

СРАВНИТЕЛЬНЫЙ АНАЛИЗ ГРАДИЕНТНЫХ  
МЕТОДОВ ОПРЕДЕЛЕНИЯ ИСТОЧНИКА  
ДИФFUЗИОННО-ЛОГИСТИЧЕСКОЙ МОДЕЛИ<sup>1)</sup>© 2022 г. Т. А. Звонарева<sup>1,\*</sup>, О. И. Криворотко<sup>1,2,\*\*</sup><sup>1</sup> 630090 Новосибирск, ул. Пирогова, 2, Новосибирский государственный университет, Россия<sup>2</sup> 630090 Новосибирск, пр-т Акад. Лаврентьева, 6, Институт вычислительной математики  
и математической геофизики СО РАН, Россия

\*e-mail: t.zvonareva@g.nsu.ru

\*\*e-mail: krivorotko.olya@mail.ru

Поступила в редакцию 10.08.2021 г.

Переработанный вариант 10.08.2021 г.

Принята к публикации 16.12.2021 г.

Проведен сравнительный анализ численного решения задачи определения источника диффузионно-логистической модели по данным о диффузионном процессе в фиксированные моменты времени и точках пространства градиентными методами для случаев непрерывной и дискретной постановок. Получены выражения вычисления градиента целевого функционала для двух постановок, связанных с решением соответствующих сопряженных задач. Показано, что в случае аппроксимации дискретных функций модели кубическими сплайнами точность восстановленного источника совпадает с решением в непрерывной постановке. Численные эксперименты решения задачи об источнике для дискретной модели распространения информации в онлайн социальных сетях показали, что применение дискретного подхода в разы увеличивает вычислительное время по сравнению с использованием непрерывного подхода. Библиография: 23. Фиг. 1. Табл. 5.

**Ключевые слова:** диффузионно-логистическая модель, задача об источнике, обратная задача, градиентные методы, сопряженная задача, сравнительный анализ, регуляризация, оптимизация, социальные процессы.

DOI: 10.31857/S0044466922040147

## ВВЕДЕНИЕ

Градиентные методы решения задач об источнике для дифференциальных уравнений применялись в [1]–[7]. Основная их идея состоит в последовательном уменьшении значения целевого функционала  $J(q) = \|A(q) - f\|^2$  в виде

$$q^{m+1} = q^m - \alpha_m J'(q^m), \quad q^0 \in Q, \quad \alpha_m > 0, \quad (1)$$

где  $\alpha_m$  – параметр спуска, характеризующий тот или иной градиентный метод,  $J'(q^m)$  – градиент целевого функционала  $J(q)$ ,  $A : Q \rightarrow F$  – оператор обратной задачи.

В зависимости от типа пространства  $Q$  (гильбертово, соболевское, евклидово и т.п.) области определения функционала  $J(q)$  и его градиента  $J'(q)$  могут быть интегрируемые, непрерывно-дифференцируемые или дискретные. В связи с этим возможно использование двух подходов: 1) непрерывный, который состоит в формулировке используемых функций в непрерывных пространствах с последующим представлением задачи в конечно-разностном виде и 2) дискретный, в котором сначала переходят к конечно-разностному аналогу, для которого формулируют дискретные функции  $J(q)$  и  $J'(q)$  (см. [8]).

<sup>1)</sup> Работа выполнена при финансовой поддержке РФФИ (проект № 18-71-10044-П).

В данной работе для задачи об источнике для диффузионно-логистической модели, возникающей при описании распространения информации в онлайн социальных сетях и сформулированной в непрерывном и дискретном видах, применены оба подхода. Выявлены области их применимости. Получены градиенты целевых функционалов в непрерывной и дискретной постановках, связанных с решениями соответствующих сопряженных задач. Отметим, что дискретный вариант градиента не является следствием дискретизации непрерывного своего аналога, а получен путем минимизации функции Лагранжа (см. [9]–[12]).

Статья организована следующим образом. Формулировки прямой и обратной задач приведены в разд. 1. Вывод градиента целевого функционала в дискретной форме приведен в п. 2.1. Выражение для градиента целевого функционала в непрерывной форме и соответствующая постановка сопряженной задачи приведены в п. 2.2. Алгоритмы градиентных методов для разных параметров спуска, а также сравнительный анализ результатов численных расчетов для решения задачи об источнике для диффузионно-логистической модели приведены в разд. 3.

## 1. ПОСТАНОВКИ ПРЯМОЙ И ОБРАТНОЙ ЗАДАЧ

### 1.1. Постановка прямой задачи

Под прямой задачей в данной работе понимается задача моделирования процесса распространения информации в онлайн социальных сетях, в которой требуется найти функцию плотности активных пользователей  $u(x_j, t_n)$  в каждой точке пространства и в момент времени. За расстояние  $x$  принимается целочисленная величина, описывающая минимальное количество дружеских связей между пользователем и источником информации и измеряемая в единицах. Время  $t$  измеряется в часах, а плотность активных пользователей  $u(x_j, t_n)$  – в количестве человек на единицу расстояния.

Данные, представленные в [13], иллюстрируют, что интерес к информации проявляется у пользователей с расстоянием  $x$  в пределах от 1 до 6. Именно в этих границах происходит значительный вклад в изменение плотности активных пользователей  $u(x_j, t_n)$ . Также для социальных сетей справедлива теория шести рукопожатий, согласно которой большинство агентов находится на расстоянии  $x \leq 6$ . Поэтому в модели рассматриваются граничные условия Неймана, которые описывают отсутствие потока информации через границы при  $x = 1, 6$ .

