

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«НОВОСИБИРСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»



Новосибирск
2021

Всероссийская научная
конференция молодых ученых,
посвященная году науки и технологий в России

НАУКА ТЕХНОЛОГИИ ИННОВАЦИИ

Сборник
научных трудов

Часть 2

г. Новосибирск, 06 – 10 декабря 2021 г.

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«НОВОСИБИРСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

НАУКА ТЕХНОЛОГИИ ИННОВАЦИИ

**XV Всероссийская научная конференция молодых
ученых, посвященная Году науки и технологий в России
г. Новосибирск, 06-10 декабря 2021 г.**

**Сборник научных трудов
в 10-и частях**

Часть 2

НОВОСИБИРСК
2021

УДК 62(063)
ББК 72.5я431
Н34

Н34 НАУКА. ТЕХНОЛОГИИ. ИННОВАЦИИ: XV Всероссийская научная конференция молодых ученых, посвященная Году науки и технологий в России (г. Новосибирск, 06-10 декабря 2021 г.) // Сборник научных трудов в 10 ч. / Под ред. Соколовой Д.О. – Новосибирск: Изд-во НГТУ, 2021.

ISBN 978-5-7782-4562-4

Часть 2: – 298 с.

ISBN 978-5-7782-4564-8

В сборнике опубликованы материалы научных исследований, представленных на XV Всероссийской научной конференции молодых ученых, посвященной Году науки и технологий в России.

В части 1 сборника представлены материалы по направлению «Информатика, автоматика, вычислительная и измерительная техника». В части 2 сборника представлены материалы по направлению «Информационные технологии математического моделирования и обработки данных». В части 3 сборника представлены материалы по направлению «Технология, оборудование и автоматизация машиностроительных производств. Материаловедение, технологические процессы и аппараты». В части 4 сборника представлены материалы по направлению «Энергетика». В части 5 сборника представлены материалы по направлению «Электротехника, электромеханика и электротехнологии». В части 6 сборника представлены материалы по направлению «Электроника и биомедицинская техника». В части 7 сборника представлены материалы по направлению «Экономика и управление». В частях 8 и 9 сборника представлены материалы по направлению «Гуманитарные науки и современность». В части 10 сборника представлены материалы по направлению «Актуальные проблемы авиаракетостроения».

ОРГАНИЗАЦИОННЫЙ КОМИТЕТ

Председатель Оргкомитета:

Батаев А.А., д.т.н., профессор, ректор НГТУ

Зам. председателя Оргкомитета:

Брованов С.В., д.т.н., доцент, проректор по научной работе НГТУ

Программный комитет:

Драгунов В.П., д.т.н., профессор, начальник ОПК ВК

Корель И.И., к.ф.-м.н., доцент, декан ФТФ

Осьмук Л.А., д.соц.н., профессор, директор ИСТР

Рева И.Л., к.т.н., доцент, декан АВТФ

Мелехина Е.А., к.п.н., доцент, декан ФГО

Саленко С.Д., д.т.н., профессор, декан ФЛА

Тимофеев В.С., д.т.н., доцент, декан ФПМИ

Овчинникова Н.Н. к.э.н., декан ФБ

Стрельцов С.А., к.т.н., доцент, декан РЭФ

Вильбергер М.Е., к.т.н., доцент, декан ФМА

Тюрин А.Г., к.т.н. доцент, декан МТФ

Русина А.Г., д.т.н. доцент, декан ФЭН

УДК 62(063)

ББК 72.5я431

ISBN 978-5-7782-4564-8 (Ч.2)

© Коллектив авторов, 2021

ISBN 978-5-7782-4562-4

© Новосибирский государственный
технический университет, 2021

Секция МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ,
АНАЛИЗ И ОБРАБОТКА ДАННЫХ

**УСТОЙЧИВЫЙ АЛГОРИТМ
НЕПАРАМЕТРИЧЕСКОЙ ИДЕНТИФИКАЦИИ
ЛИНЕЙНЫХ ТЕХНИЧЕСКИХ СИСТЕМ**

В.А. Боева

**Новосибирский государственный архитектурно-строительный
университет, г. Новосибирск, v.boyeva@sibstrin.ru
Научный руководитель: Воскобойников Ю.Е., д-р физ.-мат.
наук, профессор**

Исследование посвящено построению устойчивого алгоритма непараметрической идентификации линейных стационарных динамических систем при произвольном входном воздействии на основе аппарата сглаживающих кубических сплайнов с особым типом краевых условий. Большое внимание уделяется выбору параметра сглаживания исходя из наличия или отсутствия априорной информации о характеристиках шума измерений.

The research presented is focused on the construction of stable non-parametric identification algorithm for stationary linear dynamic systems when an input signal is arbitrary. The algorithm based on smoothing cubic splines with special boundary conditions. Much attention is given to smoothing parameter evaluation depending on the availability or the lack of a priori information about measurement noise characteristics.

Для описания математических моделей линейных стационарных динамических систем часто применяют интегральное уравнение Вольтерра I рода, решение которого по Адамару является некорректно поставленной задачей [1]. Дифференцируя это уравнение по переменной t и выполняя несложные преобразования, можно получить интегральное уравнение Вольтерра II рода [2]:

$$k(t) + \frac{1}{\varphi(0)} \cdot \int_0^t \varphi'_i(t - \tau) \cdot k(\tau) d\tau = \frac{f'(t)}{\varphi(0)}, t \in [0, T], k(t) = 0 \text{ при } t < 0, \quad (1)$$

где $k(t)$ – импульсная переходная функция (ИПФ) системы, $\varphi(t, \tau)$ – входной сигнал системы (ядро уравнения), $f(t)$ – выходной сигнал системы (правая часть). Задача непараметрической идентификации для модели, описываемой уравнением (1), заключается в построении оценки ИПФ $k(t)$ по зарегистрированным

(или заданным) значениям входного сигнала $\varphi(t, \tau)$ и выходного сигнала $\varphi(t, \tau)$.

На практике значения исходных сигналов системы искажаются шумами измерений и допускают представление [3]:

$$\tilde{f}(t_i) = f(t_i) + \eta_i, \tilde{\varphi}(\tau_i) = \varphi(\tau_i) + \xi_i, i = 1 \dots N, \quad (2)$$

где $f(t)$, $\varphi(\tau)$ – выходной и входной сигналы соответственно, измеряемые в узлах t_i и τ_i со случайными погрешностями η_i и ξ_i , N – количество узлов измерений. Очевидно, что при вычислении производных от зашумлённых сигналов $\tilde{f}(t)$, $\tilde{\varphi}(\tau)$ решение уравнения (1) может не существовать или быть неустойчивым относительно правой части [1,4]. Для вычисления устойчивых производных от зашумлённых сигналов (2) и вычисления оценки искомой ИПФ $\hat{k}(t)$ на основе уравнения (1) предлагается использовать аппарат сглаживающих кубических сплайнов (СКС), подробно описанный в работах [5,6].

Алгоритм идентификации ИПФ $k(t)$ при произвольном входном воздействии $\tilde{\varphi}(\tau)$ будет включать следующие этапы.

Этап 1. Предварительная обработка зашумлённых значений входного сигнала $\tilde{\varphi}(\tau)$ и выходного сигнала $\tilde{f}(t)$.

