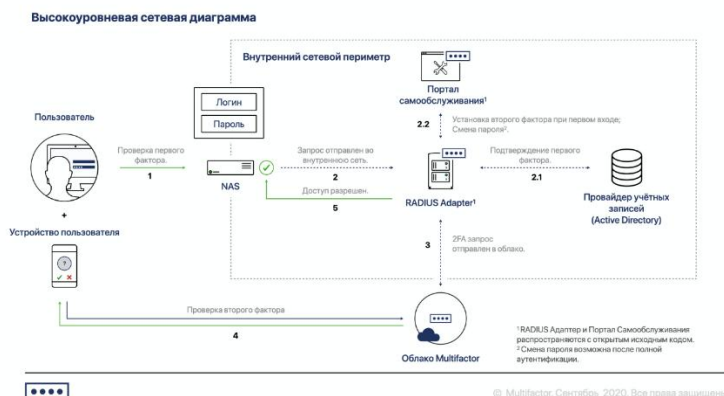


Рисунок 6 – Принцип работы системы
многофакторной аутентификации

Примеры решений для MFA Google Authenticator – приложение для создания одноразовых паролей и Duo Security – решение для многофакторной аутентификации[19].

2. DLP-системы (Data Loss Prevention)

Системы предотвращения утечек данных (DLP) предназначены для обнаружения и предотвращения несанкционированной передачи конфиденциальной информации за пределы организации.

Примеры DLP-решений

- Symantec Data Loss Prevention – комплексное решение для защиты данных.
- Digital Guardian – платформа для предотвращения утечек данных.

3. SIEM-системы (Security Information and Event Management)

Системы управления информацией и событиями безопасности (SIEM) собирают, анализируют и предоставляют отчеты о событиях безопасности в реальном времени.

Наиболее распространенными примерами SIEM систем являются: Splunk – платформа для мониторинга и анализа данных, IBM QRadar – решение для управления информацией и событиями безопасности и др.

4. Обучение и повышение осведомленности сотрудников

Одним из ключевых аспектов защиты персональных данных является обучение сотрудников принципам информационной безопасности. Регулярные тренинги и курсы повышают осведомленность персонала о потенциальных угрозах и мерах защиты.

5. Учет и аудит доступа к данным

Для обеспечения безопасности персональных данных необходимо регулярно проводить учет и аудит доступа к данным. Это позволяет отслеживать, кто и когда имел доступ к конфиденциальной информации, и выявлять возможные нарушения.

6. Управление инцидентами безопасности

Эффективное управление инцидентами безопасности включает в себя разработку процедур и планов действий на случай нарушения безопасности данных. Это позволяет минимизировать последствия инцидентов и быстро восстанавливать нормальную работу системы.

В таблице 1 представлен сравнительный анализ характеристик современных технологий и инструментов при создании СЗПД на предприятии[20].

Таблица 1– Сравнительный анализ характеристик современных технологий и инструментов при создании СЗПД на предприятии

Технология/Метод	Преимущества	Недостатки
Системы обнаружения и предотвращения вторжений	- Высокий уровень обнаружения угроз	- Высокая стоимость

(IDS/IPS)	- Возможность автоматического реагирования на инциденты безопасности	- Возможные ложные срабатывания
Шифрование данных	- Высокий уровень защиты данных	- Затраты на вычислительные ресурсы
	- Возможность использования в различных средах	- Сложности в управлении ключами
Управление доступом к информации (IAM)	- Централизованное управление доступом	- Сложность внедрения
	- Снижение риска несанкционированного доступа	- Необходимость постоянного администрирования
Антивирусное программное обеспечение (ПО)	- Обнаружение и удаление большинства вредоносных программ	- Возможность пропуска новых угроз
		- Замедление работы системы
Файрволлы	- Эффективная защита сети	- Сложность настройки и администрирования
	- Возможность настройки правил доступа	- Не защищают от внутренних угроз
Системы резервного копирования	- Восстановление данных в случае сбоев	- Требуют значительных ресурсов для хранения
	- Защита от потерь данных	- Возможность устаревания резервных копий
Многофакторная аутентификация (MFA)	- Повышение уровня безопасности за счет дополнительных факторов аутентификации	- Увеличение времени доступа
		- Необходимость дополнительного оборудования или ПО
DLP-системы	- Защита конфиденциальной информации	- Сложность настройки и управления
	- Предотвращение утечек данных	- Высокая стоимость
SIEM-системы	- Мониторинг безопасности в реальном времени	- Высокая стоимость внедрения и эксплуатации
	- Возможность анализа и корреляции событий	- Необходимость квалифицированного персонала

Обучение и повышение осведомленности сотрудников	- Уменьшение числа инцидентов, связанных с человеческим фактором	- Необходимость регулярного проведения тренингов
	- Повышение уровня ответственности сотрудников за защиту данных	- Требование ресурсов для разработки и проведения обучающих программ
Учет и аудит доступа к данным	- Возможность быстрого выявления несанкционированного доступа	- Требуется внедрения специализированного ПО
	- Учет всех действий, связанных с обработкой данных	- Необходимость в ресурсах для проведения регулярных аудитов
Управление инцидентами безопасности	- Сокращение времени реагирования на инциденты	- Необходимость постоянного обновления планов и процедур
	- Снижение ущерба от потенциальных утечек данных	- Требуется квалифицированного персонала для управления инцидентами

Таким образом из проведенного анализа технологий решения и методов построения ПЗПД с целью обеспечения надежной защиты персональных данных на предприятии рекомендуется использовать комплексный подход, включающий мероприятия, реализующие следующие обязательные функции ПЗПД:

- Внедрение IDS/IPS для мониторинга и защиты от сетевых атак.
- Использование шифрования данных для защиты информации как в покое, так и при передаче.
- Реализация IAM для управления доступом к информационным ресурсам на основе политики доступа.
- Установка антивирусного ПО для защиты от вредоносных программ.
- Настройка фаерволлов для ограничения несанкционированного доступа к сети предприятия.
- Регулярное создание резервных копий данных и проверка их на целостность.
- Внедрение многофакторной аутентификации для повышения уровня безопасности.
- Использование DLP-систем для предотвращения утечек данных.
- Внедрение SIEM-систем для мониторинга и анализа событий безопасности в реальном времени.

На основании проведенного анализа можно сделать вывод о том, что для обеспечения эффективной защиты персональных данных на предприятиях необходимо применять комплексный подход, сочетающий технические, организационные и правовые меры. Современные технологии, такие как шифрование, многофакторная аутентификация и системы мониторинга, играют ключевую роль в предотвращении утечек данных и защите от несанкционированного доступа. Внедрение риск-ориентированных подходов и адаптация систем безопасности к новым вызовам позволит повысить устойчивость предприятий к внешним и внутренним угрозам. Следует отметить, что успешная реализация системы защиты данных невозможна без учета специфики бизнеса и постоянного мониторинга

актуальных угроз и уязвимостей. В этом контексте важную роль играет внедрение систем управления информационной безопасностью, основанных на международных стандартах, таких как ISO/IEC 27001.

Список литературы

1. Закон Кыргызской Республики «О персональных данных» от 14 апреля 2008 года № 58.
2. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation).
3. Указ президента Кыргызской Республики «О мерах по защите персональных данных» от 1 марта 2017 года.
4. Лившиц В.М., Шестаков Д.В. Информационная безопасность организаций. – М.: Юнити-Дана, 2020. – 320 с.
5. Сорокин В.А. Методы защиты информации. – СПб.: Питер, 2019. – 456 с.
6. Stallings W. Cryptography and Network Security: Principles and Practice. – Pearson, 2020. – 816 p.
7. Журавлев В.П. Защита информации в корпоративных системах. – М.: Горячая Линия – Телеком, 2019. – 384 с.
8. Williams, P. and Wills, L. Security Information and Event Management (SIEM) Implementation. – McGraw-Hill, 2018.
9. Cuppens, F., & Mieke, A. Intrusion Detection Systems. – Springer, 2021. – 512 p.
10. Kizza, J.M. Guide to Computer Network Security. – Springer, 2020. – 535 p.
11. ISO/IEC 27001:2022. Information security, cybersecurity and privacy protection — Information security management systems — Requirements.
12. Harris S. CISSP All-in-One Exam Guide. – McGraw-Hill Education, 2019. – 1456 p.
13. Anderson, R. Security Engineering: A Guide to Building Dependable Distributed Systems. – Wiley, 2020. – 1088 p.
14. Bishop, M. Introduction to Computer Security. – Addison-Wesley, 2020. – 786 p.
15. Бойко, В. В. (2021). Информационная безопасность. М.: Юрайт. <http://www.publishing-vak.ru/file/archive-economy-2021-2/16-ershova-boiko.pdf>
16. Снитко, С. А. (2020). Управление информационной безопасностью. https://rusneb.ru/catalog/000199_000009_010930271/
17. Kaufman, C., Perlman, R., & Speciner, M. (2016). Network Security: Private Communication in a Public Internet. Prentice Hall. (<https://www.pearson.com/store/p/network-security-private-communication-in-a-public-internet/P100000323269>)
18. Symantec Corporation. (2020). Symantec Data Loss Prevention. [https://cortex.marketplace.pan.dev/marketplace/details/SymantecDLP/#:~:text=Symantec%20Data%20Loss%20Prevention%20\(DLP\),of%20your%20sensitive%20corporate%20data](https://cortex.marketplace.pan.dev/marketplace/details/SymantecDLP/#:~:text=Symantec%20Data%20Loss%20Prevention%20(DLP),of%20your%20sensitive%20corporate%20data)
19. Cisco Systems. (2019). Cisco Firepower Next-Generation Firewall. https://www.cisco.com/c/dam/m/ru_ua/campaigns/security-hub/pdf/datasheet-c78-736661.pdf
20. Microsoft Corporation. (2021). Microsoft Azure Active Directory. <https://www.microsoft.com/en-us/security/business/identity-access/microsoft-entra-id>

УДК: 621.224.7:621.311.214.1

*Т.Т. Медеров, mtt-kg@mail.ru,**Р.А. Акпаралиев, ruslan.akparaliev@gmail.com**Институт машиноведения и автоматизации НАН КР*

ИССЛЕДОВАНИЕ ГИДРОДИНАМИКИ БИРОТОРНОЙ ГИДРОТУРБИНЫ ПРИ ПОМОЩИ ПРОГРАММЫ FLOWVISION

В статье представлено исследование и моделирование гидротурбины бироторной микро-ГЭС с использованием современных численных методов гидродинамики. Принцип работы данной микро-ГЭС основан на использовании двух рабочих колес, вращающихся в противоположных направлениях в одном потоке воды, что позволяет увеличить эффективность преобразования водной энергии. Отмечаются результаты численного моделирования, проведенного с использованием программного комплекса FlowVision, включая распределение скорости, направление векторов и линии тока, а также распределение давлений на лопастях гидротурбин. Подчеркивается необходимость дальнейших исследований, включая анализ гидродинамики в переходных режимах, взаимодействия двух турбин и оптимизации процессов, а также разработку надежных систем управления и безопасности. Исследование бироторной микро-ГЭС представляет собой важное направление и может способствовать развитию устойчивых, возобновляемых источников энергии.

Ключевые слова: микро-ГЭС; бироторная микро-ГЭС; возобновляемые источники энергии; гидротурбина; гидродинамика; численные методы; программный комплекс FlowVision; напор; расход; мощность; вычислительная гидродинамика (ВГД).

Введение

В большинстве стран уже возобновляемые источники энергии (ВИЭ) стали основным ресурсом для расширения мощностей в электроэнергетике и в настоящее время привлекают наибольшее внимание инвесторов. Средневзвешенная стоимость производства электроэнергии на новых промышленных солнечных фотоэлектрических станциях в мире снизилась на 85% за последние десять лет. Для концентрированной солнечной энергии (КСЭ) это снижение составило 68%, для наземной ветроэнергетики — 56%, а для морской ветроэнергетики — 48%. В условиях истощения и роста цен на ископаемое топливо экономические перспективы энергетики на основе ВИЭ становятся все более очевидными [1].