Задача моделирования процесса распространения информации основана на законе сохранения информационного потока

$$\int_a^b \left( \frac{\partial u}{\partial t} + \frac{\partial H}{\partial x} - g(u, x, t) \right) dx = 0.$$

А именно, для фиксированного участка  $[a, b]$  скорость изменения общего количества вовлеченных пользователей  $\frac{\partial}{\partial t} \int_a^b u(x, t) dx$  на этом участке должна равняться сумме потока  $H(a, t) - H(b, t)$ ,

с которой они поступают, и скорости  $\int_a^b g(u, x, t) dx$ , с которой в пределах  $a < x < b$  появляются новые вовлеченные пользователи. Информация в социальной сети распространяется от высокой плотности вовлеченных пользователей к низкой, поэтому поток может быть представлен в виде  $H = -d \frac{\partial u}{\partial x}$ . В [13] функция  $g(u, x, t)$  выбирается в виде  $(1 - u/K)r(t)u$  и описывает динамику изменения численности активных пользователей. Неопределенность построенной модели заключается в том, что распределение плотности активных пользователей в начальный момент времени неизвестно и зависит от структуры социальной сети. Поэтому актуальной становится задача идентификации начального распределения пользователей с целью корректного описания распространения информации в конкретной сети и ее дальнейшего контроля/управления.

Обозначим начальные значения плотности вовлеченных пользователей  $u(x_i, 1) = q_i$ ,  $i = 0, \dots, 5$ . Интерполируя вектор значений  $q = (q_0, \dots, q_5)$  кубическими сплайнами, переходим к

**Таблица 1.** Параметры модели (2), описывающие распространение информации в онлайн социальной сети

Параметр	Описание	Среднее значение	Единицы измерения
$d$	популярность информации, которая способствует распространению информации через неструктурные действия, такие как целенаправленный поиск пользователем рассматриваемой информации	0.01	расст. <sup>2</sup> /ч
$K$	пропускная способность, которая является максимально возможным числом активных пользователей	25	чел./расст.
$r(t)$	$= \frac{\beta_2}{\beta_1} - e^{-\beta_1(t-1)} \left( \frac{\beta_2}{\beta_1} - \beta_2 \right)$ скорость роста числа вовлеченных пользователей	–	1/ч
$\beta_1$	скорость снижения интереса к информации с течением времени	1.5	–
$\beta_2$	остаточная скорость	0.375	–
$\beta_3$	начальная скорость роста числа активных пользователей	1.65	–

непрерывной постановке начально-краевой задачи для диффузионно-логистической модели, описываемой уравнением в частных производных параболического типа (см. [13]):

$$\begin{aligned}
 u_t &= du_{xx} + \left(1 - \frac{u}{K}\right)r(t)u, \quad l \leq x \leq L, \quad t \geq 1, \\
 u(x, 1) &= Q(x), \quad l \leq x \leq L, \\
 u_x(l, t) &= u_x(L, t) = 0, \quad t \geq 1.
 \end{aligned} \tag{2}$$

Здесь  $Q(x) \geq 0$  – начальная функция плотности. Согласно [13], если  $Q(x) \not\equiv 0$  и удовлетворяет следующим условиям корректности: 1)  $Q(x) \in C^2(l, L)$ , 2)  $Q'(l) = Q'(L) = 0$ , 3)  $dQ'' + (1 - Q/K)rQ \geq 0$ , то по принципу максимума существует единственное положительное решение  $u(x, t) \in C^{2,1}((l, L) \times (1, +\infty)) \cap C^{1,0}([l, L] \times [1, +\infty))$  прямой задачи (2).

Средние значения и описания параметров модели представлены в табл. 1. Параметры  $\beta_1$ ,  $\beta_2$  и  $\beta_3$  в рассматриваемой модели соответствуют представленным в [13].

### 1.2. Постановка обратной задачи

Предположим, что известна дополнительная информация следующего вида:

$$u(x_i, t_k; q) = f_{ik}, \quad i = 1, \dots, N_1, \quad k = 1, \dots, N_2, \tag{3}$$

где  $u(x, t; q)$  – решение прямой задачи для начальной функции плотности  $Q(x)$ , определяемой из набора параметров  $q$ ,  $N_1 \leq 6$ . Обратная задача (2)–(3) состоит в определении набора параметров  $q$  по данным  $f_{ik}$  вида (3). Обратная задача может быть записана в виде  $Aq = f$ , где  $f = (f_{11}, \dots, f_{1N_2}, f_{21}, \dots, f_{2N_2}, f_{N_11}, \dots, f_{N_1N_2}) \in \mathbb{R}^{N_1N_2}$ ,  $A$  – оператор обратной задачи.

Отметим, что в линейном приближении при  $N_1N_2 = 6$  и  $\det A \neq 0$  существует единственное решение системы  $Aq = f$ . В случае  $N_1N_2 > 6$  вводится понятие нормального псевдорешения, т.е. решения, реализующего минимум нормы невязки

$$J(q) = \|Aq - f\|^2.$$

В [14] в линейном приближении решение обратной задачи (2)–(3) исследовано на устойчивость, и показано, что число обусловленности матрицы  $A$  имеет порядок  $10^{16}$ , что свидетельствует о неустойчивости решения обратной задачи.

В данной работе невязка имеет вид

$$J(q) = w \sum_{i=1}^{N_1} \sum_{k=1}^{N_2} \int_1^L \int_1^T |u(x, t; q) - f(x, t)|^2 \delta(x - x_i) \delta(t - t_k) dt dx, \tag{4}$$

здесь  $w = (L - l)(T - 1)/(N_1 N_2)$ .