Как показано в [7], в случае искажения исходных данных импульсными шумами или однородным шумом уровня более 10% целесообразно проводить предварительную фильтрацию таких сигналов с целью снижения ошибки искомого решения. Если шумы измерений η_i , ξ_i содержат импульсные составляющие, в несколько раз превосходящие по амплитуде уровень точного сигнала, предварительная обработка зашумлённых сигналов $\tilde{f}(t)$, $\tilde{\varphi}(\tau)$ осуществляется с помощью комбинированного локально-пространственного фильтра [7]. Если шумы измерений η_i , ξ_i однородные, а их относительный уровень более 10%, то предварительная обработка зашумлённых сигналов $\tilde{f}(t)$, $\tilde{\varphi}(\tau)$ осуществляется с помощью пороговых алгоритмов вейвлет-фильтрации [7].

Этап 2. Вычисление СКС с их устойчивыми производными $S'_{\tilde{\varphi}, \alpha_w}(t)$, $S'_{\tilde{f}, \alpha_w}(t)$ по отфильтрованным на предыдущем этапе значениям $\hat{f}(t)$, $\hat{\varphi}(\tau)$.

Построение СКС для этого алгоритма подобно рассматривается в [5,6] и включает следующие моменты.

Во-первых, это задание комбинированных краевых условий [8], которые актуальны при решении практических инженерных задач и задаются исходя из их специфики, в вариациях:

- слева – значение первой производной, справа – нулевая вторая производная;
- слева – нулевая вторая производная, справа – значение первой производной.

Во-вторых, это вычисление оценки α_w параметра сглаживания СКС. При решении модельных задач, когда дисперсия шумов измерений достоверно известна, значение α_w вычисляется на основе статистического критерия оптимальности [9]. При решении практических задач по экспериментальным данным, когда информация о числовых характеристиках шумов измерений отсутствует, значение α_w вычисляется на основе метода L -кривой [10].

С учётом вычисленных СКС $S'_{\hat{\phi},\alpha_w}(t)$, $S'_{\hat{f},\alpha_w}(t)$ по отфильтрованным входному и выходному сигналам уравнение (1) примет вид:

$$k(t) + \frac{1}{S_{\hat{\phi},\alpha_w}(t_1)} \cdot \int_0^t S'_{\hat{\phi},\alpha_w}(t-\tau) \cdot k(\tau) d\tau = \frac{S'_{\hat{f},\alpha_w}(t)}{S_{\hat{\phi},\alpha_w}(t_1)}, t \in [0, T]. \quad (3)$$

Этап 3. Вычисление интеграла свёртки $\int_0^t S'_{\hat{\phi},\alpha_w}(t-\tau) \cdot k(\tau) d\tau$ из уравнения (3).

При аппроксимации интеграла свёртки получим:

$$\int_0^{t_j} S'_{\hat{\phi},\alpha_w}(t_j-\tau) \cdot k(\tau) d\tau \approx \sum_{i=1}^{j-1} k_i \cdot \left[\int_{t_i}^{t_{i+1}} S'_{\hat{\phi},\alpha_w}(t_j-\tau) d\tau \right], i = 1 \dots N, j = 2 \dots N. \quad (4)$$

Далее вводится матрица Φ' размерности $(N-1) \times (N-1)$, элементы которой для $i = 1 \dots N-1, j = 1 \dots N-1$ определяются по правилу:

$$\Phi'_{i,j} = \begin{cases} \int_{t_{i+1-j}}^{t_{i+2-j}} S'_{\hat{\phi},\alpha}(\tau) d\tau, & \text{если } j \leq i, \\ 0, & \text{если } j > i. \end{cases} \quad (5)$$

Этап 4. Вычисление оценки ИПФ $\hat{k}(t)$:

$$k = \left(I + \frac{1}{S_{\hat{\phi},\alpha_w}(t_1)} \cdot \Phi' \right)^{-1} \cdot \frac{\hat{f}'}{S_{\hat{\phi},\alpha_w}(t_1)}, \quad (6)$$

где I – единичная матрица размерности $(N-1) \times (N-1)$, а проекции вектора k , являются оценками для значений ИПФ идентифицируемой системы.

В работах автора [2,6-10] приводятся результаты многочисленных вычислительных экспериментов, посвящённых исследованию эффективности построенного алгоритма. Можно наблюдать, что даже в случаях сильно зашумлённых исходных данных (при относительном уровне исходного шума 20% и более), или искажения импульсными и сильно коррелированными шумами, относительная ошибка идентификации не превышала 15%.

Высокая точность вычисленной оценки ИПФ достигается благодаря нескольким факторам. Во-первых, предварительная фильтрация зашумлённых исходных сигналов системы позволяет снизить значение относительной ошибки идентификации в среднем на 5-7%. Во-вторых, комбинированные краевые условия учитывают специфику каждой конкретной задачи. В-третьих, эффективные подбор наилучшей оценки параметра сглаживания помогает снизить уровень среднеквадратической ошибки сглаживания, что также повышает точность искомой оценки ИПФ.

Благодаря используемому подходу на основе обращения интегрального уравнения Вольтерра I рода имеем хорошо обусловленную матрицу СЛАУ (6), и на точность идентификации оказывают влияние только ошибки дифференцирования входного сигнала, определяемые элементами матрицы Φ' , и ошибки дифференцирования выходного сигнала, определяемые проекциями вектора \hat{f}' . Предложенный алгоритм апробирован и доказал свою эффективность при решении практических прикладных задач

непараметрической идентификации для теплофизических, теплоэнергетических систем и систем вентилирования.

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 20-38-90041. Funding: The reported study was funded by RFBR, project number 20-38-90041.

Литература:

1. *Hadamard J.* Le probleme de Cauchy et les equations aux derives particlee lineaires hyperbolique. – Paris: Hermann, 1932.

2. *Боева В.А.* Алгоритм устойчивого обращения уравнения Вольтерра в задачах идентификации // Материалы XII Всероссийской научно-технической конференции «Актуальные вопросы архитектуры и строительства». – Новосибирск: Изд-во НГАСУ (Сибстрин), 2019. – С. 376-381.

3. *Сизиков В.С.* Математические методы обработки результатов измерений. – СПб.: Политехника, 2001. – 240 с.

4. *Васин В.В.* Методы решения неустойчивых задач. – Екатеринбург: Наука, 1989. – 248 с.

5. *Завьялов Ю.С., Квасов Б.И., Мирошниченко В.Л.* Методы сплайн-функций. – Москва: Наука, 1980. – 345 с.

6. *Воскобойников Ю.Е., Боева В.А.* Исследования эффективности использования сглаживающих кубических сплайнов в задачах непараметрической идентификации // Автоматика и программная инженерия. – 2019. – № 4 (30). – С. 58-64.

7. *Боева В.А.* О целесообразности предварительной фильтрации зашумлённых сигналов в задачах идентификации // Инженерно-строительный вестник Прикаспия. – 2019. – № 4 (30). – С. 141-145.

8. *Воскобойников Ю.Е., Боева В.А.* Дескриптивное сглаживание сигнала в одном алгоритме непараметрической идентификации технических систем // Современные наукоёмкие технологии. – 2020. – № 7. – С. 24-28. – doi: 10.17513/snt.38128.