В Кыргызской Республике, где более 93% территории покрыто горами и множество рек извилисто протекают через ущелья, использование гидроэнергии этих малых рек считается одним из наиболее перспективных источников возобновляемой энергии. По оценкам, потенциал этих рек составляет примерно 5–8 миллиардов киловатт-часов в год [2]. В исследованиях, проведенных Программой развития ООН (ПРООН) и Организацией промышленного развития ООН (ЮНИДО), было установлено, что возможно создание или восстановление от 87 до 92 малых гидроэлектростанций. По предварительным расчетам, их совокупная установленная мощность составит около 180 МВт, при этом коэффициент использования установленной мощности в среднем составит 63%. Годовой потенциал производства энергии этих объектов оценивается примерно в 1 миллиард кВт.ч, что эквивалентно 1000 ГВт.ч [3]. Строительство крупных гидроэлектростанций требует значительных временных, трудовых и финансовых затрат. В то же время создание и внедрение малых гидроэнергетических установок значительно проще и быстрее, особенно в удаленных регионах, где другие виды энергии могут быть непрактичными или экономически невыгодными. Малые и микро-ГЭС представляют собой эффективное решение для электрификации сельских территорий. Благодаря способности оперативно реагировать на изменения нагрузки пиковая мощность таких ГЭС играет важную роль в обеспечении стабильного электроснабжения. Они могут снижать частоту запусков и остановок тепловых электростанций, что уменьшает нагрузку на них и сокращает выбросы углекислого газа [4]. Также гидроэлектростанции могут способствовать решению вопросов, связанных с интеграцией источников энергии, работающих на

возобновляемых ресурсах, таких как солнечные и ветровые электростанции. В Кыргызстане вырабатывается около 15 млрд кВт.ч электроэнергии каждый год, причем потребление растет в среднем на 5–10% [5]. Рост потребления связан с тем, что в стране установлены одни из самых низких тарифов на электроэнергию для бытовых потребителей в мире. Благодаря социально ориентированным тарифам электроэнергия широко используется для отопления, что значительно увеличивает объемы ее потребления. По данным энергокомпаний, последние годы в стране фиксируются рекордные значения потребления электрической энергии за сутки в осенне-зимний период (ОЗП), и энергетическая система работает в пределах существующих мощностей [6]. Текущая ситуация указывает на острую необходимость ввода новых генерирующих мощностей. Для решения одной части задач ввода новых мощностей эффективно может быть использован потенциал малых рек. Использование энергии малых водотоков в значительной мере могут решить проблему электроснабжения маломощных потребителей, расположенных в предгорных и горных районах, тем самым повысив их социально-экономическое положение [7, 8]. По этой причине исследования в данной сфере представляют собой актуальную задачу.

Сегодня разработано и успешно используется большое разнообразие малых и микро-ГЭС для электроснабжения таких потребителей. В зависимости от условий эксплуатации они имеют разные формы и компоненты. Все они, помимо своих преимуществ, обладают и рядом недостатков, таких как громоздкость, высокая металлоемкость, значительная стоимость и т.д. В связи с этим группой исследователей из технического университета была разработана и создана микро-ГЭС бироторного типа с совершенно новым дизайном, избавленная от упомянутых недостатков, а также получен ряд патентов на изобретение. Были проведены лабораторные и натурные эксперименты, в результате которых были доказаны верности основных положений концепции бироторных микро-ГЭС и их высокой эффективности. Принцип работы бироторной микро-ГЭС заключается в использовании двух рабочих колес на одной оси, расположенных последовательно и вращающихся в противоположных направлениях в одном потоке воды. Это техническое решение позволяет вращать и статор генератора, снизить размеры и частоту вращения генератора, что в конечном итоге повышает эффективность преобразования водной энергии и делает микро-ГЭС более экономически выгодной.

Создание бироторной микро-ГЭС представляет собой важное техническое достижение, тем не менее следует отметить, что эта технология находится на стадии развития и требует дальнейших научных исследований и оптимизации. Необходимо провести более глубокое научное исследование, в рамках которого будут изучены различные аспекты, оптимизированы процессы и улучшены характеристики устройства. Главными направлениями исследования следует считать анализ гидродинамики в потоке, изучение взаимодействия двух турбин в одном гидравлическом потоке, а также анализ переходных процессов в системе. Первоначально необходимо более подробно исследовать гидродинамические особенности данного процесса, включая распределение скоростей, давления и потери энергии в гидравлическом потоке. Это позволит оптимизировать форму и параметры рабочих колес, а также выбрать оптимальный режим работы для достижения максимальной эффективности. Далее необходимо более подробно изучить взаимодействие двух турбин, включая анализ вихревых явлений, уделить внимание исследованию переходных процессов в системе, таких как пуск и остановка микро-ГЭС, а также реакция оборудования на изменения в гидравлических условиях, что позволит создать надежные и эффективные системы управления и безопасности для таких станций. Это будет способствовать более эффективной интеграции микро-ГЭС в существующие энергетические сети и гарантировать стабильность электроснабжения.

Постановка задачи

- Определение распределения скорости и давления внутри бироторной гидротурбины как в стационарном, так и в переходных режимах.
- Оптимизация положения профиля лопастей на рабочих колесах.
- Моделирование турбулентного вязкого движения несжимаемой жидкости между лопастями двух рабочих колес.
- Определение эффекта взаимодействия между двумя рабочими колесами и их воздействия на гидравлический поток.
- Расчет общей производительности бироторной гидротурбины, включая оценку КПД как в стационарном, так и в переходных режимах.
- Исследование гидродинамических потерь в турбине как в стационарных, так и в переходных режимах.
- Оценка динамического поведения системы при изменении нагрузки или других внешних воздействиях.

Метод моделирования

Использовать численные методы для решения уравнений Навье-Стокса, адаптированные к гидродинамике турбин, включая учет вращения рабочих колес и междутурбинного расстояния. Для учета переходных процессов использовать методы транзитного моделирования, учитывая изменения параметров со временем. Программа FlowVision обладает необходимыми инструментами для проведения такого моделирования.

Основные физические процессы, уравнения и обозначения

В программном комплексе FlowVision реализованы математические модели физических процессов, которые задаются и настраиваются в окнах свойств различных элементов деревьев Препроцессора и Решателя. Физические процессы, протекающие в непрерывной *Фазе*, активируются в свойствах элемента *Препроцессор > Фазы > Фаза №i* (непрерывная) > *Физические процессы*. Это: Движение; Теплопередача; Массообмен; Турбулентность; Фазовый перенос; Излучение.

Физические процессы, протекающие в дисперсной *Фазе* типа *Частицы*, активируются в свойствах элемента *Препроцессор > Фазы > Фаза №i* (дисперсная типа Частицы) > *Физические процессы*. Это: Теплопередача; Фазовый перенос; Движение; Массоперенос.

Физический процесс *Движение* присутствует в позиции *Движение* в свойствах элемента *Препроцессор > Фазы > Фаза № i > Физические процессы*.

Описание модели течения, основанное на уравнении Навье-Стокса:

Уравнение неразрывности:

$$\frac{\partial p}{\partial t} + \nabla(pV) = 0. \quad (1)$$

Уравнение количества движения:

$$\frac{\partial pV}{\partial t} + \nabla(pV \otimes V) = -\nabla p + \nabla \cdot \hat{\tau}_{eff} - DV - \hat{D}V + F + F_L + F_{user} \quad (2)$$

$$\hat{\tau}_{eff} = (\mu + \mu_t) \left(2\hat{S} - \frac{2}{3}(\nabla \cdot V)\hat{I} \right) \quad (3)$$

$$\hat{S}_{ij} = \frac{1}{2} \left(\frac{\partial V_i}{\partial x_j} + \frac{\partial V_j}{\partial x_i} \right) \quad (4)$$

$$D = \left(\frac{\mu}{k_d} + C \frac{1}{2} p |V| \right) \varphi_c \quad (5)$$

$$F = (p - p_g)g - p(2\omega \times V + \omega \times \omega \times r). \quad (6)$$

Здесь:

D — коэффициент, характеризующий изотропное сопротивление.

$D \square$ — матрица коэффициентов анизотропного сопротивления.

k_d — проницаемость твердого каркаса (неподвижной матрицы).

φ_c — относительный объем, занимаемый жидкостью в ячейке (объем пор).

F_L — сила Лоренца, действующая на единицу объема среды со стороны электромагнитного поля.

F_{user} — объемная сила, приложенная пользователем.

$I_{\square}$ — единичный тензор.

ω — угловая скорость вращающейся подобласти.

r — радиус-вектор, исходящий из оси вращения.

Уравнения (1) и (2) вместе обычно называются уравнениями Навье-Стокса. Если коэффициент D в интерфейсе отличается от нуля, уравнение (2) описывает движение жидкости или газа в изотропной пористой среде. Реализованный численный алгоритм не зависит от локального числа Маха. Он позволяет рассчитывать до-, транс-, сверх- и гиперзвуковые течения.

Геометрия и начальные условия

Рабочие колеса бироторной гидротурбины были спроектированы в программе автоматизированного проектирования КОМПАС 3D, вне программного комплекса FlowVision. Диаметр 1-го рабочего колеса 0,2 м, диаметр втулки 0,07 м, количество лопастей 3. Диаметр 2-го рабочего колеса 0,2 м, диаметр втулки 0,07 м, количество лопастей 3. Межтурбинное расстояние 0,01 м. 3D-модель бироторной турбины представлена на рисунке 1.

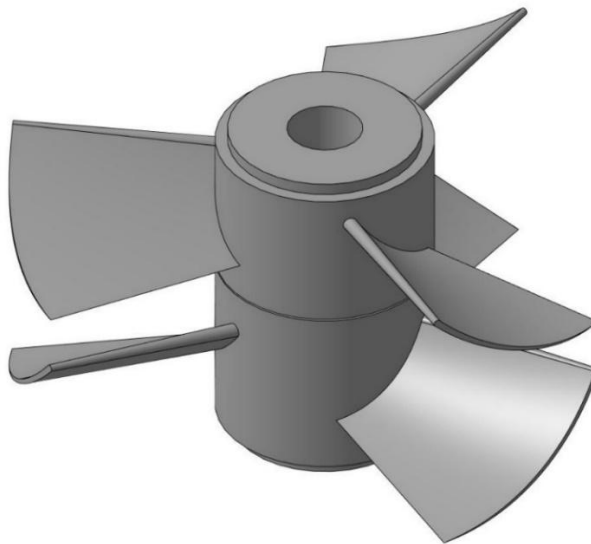


Рисунок 1 – 3D-модель бироторной гидротурбины

На следующем этапе методом вычитания 3D-модели бироторной гидротурбины из замкнутого объема была создана геометрия расчетной области, т.е. модель жидкости, заполнившей бироторную турбину, которая представлена на рисунке 2. Данная геометрия импортирована во FlowVision, и в ней создается проект.