По аналогии сформулируем дискретную постановку целевого функционала

$$I(q) = w \sum_{i=1}^{N_1} \sum_{k=1}^{N_2} |v_i^k - f_{ik}|^2, \tag{5}$$

где  $v_i^k := u(x_i, t_k; q)$ .

## 2. НЕПРЕРЫВНАЯ И ДИСКРЕТНАЯ ПОСТАНОВКИ ГРАДИЕНТОВ ФУНКЦИОНАЛА

Исследование непрерывной и дискретной постановок прямой и обратной задач для диффузионно-логистической модели было навеяно работой А.Л. Карчевского [8], в которой похожий анализ был проделан для решения обратной задачи для гиперболического уравнения.

Существуют два подхода к численному решению поставленной обратной задачи (2)–(3). Первый состоит из следующих шагов:

- перейти от непрерывной постановки прямой задачи  $L_q u = 0$  ( $L_q$  – оператор прямой задачи) к дискретной  $\Lambda_p v = 0$  ( $\Lambda_p$  – конечно-разностный аналог оператора прямой задачи, а функция  $p$  является некоторым приближением функции  $Q$  и  $p_j = q_{s-1}$  при  $x_j = s, s = 1, \dots, 6$ );

- выписать целевой функционал в дискретной постановке  $I(q)$ ;

- получить постановку сопряженной задачи  $\Lambda_p^* \phi = 0$  и градиент целевого функционала в дискретной форме  $I'(q)$ ;

- решить задачу минимизации функционала  $I(q)$ .

Второй подход подразумевает следующую схему действий:

- выписать целевой функционал в непрерывной постановке  $J(q)$ ;

- получить постановку сопряженной задачи  $L_q^* \psi = 0$  и выражение градиента целевого функционала в непрерывной форме  $J'(q)$ ;

- перейти к задаче  $\Lambda_p v = 0$ ;

- выписать целевой функционал  $I(q)$ , аппроксимирующий  $J(q)$ ;

- от постановки сопряженной задачи  $L_q^* \psi = 0$  перейти к задаче  $\Lambda_p^* \phi = 0$ ;

- получить соотношение  $I'(q)$ , которое аппроксимирует градиент целевого функционала  $J'(q)$ ;

- решить задачу минимизации функционала  $J(q)$ .

Особенность постановки обратной задачи (2)–(3) состоит в дискретном задании дополнительных измерений, а также в определении набора параметров  $q = (q_0, \dots, q_5)$  вместо функции  $Q(x)$ .

### 2.1. Градиент функционала в дискретной постановке

Дискретный градиент, полученный с помощью подхода из работ Ю.Г. Евтушенко и Ф.Л. Черноусько [9]–[12], формулируется для дискретной постановки рассматриваемой модели, т.е. для задачи

$$\begin{aligned} \frac{v_j^{n+1} - v_j^n}{\tau} &= d \frac{v_{j+1}^n - 2v_j^n + v_{j-1}^n}{h^2} + \left(1 - \frac{v_j^n}{K}\right) r^n v_j^n, \quad j = 1, \dots, N_x - 1, \quad n = 0, \dots, N_t - 1, \\ v_j^0 &= p_j, \quad j = 0, \dots, N_x, \\ v_0^{n+1} &= \frac{4v_1^{n+1} - v_2^{n+1}}{3}, \quad n = 0, \dots, N_t - 1, \\ v_{N_x}^{n+1} &= \frac{4v_{N_x-1}^{n+1} - v_{N_x-2}^{n+1}}{3}, \quad n = 0, \dots, N_t - 1. \end{aligned} \tag{6}$$

Преобразуем ее в условие связи

$$\Phi(n, v^n; q) = 0, \quad n = 0, \dots, N_t, \tag{7}$$

где  $v^n$  – набор векторов  $v_j, j = 0, \dots, N_x$ .

Каждому условию в (7) ставится в соответствие вектор  $\phi^n \in \mathbf{R}^{N_x+1}$ , их объединение есть вектор  $\phi^T = (\phi_0^T, \dots, \phi_{N_t}^T), \phi \in \mathbf{R}^{(N_t+1) \times (N_x+1)}$ . И рассматривается аналог функции Лагранжа для многошагового процесса (7):

$$E(v, \phi; q) = I(q) - \sum_{n=0}^{N_t} \Phi^T(n, v^n; q) \phi^n. \tag{8}$$

Тогда (см. [9])

$$\phi^{i_n} = E_{v^{i_n}}(v, \phi; q) = I_{v^{i_n}}(q) - \sum_{n=0}^{N_t} \Phi_{v^{i_n}}^T(n, v^n; q) \phi^n, \quad i_n = 0, \dots, N_t, \tag{9}$$

$$\frac{\partial I(q)}{\partial q_{i_s}} = E_{q_{i_s}}(v, \phi; q) = I_{q_{i_s}}(q) - \sum_{n=0}^{N_t} \Phi_{q_{i_s}}^T(n, v^n; q) \phi^n, \quad i_s = 0, \dots, 5. \tag{10}$$

И справедлива следующая лемма.