9. *Voskoboynikov Yu., Boeva V.A.* Synthesis of smoothing cubic spline in non-parametric identification technical systems' algorithm // IOP Conference series: Materials Science and Engineering XIII International Scientific Conference Architecture and Construction 2020. – 953 (2020) 012035. – doi: 10.1088/1757-899X/953/1/012035.

10. *Воскобойников Ю.Е., Боева В.А.* Выбор параметра сглаживающего кубического сплайна при неизвестных характеристиках шума измерений // Современные наукоёмкие технологии. – 2021. – № 4. – С. 25-29. DOI: 10.17513/snt.38729

РЕЗУЛЬТАТЫ КОМПЬЮТЕРНОГО ЧИСЛЕННОГО ЭКСПЕРИМЕНТА ПО КИНЕМАТИЧЕСКОЙ ДИАГНОСТИКЕ ВЕРТИКАЛЬНО-НЕОДНОРОДНЫХ СРЕД В УСЛОВИЯХ ЛИНЕАРИЗАЦИИ ФУНКЦИИ СКОРОСТИ

А.А. Вохмин

**Новосибирский государственный технический университет,
г. Новосибирск, alexandervokhmin@gmail.com
Научный руководитель: Зеркаль С.М., д.т.н., профессор**

Изучается решение обратной кинематической задачи рассеяния в линеаризованной постановке для вертикально-неоднородных сред. Приводится вывод формулы обращения для вычисления показателя преломления. По составленному алгоритму выполнено тестовое компьютерное моделирование численного эксперимента, результаты которого приводятся в работе.

We study the solution of the inverse kinematic scattering problem in a linearized formulation for vertically inhomogeneous media. The derivation of the inversion formula for calculating the refractive index is given. A test computer simulation of a numerical experiment was carried out according to the developed algorithm, the results of which are presented in the work.

В данной работе рассматривается обратная кинематическая задача, использующая кинематическую информацию о временах прихода волн, рассеянных в заданных направлениях. Эту задачу естественно назвать обратной кинематической задачей рассеяния.

Используется следующая система наблюдений: в плоскости xz в точке $(0,0)$ находится источник сигнала, в точке $(1,0)$ регистрируется время прихода рассеянной на неоднородностях среды волны, причем известно α_1 – угол входа луча из источника и α_2 – угол прихода регистрируемого луча (рисунок 1). Среда имеет скорость распространения сигнала $c(z)$, которая подлежит определению.

В отличие от общеизвестной обратной кинематической задачи сейсмологии, данная постановка применима в тех случаях, когда возможно генерирование узколокализованных «шнуровых» решений волнового уравнения. Компенсацией этого недостатка является возможность однозначного определения волноводов. Измерение такой информации экспериментально следует из работы [1].

В математическом отношении задача сводится к решению следующего нелинейного интегрального уравнения первого рода:

$$A(c) = \int_{\Gamma(x)} \frac{dS}{c(z)} = T(x), \quad \Gamma(x) = \Gamma_1(x) \cup \Gamma_2(x)$$

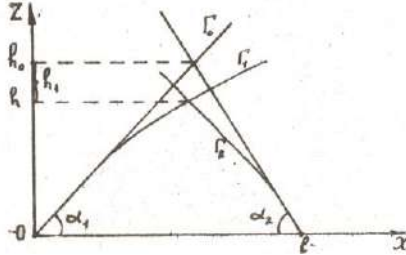


Рис.1 – Система наблюдений обратной кинематической задачи рассеяния

$\Gamma_0(x)$ – лучевая траектория при отсутствии неоднородности

$\Gamma_1(x)$ – лучевая траектория зондирующего сигнала

$\Gamma_2(x)$ – лучевая траектория рассеянного сигнала

Рассмотрим случай $\alpha_1 = \alpha_2 = \alpha$ и будем рассматривать линейризованную постановку. Тогда справедливо

$$T - T_0 = z \int_{\Gamma_0} a(z) dS.$$

Здесь T_0 – время прихода рефрагированной волны при отсутствии неоднородности, T – фактическое время, а Γ_0 выражается уравнением:

$$\Gamma_0 = z \cdot ctg\alpha.$$

Тогда

$$T(\alpha) - T_0(\alpha) = z \int_0^{\frac{l}{2}tg\alpha} a(z) \sqrt{1 + x_0'^2} dz$$

Отсюда

$$a\left(\frac{l}{2}tg\alpha\right) = \frac{\cos^2\alpha}{l} \frac{d}{d\alpha} [\sin\alpha (T(\alpha) - T_0(\alpha))]. \quad (1)$$

Таким образом, меняя угол α , можно определить показатель преломления в исследуемой среде.

Прямая кинематическая задача сейсмоки рассматривается в рамках лучевого метода [2]. В результате ее решения по заданному изменению скорости распространения акустической волны в среде $V(x,y,z)$ строятся лучевые траектории, вдоль которых распространяется исследуемая волна и вычисляются времена пробега волны вдоль луча.

Рассмотрим прямую задачу для модели градиентной среды с линейным законом изменения скорости от глубины:

$$V(z) = l + Bz, \quad \text{где } B = \text{const}, B > 0.$$

Соответственно, коэффициент преломления также меняется линейно:

$$n(z) = l + a(z), \quad \text{где } a(z) = -0,08z.$$

Из рисунка 2 можно заметить, что

$$R^2 = \left(z + \frac{1}{\lambda}\right)^2 + \left(x - \frac{L}{2}\right)^2 \Rightarrow z = -\frac{1}{\lambda} + \sqrt{R^2 - \left(x - \frac{L}{2}\right)^2}. \quad (2)$$

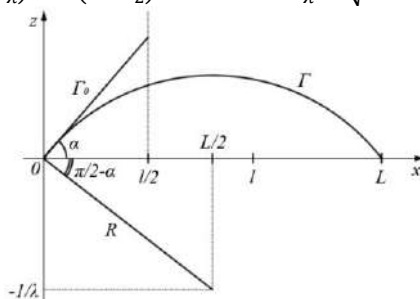


Рис.2 – Иллюстрация к выполнению вычислений

Рассчитываем времена прихода рассеянной волны из источника с координатой $x_0 = 0$ в точку на поверхности с координатой $x_1 = 1$ по формуле [3] с учетом (2). Получаем:

$$T\left(0, \frac{l}{2}\right) = \frac{1}{2\lambda} \ln \frac{\left(R + \frac{l}{2} - \frac{L}{2}\right)\left(R + \frac{L}{2}\right)}{\left(R - \frac{l}{2} + \frac{L}{2}\right)\left(R - \frac{L}{2}\right)}.$$

Для расчета T_0 используется формула:

$$T_0(\alpha) = \frac{l}{\cos \alpha}.$$

В результате проведения расчетов времен прихода волны удалось получить графики, представленные на рисунке 3.

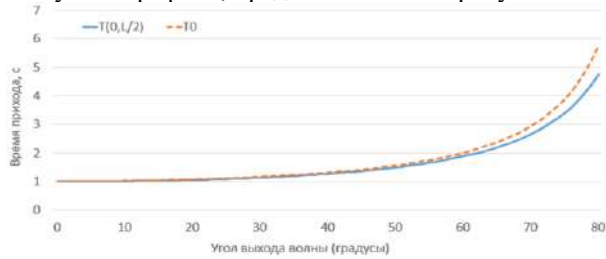


Рис.3 – Сравнение графиков времен прихода волны

Решая обратную задачу вспомним, что для таких сред справедлива следующая формула, связывающая $a(z)$ с T , α и l :

$$a(z(\alpha)) = \frac{\cos^2 \alpha}{2l} \frac{d}{d\alpha} [\sin \alpha \cdot (T(\alpha) - T_0(\alpha))], \quad (3)$$

Проведя математические расчеты удалось получить график изменения коэффициента преломления с изменением величины z и сравнить его с точным решением (рисунок 4).