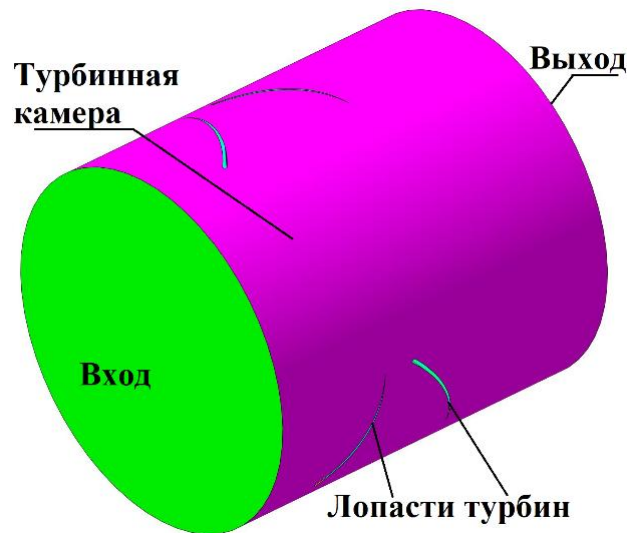


Рисунок 2 – Геометрия расчетной области бироторной гидротурбины

Далее в древе проекта создается физическая модель, где устанавливаются вещество, опорные величины и решаемые уравнения. Затем расставляются граничные условия, создается расчетная сетка, слои и результаты.

На рисунке 3 представлены все результаты визуализационных слоев моделирования бироторной гидротурбины. Как видно из рисунка, были получены данные о распределении скорости в бироторной гидротурбине, направления векторов и линий тока, а также распределения давлений на лопастях гидротурбин. Эти результаты получены в режиме стационарной гидродинамики, что означает, система рассматривается как статическая или неподвижная в течение всего расчетного периода, изменения параметров жидкости, такие как скорость и давление, могут быть учтены только как результаты стационарного расчета. В дальнейшем наши исследования будут направлены на решения поставленных выше задач в переходных режимах.

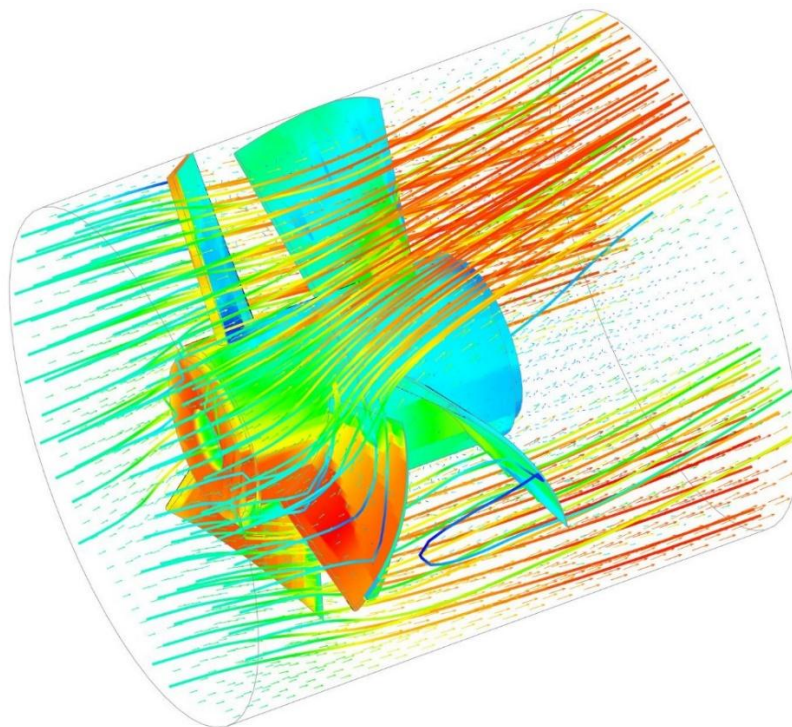


Рисунок 3– Результаты моделирования

Заключение

В заключение можно сделать следующие выводы: в контексте мировой тенденции к развитию ВИЭ использование гидроэнергии малых рек, особенно в странах с гористым рельефом, становится важным и перспективным источником электроэнергии. Малые и микро-ГЭС обладают рядом преимуществ, включая высокую эффективность, способность к снижению выбросов углерода и возможность решения проблем электроснабжения. Для эффективного использования потенциала малых гидроресурсов необходимо проводить исследования в данной области с целью улучшения и оптимизации параметров. Для эффективного использования потенциала малых ГЭС необходимо проводить исследования, такие как гидродинамический анализ бироторных микро-ГЭС с использованием современных методов, включая вычислительную гидродинамику. Эти исследования помогут оптимизировать производительность гидротурбин и улучшить их параметры. Эти данные представляют собой важный базис для дальнейших исследований и оптимизации. Дальнейшие научные исследования и оптимизация технологии позволят улучшить производительность и надежность бироторных микро-ГЭС и внедрить их в практику для обеспечения устойчивого энергоснабжения.

Литература:

1. IRENA (2022), Прогноз преобразования мировой энергетической системы 2022 года: стратегия по ограничению глобального потепления 1,5°C, Международное агентство по возобновляемым источникам энергии, Абу-Даби.
2. LIU, D., LIU, H., WANG, X., and Kremere, E., eds. (2019). World Small Hydropower Development Report 2019. United Nations Industrial Development Organization; International Center on Small Hydro Power. Available from www.smallhydropower.org.
3. Малые ГЭС в Кыргызской Республике: Оценка потенциала и сложностей развития/ Всемирный банк, Глобальная практика в области энергетики и добывающих отраслей/ Итоговый отчет/ Июнь 2017 г.
4. Kumar, A., T. et al. 2011: Hydropower. In IPCC Special Report on Renewable Energy Sources and Climate Change Mitigation/Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.
5. <http://www.nesk.kg/ru/9-o-kompanii/> (Дата обращения: 24.08.23)
6. <https://nehk.energo.kg/> (Дата обращения: 24.08.23)
7. Медеров Т.Т. Бироторная микрогидроэлектростанция / Т.Т. Медеров, Р.У. Ураимов // Материалы международной научно-практической конференции на тему «Качественное образование, передовая наука, зеленая экономика – будущее планеты». – Алматы, 2014. – С. 279 – 282.
8. Медеров Т.Т. Особенности работы бироторных микрогидроэлектростанций / Т.Т. Медеров, А. Дж. Обозов // Известия Кыргызского государственного технического университета им. И. Раззакова, №31. – Бишкек, 2014. – С. 205–208.
9. Викторов, Г.В. Гидродинамическая теория решеток. – М.: «Высшая школа», 1969. – 368 с.
10. Лойцянский, Л.Г. Механика жидкости и газа/Учеб. для вузов. – Изд. 7-е / Л.Г. Лойцянский. – М.: Дрофа, 2003. – 840 с.

УДК 004.75

*В. В. Гайдамако,**Кыргызско-Российский (Славянский) университет имени Б.Ельцина, Бишкек**Институт машиноведения, автоматизации и геомеханики Национальной академии наук, Бишкек*

О МОДЕЛИРОВАНИИ ДИНАМИЧЕСКОЙ МИГРАЦИИ ВИРТУАЛЬНЫХ МАШИН В ПЛАТФОРМЕ SIMGRID

В статье рассматриваются методы динамической миграции виртуальных машин в облачных системах и реализация живой миграции в расширении фреймворка моделирования Simgrid. Рассматриваются методы динамической миграции с пред- и пост-копированием, влияние скорости изменения страниц оперативной памяти на время миграции. Приведены примеры динамической миграции в Simgrid-модели облачной системы.

Ключевые слова: облачные вычисления, центр обработки данных (ЦОД), виртуальная машина (ВМ), гипервизор, миграция ВМ, динамическая миграция, моделирование, Simgrid.

Введение

Облачные вычисления предоставляют пользователю доступ к ресурсам через сеть по запросу в требуемом объеме независимо от местоположения. Ресурсами могут быть не только приложения, но и вычислительные мощности, системы хранения, сети, то есть целая инфраструктура. Основой облачных приложений является технология виртуализации, виртуальные машины (ВМ) с необходимым для обслуживания запроса программным обеспечением размещаются на физических машинах (серверах, хостах). При планировании этого размещения должны учитываться разные факторы – требования соглашения об уровне услуг (Service Level Agreement), показатели производительности, затраты провайдера облака на обслуживание запроса. Потребление электроэнергии и углеродный след [1] должны быть минимальными. В некоторых случаях требуется переместить виртуальные серверы на другую физическую машину – для экономии электроэнергии, балансировки нагрузки, обеспечения доступности сервиса, причем перемещение может производиться как внутри, так и между центрами обработки данных (ЦОД). Так как реальные облачные системы сложны, велики и часто недоступны для натурных экспериментов, в исследовательском сообществе используются фреймворки (платформы) моделирования распределенных и облачных систем. Актуальность, достоверность, гибкость, возможность создания крупномасштабных моделей с использованием открытого фреймворка SimGrid доказана многочисленными публикациями [2], поэтому именно он был выбран для моделирования облачных информационно-измерительных систем [3], включающих, помимо обычных серверов, физические и виртуальные датчики.

Виртуализация и гипервизоры

Виртуализация – процесс, позволяющий нескольким изолированным средам с собственной операционной системой совместно использовать аппаратные ресурсы компьютера. Каждая из таких виртуализированных сред работает с выделенными ей ресурсами – оперативной памятью, вычислительной мощностью, хранилищем, внешними устройствами. Благодаря виртуализации можно переключаться между разными операционными системами на одном сервере без перезагрузки [4]. Виртуализация может быть аппаратной или программной.

Виртуальная машина (ВМ) – это программно-определяемый компьютер, изолированно работающий на физическом компьютере, с установленной отдельно операционной системой и выделенными ресурсами. Физический компьютер, на котором создается ВМ, называется хост-машиной, а виртуальная машина – гостевой. На одной

физической машине может работать несколько виртуальных машин. Виртуальные машины получают доступ к аппаратуре хоста через гипервизор,

Гипервизор, или монитор виртуальных машин – программа, или аппаратная схема, позволяющая одновременно параллельно запускать несколько виртуальных сред (операционных систем) на одном хост-компьютере. Гипервизор обеспечивает изоляцию операционных систем друг от друга, защиту и безопасность, разделение ресурсов между различными запущенными ОС и управление ресурсами [4].

Аппаратная виртуализация – виртуализация с поддержкой специальной процессорной архитектуры. В отличие от программной виртуализации с помощью данной техники возможно использование изолированных гостевых операционных систем, управляемых гипервизором напрямую [5]. Аппаратная виртуализация обеспечивает производительность, сравнимую с производительностью неvirtуализованной машины, что дает виртуализации возможность практического использования и влечет ее широкое распространение. Наиболее распространены технологии виртуализации Intel-VT и AMD-V.

Взаимодействия между гостевыми ОС, работающими на одной хост-машине, происходит так, как если бы они выполнялись на разных физических машинах, то есть по сети. Существует два основных типа гипервизора – тип 1 (X), автономный гипервизор, аппаратный гипервизор и тип 2 (V) – гипервизор на основе ОС.

Гипервизоры первого типа устанавливаются непосредственно на аппаратном обеспечении компьютера (bare metal), имеют свои драйверы устройств, свой планировщик и не зависят от хостовой ОС. Обеспечивают высокую производительность, именно они используются в ЦОД и корпоративных приложениях. К этому типу относятся гипервизоры Xen, VMware ESX, Citrix XenServer. KVM использует гипервизор первого типа для хостинга нескольких виртуальных машин в операционной системе Linux.

Гипервизоры второго типа устанавливаются на хостовой ОС. Гостевой код может выполняться прямо на физическом процессоре, но доступ к устройствам ввода-вывода компьютера из гостевой ОС осуществляется через второй компонент, обычный процесс основной ОС – монитор уровня пользователя [4,5]. К таким гипервизорам относятся Microsoft Virtual PC, VMware Workstation, QEMU, Oracle VirtualBox.

Существуют также гибридный (тип 1+). Гибридный гипервизор состоит из двух частей: из тонкого гипервизора, контролирующего процессор и память, а также специальной служебной ОС, работающей под его управлением. Через служебную ОС гостевые ОС получают доступ к физическому оборудованию. Это гипервизоры Microsoft Virtual Server, Sun Logical Domain, OVirt [5].