**Лемма.** *Градиент функционала (5) имеет вид*

$$\begin{aligned} I'(q)[0] &= -\frac{d}{h^2} \phi_1^0, \\ I'(q)[s] &= -\frac{\phi_{j_s}^0}{\tau} - d \frac{\phi_{j_{(s-1)}}^0 - 2\phi_{j_s}^0 + \phi_{j_{(s+1)}}^0}{h^2} - r^0 \phi_{j_s}^0 + \frac{2r^0 q_s}{K} \phi_{j_s}^0, \quad s = 1, 2, 3, 4, \\ I'(q)[5] &= -\frac{d}{h^2} \phi_{N_x-1}^0, \end{aligned} \tag{11}$$

где  $\phi_{j_s}^0$  соответствует значению  $\phi$  в точке  $(x_j = s + 1, t_0)$  и функция  $\phi_j^n$  удовлетворяет сопряженной задаче

$$\begin{aligned} \frac{\phi_j^{n-1} - \phi_j^n}{\tau} &= d \frac{\phi_{j-1}^n - 2\phi_j^n + \phi_{j+1}^n}{h^2} + r^n \phi_j^n - \frac{2r^n v_j^n}{K} \phi_j^n - \frac{[\phi]}{\tau h}, \quad j = 1, \dots, N_x - 1, \quad n = 1, \dots, N_t, \\ \phi_j^{N_t} &= 0, \quad j = 0 \dots, N_x, \\ \phi_0^n &= \frac{d}{h^2} \phi_1^n, \quad n = 0 \dots, N_t - 1, \\ \phi_{N_x}^n &= \frac{d}{h^2} \phi_{N_x-1}^n, \quad n = 0 \dots, N_t - 1, \end{aligned} \tag{12}$$

где  $[\phi] = 2w(v_i^k - f_{ik})$  при  $(x_j, t_n) = (x_i, t_k)$  и  $[\phi] = 0$  при  $(x_j, t_n) \neq (x_i, t_k)$ .

**Доказательство.** Функция  $E$ , определенная по формуле (8), имеет вид

$$\begin{aligned} \frac{E(v, q, \phi)}{\tau h} &= \frac{I(q)}{\tau h} + \sum_{j=1}^{N_x-1} \sum_{n=1}^{N_t-1} \left[ \frac{v_j^{n+1} - v_j^n}{\tau} - d \frac{v_{j+1}^n - 2v_j^n + v_{j-1}^n}{h^2} - r^n v_j^n + \frac{r^n (v_j^n)^2}{K} \right] \phi_j^n + \\ &+ \sum_{j=1}^{N_x-1} \left[ \frac{v_j^1 - p_j}{\tau} - d \frac{p_{j+1} - 2p_j + p_{j-1}}{h^2} - r^0 p_j + \frac{r^0 (p_j)^2}{K} \right] \phi_j^0 + \\ &+ \sum_{n=0}^{N_t-1} \left[ v_0^{n+1} - \frac{1}{3} (4v_1^{n+1} - v_2^{n+1}) \right] \phi_0^n + \sum_{n=0}^{N_t-1} \left[ v_{N_x}^{n+1} - \frac{1}{3} (4v_{N_x-1}^{n+1} - v_{N_x-2}^{n+1}) \right] \phi_{N_x}^n = 0. \end{aligned}$$

Находим производную функции  $E$  по  $v_j^n$ , что из (9) есть формула для определения функции  $\phi_j^n$ :

$$\frac{[\phi]}{\tau h} + \frac{\phi_j^{n-1} - \phi_j^n}{\tau} - d \frac{\phi_{j-1}^n - 2\phi_j^n + \phi_{j+1}^n}{h^2} - r^n \phi_j^n + \frac{2r^n u_j^n}{K} \phi_j^n = 0,$$

при  $j = 1, \dots, N_x - 1, n = 1, \dots, N_t$ . В случаях  $j = 0$  и  $j = N_x$  для всех  $0 \leq n \leq N_t$

$$\phi_0^{n-1} = \frac{d}{h^2} \phi_1^n, \quad \phi_{N_x}^{n-1} = \frac{d}{h^2} \phi_{N_x-1}^n.$$

При всех  $0 \leq j \leq N_x$  составляющая  $v_j^0$  не входит в выражение для функции  $E$ , поэтому можно положить

$$\phi_j^{N_t} = 0, \quad j = 0, \dots, N_x.$$

Таким образом, получили постановку сопряженной задачи (12).

Так как  $p_j = q_{s-1}$  при  $x_j = s$ , находим производные функции  $E$  по  $q_s$ :

$$-\frac{\phi_{j_s}^0}{\tau} - d \frac{\phi_{j(s-1)}^0 - 2\phi_{j_s}^0 + \phi_{j(s+1)}^0}{h^2} - r^0 \phi_{j_s}^0 + \frac{2r^0 q_s}{K} \phi_{j_s}^0,$$

при  $s = 2, 3, 4, 5$ , где  $\phi_{j_s}^0$  соответствует значению  $\phi$  в точке  $(x_j = s, t_0)$ . При  $s = 1$  и  $s = 6$  производные соответственно

$$E_{q_0} = -\frac{d}{h^2} \phi_1^0, \quad E_{q_5} = -\frac{d}{h^2} \phi_{N_x-1}^0.$$

Тогда из (10) следует (11), что завершает доказательство леммы.

### 2.2. Градиент функционала в непрерывной постановке

Градиент функционала (4) имеет вид

$$J'(q) = -\psi(x, 1), \tag{13}$$

где функция  $\psi(x, t)$  удовлетворяет решению сопряженной задачи

$$\begin{aligned} \psi_t &= -d\psi_{xx} - r(t)\psi + \frac{2r(t)u}{K}\psi + \xi, \quad 1 \leq t \leq T, \quad l \leq x \leq L, \\ \psi(x, T) &= 0, \quad l \leq x \leq L, \\ \psi_x(l, t) &= \psi_x(L, t) = 0, \quad 1 \leq t \leq T, \end{aligned}$$

где  $\xi = 2w \sum_{i=1}^{N_1} \sum_{k=1}^{N_2} \int_l^L \int_1^T |u(x, t; q) - f(x, t)|^2 \delta(x - x_i) \delta(t - t_k) dt dx$ .

В [15] приведен вывод формулы (13) для более общей постановки задачи (2) с произвольной правой частью.