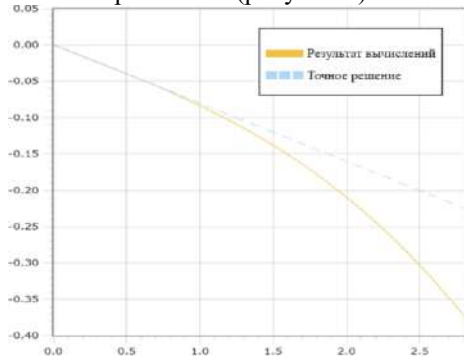


Рис.4 – Сравнение результатов вычисления с точным решением

Проведенные численные эксперименты позволили установить необходимость введения в знаменатель формулы (1 и 3) обращения

эмпирического коэффициента равного двум. Введение этого изменения существенно улучшило результат решения обратной задачи. По-видимому, это связано с процедурой линеаризации функции скорости естественно вносящей отклонение в точность определения функции $a(z)$. Предполагается уточнить формулу обращения выполнением более детального подхода при осуществлении вывода формулы обращения что позволит повысить устойчивость и достоверность результатов ее использования.

Литература:

1. Машошин А.И. Практические задачи гидроакустики, решаемые с использованием алгоритмов обработки сигналов, согласованных со средой их распространения (обзор). Фундаментальная и прикладная гидрофизика, 2017, Т. 10, № 1
2. Lavrent'ev M.M., Zerkal S.M., Trofimov O.E. Computer Modeling in Tomography and Ill-Posed Problems. Utrecht – Boston – Koln – Tokyo: VSP, 2001
3. Смирнов В.И., Курс высшей математики. – М.: Наука, 1968, Т. 1. – 400 с.

ИССЛЕДОВАНИЕ ВЛИЯНИЯ НАРУШЕНИЯ ГОМОСКЕДАСТИЧНОСТИ НА РАСПРЕДЕЛЕНИЕ ROC AUC И СВЯЗАННЫХ С НИМ ХАРАКТЕРИСТИК, ПОЛУЧАЕМЫХ В РЕЗУЛЬТАТЕ ПРОВЕДЕНИЯ ЛИНЕЙНОГО ДИСКРИМИНАНТНОГО АНАЛИЗА

И.В. Грициенко

**Новосибирский государственный технический университет,
г. Новосибирск, gricienko.2016@stud.nstu.ru
Научный руководитель: Волкова В.М., к.т.н.**

В статье рассматривается проблема выявления устойчивости критериев, использующих значения специфичности, чувствительности и ROC AUC к изменению характеристик входных данных. Для решения задачи классификации используется линейный дискриминантный анализ, результаты которого в медицине оцениваются в первую очередь по критериям качества, основанных на ROC AUC. Исследования проводятся с помощью метода Монте-Карло, строятся модели распределений рассматриваемых характеристик.

The article discusses the problem of identifying the stability of criteria using the values of specificity, sensitivity and ROC AUC to changes in the characteristics of the input data. To solve the classification problem, linear discriminant analysis is used, the results of which in medicine are evaluated primarily according to quality criteria based on ROC AUC. Research is carried out using the Monte Carlo method, models of the distributions of the characteristics under consideration are built.

Предметная область работы связана с медицинскими данными, практическая задача была вызвана необходимостью обоснованного прогноза наличия анемии плода по значению кардиоторакального индекса (КТИ) плода у беременных женщин на разных сроках беременности. В качестве инструмента анализа было выбрано построение дискриминантной модели, которая продемонстрировала высокую прогнозную ценность. Но одним из базовых предположений дискриминантного анализа является равенство ковариационных матриц дискриминируемых совокупностей. Однако ожидать, что при наличии анемии разброс по значениям решающего признака КТИ будет такой, как и для плодов, находящихся в пределах нормы, не приходится. Поэтому было принято решение провести средствами компьютерного моделирования по методике Монте-Карло исследование границ корректности построенной модели с точки зрения нарушения предположения о гомоскедастичности и влияния этого условия на распределения характеристик, по которым оценивается прогностическая значимость модели в медицине.

При проведении исследования отталкивались от структуры данных, соответствующей реальной практической задаче, поэтому в качестве входных параметров для моделирования наборов данных рассматриваются модели КТИ, полученные путем регрессионного анализа на основе предоставленных реальных данных.

В условиях проведения исследования для достаточного количества данных используется их моделирование по полученным моделям:

$$CTR_n = 0.206 + 0.0061 * W + e_1, \quad (1)$$

где: W - номер недели, на котором производится осмотр (находится в диапазоне от 10 до 35); e_1 - шум, $e_1 \in N(0, 0.032^2)$.

$$CTR_a = 0.39 + 0.0042 * W + e_2, \quad (2)$$

где: W - номер недели, на котором производится осмотр (находится в диапазоне от 10 до 35); e_2 - шум, $e_2 \in N(0, 0.059^2)$.

Для определения принадлежности результатов генерации к классам используется алгоритм бинарной классификации – линейный дискриминантный анализ [1, 2]. Он разделяется на две стадии: обучения модели и прогнозирования. После стадии прогнозирования имеем матрицу ошибок, которая отображает верные прогнозы и ошибки первого и второго рода. На ее основе вычисляются важные для медицинской отрасли показатели [3] чувствительности (Sn) и специфичности (Sp), а также ROC AUC. Ниже приведены формулы для вычисления данных показателей:

$$Sn = \frac{TP}{TP + FN}, \quad (3)$$

где: TP - объекты, верно отнесенные классификатором к рассматриваемому классу; FN - объекты, принадлежащие к рассматриваемому классу, но не отнесенные к нему классификатором (ошибка первого рода).

$$Sp = \frac{TN}{TN + FP}, \quad (4)$$

где TN - объекты, не принадлежащие к рассматриваемому классу, но не отнесенные к нему классификатором, FP - объекты, не принадлежащие к рассматриваемому классу, но отнесенные к нему классификатором (ошибка второго рода).

ROC-кривая представляет собой график, позволяющий оценить бинарную классификацию. Он отображает соотношение между величинами Sn и $1 - Sp$. Количественная интерпретация ROC-кривой (AUC) – это площадь под кривой, ограниченная снизу осью $1 - Sp$.

Были проведены исследования (см. табл. 1), определяющие степень зависимости результирующих критериев от входных показателей шума. Для этого изменялась величина среднеквадратичного отклонения, используемого для генерации шума. Ниже приведены усредненные значения, полученные при использовании метода Монте-Карло.

Исходное соотношение, наблюдаемое для предоставленных данных, $\sigma_n/\sigma_a = 0.535$. При этом $\sigma_n + \sigma_a = 0.091$.