Контейнерная виртуализация, контейнеризация, виртуализация на уровне операционной системы – метод виртуализации, при котором ядро операционной системы поддерживает несколько изолированных экземпляров пространства пользователя вместо одного. Эти экземпляры с точки зрения выполняемых в них процессов идентичны отдельному экземпляру операционной системы. Ядро обеспечивает полную изолированность контейнеров, поэтому программы из разных контейнеров не могут воздействовать друг на друга [6].

Контейнеры используют ядро хостовой ОС, в контейнере можно запустить ОС только с тем же ядром, что у хостовой ОС, поэтому не нужны дополнительные ресурсы на эмуляцию оборудования и полноценную ОС. Пример – LXD Linux контейнеры, контейнеры Docker [7].

Миграция виртуальных машин в облачной среде

В работе облака может возникнуть необходимость переместить ВМ на другой сервер, иногда в другой ЦОД. Миграция ВМ может понадобиться при балансировке нагрузки, для обеспечения выполнения условий соглашения об уровне обслуживания, в случаях, когда серверу не хватает вычислительной мощности, для экономии

электроэнергии и уменьшения углеродного следа, а также в случае сбоев в работе оборудования. Для экономии электроэнергии ВМ с недогруженных серверов перемещаются на другой сервер, а освободившийся сервер переводится в спящий режим.

Для осуществления миграции ВМ можно или полностью остановить, или приостановить – то есть остановить с сохранением состояния всех работающих приложений, переместить на другой сервер файл ВМ и запустить ее уже на новом сервере. В этом случае ВМ остается недоступной на все время перемещения. Это off-line или холодная миграция.

Динамическая, живая, или онлайн-миграция (Live Virtual Machine Migration) – метод перемещения ВМ, то есть ОС и связанных с ней приложений с одной физической машины на другую без прекращения работы запущенных приложений, то есть приложение остается доступным во время миграции. При осуществлении живой миграции ВМ тоже приостанавливается, но время останова и недоступности значительно меньше.

Для оценки процесса миграции используются следующие показатели:

- общее время миграции – общее время, необходимое для миграции виртуальной машины с ее хост-машины на целевую машину (total migration time);
- время освобождения сервера-источника (eviction time);
- время простоя (downtime). При выполнении динамической миграции виртуальная машина все же будет какое-то время недоступна (время простоя), так как перемещается ВМ в состоянии приостановки (suspended). Это время должно быть минимальным [8].

Производительность миграции (Migration performance) – это общее время миграции и общее время простоя. Факторы, влияющие на производительность миграции, – ширина полосы пропускания канала, объем передаваемых данных, затраты системных ресурсов на миграцию (migration overhead). Также может рассматриваться изменение производительности приложений, выполняющихся на мигрировавшей машине.

Для осуществления живой миграции необходимо переместить: состояние устройств, сетевых подключений, ОС и всех приложений, страницы оперативной памяти (Memory) и, если нужно, данные на внешнем носителе (Storage). Данные на внешнем носителе не переносятся, если они доступны с целевого хоста, на который перемещается ВМ, то есть исходный и целевой хост находятся в кластере, используется сетевое хранилище – Network Access Storage или Storage Area Network (NAS, SAN), к которому оба хоста имеют доступ [9].

Существуют различные подходы к выполнению живой миграции – с предварительным копированием страниц памяти, с пост-копированием и различные их комбинации.

Подход с предварительным копированием (Pre-copy)

Первыми перемещаются страницы памяти, затем все остальное, этапы миграции:

1. Выбор перемещаемой ВМ и целевого хоста.
2. Предварительное копирование состояния памяти ВМ на целевой хост, ВМ не прекращает работы на исходном хосте. Так как ВМ не прекращает работу, во время перемещения некоторые страницы могут быть изменены. Измененные («грязные») страницы также перемещаются, но во время перемещения они также могут быть изменены. Поэтому этот шаг итеративно повторяется до тех пор, пока объем страниц для перемещения не станет достаточно небольшим.
3. Приостановка работы виртуальной машины и пересылка состояния (не связанного с памятью). Гипервизор старается минимизировать это время.

4. Пересылка оставшегося состояния памяти (оставшееся небольшое количество страниц, измененных во время выполнения стадии 2 и удаление зависимостей на исходном хосте) [10, 11].
5. Передача управления ВМ на целевом хосте и возобновление работы.

«Достаточно небольшой» объем для передачи на шаге 2 означает такой объем, который может быть передан за заданное время приостановки ВМ. Этот объем зависит от пропускной способности сети, а также от интенсивности обновления оперативной памяти при работе приложений.

Подход с пост-копированием (Post-copy)

В подходе с пост-копированием сначала переносятся состояния виртуальной машины, не относящиеся к памяти (состояние ЦП), а затем передаются страницы памяти виртуальной машины. Главное преимущество этого подхода заключается в том, что каждая страница памяти передается не более одного раза, что позволяет избежать дублирующих накладных расходов на передачу при предварительном копировании.

Стратегия пост-копирования может обеспечить выигрыш за счет сокращения общего времени миграции при сохранении жизнеспособности виртуальной машины во время миграции, но время простоя при таком подходе будет больше, чем при стратегии с предварительным копированием [10, 11].

Влияние скорости обновления оперативной памяти ВМ на время миграции

Гипервизор пытается минимизировать время останова ВМ, чтобы оно было незаметно для пользователей и приложений. Например, максимальное время простоя для гипервизора Qemu/KVM (де-факто ВМ менеджер Linux) по умолчанию составляет 30 мс [12]. На этапе переноса страниц памяти при предварительном копировании гипервизор Qemu/KVM итеративно копирует обновленные страницы памяти в место назначения, пока размер оставшихся страниц памяти не станет меньше порогового значения, которое позволит достичь времени простоя в 30 мс. Следовательно время миграции в основном зависит от скорости обновления памяти ВМ и сетевой скорости трафика миграции. ВМ, интенсивно обновляющая страницы памяти, потребует более длительного времени миграции, и в худшем случае (т.е. скорость обновления памяти выше пропускной способности сети) миграция не завершится (т. е. не сойдется в техническом плане).

В [11] показаны результаты экспериментов, исследующих скорость обновления памяти в облачных приложениях (веб-сервер и сервер базы данных) в зависимости от рабочих нагрузок. Было показано, что скорость обновления памяти может быть весьма значительной по сравнению с пропускной способностью сети. Предоставление «наивной» модели, которая просто получает стоимость живой миграции путем деления размера памяти виртуальной машины на доступную пропускную способность сети, является неверным, поскольку скорость обновления памяти определяет продолжительность этапа 2 (перенос страниц памяти ВМ) алгоритма предварительного копирования. Критически важно учитывать влияние обновлений памяти, чтобы точно оценить время миграции и генерируемый сетевой трафик.

Кроме того, время выполнения первых двух этапов зависит от активности других виртуальных машин и рабочих нагрузок, запущенных в системе:

- сеть – кроме трафика живой миграции, по сети передаются другие данные. Это может быть рабочая нагрузка перемещаемой ВМ или других ВМ, работающих на том же сервере или в том же сегменте сети, трафик от других живых миграций. Всегда надо учитывать, что доступная для живой миграции пропускная способность сети динамически изменяется;

- когда несколько ВМ размещены на одном хосте, они конкурируют друг с другом за получение ресурсов ЦП. Это соперничество может привести к снижению производительности каждой ВМ. Скорость изменения памяти (то есть появление «грязных» страниц, которые перемещаются несколько раз во время живой миграции) виртуальной машины динамически изменяется из-за активности других ВМ, размещенных на том же сервере.

«Наивная» модель вычисляет время миграции как соотношение объема перемещаемых данных и пропускной способности канала миграции, и это время без учета динамических изменений будет вычислено неправильно. Следовательно такой способ вычисления времени миграции не может использоваться при моделировании процесса живой миграции, так как приведет к ошибочным результатам моделирования, далеким от реального поведения.

Фреймворк моделирования должен учитывать конкуренцию в виртуализированных и не виртуализированных системах: если виртуальные машины размещены на одной физической машине, система должна вычислять правильное время ЦП для каждой виртуальной машины. Если сетевое соединение используется совместно с несколькими передачами данных, включая миграции, системе необходимо назначать правильную полосу пропускания для каждой передачи [11].

Представление виртуальных машин в Simgrid

Рассмотрим интерфейс программирования S4U (SimGrid for you), предпочтительный, а с 2023 года главный интерфейс программирования [13]. Основными языками S4U являются C, C++. Модель Simgrid состоит из Actov, выполняющих функции, заданные пользователем. Actor выполняет активности (Activity), представляющие вычисления, коммуникации, использование диска так, что они учитываются в модели. Активности используют ресурсы – хосты (физические машины – Host), связи (линия связи и сетевые устройства, Link), диски (Disk). Simgrid вычисляет время выполнения активности и управляет запуском акторов. Коммуникации осуществляются через почтовые ящики (MailBox). Акторы могут выполняться не только на физических, но и на виртуальных машинах (VirtualMachine). Виртуальная машина создается на хосте, при этом можно указать количество процессоров и размер оперативной памяти.

```
VirtualMachine *simgrid::s4u::Host::create_vm(const std::string &name,
                                              int core_amount)
VirtualMachine *simgrid::s4u::Host::create_vm(const std::string &name,
                                              int core_amount, size_t ramsize)
```

Можно получить информацию о работающей ВМ, определить и изменить состояние, то есть стартовать start() и остановить shutdown(), приостановить suspend() и запустить resume(), так же, как это делается в реальном мире.

Динамическая миграция виртуальных машин в Simgrid

Модель живой миграции, реализованная в SimGrid, основана на подходе предварительного копирования. В [11] описано расширение гипервизора Qemu, созданное для использования фреймворком Simgrid при управлении работой ВМ в модели. Для вычисления времени миграции используются параметры:

- R – размер памяти виртуальной машины;
- α (байт/флос) – интенсивность обновления памяти виртуальной машины, обозначает, сколько страниц памяти помечаются как «грязные» при выполнении одной операции ЦП (в мире моделирования). Но для конечного пользователя указание этого параметра может быть затруднительно, поэтому его можно указать

через объем изменяемых данных (т. е. Мб/с) при загрузке ЦП на 100% (преобразование затем вычисляется внутренне, связь показана в [11]).

Расширение Qemu позволяет получать интенсивность обновления памяти конкретной виртуальной машины, следовательно можно определить интенсивность обновления памяти для изучаемой рабочей нагрузки. Во время живой миграции модель повторяет передачу данных до конца этапа 5 (передача оставшейся небольшой порции страниц памяти).

Размер данных i -й передачи определяется как X_i байт. На этапе 1 алгоритм предварительного копирования отправляет все страницы памяти, т. е. $X_1 = R$. На этапе 2 алгоритм отправляет страницы памяти, обновленные во время предыдущей передачи данных. В [11] показано, что скорость обновления памяти пропорциональна уровню загрузки ЦП, то есть размер обновленных страниц памяти во время передачи данных пропорционален общему количеству долей ЦП, назначенных виртуальной машине в течение этого периода. Таким образом, размер передачи данных следующей фазы можно определить как $X_{i+1} = \min(\alpha S_i; R_0)$, где S_i (плавающие операции, flops) – это общее количество долей ЦП, назначенных виртуальной машине во время i -й передачи данных, а R_0 – размер памяти рабочего набора (набор страниц памяти, используемых во время выполнения операции), используемого приложениями на виртуальной машине. Размер обновленных страниц памяти никогда не превышает размер рабочего набора [11].

Фреймворк моделирования формулирует набор ограничений для определения продолжительности каждой передачи данных, исходя из скорости сети, задержки и других задач связи, использующих то же сетевое соединение. Если задана максимальная пропускная способность трафика миграции, модель контролирует скорость передачи данных миграции так, чтобы не превышать эту пропускную способность. Этот параметр соответствует migrate-set-speed API libvirt [14].