Отметим, что в силу свойств дельта-функции сопряженная задача может быть записана в виде (см. [16])

$$\begin{aligned} \psi_t &= -d\psi_{xx} - r(t)\psi + \frac{2r(t)u}{K}\psi, \quad 1 \leq t \leq T, \quad l \leq x \leq L, \\ \psi(x, T) &= 0, \quad l \leq x \leq L, \\ \psi_x(l, t) &= \psi_x(L, t) = 0, \quad 1 \leq t \leq T, \\ [\psi]_{x=x_i} &= 2w(u(x_i, t_k; q) - f_{ik}), \quad i = 1, \dots, N_1, \quad k = 1, \dots, N_2. \end{aligned}$$

### 3. ЧИСЛЕННЫЕ ЭКСПЕРИМЕНТЫ

Для решения обратной задачи (2)–(3) в качестве синтетических данных  $f_{ik}$  были взяты значения решения прямой задачи при значениях параметров, представленных в табл. 1 и 2, в каждой

**Таблица 2.** Дискретизованная функция начальной плотности  $Q(x)$ , используемая как точное решение  $q_{\text{ex}}$  обратной задачи

Параметр	$q_0$	$q_1$	$q_2$	$q_3$	$q_4$	$q_5$
Значение	5.8	1.7	1.9	1	0.95	0.7

десятой точке по  $x < N_x$  и в каждой двадцать пятой точке по  $t \geq 50$ , т.е.  $N_1 = 5$  и  $N_2 = 22$ . Значения параметров в табл. 2 эмитируют реальные данные новостного сайта Digg.com, представленные в [13].

Кроме значения функционала также измерялась относительная ошибка

$$\delta = \frac{\|q_{\text{ex}} - \tilde{q}\|^2}{\|q_{\text{ex}}\|^2}.$$

Здесь  $q_{\text{ex}}$  – точные значения дискретизованной функции  $Q(x)$ , а  $\tilde{q}$  – найденное решение обратной задачи, которое соответствует минимуму функционала  $J(q)$  (5).

### 3.1. Градиентные методы

Решение задачи минимизации целевых функционалов  $J(q)$  и  $I(q)$  было получено с использованием двух градиентных методов:

1. Метод градиентного спуска (МГС) (см. [17], [18]) – классический одношаговый градиентный метод (1) с постоянным параметром спуска  $\alpha_k = \alpha$ . Скорость сходимости по функционалу (см. [18])

$$J(q^m) \leq \frac{\|q^0 - q_{\text{ex}}\|^2}{m\alpha(1 - \alpha\|A\|^2)}.$$

2. Многоуровневый градиентный метод (МГМ) (см. [19], [20]) – это метод локальной оптимизации функции от нескольких переменных. Он является модификацией метода градиентного спуска, что позволяет значительно увеличить скорость сходимости. Алгоритм метода имеет следующий вид:

$$\begin{aligned} q^{m+1} &= \theta_{m+1} z^m + (1 - \theta_{m+1}) y^m, \quad \theta_{m+1} \in \arg \min_{\theta \in [0;1]} J(\theta z^m + (1 - \theta) y^m), \\ y^{m+1} &= q^{m+1} - \zeta_{m+1} J'(q^{m+1}), \quad \zeta_{m+1} \in \arg \min_{\zeta \geq 0} J(q^{m+1} - \zeta J'(q^{m+1})), \\ z^{m+1} &= z^m - \eta_{m+1} J'(q^{m+1}), \quad \eta_{m+1} = \frac{1}{2L_{m+1}} + \sqrt{\frac{1}{4L_{m+1}^2} + \eta_m^2}, \quad \eta_0 = 0. \end{aligned}$$

В случае МГМ скорость сходимости по функционалу оценивается следующим образом (см. [21]):

$$J(q^m) \leq \frac{4L\|q^0 - q_{\text{ex}}\|^2}{(m+1)^2},$$

где  $L$  – константа Липшица для градиента функционала

$$\|J'(q^{m+1}) - J'(q^m)\|_2 \leq L\|q^{m+1} - q^m\|_2.$$

Были применены и проанализированы МГС и МГМ с непрерывным и дискретным видами градиентов. Причем в случае дискретного градиента МГС имеет параметр спуска  $\alpha^{(1)} = 10^{-7}$  и минимум по  $\zeta$  в МГМ находится на отрезке  $[0, 10^{-6}]$  с шагом  $10^{-7}$ . В случае непрерывного градиента  $\alpha^{(2)} = 10^{-3}$ , а минимум по  $\zeta \in [0, 10^{-5}]$  определяется с шагом  $10^{-6}$ .

**Таблица 3.** Решение обратной задачи, полученное методом роя частиц

Параметр	$\hat{q}_0$	$\hat{q}_1$	$\hat{q}_2$	$\hat{q}_3$	$\hat{q}_4$	$\hat{q}_5$
Значение	5.809	1.697	1.901	0.999	0.949	0.701

**Таблица 4.** Результаты, полученные методами градиентного спуска и его модификаций

Метод	$q^0$	$J(q), \times 10^{-4}$	$\delta, \times 10^{-4}$	$N_{iter}$	$t_{ЭВМ}$
МГС I	$\hat{q}$	9.604	6.274	5380	3
МГМ I	$\hat{q}$	9.544	6.273	397	3
МГМ II	0	9.029	6.262	245	2
МГС II	0	9.032	6.262	16374	9
МГС III	$\hat{q}$	9.847	6.276	166862	93
МГМ III	$\hat{q}$	9.793	6.276	1372	11
МГМ IV	0	9.047	6.261	1630	12
МГС IV	0	9.101	6.266	1005999	569
МГМ V	0	11.26	6.254	1586	25

Примечание. В столбце начального приближения  $q^0$  под  $\hat{q}$  понимается вектор решения обратной задачи, полученный с помощью метода роя частиц (представленный в табл. 3), а 0 – нулевой вектор  $N_{iter}$  – число итераций,  $t_{ЭВМ}$  – время работы программы.