Таблица 1 – Зависимость критериев качества от отношения среднеквадратичных отклонений

σ_n/σ_a	Sp	Sn	AUC
0.1	0.843	1.000	0.953
0.2	0.863	0.999	0.964
0.4	0.894	0.981	0.976
0.535 (исходное)	0.923	0.952	0.980
0.8	0.959	0.901	0.985
1	0.978	0.858	0.985
1.2	0.988	0.825	0.985

Как видно из таблицы, при увеличении разброса отклонения для модели нормального типа и соответствующем этому уменьшении разброса отклонения для модели анемичного типа, увеличивается значение специфичности. Данная зависимость является обратно пропорциональной для чувствительности.

Следующим этапом является исследование принадлежности полученных распределений Sn , Sp и ROC AUC к существующим классам распределений, и влияния на это отношения дисперсий.

Для осуществления этого используется подход, при котором исходная выборка разделяется на три. На первой выборке специальный алгоритм посредством перебора подбирает наилучшие параметры для каждого из предоставленных ему распределений [4, 5], основываясь на методе максимального правдоподобия. На второй выборке происходит проверка принадлежности к распределению по критерию согласия Колмогорова-Смирнова. Результаты p-value сравниваются с выбранным уровнем значимости ($\alpha = 0.05$), после чего отсеиваются неудовлетворяющие этому условию распределения. На третьей выборке вычисляется расстояние Колмогорова, на основании которого и отбираются наиболее подходящие распределения.

Наиболее важным и определяющим фактором для принятия того или иного решения является показатель ROC AUC. Ниже приведена таблица 2 с наилучшими найденными распределениями в условиях гомоскедастичности.

Таблица 2 – Распределения, описывающие данные ROC-AUC наилучшим образом в условиях гомоскедастичности

Распределение	Параметры	p-value	Расстояние Колмогорова
Св-Джонсона	(-1.54, 1.51, 0.92, 0.08)	0.64	0.042
Экспоненциальное Вейбулла	(0.75, 6.36, 0.93, 0.05)	0.62	0.035
SL-Джонсона	(9.99, 3.45, 1.01, 0.01)	0.55	0.041
Обобщенное нормальное	(4.18, 0.93, 0.06)	0.52	0.045
Пирсона III	(-0.88, 0.98, 0.01)	0.51	0.043
Гумбеля	(0.98, 0.01)	0.29	0.051

Дальнейшие исследования показали, что в условиях гетероскедастичности согласие с предложенными распределениями не выполняется. Однако, вышеупомянутые классы распределений все еще наилучшим образом описывают рассматриваемые данные при условии подобранных параметров под каждый случай гетероскедастичности.

Была реализована процедура линейного дискриминантного анализа и с помощью метода Монте-Карло были исследованы зависимости значений специфичности, чувствительности и ROC AUC от входных значений параметров. Проведен поиск принадлежности результатов критериев к известным классам распределений, исследованы и зафиксированы аналогичные прошлому пункту зависимости.

Литература:

1. Сошникова Л.А., Тамашевич В.Н. Многомерный статистический анализ в экономике. – М.: ЮНИТИ-ДАНА, 1999. – 599 с.
2. Ким Дж.О., Мьюллер Ч.У. Факторный, дискриминантный и кластерный анализ. – М.: Финансы и статистика, 1989. – 215с.
3. Гланц С. Медико-биологическая статистика. – М.: Практика, 1998. – 459 с.
4. Лемешко Б.Ю., С.Н. Постовалов. Компьютерные технологии анализа данных и исследование статистических законов. – Новосибирск: Изд-во НГТУ, 2007. – 57 с.
5. Губарев В.В. Вероятностные модели: Справочник. В 2 ч. Ч. 2. – Новосибирск: Изд-во НЭТИ, 1992. – 232 с.

РАЗРАБОТКА ЭЛЕМЕНТОВ ОБРАТНОЙ СВЯЗИ С ВЫПУСКНИКАМИ ФАКУЛЬТЕТА

В.К. Дворецкая

**Новосибирский государственный технический университет,
г. Новосибирск, *dvoreczkaya.2018@stud.nstu.ru***

Научный руководитель: Стасышин М. В., к.т.н., доцент

Обратная связь с выпускниками является одним из инструментов, помогающих повысить уровень и качество учебного процесса. В статье содержится предварительный анализ полученных данных от выпускников определенных годов и описание архитектуры подсистемы обратной связи с выпускниками.

Feedback from alumni is one of the tools to help improve the level and quality of the educational process. The article contains a preliminary analysis of the data obtained from graduates and a description of the subsystem of feedback to graduates.

Государство в лице Министерства науки и высшего образования РФ в течение последних нескольких лет последовательно добивается заявленной им цели – избавиться от слабых вузов на рынке образовательных услуг. В этой конкурентной борьбе способны победить только те вузы, которые в состоянии обеспечить высокое качество и высокий уровень обучения в вузе.

Качество образования в университете во многом зависит от уровня образовательных программ и содержания учебных дисциплин. Только действительно современные учебные программы могут обеспечить конкурентоспособность университета. Важным инструментом, помогающим повысить уровень и качество учебного процесса, является система обратной связи с выпускниками. Эта система, в частности, позволяет получать информацию не только о карьерном росте выпускников и технологиях, которыми они владеют, но и об IT-технологиях, которые востребованы и актуальны в современном быстро меняющемся мире. Анализ этой информации позволяет критически оценить учебную деятельность на факультете и, при необходимости, скорректировать учебные планы рабочих программ.

Идея проведения работы по сбору и анализу данных о выпускниках факультета возникла в рамках учебной дисциплины «Проектная деятельность» полтора года назад. По счастливой случайности эта работа совпала с приказом по университету № 492 от 20.04.2021г. «О создании единой базы выпускников НГТУ

НЭТИ». Следует отметить, что подобные системы существуют в большинстве российских вузов.

Как и любая другая работа по анализу данных работа по сбору и анализу данных о выпускниках факультета включает следующие этапы:

- решение методологических вопросов по процедуре организации сбора данных;
- подготовка инструментальных средств по сбору данных;
- проектирование и создание базы данных;
- непосредственный сбор данных;
- очистка данных;
- загрузка данных в созданную базу данных;
- анализ данных.

Данные для анализа собираются путем анкетирования выпускников факультета. Помимо контактных данных и сведений о карьерном росте выпускников собираются данные о направлении их профессиональной деятельности, IT-технологиях, которыми выпускники владеют, выпускники дают оценку качества обучения на факультете, учебных дисциплин, которые, с их точки зрения, были наиболее (наименее) полезны для их производственной деятельности, преподавателей, оставивших у них позитивные (негативные) воспоминания. При установлении контактов используются адреса в социальных сетях ВКонтакте, Facebook, LinkedIn, Instagram, телефонная связь, электронная почта. Процесс сбора данных от выпускников факультета разбит на два этапа:

- выпускники, поступавшие на факультет ФПМИ на направления бакалавриата с 2008 по 2015 годы; данный этап завершен, за указанный период обучение факультет успешно закончили 557 выпускников, в опросе участвовало 154 выпускников, что составляет 27,6%; данный процент, а также охват выпускников восьми потоков говорит о достаточно высокой репрезентативности выборки;
- выпускники, поступавшие на факультет ФПМИ на направления бакалавриата с 1998 по 2007 годы; за указанный период обучение факультет успешно закончили более 1180 выпускников; в настоящее время идет работа с выпускниками этой группы.