При передаче данных миграции модель оценивает доступную пропускную способность для текущей живой миграции так же, как это делает гипервизор в реальном мире. Текущая пропускная способность трафика миграции вычисляется путем деления размера отправляемых данных на время, необходимое для отправки данных. Это расчетное значение используется для определения приемлемого размера оставшихся данных при миграции на 2 этапе (передача страниц памяти). Если максимальное время простоя составляет 30 мс (т. е. значение по умолчанию Qemu), а пропускная способность трафика миграции оценивается в 1 Гбит/с, на этапе 3 должно оставаться менее 3750 Кб (измененные страницы). Механизм миграции повторяет итерацию этапа 2 (передачу страниц памяти), пока это условие не будет выполнено. На этапе 4 модель [11] создает задачу связи размером X_n байт для передачи оставшихся страниц памяти. После завершения этой последней задачи система переключает хост выполнения виртуальной машины на пункт назначения.

Расширение (плагин) live_migration

Для моделирования живой миграции в Simgrid используется плагин live_migration, его необходимо инициализировать с помощью функции void sg_vm_live_migration_plugin_init(). Плагин позволяет установить и получить интенсивность изменения памяти – функции sg_vm_set_dirty_page_intensity, sg_vm_get_dirty_page_intensity; отслеживать изменение страниц – функции sg_vm_dirty_page_tracking_init, старт отслеживания sg_vm_start_dirty_page_tracking, и останов sg_vm_stop_dirty_page_tracking; устанавливать и получать размер рабочего набора – функции sg_vm_set_working_set_memory, sg_vm_get_working_set_memory; устанавливать и получать скорость миграции – sg_vm_set_migration_speed, sg_vm_get_migration_speed(const sg_vm_t vm); получать время останова sg_vm_get_max_downtime(const sg_vm_t vm).

Виртуальная машина для миграции создается с помощью функции `sg_vm_t sg_vm_create_migratable(sg_host_t pm, const char* name, int coreAmount, int ramsize, int mig_netspeed, int dp_intensity)`.

Параметры – физическая машина (хост), на которой создается ВМ, имя ВМ, количество процессорных ядер, размер оперативной памяти (ОП), скорость сети, интенсивность обновления ОП.

Миграция осуществляется функцией `sg_vm_migrate(sg_vm_t vm, sg_host_t dst_pm)`, параметрами являются имя ВМ и хост, на который ее перемещают.

Рассмотрим пример использования плагина на платформе, показанной на рисунке 1.

При запуске модели в ЦОД включается только один сервер – контроллер ЦОД `cb1`, остальные серверы запускаются в зоне физических машин. Машины внутри зоны физических машин связаны между собой сетью с пропускной способностью 1Гб/с, как и контроллер ЦОД с зоной физических машин. Зона пользователей связывается с ЦОД по медленной линии с пропускной способностью 10Мб/с.

В примере на физической машине создаются ВМ и мигрируют на разные физические машины, сначала из зоны физических машин в том же ЦОД, потом на машину из зоны пользователей, отслеживается время простоя и время миграции.

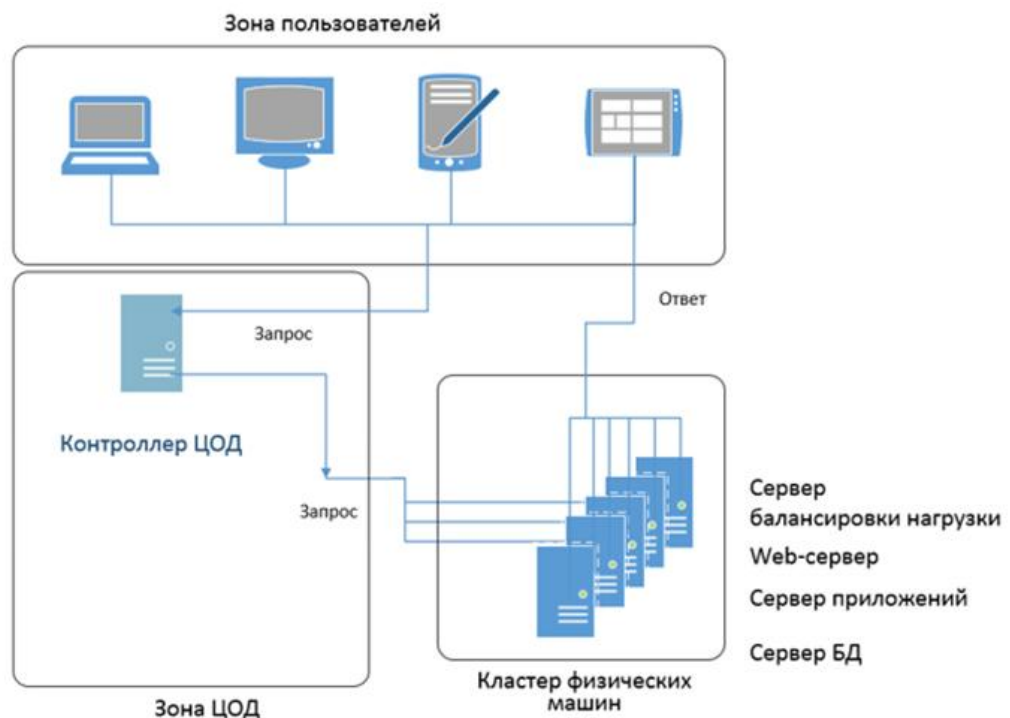


Рисунок 1 – Центр обработки данных с зоной физических машин и зона пользователей [15]

Функция миграции крайне проста [13]:

```
static void vm_migrate(sg4::VirtualMachine* vm, sg4::Host* dst_pm)
{
    const sg4::Host* src_pm = vm->get_pm();
    double mig_sta = sg4::Engine::get_clock();
    sg_vm_migrate(vm, dst_pm);
    double mig_end = sg4::Engine::get_clock();
    XBT_INFO("%s migrated: %s->%s in %g s, speed %g downtime %g",
            vm->get_cname(), src_pm->get_cname(),
            dst_pm->get_cname(), mig_end - mig_sta,
            sg_vm_get_migration_speed(vm), sg_vm_get_max_downtime(vm))
}
```

В таблице 1 представлено время миграции с одной физической машины на другую внутри ЦОД. Мигрируют по очереди: одна ВМ с оперативной памятью 1 Гб (строка 1) и одна с памятью 100 Мб (2), две ВМ с одной машины на другую (3), две ВМ с одной машины на две разные машины (4,5). Увеличение времени миграции в случаях 3-5 обусловлено одновременной передачей страниц памяти по одной линии связи (от одной физической машины).

Таблица 1 – Время миграции ВМ в сети 1 Гб/с

№	Источник			Приемник		Сеть	Время простоя, с	Время миграции, с
	ФМ	Кол-во ВМ	ОП	ФМ	Кол-во ВМ			
1	pm1-0.pm	1	1Гб	pm1-1.pm	1	1Гб/с	0.03	1.08716
2	pm1-0.pm	1	100Мб	pm1-1.pm	1	1Гб/с	0.03	0.112931
3	pm1-0.pm	2	1Гб	pm1-1.pm	2	1Гб/с	0.03	2.16963
4	pm1-0.pm	2	1Гб	pm1-1.pm	1	1Гб/с	0.03	2.16963
5				pm1-2.pm	1	1Гб/с	0.03	2.16963

В таблице 2 представлено время миграции с одной физической машины внутри ЦОД на другую машину из зоны пользователей с медленной линией связи.

Таблица 2 – Время миграции ВМ в сети 10 Мб/с

№	Источник			Приемник		Сеть	Время простоя, с	Время миграции, с
	ФМ	Кол-во ВМ	ОП	ФМ	Кол-во ВМ			
1	pm1-1.pm	1	1 Гб	u-0.uz	1	10 Мб/с	0.03	108.641
2	pm1-1.pm	1	100Мб	u-0.uz	1	10 Мб/с	0.03	11.2183
3	pm1-1.pm	2	1 Гб	u-0.uz	2	10 Мб/с	0.03	216.888
4	pm1-1.pm	2	1 Гб	u-0.uz	1	10 Мб/с	0.03	216.888
5				u-0.uz	1	10 Мб/с	0.03	216.888

Результат очевиден, так как на ВМ нет акторов, они не изменяют страницы, поэтому во время миграции просто один раз копируется вся память. Время простоя (30 мс), или время недоступности ВМ, не меняется, так как нет измененных страниц.

Заключение

В облачных вычислениях миграция виртуальных машин между серверами часто становится необходимой для адаптации к изменяющимся условиям. Так как время простоя и недоступности виртуальной машины должно быть минимальным, осуществляется динамическая (живая) миграция с пред- или пост-копированием страниц оперативной памяти. Для исследования облачных систем из-за дороговизны или недоступности натурных экспериментов используются имитационные модели облачных систем, в которых должны быть представлены и физические, и виртуальные ресурсы, а следовательно, и миграция виртуальных машин. В статье рассмотрено расширение для моделирования живой миграции фреймворка Simgrid, использующее подход с предварительным копированием, описан расчет времени миграции с учетом скорости изменения оперативной памяти и конкурирующего сетевого трафика, приведены примеры использования плагина.

Литература

1. Что такое углеродный след? URL: <https://www.microsoft.com/ru-ru/sustainability/learn/reduce-carbon-footprint> (дата обращения 25.10.2024)
2. SimGrid Publications. URL: <https://simgrid.org/publications.html> (дата обращения 25.10.2024)
3. Гайдамако В.В. Инфраструктура Sensor-Cloud – облачные информационно-измерительные системы. // Проблемы автоматизации и управления – 2018. – №2 (35). – С. 109–118.

4. Что такое виртуализация? URL: <https://aws.amazon.com/ru/what-is/virtualization/> (дата обращения 25.10.2024)
5. Oracle VirtualBox. URL: <https://www.virtualbox.org/> (дата обращения 25.10.2024)
6. <https://ru.wikipedia.org/wiki/Контейнеризация> (дата обращения 25.10.2024)
7. Docker: Accelerated Container Application Development. URL: <https://www.docker.com/> (дата обращения 25.10.2024)
8. Soni G, Kalra M (2013) Comparative Study of Live Virtual Machine Migration Techniques in Cloud. *Int J Comput Appl* 84(14):19–25
9. Костенко В.А., Чупахин А.А. Схемы организации «живой» миграции в центрах обработки данных // Программирование. – 2020. – №5. URL: <https://www.elibrary.ru/item.asp?id=43182341>
10. Алексанков С. М. Модель процесса динамической миграции с копированием данных после остановки виртуальных машин // Приборостроение. – 2016. – №5. URL: <https://cyberleninka.ru/article/n/model-protsessa-dinamicheskoy-migratsii-s-kopirovaniem-dannyh-posle-ostanovki-virtualnyh-mashin> (дата обращения: 06.11.2024).
11. Pouilloux, L., Hirofuchi, T., Lebre. SimGrid VM: Virtual Machine Support for a Simulation Framework of Distributed Systems. *IEEE Transactions on Cloud Computing*, 2018. (1) WWW doi:10.1109/TCC.2015.2481422
12. QEMU. A generic and open source machine emulator and virtualizer. URL: <https://www.qemu.org/> (дата обращения 25.10.2024)
13. SimGrid Tutorials. URL: https://simgrid.org/doc/latest/intro_concepts.html (дата обращения 25.10.2024)
14. libvirt: The virtualization api. URL: <http://libvirt.org> (дата обращения 25.10.2024)
15. Гайдамако, В.В. 2023. Моделирование сервера в инфраструктуре облачного центра обработки данных на базе платформы SIMGRID // Проблемы автоматизации и управления. Вып. 3 (декабрь):81–92. URL: <https://pau.imash.kg/index.php/pau/article/view/443>

УДК 681.3

Лыченко Н.М.*Кыргызско-Российский славянский университет, Бишкек, Кыргызстан**E-mail: nlychenko@mail.ru*

МОДЕЛИ ПРОГНОЗА КАЧЕСТВА ВОЗДУХА НА ОСНОВЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ С УЧЕТОМ ВРЕМЕННЫХ И ПРОСТРАНСТВЕННЫХ ЗАВИСИМОСТЕЙ

Представлен аналитический обзор наиболее используемых моделей прогноза процессов загрязнения атмосферного воздуха, характеризующихся сложными и нестационарными временными зависимостями. Пространственное положение, в дополнение к динамическим изменениям с течением времени, также оказывает влияние на качество воздуха. Относительно новый инструмент для решения задач прогноза качества воздуха с учетом не только временных, но и пространственных зависимостей - графовые нейронные сети. Основная характеристика этого типа сетей – использование информации, полученной в результате динамического взаимодействия между соседними географическими точками (городами, районами, улицами), которые взвешиваются в зависимости от расстояния между ними.