**3.2. Конечно-разностные схемы решения прямой и сопряженной задач**

Для построения разностных схем введем равномерную сетку в замкнутой области  $\bar{D} = \{(x, t) | l \leq x \leq L, 1 \leq t \leq T\}$ :

$$\bar{\omega} = \{(x_j, t_n) | x_j = l + jh, t_n = 1 + n\tau, j = 0, \dots, N_x, n = 0, \dots, N_t\},$$

где  $h = (L - l)/N_x$  и  $\tau = (T - 1)/N_t$ .

Для применения классического непрерывного подхода функция начального приближения  $Q(x)$  определяется из вектора  $q$  (см. табл. 2) интерполяцией кубическими сплайнами.

Прямая задача (2) решается с помощью явной конечно-разностной схемы с порядком аппроксимации  $O(\tau + h^2)$  в виде (6).

Положим в численных расчетах  $l = 1, L = 6, T = 24, N_x = 50$  и  $N_t = 575$ . Такие значения  $l, L$  были выбраны в соответствии с данными, представленными в [13], которые иллюстрируют, что интерес к информации проявляется у пользователей с расстоянием  $x$  в пределах от 1 до 6. Именно в этих границах происходит значительный вклад в изменение плотности активных пользователей  $u(x, t)$ . Значения разбиения сетки  $N_x$  и  $N_t$  выбраны таким образом, чтобы выполнялось условие Куранта–Фридрихса–Леви

$$\tau \leq \frac{2h^2}{4d + r_c h^2}$$

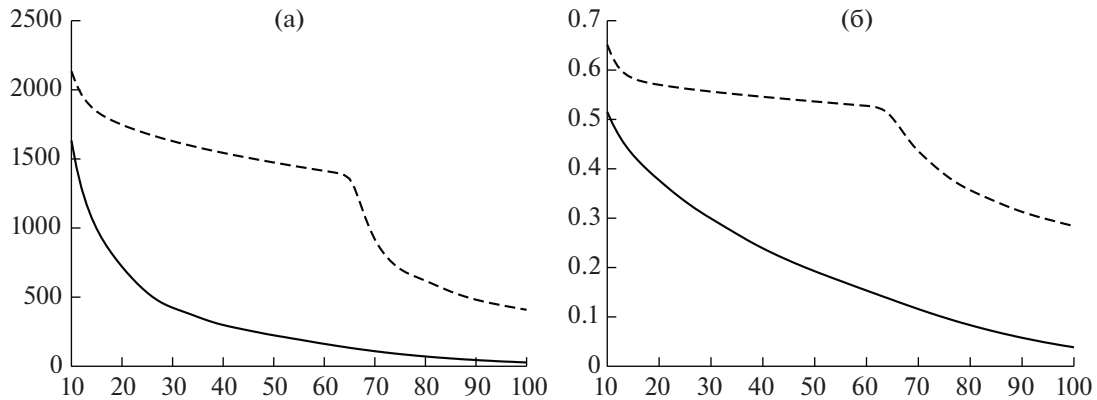
с учетом того, что  $r_c = \max_t r(t) = 0.44, h = 0.1$  и  $\tau = 0.04$ .

Для анализа устойчивости решения прямой задачи была реализована схема Кранка–Николсон с порядком аппроксимации  $O(\tau^2 + h^2)$ .

**3.3. Сравнительный анализ непрерывной и дискретной постановок**

Проведем сравнительный анализ подходов решения обратной задачи: с точным подсчетом градиента для дискретной задачи (6)–(3) и с непрерывной постановкой и дискретизацией при вычислении градиента (2)–(3). Этапы сходства и различия алгоритмов решения указанных обратных задач приведены в таблице 3. Отметим, что в обоих случаях идентично решена прямая задача, а сопряженные задачи решены одной конечно-разностной схемой, которые входят в соответствующие выражения для вычисления градиентов целевого функционала. К решению





**Фиг. 1.** Графики убывания функционалов  $J(q)$  и  $I(q)$  (а) и относительной погрешности  $\delta$  (б) на итерациях с 10-й по 100-ю. Сплошная линия характеризует использование непрерывного градиента, а пунктирная — дискретного градиента целевого функционала.

соответствующих задач минимизации применены одинаковые алгоритмы. В качестве сравнения двух подходов анализируется значение целевого функционала и относительная погрешность восстановленных решений. Решение, полученное с помощью градиентных методов, зависит от начального приближения. Было рассмотрено два варианта выбора начального приближения (см. [19], [22], [23]): решение обратной задачи, полученное с помощью метода глобальной оптимизации роя частиц (МРЧ), представленное в табл. 3, и нулевое начальное приближение. Нулевое начальное приближение выбрано из физической постановки, так как  $q^0$  описывает первую реакцию пользователей на новость в отсутствие дополнительной информации об этой реакции. В приведенной реализации МРЧ функционал  $J(\hat{q}) = 1.114 \times 10^{-3}$  и относительная погрешность  $\delta = 1.533 \times 10^{-3}$ .