Подсистема обратной связи с выпускниками является компонентом Информационной системы (ИС) университета. Это означает, что в ней используется единая с ИС система безопасности и авторизации пользователей. Последнее особенно важно,

поскольку собираемые сведения подпадают под категорию персональных данных, и защищаются наряду с другими персональными данными в ИС.

Собранные через подготовленную в корпоративной подсистеме опросов форму сведения хранятся в структурах Информационной системы университета (рис. 1). Далее посредством разработанного программного приложения эти данные очищаются, структурируются и заносятся в разработанную на платформе Oracle базу данных подсистемы обратной связи с выпускниками. Для проведения анализа данные в обезличенном виде выгружаются в базу данных СУБД PostgreSQL. При необходимости с помощью специально разработанного API из базы данных Информационной системы в формате JSON могут быть запрошены недостающие для анализа данные.

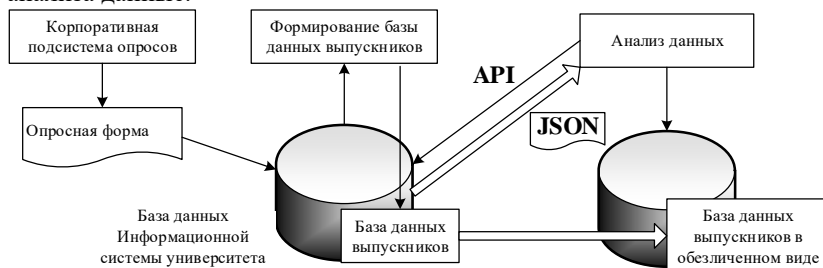


Рис. 1. Архитектура подсистемы обратной связи с выпускниками

Предварительный анализ данных, полученных на первом этапе опроса выпускников, позволяет сделать следующие выводы:

1. *Карьера и опыт работы. Места работы выпускников.*

Из 154 выпускников, участвующих в опросе, сведения о месте своей работы предоставили 93,5% выпускников. Собранные данные позволили

- составить ранжированный перечень наиболее популярных мест работы выпускников; в общей сложности выпускники указали более 180 организаций, где они работали (работают).

- выделить категории программного обеспечения, на разработке которого специализируются организации, где работают (работали) выпускники;

- составить ранжированный список направлений специализации выпускников; отвечая на вопрос о направлении производственной деятельности, в которой они специализируются,

выпускники чаще всего отмечали проектирование и разработка программного обеспечения.

- составить ранжированный перечень наиболее востребованных выпускниками технологий; перечисляя технологии, которыми они владеют, выпускники, участвовавшие в опросе, указали список более чем из 150 технологий;

2. Оценка выпускниками качества обучения в университете

В большинстве выпускники, участвующие в опросе, достаточно высоко оценили уровень и качество обучения факультета. В большей степени это относится к обучению в бакалавриате, обучение в магистратуре выпускниками оценивается ниже.

Доля тех, кто по оценивает обучение на «Хорошо» и «Отлично» составляет в бакалавриате 91%, в магистратуре – 82%. Вместе с тем выпускники зачастую указывают на недостаточно высокий уровень IT-составляющей учебных дисциплин. Данная информация, а также положительные и негативные и критические отзывы выпускников об обучении на факультете должны быть очень важны для руководства факультета и выпускающих кафедр, поскольку именно уровень преподавания IT-дисциплин позволяет обеспечить конкурентноспособность образовательных программ факультета, и, как следствие, профессиональный уровень выпускников на рынке.

Полученные от выпускников сведения позволили сформировать ранжированный перечень дисциплин, которые, по мнению выпускников, оказались наиболее полезными им в их профессиональной деятельности, а также ранжированный список дисциплин, которые у выпускников получили негативную оценку.

Заслуживают внимания также мнения выпускников относительно дисциплин, разделов, технологий, навыков и компетенций, которые они считают целесообразно включить в учебные планы.

3. Оценка работы преподавателей факультета.

Выпускникам предлагалось оценить работу преподавателей факультета, отметив преподавателей, которые оставили у них позитивные и, наоборот, негативные воспоминания. Ранжированные списки преподавателей по этому показателю представлены руководителям выпускающих кафедр.

Следует отметить, что одна количественная оценка (позитивная или негативная) не совсем точно оценивает преподавателя. Соотношение количества позитивных и негативных оценок (их разность, например) является, видимо, более точной мерой, которую можно было бы использовать для оценки преподавателя.

СОЗДАНИЕ ИМИТАЦИОННОЙ МОДЕЛИ В СРЕДЕ SIMINTECH ДЛЯ ИССЛЕДОВАНИЯ ХАРАКТЕРИСТИК СИНХРОННОГО ДВИГАТЕЛЯ С ПОСТОЯННЫМИ МАГНИТАМИ

Н.И. Добшинский

Национальный исследовательский ядерный университет

«МИФИ», г. Москва

Северский технологический институт, г. Северск,

dobschinsky@yandex.ru

Научный руководитель: Федянин А.Л., доцент

В данной статье исследуются характеристики синхронного двигателя с постоянными магнитами с помощью среды моделирования SimInTech.

This article examines the characteristics of a synchronous motor with permanent magnets using the SimInTech simulation environment.

В настоящее время СДПМ применяются в регулируемых электроприводах в системах автоматического регулирования координат промышленных установок и технологических комплексов, робототехнике, регулируемых приводах электротранспорта, где требуется сочетание таких качеств, как высокий момент, малые масса и габариты, высокое быстродействие.

СДПМ обладают рядом преимуществ перед другими типами двигателей, таких как: отсутствие обмотки возбуждения на роторе, что обеспечивает уменьшение электрических потерь, высоким КПД, что обуславливается отсутствием потерь в роторе, лучшим отношением максимальной полезной мощности к массе по сравнению с другими двигателями аналогичной мощности, а также высокими эксплуатационными характеристиками из-за отсутствия скользящих контактов.

Математическое моделирование позволяет рассмотреть возможные режимы работы, а также составить объективный прогноз будущих состояний системы.

При разработке электрических машин, в том числе двигателей используются математические модели, которые позволяют исследовать и улучшать характеристики подобных синхронных машин.

В основе модели любой электрической машины лежат дифференциальные уравнения электрического состояния ветвей,

электромагнитного момента и уравнения движения вращающихся частей.

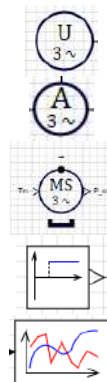
Создание математической модели завершается исследованием её выходных характеристик. Для этого применим систему моделирования SimInTech.

Моделирование СДПМ в SimInTech

С помощью среды моделирования SimInTech я создал имитационную модель двигателя, которая на основе математической модели позволяет исследовать такие характеристики, как: скорость вращения, электромагнитный момент.

Для создания имитационной модели необходимы следующие элементы:

- Источник питания трёхфазный
- Амперметр идеальный трёхфазный
(для снятия показаний токов)
- Двигатель синхронный с постоянными магнитами
- Ступенька
(для создания момента сопротивления)
- Временные графики
(для снятия основных выходных характеристик)



В результате соединения данных элементов получим схему (рисунок 1).