Ключевые слова: модели прогноза качества воздуха, временные ряды, пространственно-временные зависимости, машинное обучение, графовые нейронные сети.

Введение. Загрязнение воздуха является актуальной проблемой, которая представляет значительную угрозу здоровью человека и окружающей среде, приводит к дисбалансу в экосистеме. Термин «качество воздуха» относится к количеству и составу различных загрязняющих веществ, присутствующих в атмосфере, включая угарный газ, диоксид серы, диоксид азота, твердые частицы PM_{2.5}, PM₁₀ и другие. Индекс качества воздуха (AQI) рассчитывается на основе уровня концентрации загрязняющих веществ и степени их воздействия на здоровье. Для прогнозирования качества воздуха используется профессиональное оборудование и технологии мониторинга для сбора и анализа атмосферных параметров (таких как температура, влажность, скорость ветра, атмосферное давление, осадки и др.), географического положения, индекса качества воздуха и других данных, чтобы предсказывать и предупреждать о качестве воздуха в будущем. Точный прогноз качества воздуха имеет большое значение для принятия мер по минимизации воздействия опасных загрязняющих веществ на население и поддержания баланса экосистемы. Хотя для прогнозирования качества воздуха было разработано множество методов, прогнозирование качества воздуха остается непростой задачей, поскольку процессы, протекающие в атмосферном воздухе, характеризуются сложными и нестационарными временными и пространственными зависимостями [1].

Временную зависимость характеризуют периодичность и тренд. Периодичность — это появление схожих паттернов или регулярных изменений в течение определенного времени. Для индекса качества воздуха г. Бишкек, например, характерна явно выраженная периодичность сезонная (определяемая холодным и теплым временем года), явно выраженная суточная и – слабо выраженная недельная периодичность [2].

Пространственное положение, в дополнение к динамическим изменениям с течением времени, также оказывает влияние на качество воздуха. Качество воздуха в одной географической точке (разные города, разные районы, разные улицы) будет зависеть от атмосферных условий в другой географической точке. В сочетании с динамическими процессами в атмосфере, учет пространственных зависимостей в процессах загрязнения представляет собой очень сложную задачу.

Существует множество методов оценки качества воздуха, которые учитывают корреляцию со временем, включая авторегрессионную интегрированную модель скользящего среднего, модель рекуррентных нейронных сетей и модель долгосрочной и краткосрочной памяти (LSTM). Как правило, эти подходы учитывают динамические изменения качества воздуха, но не учитывают пространственные отношения между качеством воздуха и географическим положением.

В последние годы в прогнозировании качества воздуха стали использоваться графовые нейронные сети (Graph Neural Networks, GNN) [3-5]. Основная характеристика этого типа сетей – это использование информации, полученной в результате динамического взаимодействия между соседними географическими точками (городами, районами, улицами), которые взвешиваются в зависимости от расстояния между ними.

В настоящей работе приведен аналитический обзор наиболее используемых моделей прогноза процессов загрязнения атмосферного воздуха, характеризующихся сложными и нестационарными временными и пространственными зависимостями. Особое внимание уделено описанию графовых нейронных сетей как инструменту для решения задач прогноза качества воздуха с учетом не только временных, но и пространственных зависимостей.

1. Машинное обучение и классические регрессионные алгоритмы. Методы машинного обучения являются наиболее распространенными методами прогнозирования качества воздуха. С начала XXI века в литературе можно найти сотни произведений, которые предлагают реализации различных моделей для достижения максимальной точности прогнозирования индекса AQI или прогноза концентрации загрязняющих веществ. В [1] было проведено подробное исследование по применению различных алгоритмов машинного обучения как для прогнозирования AQI так и уровня концентрации конкретных загрязняющих веществ, связанных с качеством воздуха..

1.1 Классические статистические алгоритмы. Регрессионный анализ используется для установления отношения между зависимой переменной и набором независимых переменных. На основе этого отношения и с использованием значений независимых переменных оценивается значение зависимой переменной. Далее рассмотрим классические алгоритмы для построения регрессии.

Множественная линейная регрессия (Multiple Linear Regression, MLR). Пусть y и $x_1, \dots, x_p$ будут зависимой переменной и независимыми переменными соответственно. Цель линейной регрессии — определить линейную функцию $f(x_1, \dots, x_p)$, которая минимизирует среднюю квадратичную ошибку, то есть:

$$\min\{(y - f(x_1, \dots, x_p))^2\}.$$

Примеры применения множественной линейной регрессии в задачах прогноза качества воздуха можно найти в литературе, например [6, 7].

Авто-регрессионное интегрированное скользящее среднее (Auto-Regressive Integrated Moving Average, ARIMA). Эта модель использует временные ряды для прогнозирования, то есть она делает предсказания на основе прошлых и настоящих данных. Модель ARIMA включает три компонента:

1) авто-регрессионный компонент - характеризует количество задержек, используемых в модели. Например, авто-регрессионное значение 3 устанавливает, что только три предыдущих значения будут использоваться для объяснения текущего значения;

2) интегрированный компонент - представляет собой степень дифференцирования, необходимую для преобразования временного ряда в стационарный ряд;

3) скользящее среднее - относится к количеству прошлых ошибок, необходимых для объяснения текущей ошибки. Снова, если значение скользящей средней равно 3, то только три предыдущих значения ошибок могут быть использованы для объяснения текущей ошибки. Когда ВР содержит сезонность, используется вариант ARIMA, называемый SARIMA.

Модель ARIMA является одной из репрезентативных статистических моделей для прогнозирования качества воздуха. Например, ARIMA использовалась для прогнозирования загрязняющих веществ и показала хорошие результаты [8, 9, 10]. Однако определение параметров в статистических методах сильно зависит от теоретических предположений и предварительных знаний о данных, что делает затруднительным соответствие нелинейным и нестационарным данным о качестве воздуха. Таким образом, статистические методы могут привести к смещению прогноза AQI, особенно для точного долгосрочного прогноза AQI [4].

1.2 Алгоритмы машинного обучения на основе регрессионного анализа

Метод опорных векторов для регрессии (Support vector regression, SVR). Методы опорных векторов в основном применяются в задачах классификации. Однако их также можно использовать для регрессии. В этом случае подход называется регрессией на основе опорных векторов. Пусть y и $x_1, \dots, x_p$ будут зависимой переменной и независимыми переменными соответственно. В основном, SVR работает следующим образом. Сначала нужно определить линейную регрессионную функцию, то есть гиперплоскость $h(x) = w_1x_1 + \dots + w_px_p + b$. Затем предполагается, что все данные будут находиться на расстоянии не более чем ε от гиперплоскости. Если отклонение некоторых точек превышает это значение, можно ввести свободные переменные $\xi, \xi' \geq 0$ для работы с ними. Конечная цель — найти минимум функции:

$$\min \left(\frac{1}{2} \|w\|^2 + C \sum_{i=1}^N (\xi + \xi') \right)$$

при ограничениях:

$$y - wx_i - b \leq \varepsilon + \xi$$

$$wx_i + b - y_i \leq \varepsilon + \xi'.$$

В случае нелинейных функций SVR использует ядерные функции для преобразования данных в пространство более высокой размерности, чтобы разработать линейную регрессионную трансформацию

Деревья решений (Decision trees, DT). Цель этого алгоритма — разработать модель для прогнозирования количественной переменной на основе набора независимых переменных. Алгоритм основан на рекурсивном делении. Деревья состоят из узлов решений и листьев. Регрессия DT обычно строится с учетом уменьшения стандартного отклонения для определения того, как разделить узел на две или более ветки. Корневой узел — это первый узел решения, который делится на основе наиболее релевантной независимой переменной. Узлы снова делятся, принимая во внимание переменную с наименьшей суммой квадратов оценки ошибок (SSE) в качестве узла решения. Набор данных делится на основе значений выбранной переменной. Процесс завершается, когда выполнено заранее установленное условие остановки. Последние узлы известны как

листовые узлы и предоставляют предсказание зависимой переменной. Это значение соответствует среднему значениям, связанным с листьями. Рисунок 3а [11] показывает графическое представление общей структуры стандартного DT.

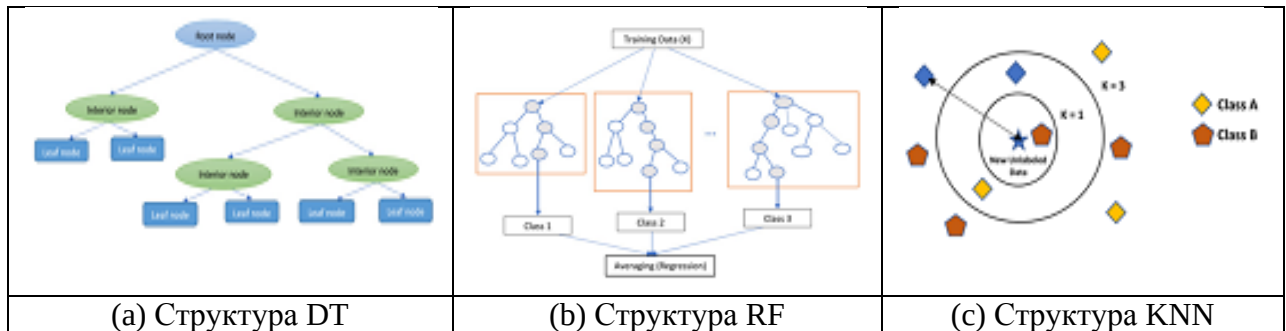


Рисунок 3- Графическое представление алгоритмов на основе регрессионного анализа.

Случайный Лес (Random Forest, RF). Случайный Лес основан на генерации нескольких деревьев решений. Прогноз будет средним значением прогнозов, предоставленных различными деревьями. Для построения каждого дерева решений выбирается выборка данных из обучающего набора данных. Остальные данные будут использованы для оценки ошибки дерева решений. Подмножество независимых переменных, которое может быть использовано для разделения каждого узла, выбирается случайным образом. Экстремально рандомизированные деревья (Extremely randomised trees, ERT) — это слегка модифицированный алгоритм случайного леса. Рисунок 3b [11] показывает графическое представление общей структуры регрессора RF.

Регрессия по k-ближайшим соседям (K-nearest neighbours regression, KNN). Алгоритм k -ближайших соседей часто применяется к задачам классификации, хотя его также можно применить и к задачам регрессии. Идея этого алгоритма проста. Учитывая расстояние (евклидово расстояние, например) и значение k , алгоритм вычисляет расстояние между точкой данных и точками обучающего набора данных для выбора k ближайших и устанавливает их среднее значение в качестве прогноза. Улучшением этого алгоритма является алгоритм, известный как взвешенные k -ближайшие соседи (WKNN). В этом случае прогноз учитывает взвешенное арифметическое среднее для расчета прогноза. Рисунок 3с [12] показывает графическое представление общей структуры модели KNN.