В табл. 4 представлены результаты, полученные следующими методами:

- МГС I и МГМ I — методы с непрерывным градиентом и начальным приближением в виде решения, полученного МРЧ;
- МГС II и МГМ II — методы с непрерывным градиентом и нулевым начальным приближением;
- МГС III и МГМ III — методы с дискретным градиентом и начальным приближением в виде решения, полученного МРЧ;
- МГС IV и МГМ IV — методы с дискретным градиентом и нулевым начальным приближением;
- МГМ V — метод с нулевым начальным приближением и дискретным градиентом в случае, когда прямая и сопряженная задачи численно решались конечно-разностной схемой Кранка—Николсон.

Из табл. 4 видно, что МГМ находит решение с заданной точностью, производя более чем в 10 раз меньше итераций, чем МГС при одинаковых начальных условиях и формах нахождения градиента целевого функционала. В случае начального приближения  $\hat{q}$  и непрерывного градиента (т.е. МГС I и МГМ I) времена работ программ не отличаются, однако при других вариантах реализации градиентных методов МГМ находит решение с заданной точностью в несколько раз быстрее МГС. При нулевом начальном приближении методы работают в несколько раз дольше, но достигают большей точности только на шестом знаке после запятой. В случае дискретного градиента МГМ достигает решения того же порядка в несколько раз дольше, а МГС — более чем в 30 раз дольше по времени. При этом МГМ V имеет наименьшее среди всех рассмотренных методов значение погрешности, однако при этом метод работает в несколько раз дольше других реализаций МГМ. На фиг. 1 непрерывными линиями изображены кривые, полученные МГМ II в случае непрерывного вычисления градиента функционала, а пунктирными линиями — кривые, полученные МГМ IV в случае дискретной реализации градиента. Из фиг. 1 следует, что значение функционала и погрешность в случае МГМ II убывают быстрее.

**Таблица 5.** Результаты решения обратной задачи с уровнем зашумленности в данных  $\varepsilon = 1, 5, 10\%$ , полученные многоуровневым градиентным методом с градиентом в дискретной и непрерывной постановках

Начальное приближение	$\varepsilon, \%$	Дискретный градиент (11)		Непрерывный градиент (13)	
		$I(q)$	$\delta$	$J(q)$	$\delta$
Нулевое	1	$1.133 \times 10^{-2}$	$6.706 \times 10^{-4}$	$9.628 \times 10^{-2}$	$7.356 \times 10^{-4}$
	5	$1.475 \times 10^{-1}$	$8.447 \times 10^{-4}$	$3.272 \times 10^{-3}$	$6.399 \times 10^{-4}$
	10	$3.461 \times 10^{-2}$	$7.305 \times 10^{-4}$	$3.286 \times 10^{-3}$	$6.407 \times 10^{-4}$
Решение МРЧ	1	$1.242 \times 10^{-3}$	$6.271 \times 10^{-4}$	$1.403 \times 10^{-3}$	$6.262 \times 10^{-4}$
	5	$3.165 \times 10^{-2}$	$6.431 \times 10^{-4}$	$1.865 \times 10^{-3}$	$6.261 \times 10^{-4}$
	10	16.795	$1.707 \times 10^{-2}$	11.289	$1.388 \times 10^{-2}$

Был проанализирован случай зашумленных данных:

$$f_{ik}^\varepsilon = f_{ik} + \varepsilon \gamma \max |f_{ik}|, \quad i = 1, \dots, N_1, \quad k = 1, \dots, N_2,$$

где  $\varepsilon \in (0, 1)$  – уровень шума, а  $\gamma$  – случайная величина, имеющая стандартное нормальное распределение.

Для зашумленных данных обратной задачи был реализован МГМ с различными начальными приближениями, а прямая задача решалась методом конечных разностей схемой Кранка–Николсон. Результаты численных экспериментов для такой постановки обратной задачи в случае дискретного и непрерывного градиентов приведены в табл. 5. Можно видеть, что в случае нулевого начального приближения метод с дискретным градиентом имеет наименьшее значение относительной ошибки при шуме в 1%, в то время как метод с непрерывным градиентом – при шуме в 5 и 10%. В случае начального приближения в виде решения, полученного с помощью МРЧ, точности восстановленных решений обратной задачи  $\delta$  различаются незначительно.

В работе проведено исследование двух подходов к решению задачи восстановления источника, в рамках которых были реализованы две конечно-разностные схемы, в которых порядок аппроксимации по пространственной переменной совпадает и равен 2, а по временной имеют 1-й и 2-й (Кранка–Николсон) порядки. В таблицах 5 и 6 экспериментально показано, что результат сходимости не чувствителен к порядку аппроксимации по временной переменной. Исходя из оценок сходимости и итерационных выражений определения решения обратной задачи, порядок аппроксимации по пространственной переменной при решении прямой и сопряженной задач имеет большое влияние на точность полученного решения и, скорее всего, на скорость сходимости градиентных методов. В случае определения плотности начального распределения пользователей относительно популярной новости в онлайн социальных сетях с большой пропускной способностью (Twitter, Facebook, Reddit и др.) требуется качественно определить тип вовлеченности пользователей и структуру распределения. Для получения более точных оценок взаимосвязи скорости сходимости градиентных методов с точностью аппроксимации необходимы отдельные исследования.

## ЗАКЛЮЧЕНИЕ

В работе проведен сравнительный анализ численного решения задачи об источнике для диффузионно-логистической модели по данным о диффузионном процессе в фиксированные моменты времени и точках пространства градиентными методами для случаев непрерывной (классической) и дискретной постановок. Получены выражения для вычисления градиента целевого функционала в случае двух постановок, связанных с решением соответствующих сопряженных задач.