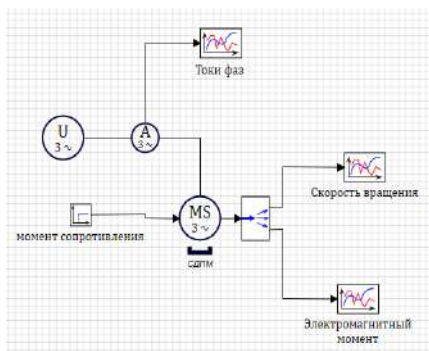


Рис.1 – имитационная модель СДПМ

Для исследования характеристик двигателя необходимо задать следующие параметры:

Таблица 1 – Задаваемые параметры

Наименование параметра	Значение
Напряжение источника напряжения, В	380
Частота источника напряжения, Гц	50
Момент сопротивления, Н·м	2
Номинальная частота вращения двигателя, об/мин	3000
Число пар полюсов	1
Момент инерции, кг·м ²	$2,5 \cdot 10^{-5}$
Сопротивление статора, Ом	15,6

Результаты математического моделирования

На рисунке 2 представлены результаты моделирования СДПМ в среде SimInTech.

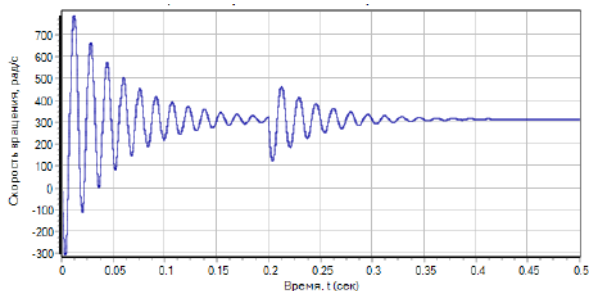


Рис.2 – скорость вращения электродвигателя с подключаемой в момент времени ($t=0,2$) секунды нагрузкой

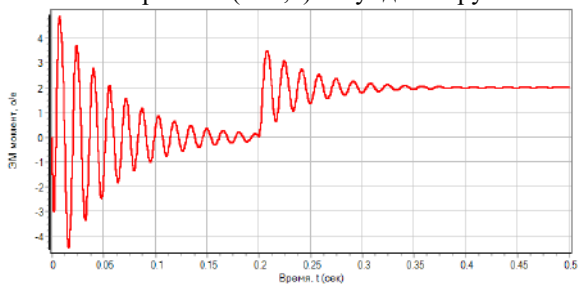


Рис.3 – электромагнитный момент электродвигателя с подключаемой в момент времени ($t=0,2$) секунды нагрузкой

На графиках изображены основные выходные характеристики синхронного двигателя с постоянными магнитами, такие как: зависимости скорости вращения и электромагнитного момента от времени.

По графикам видно, в момент подключения нагрузки (0,2 секунды) происходит отклонение скорости вращения от номинальной, а также мгновенное увеличение электромагнитного момента с его последующим переходом в нормальный режим.

Исследуем зависимость частоты вращения электродвигателя от частоты тока при неизменной нагрузке.

Момент подключения нагрузки установим 0,3 секунды.

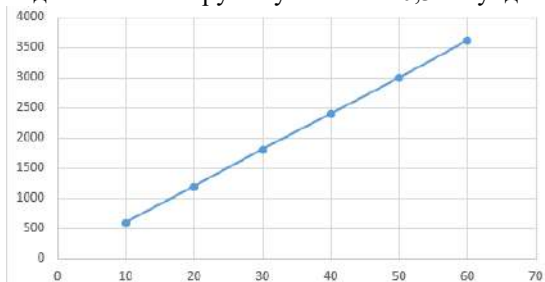


Рис.4 – зависимость частоты вращения вала ЭД от частоты тока

При увеличении частоты тока происходит линейное увеличение количества оборотов.

Таким образом данная среда моделирования позволяет создать имитационную модель, а также снять её основные характеристики.

Литература:

1. Калачёв, Ю. Н. SimInTech : моделирование в электроприводе / Калачёв Ю. Н. - Москва : ДМК Пресс, 2019. - 98 с. - ISBN 978-5-97060-766-4. - Текст : электронный // ЭБС "Консультант студента" : [сайт]. - URL : <https://www.studentlibrary.ru/book/ISBN9785970607664.html> (дата обращения: 02.11.2021).

2. Поляков В.Н. Энергоэффективные режимы двигателей переменного тока в системах частотного управления : учеб. пособие / В.Н. Поляков, Р.Т. Шрейнер ; под общ. ред. Р.Т. Шрейнера.— Екатеринбург : УрФУ, 2017 256 с.

ИМИТАЦИОННОЕ МОДЕЛИРОВАНИЕ СИСТЕМ МАССОВОГО ОБСЛУЖИВАНИЯ С ОБРАТНОЙ СВЯЗЬЮ

Е.А. Дорошенко, Е.Ю. Титаренко
Национальный исследовательский Томский политехнический
университет, г. Томск, ead26@tpu.ru

Имитационное моделирование позволяет эффективно исследовать системы различной сложности. Оно позволяет проводить наблюдения за работой системы в течении необходимого периода времени, определять закономерности и показатели, характеризующие работу системы. В работе было проведено исследование RQ-систем массового обслуживания вида $M^n/M/1$ и $MMPP/M/1$. Определяется стационарное распределение вероятностей числа заявок на орбите, состояние орбиты. Получена зависимость числа заявок на орбите от времени работы системы.

Simulation modeling is an effective method of investigation systems of varying complexity. It allows to monitor the system for the required period of time, to determine patterns and indicators that characterize the operation of the system. The paper studies $M^n/M/1$ and $MMPP/M/1$ RQ-systems. The stationary distribution of the probabilities of the customers in the orbit and the state of the orbit are determined. The dependence of the number of customers in the orbit on the operating time of the system was also obtained.

Различные модели теории массового обслуживания часто используются как математические модели экономических, технических и информационных систем. В частности, телекоммуникационные системы моделируются как системы с повторными вызовами (Retrial Queue Systems или RQ-системы). Такие системы используют для анализа и исследования процессов работы телекоммуникационных и компьютерных систем, проектирования мобильных сотовых радиосетей и др.

Математические модели RQ-систем

Рассмотрим систему массового обслуживания $M^n/M/1$ (рис. 1) с одним обслуживающим устройством, на вход которой поступает пуассоновский неординарный поток заявок с параметром λ . Пусть события появляются в моменты t_1, t_2, t_3, \dots «пачками», так что в момент t_k появляются сразу η_k событий. Будем считать, что η_k событий есть последовательность независимых одинаково

распределенных случайных величин с распределением $P\{\eta_k = n\} = q_n$. Если обслуживающий прибор свободен, то заявка из пачки поступает на обслуживание, в противном случае заявка попадает на орбиту. Время обслуживания заявки является случайной величиной, имеющей экспоненциальный закон распределения с параметром μ . Заявка, обслуживание которой завершено, может покинуть систему с вероятностью r_0 , повторно поступить на обслуживание с вероятностью r_1 или перейти на орбиту с вероятностью r_2 . Заявки ожидают повторного обслуживания в течение времени, распределенного по экспоненциальному закону с параметром σ .