Традиционные методы машинного обучения совместно со статистическими моделями достаточно широко применяются к решению задачи прогноза качества воздуха. Так, например, методы K -ближайших соседей KNN, опорных векторов SVR, деревьев решений DT, RandomForest RF были реализованы для создания системы прогнозирования статуса индекса качества воздуха [13], а также были исследованы в [14] (KNN), [15] (SVM), [16] (DT, RF), и многих других. Основываясь на ограниченных наборах данных, эти методы машинного обучения могут лишь частично уловить ограниченные нелинейные временные и пространственные корреляции, влияющие на качество воздуха. Они больше подходят для поверхностных скрытых признаков, как показывают результаты. Учет очень сложных динамических нелинейных пространственно- временных связей им не поддается. Для эффективного преодоления этой проблемы создаются многочисленные комбинированные методы, использующие преимущества статистических методов и методов машинного обучения (см., например [17]).

В отличие от статистических методов, традиционные методы машинного обучения и методы глубокого обучения не зависят от предварительных физических знаний о данных и теоретических предположений. Они автоматически обучают свои параметры модели на

основе исторических данных, чтобы уловить нелинейные взаимосвязи и знания о данных качества воздуха. По сравнению со статистическими методами, традиционные методы машинного обучения, как правило, обеспечивают лучшую производительность в прогнозировании загрязняющих веществ в воздухе. Однако их эффективность, как правило, ограничена в практическом применении. Это связано с тем, что поверхностные структуры традиционных моделей машинного обучения часто не могут уловить сложные и нестационарные колебания данных о загрязнителях воздуха, что приводит к их плохой производительности в долгосрочном прогнозировании загрязняющих веществ в воздухе, а это имеет решающее значение для превентивных мер по предотвращению загрязнений воздуха и управления ими [4].

2. Алгоритмы глубокого обучения. Алгоритмы глубокого обучения используют искусственные нейронные сети (ИНС). ИНС — это алгоритм, основанный на связях, состоящих из нейронов или узлов. Эти связи организованы в три типа слоев (рисунок 4). Входной слой получает на вход исходные предикторные переменные. Выходной слой производит предсказанные значения для заданных входов. Эти два слоя соединены скрытыми слоями. Скрытые слои (более одного в случае глубокого обучения) содержат ненаблюдаемые нейроны, отвечающие за вычисления. Каждый узел в слое соединен с узлами в следующем слое. Каждое соединение имеет присвоенный вес, который используется для комбинации входных данных. Каждый узел или нейрон в следующем слое получает взвешенное значение и преобразует его с помощью активационной функции. Наиболее популярная функция - сигмоида. Полученный результат — это значение, которое передается в качестве входа узлам следующего слоя. Этот процесс продолжается, пока не будет достигнут выходной слой. На этом этапе производится выходное предсказание.

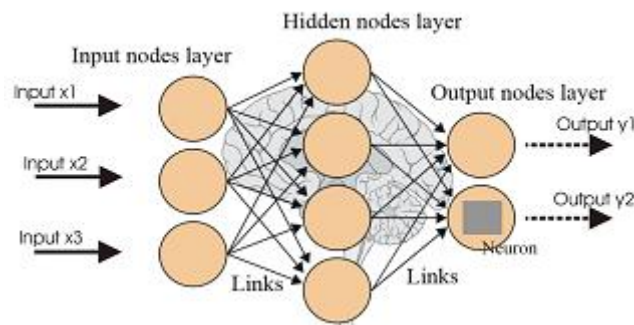


Рисунок 4- Структура ИНС [17]

Конечная цель ИНС — настроить веса для минимизации функции ошибки, обычно квадратичной функции. Для этого ИНС использует алгоритм обратного распространения ошибки. Этот алгоритм применяет метод градиентного спуска, используя частные производные слоев, чтобы найти оптимальный вес для каждого узла.

Типы ИНС, наиболее часто используемые при прогнозе качества воздуха с учетом временных зависимостей, — следующие.

Многослойные перцептронные нейронные сети (Multi-layer perceptron neural networks, MLP). MLP — это классические нейронные сети с одним или несколькими скрытыми слоями. Широко используемыми моделями в задачах прогноза качества воздуха являются нейронные сети с обратным распространением (BPNN) [18, 19] и обобщенно-регрессионные нейросетевые модели (GRNN) [20, 7]. Эти методы машинного обучения обладают четкой математической логикой, в которой связь между входными и выходными данными относительно определенная. Кроме того, у них относительно

простая структура сети, что приводит к ограниченной способности моделирования зависимостей во временных рядах данных PM2.5 и AQI.

Сверточные нейронные сети (Convolutional neural networks, CNN). CNN обычно применяются к изображениям, однако также используются в задачах прогноза качества воздуха. Это расширение MLP, в котором чередуются два типа слоев: сверточные и слой пуллинга (подвыборки). Цель сверточных слоев — извлечение признаков из входного изображения. Сверточная операция, которая выводит свёрнутые признаки, выполняется с помощью матрицы, называемой ядром или фильтром. Эти признаки становятся входом для подвыборочных слоев. Задача подвыборочного слоя — уменьшить размер свёрнутых признаков с целью снижения вычислительных ресурсов, необходимых для обработки изображения. В частности, в [21] представлен метод прогнозирования AQI с помощью одномерной сверточной нейронной сети и оптимизации экспоненциальных адаптивных градиентов.

Рекуррентные нейронные сети (Recurrent neural networks, RNN). RNN работают с временными рядами или последовательными данными, то есть с информацией, упорядоченной и взаимосвязанной. RNN имеют внутреннюю память в том смысле, что нейрон может обратным образом получать на вход выходные данные, которые он ранее произвел. Это позволяет модели приобретать краткосрочную память, которая необходима для прогнозирования временных рядов. Модели на основе RNN обладают отличной способностью улавливать временные зависимости во входных временных рядах данных PM2.5 и AQI [22].

Нейронные сети с долгосрочной и краткосрочной памятью (Long-short term memory neural networks, LSTM). LSTM — это расширение RNN. Они имеют расширенную память, которая позволяет работать с долгосрочными зависимостями. LSTM могут запоминать информацию на произвольных временных интервалах. Основным компонентом является состояние ячейки, которое переносит информацию на протяжении обработки данных. Информация обновляется на основе трех слоев (рисунок 5). Каждое из них контролирует информацию, которая должна находиться в состоянии ячейки, с использованием сигмоидной функции активации. Слой забывания (forget gate) определяет, какая часть информации из предыдущего состояния должна быть забыта. Слой ввода (input gate) решает, какая новая информация будет использована для обновления памяти. Используя гиперболическую тангенс-функцию, он создает вектор-кандидат для добавления к состоянию ячейки. Последний слой, известный как выходной слой (output gate), использует функцию активации tanh для определения, какая часть обновленного состояния ячейки будет использоваться в качестве выхода.

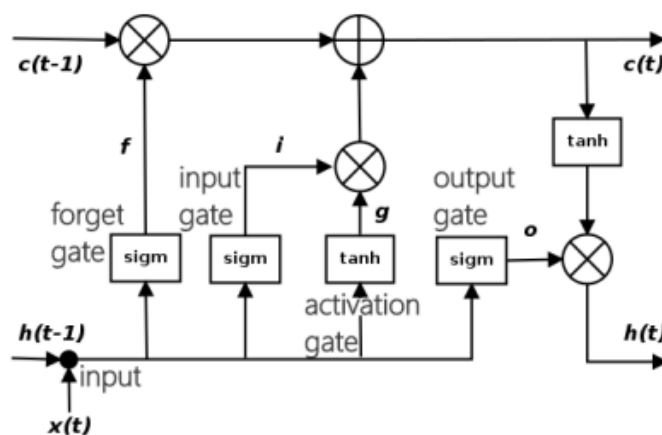


Рисунок 5 - LSTM - ячейка: [22]

Модели LSTM успешно применяются для предсказания концентраций загрязняющих веществ в воздухе, улавливая долговременные зависимости и паттерны во временных данных (см., например, [23, 24]) и являются наиболее часто используемыми среди методов глубокого обучения (в 36.13% статей [1]).

Упрощенная рекуррентная единица (Gated recurrent unit GRU). GRU — это упрощенная версия LSTM, которая объединяет слой забывания и слой ввода. Они дают результаты, похожие на те, что получаются с помощью LSTM.

Нейронные сети типа "кодировщик-декодировщик" (Encoder-Decoder neural networks, EDNN). Модель кодировщик-декодировщик — это рекуррентная нейронная сеть, используемая для задач предсказания последовательностей. Ее архитектура состоит из трех компонентов: кодировщика, промежуточного вектора и декодировщика. Кодировщик и декодировщик состоят из набора рекуррентных единиц (обычно LSTM или GRU). В случае кодировщика каждая единица обрабатывает элемент входной последовательности и пытается запечатлеть связанную информацию в промежуточном векторе, чтобы повысить точность предсказания декодировщика. Декодировщик использует этот вектор для формирования предсказания.

3. Графовые нейронные сети (Graph Neural Network, GNN). В последнее время все более популярными в прогнозировании качества воздуха становятся графовые нейронные сети [3 – 5]. Основная характеристика этого типа сетей заключается в информации, полученной от динамического взаимодействия между соседями (например, разные города, разные районы, разные улицы), которая взвешивается в зависимости от расстояния. Например, в [5] для прогноза концентрации PM2.5 на основе метеорологических данных и географического положения трех станций мониторинга применена графовая нейронная сеть на реальных наборах данных, которая показала лучшие результаты по сравнению с базовыми моделями [1].

Подобно представленным в [3] определениям определим некоторые понятия для описания проблемы прогноза качества воздуха для случая, когда информация о загрязнении воздуха приходит не с одного, а с некоторого множества датчиков.

Определение 1: датчики и группы датчиков. Для представления некоторого набора датчиков используется множество $S = \{s_i\}_{i=1}^{N_s}$, где N_s представляет число (количество) датчиков, а местоположение датчика задаётся его долготой и широтой, и матрица локации расположения датчиков выражается как $L \in R^{N_s \times 2}$. Кроме того, множество $U = \{u_i\}_{i=1}^{N_u}$ используется для представления множества групп датчиков, где N_u представляет количество групп датчиков.

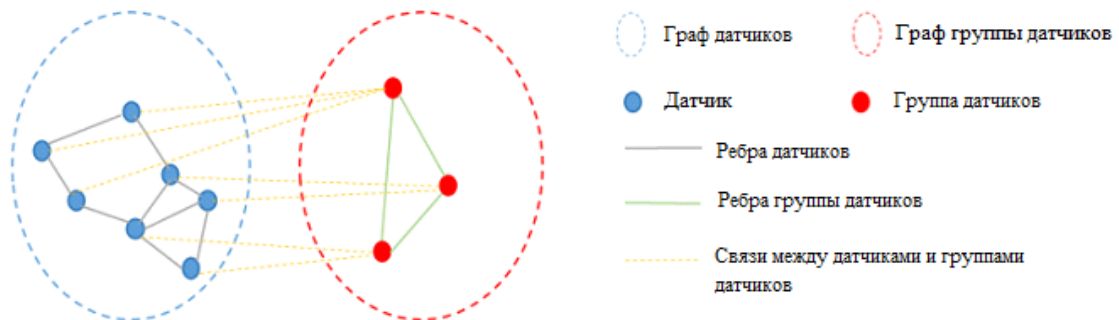


Рисунок 6 – Графы датчиков и групп датчиков.