Прямая и сопряженные задачи были решены с помощью явной схемы с порядком аппроксимации  $O(\tau + h^2)$  и соблюдением условия Куранта–Фридрихса–Леви и полунявной схемы Кранка–Николсон, имеющей порядок аппроксимации  $O(\tau^2 + h^2)$ . Задача минимизации целевого функционала была решена методами градиентного спуска с постоянным параметром спуска и многоуровневым градиентным методом. Были проанализированы два варианта начального приближения решения обратной задачи: нулевое и полученное методом роя частиц. Показано, что в случае ненулевого начального приближения с использованием выражения для градиента целевого

функционала в непрерывной постановке времена работы программ не отличаются, однако при других вариантах реализации градиентных методов многоуровневый градиентный метод находит решение с заданной точностью в несколько раз быстрее метода градиентного спуска. При нулевом начальном приближении методы работают в несколько раз дольше, но достигают большей точности. В случае дискретного градиента многоуровневый градиентный метод достигает решения того же порядка в несколько раз дольше, а метод градиентного спуска — более чем в 30 раз дольше по времени.

Показано, что в случае аппроксимации кубическими сплайнами дискретных функций в прямой задаче различия в точности полученных решений обратных задач для непрерывной и дискретной постановок наблюдаются только в шестом знаке после запятой в случае незашумленных синтетических данных, описывающих динамику вовлеченных пользователей в онлайн социальной сети. Однако машинное время решения обратной задачи в дискретной форме больше, чем непрерывный аналог.

### СПИСОК ЛИТЕРАТУРЫ

1. *Hestenes M.R., Stiefel E.* Methods of conjugate gradients for solving linear systems // J. Res. Natl. Inst. Stand. Technol. 1952. V. 49. № 6. P. 409–436.
2. *Fletcher R., Reeves C.M.* Function minimization by conjugate gradients // Comput. J. 1964. V. 7. № 2. P. 149–154.
3. *Byrd R.H., Nocedal J., Yuan Y.* Global convergence of a class of quasi-Newton methods on convex problems // SIAM J. Numer. Anal. 1987. V. 24. № 5. P. 1171–1190.
4. *Hasanov A.* Simultaneous determination of source terms in a linear parabolic problem from the final overdetermination: Weak solution approach // J. Math. Anal. Appl. 2007. V. 330. № 2. P. 766–779.
5. *Prilepko A.I., Kamynin V.L., Kostin A.B.* Inverse source problem for parabolic equation with the condition of integral observation in time // J. Inverse Ill-Posed Probl. 2018. V. 26. № 4. P. 523–539.
6. *Aida-zade K.A., Rahimov A.B.* Numerical solution to inverse source problems for linear parabolic equation // IFAC-PapersOnLine. 2018. V. 51. № 30. P. 231–236.
7. *Cheng J., Liu J.* An inverse source problem for parabolic equations with local measurements // Appl. Math. Lett. 2020. V. 103. P. 106213.
8. *Карчевский А.Л.* Корректная схема действий при численном решении обратной задачи оптимизационным методом // Сиб. журн. вычисл. матем. 2008. Т. 11. № 2. С. 139–149.
9. *Евтушенко Ю.Г.* Оптимизация и быстрое автоматическое дифференцирование. М.: ВЦ РАН, 2013.
10. *Евтушенко Ю.Г., Засухина Е.С., Зубов В.И.* О численном подходе к оптимизации решения задачи Бюргера с помощью граничных условий // Ж. вычисл. матем. и матем. физ. 1997. Т. 37. № 12. С. 1449–1458.
11. *Евтушенко Ю.Г.* Приближенный расчет задач оптимального управления // Приклад. матем. и механика. 1970. Т. 34. № 1. С. 95–104.
12. *Черноусько Ф.Л., Колмановский В.Б.* Вычислительные и приближенные методы оптимального управления // Итоги науки и техн. Сер. Мат. анализ. 1977. Т. 14. С. 101–166.
13. *Wang H., Wang F., Xu K.* Diffusive logistic model towards predicting information diffusion in online social networks // Proceed. 32nd Inter. Conf. on Distribut. Comput. Syst. Workshop. 2012. P. 133–139.
14. *Krivorotko O., Zvonareva T.* Numerical solution of the inverse problem for diffusion-logistic model arising in online social networks // Commun. Comput. Info. Sci. 2021. V. 1476. P. 444–459.
15. *Krivorotko O., Kabanikhin S., Zhang S., Kashtanova V.* Global and local optimization in identification of parabolic systems // J. Inverse Ill-Posed Probl. 2020. V. 28. № 6. P. 899–913.
16. *Владимиров В.С.* Уравнения математической физики. М.: Наука, 1981.
17. *Алифанов О.М., Артюхин Е.А., Румянцев С.В.* Экстремальные методы решения некорректных задач. М.: Наука, 1988.
18. *Kabanikhin S., Penenko A.* Gradient-type methods in inverse parabolic problems // J. Phys. Conf. Ser. 2008. V. 135. P. 012054.
19. *Гасников А.В.* Современные численные методы оптимизации. Метод универсального градиентного спуска. М.: МФТИ, 2018.
20. *Nesterov Y., Gasnikov A., Guminov S., Dvurechensky P.* Primal–dual accelerated gradient methods with small-dimensional relaxation oracle // Optim. Methods Softw. 2020. P. 1–38.
21. *Gasnikov A.V., Nesterov Y.E.* Universal method for stochastic composite optimization problems // Comput. Math. Math. Phys. 2018. V. 58. № 1. P. 48–64.
22. *Kaltenbacher B., Neubauer A., Scherzer O.* Iterative regularization methods for nonlinear ill-posed problems. New York: De Gruyter, 2008.
23. *Kaltenbacher B.* All-at-once versus reduced iterative methods for time dependent inverse problems // Inverse Probl. 2017. V. 33. P. 064002.