Определение 2: граф группы датчиков. Будем рассматривать группы датчиков как датчики с сильными зависимостями и строить граф групп датчиков как полносвязный граф. Для представления топологической структуры графа групп датчиков используется неориентированный граф $g(v, \varepsilon)$, как показано на рисунке 6, где $v = \{v_1, v_2, \dots, v_n\}$ представляет собой множество всех узлов групп датчиков, при этом для любого узла v_i значения его атрибутов обозначаются как R_i , значения атрибутов всех узлов в v могут быть выражены как матрица R . $\varepsilon = \{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n\}$ представляет собой множество рёбер всех групп датчиков. Для любого ребра ε_i значения его атрибута - M_i , значения атрибутов всех рёбер в ε могут быть выражены как матрица M .

Определение 3: граф датчиков. Для представления топологии графа датчиков используется неориентированный граф $G(V, E)$, как показано на рисунке 6, где $V = \{v_1, v_2, \dots, v_n\}$ - множество всех узлов датчиков, и для любого узла V_i определены его атрибуты X_i , значения атрибутов всех узлов в V могут быть представлены в виде матрицы X . $E = \{e_1, e_2, \dots, e_n\}$ представляет собой множество всех рёбер датчиков. Для любого ребра E_i используется Y_i для обозначения значений его атрибутов, а значения атрибутов всех рёбер в E могут быть представлены в виде матрицы Y . Граф датчиков строится следующим образом:

$$Y_{i,j} = \frac{1}{\sqrt{(a_i - a_j)^2 + (b_i - b_j)^2}},$$

$$d_{i,j} = \sqrt{(a_i - a_j)^2 + (b_i - b_j)^2}, \quad 0 < d_{i,j} < r_n,$$

где $Y_{i,j}$ - атрибут ребра между датчиком S_i и датчиком S_j , $[a_i, b_i]$ представляет локацию (местоположение) датчика S_i , а $[a_j, b_j]$ представляет локацию датчика S_j . $d_{i,j}$ - евклидово расстояние между датчиком S_i и датчиком S_j , а r_n представляет собой порог расстояния. Когда расстояние между двумя датчиками очень велико, датчики почти не влияют друг на друга. Если расстояние больше r_n , можно считать, что эти датчики принадлежат разным населенным пунктам.

Определение 4: задача прогноза индекса качества воздуха. Необходимо предсказать будущий индекс качества воздуха AQI датчика, исходя из исторической последовательности наблюдений за местоположением датчика, временем, AQI и данными о погоде. То есть, для датчика S_i , учитывая местоположение датчика L , вектор времени T_t и историческую последовательность наблюдений $H = (h_i^{t_1}, h_i^{t_2}, \dots, h_i^{t_m})$ за m временных шагов в момент времени t , можно предсказать последовательность наблюдений AQI датчика $\hat{H} = (\hat{h}_i^{t_{m+1}}, \hat{h}_i^{t_{m+2}}, \dots, \hat{h}_i^{t_{m+n}})$ на n временных шагов.

Итак, рассматривая данные о качестве воздуха, зарегистрированные на мониторинговой площадке, как сигнал соответствующего узла, GNN могут автоматически извлекать пространственную зависимость из данных о качестве воздуха без каких-либо предварительных знаний. Эта особенность позволила активно использовать GNN в моделировании прогнозирования качества воздуха. Например, в [25] предложено объединить графовую свертку (GCN) с рекуррентной нейронной сетью (RNN), структурированной как Gated Recurrent Unit (GRU), чтобы захватить временные и пространственные корреляции для улучшения прогнозирования PM2.5. В [26] авторы разработали три распределенные компонента на основе GCN для извлечения недавних, ежедневных периодических и еженедельных периодических временных пространственных зависимостей данных о качестве воздуха и объединили выходы трех компонентов с компонентом извлечения внешней информации для генерации прогнозов качества воздуха.

Отличная способность GNN к извлечению пространственных особенностей обеспечила этим методам хорошую предсказательную производительность. Однако способность этих методов извлекать временную зависимость все еще нуждается в улучшении. Например, точность модели с использованием RNN [25] значительно снижается с увеличением длины времени предсказания из-за ее ограничения на извлечение глобальной информации.

В последующих исследованиях необходимо дополнительно изучить, как графовые нейронные сети могут быть интегрированы с другими моделями временных рядов, чтобы повысить предсказуемость и универсальность этих моделей.

Литература

1. Manuel Méndez, · Mercedes G. Merayo, Manuel Núñez. Machine learning algorithms to forecast air quality: a survey// Artificial Intelligence Review (2023) 56:10031–10066 <https://doi.org/10.1007/s10462-023-10424-4>
2. Лыченко Н.М. Корреляционный анализ метеопараметров и концентраций твердых частиц PM_{2.5} в атмосферном воздухе г. Бишкек и его применение в моделях машинного обучения // Проблемы автоматизации и управления. – 2024. – № 2(50). – С. 88-96. – EDN ZIWQVG
3. A Hybrid Model for Spatiotemporal Air Quality Prediction Based on Interpretable Neural Networks and a Graph Neural Network
4. Pengfei Li, Tong Zhang and Yantao Jin. A Spatio-Temporal Graph Convolutional Network for Air Quality Prediction, Sustainability 2023, 15, 7624. <https://doi.org/10.3390/su15097624>
5. Li D, Yu H, Geng Y-a, Li X, Li Q (2021) DDGNet: A dual-stage dynamic spatio-temporal graph net-work for PM_{2.5} forecasting. In: 2021 IEEE International Conference on Big Data (Big Data), pp 1679–1685. <https://doi.org/10.1109/BigData52589.2021.9671941>
6. He, Z.; Liu, P.; Zhao, X.; He, X.; Liu, J.; Mu, Y. Responses of Surface O₃ and PM_{2.5} Trends to Changes of Anthropogenic Emissions in Summer over Beijing during 2014–2019: A Study Based on Multiple Linear Regression and WRF-Chem. Sci. Total Environ.2022, 807, 150792.
7. Великанова, Л. И. Мультирегрессионные и обобщенно-регрессионные нейросетевые модели краткосрочного прогноза загрязнения PM_{2.5} в г. Бишкек с учетом метеорологических параметров / Л. И. Великанова, Н. М. Лыченко // Проблемы автоматизации и управления. – 2019. – № 2(37). – С. 16-25. – DOI 10.5281/zenodo.3594777. – EDN MNNDYU.
8. Kulkarni, G.E.; Muley, A.A.; Deshmukh, N.K.; Bhalchandra, P.U. Autoregressive integrated moving average time series model for forecasting air pollution in Nanded city, Maharashtra, India. Model. Earth Syst. Environ. 2018, 4, 1435–1444.
9. Mirche Arsov, Eftim Zdravevski, Petre Lameski, Roberto Corizzo, Nikola Koteli, Sasho Gramatikov, Kosta Mitreski, and Vladimir Trajkovik. Multi-Horizon Air Pollution Forecasting with Deep Neural Networks // Sensors (Basel). 2021 Feb; 21(4): 1235. doi: 10.3390/s21041235
10. Верзунов С.Н., Лыченко Н.М. Анализ и ARIMA- модели динамики изменения концентрации PM_{2.5} в атмосферном воздухе г. Бишкек // Проблемы автоматизации и управления.- N1. Бишкек: Илим, 2019. –С. 21-30.
11. Balogun A-L, Tella A (2022) Modelling and investigating the impacts of climatic variables on ozone concentration in Malaysia using correlation analysis with random forest, decision tree regression, linear regression, and support vector regression. Chemosphere 299:134250. <https://doi.org/10.1016/j.chemosphere.2022.134250>

12. Tella A, Balogun A-L, Adebisi N, Abdullah S (2021) Spatial assessment of PM10 hotspots using Random Forest, K-Nearest Neighbour and Naïve Bayes. *Atmos Pollut Res* 12(10):101202. <https://doi.org/10.1016/j.apr.2021.101202>
13. Morapedi TD and Obagbuwa IC (2023) Air pollution particulate matter (PM2.5) prediction in South African cities using machine learning techniques. *Front. Artif. Intell.* 6:1230087. doi: 10.3389/frai.2023.1230087
14. Keramat-Jahromi, M., Mohtasebi, S. S., Mousazadeh, H., Ghasemi-Varnamkhasti, M. & Rahimi-Movassagh, M. Real-time moisture ratio study of drying date fruit chips based on on-line image attributes using kNN and random forest regression methods. *Measurement* 172, 108899 (2021).
15. Pisner, D. A. & Schnyer, D. M. Support Vector Machine. *Machine Learning: Methods and Applications to Brain Disorders* (Elsevier Inc, 2019).
16. Sarkhosh, M. et al. Indoor Air Quality associations with sick building syndrome: An application of decision tree technology. *Build. Environ.* 188, 107446 (2021).
17. Mr. V. Devasekhar, Dr. P. Natarajan: Prediction of Air Quality and Pollution using Statistical Methods and Machine Learning Techniques. In: *International Journal of Advanced Computer Science and Applications*. Vol. 14, No. 4 (2023)
18. Wang, W., Zhao, S., Jiao, L., Taylor, M., Zhang, B., Xu, G., et al. (2019). Estimation of PM2.5 Concentrations in China Using a Spatial Back Propagation Neural Network. *Sci. Rep.* 9 (1), 13788–13810. doi:10.1038/s41598-019-50177-1
19. Великанова, Л. И. Моделирование краткосрочного прогноза загрязненности воздуха твердыми частицами на основе искусственных нейронных сетей с учетом фактора загрязнения / Л. И. Великанова, Н. М. Лыченко // *Проблемы автоматизации и управления*. – 2022. – № 3(45). – С. 110-118. – EDN TGNBLC.
20. Zhou, Q., Jiang, H., Wang, J., and Zhou, J. A Hybrid Model for PM 2.5 Forecasting Based on Ensemble Empirical Mode Decomposition and a General Regression Neural Network. *Sci. Total Environ.* 496, 264–274, 2014. doi:10.1016/j.scitotenv.2014.07.051
21. Ragab, M. G., Abdulkadir, S. J., Aziz, N., Al-Tashi, Q., Alyousifi, Y., Alhussian, H., et al. A Novel One-Dimensional CNN with Exponential Adaptive Gradients for Air Pollution Index Prediction. *Sustainability* 12 (23), 2023. doi:10.3390/su122310090
22. Stojov, V., Koteli, N., Lameski, P., Zdravevski, E.: Application of machine learning and time-series analysis for air pollution prediction. *CIIT 2018* (2018).
23. Loy-Benitez, J.; Vilela, P.; Li, Q.; Yoo, C. Sequential Prediction of Quantitative Health Risk Assessment for the Fine Particulate Matter in an Underground Facility Using Deep Recurrent Neural Networks. *Ecotoxicol. Environ. Saf.* 2019, 169, 316–324.
24. Лыченко, Н. М. Применение LSTM-нейронных сетей для классификации индекса качества воздуха г. Бишкек / Н. М. Лыченко, А. В. Сорокова // *Проблемы автоматизации и управления*. – 2020. – № 1(38). – С. 70-80. – DOI 10.5281/zenodo.3904130. – EDN QQZNFI.
25. Wang, S.; Li, Y.; Zhang, J.; Meng, Q.; Meng, L.; Gao, F. PM2.5-GNN: A Domain Knowledge Enhanced Graph Neural Network For PM2.5 Forecasting. In *Proceedings of the 28th International Conference on Advances in Geographic Information Systems*, Seattle, WA, USA, 3–6 November 2020; pp. 163–166.
26. Chang, F.; Ge, L.; Li, S.Y.; Wu, K.Y.; Wang, Y.Q. Self-adaptive spatial-temporal network based on heterogeneous data for air quality prediction. *Connect. Sci.* 2021, 33, 427–446.