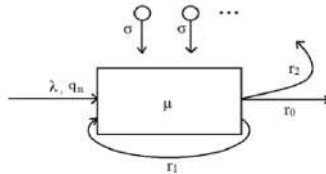


Рис.1 – Схема системы $Mn|M|1$

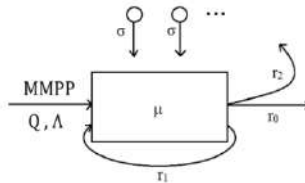


Рис.2 – Схема системы $MMPP|M|1$

Также рассмотрим систему массового обслуживания вида $MMPP|M|1$ (рис.2). Данная система является усложнением предыдущего случая. На вход данной системы поступает марковский модулированный пуассоновский поток заявок (ММРР) с конечным числом состояний K , заданный матрицей инфинитезимальных характеристик Q и диагональной матрицей Λ условных интенсивностей λ_k .

Пусть случайный процесс $i(t)$ есть число заявок на орбите в момент времени t , $k(t)$ – цепь Маркова, управляющая ММРР-потокком с

конечным числом состояний $1, \dots, K$ а $n(t)$ определяет состояние прибора следующим образом: 0 – свободен, 1 – занят.

Двумерный процесс $\{i(t), n(t)\}$ является цепью Маркова с непрерывным временем для системы $M^n|M|1$, а процесс $\{i(t), k(t), n(t)\}$ является цепью Маркова с непрерывным временем для системы $MMPP|M|1$.

Для распределения вероятностей $P_n(i)$ и $P_n(i, k)$ состояний рассматриваемых RQ-систем составим системы дифференциальных уравнений Колмогорова для стационарного режима.

Для системы $M^n|M|1$:

$$\begin{cases} -(\lambda + i\sigma)P_0(i) + r_0\mu P_1(i) + r_2\mu P_1(i-1) = 0, \\ -(r_0\mu + r_2\mu + \lambda)P_1(i) + \sum_{k=1}^{i+1} (\lambda q_k P_0(i-k+1)) + (i+1)\sigma P_0(i+1) + \\ + \sum_{k=1}^i (\lambda q_k P_1(i-k)) = 0; \end{cases}$$

Для системы $MMPP|M|1$:

$$\begin{cases} -(\lambda_k + i\sigma)P_0(i, k) + r_0\mu P_1(i, k) + \mu r_2 P_1(i-1, k) + \sum_m q_{mk} P_0(i, m) = 0, \\ -(r_0\mu + r_2\mu + \lambda_k)P_1(i, k) + \lambda_k P_1(i-1, k) + \lambda_k P_0(i, k) + (i+1)\sigma P_0(i+1, k) + \\ + \sum_m q_{mk} P_1(i, m) = 0. \end{cases}$$

Исследование RQ-систем с использованием имитационной модели

Имитационное моделирование является одним из самым эффективных методов исследований систем различной сложности. Он позволяет проводить наблюдения за работой системы в течении необходимого периода времени, определять закономерности и показатели, характеризующие работу системы.

Было разработано программное обеспечение, позволяющее моделировать поведение RQ-систем вида $M^n|M|1$ и $MMPP|M|1$ при разных значениях начальных параметров, а также проводить расчет характеристик систем (состояние обслуживающего прибора, зависимость числа заявок на орбите от времени, вероятность того,

что в некоторый момент времени на орбите находится ровно i заявок).

Описание работы алгоритма:

1. Определение вида системы.

2. Если выбрана система $M^n|M|1$:

2.1 Ввод параметров входящего потока и системы.

2.2 Пока текущее время меньше времени наблюдения за системой:

2.2.1 Случайным образом генерируется входящая «пачка» заявок, из «пачки» выбирается случайная заявка, остальные уходят на орбиту.

2.2.2 Если орбита пуста:

2.2.2.1 Если прибор свободен, то прибывшая заявка занимает прибор

2.2.2.2 Если прибор занят:

2.2.2.2.1 Если новая заявка пришла раньше конца обслуживания предыдущей заявки, то новая заявка уходит на орбиту.

2.2.2.2.2 Если новая заявка пришла после обслуживания предыдущей заявки, то обслуживавшаяся заявка со случайной вероятностью уходит из системы, поступает на повторное обслуживание или уходит на орбиту.

2.2.3 Если орбита не пуста:

2.2.3.1 Если прибор свободен:

2.2.3.1.1 Если заявка с орбиты придет раньше новой заявки, то заявка с орбиты занимает прибор.

2.2.3.1.2 Если новая заявка придет раньше заявки с орбиты, то новая заявка занимает прибор.

2.2.3.2 Если прибор занят:

2.2.3.2.1 Если заявка с орбиты придет раньше новой заявки и конца обслуживания текущей заявки, то заявка с орбиты возвращается на орбиту, то заявка с орбиты возвращается на орбиту.

2.2.3.2.2 Если новая заявка придет раньше заявки с орбиты и конца обслуживания текущей заявки, то новая заявка уходит на орбиту.

2.2.3.2.3 Если текущая заявка обслужится раньше прихода новой заявки и заявки с орбиты, то обслуживавшаяся заявка со случайной вероятностью уходит из системы, поступает на повторное обслуживание или уходит на орбиту.

3. Если выбрана система $MMPP|M|1$:

3.1 Ввод параметров входящего потока и системы.

3.2 Случайным образом изменяется интенсивность входящего потока, новая заявка приходит в систему.

3.3 Переход к проверке состояния орбиты (пункт 2.2.2)

С помощью полученной модели найдем вероятность того, что в некоторый момент времени на орбите находится ровно i заявок (рис. 5). Графики состояния орбиты для систем $M^n|M|1$ и $MMPP|M|1$ представлен на рис. 3. Графики состояния орбиты представлен на рис. 4.

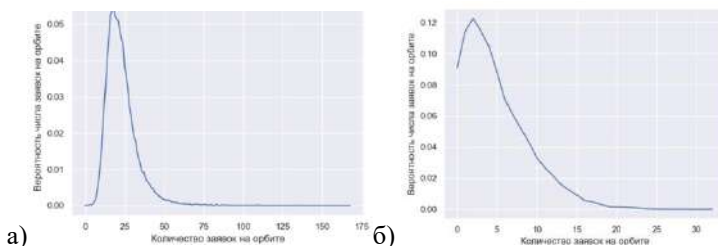


Рис. 3 – График распределения числа заявок на орбите:

а) для системы $M^n|M|1$; б) для системы $MMPP|M|1$

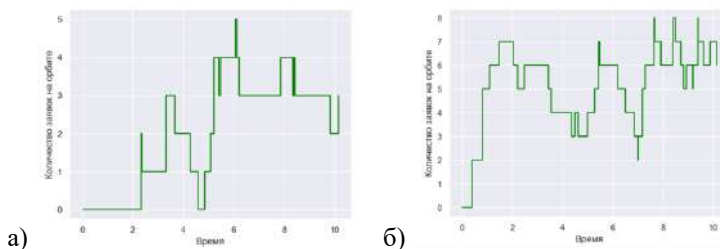


Рис. 4 – График состояния орбиты:

а) для системы $M^n|M|1$; б) для системы $MMPP|M|1$

Заключение

Была построена математическая модель для систем массового обслуживания вида $M^n|M|1$ и $MMPP|M|1$ с обратной связью, на основе математической модели был разработан и реализован алгоритм процесса работы систем. На основе результатов работы модели были получены графики распределение вероятностей числа заявок на орбите и график изменения числа заявок на орбите во времени.

Литература:

1. Шеннон Р.Е. Имитационное моделирование систем - искусство и наука. - М.: Мир, 1978. – 418 с.

2. Ивченко Г.И., Коваленко И.Н. Теория массового обслуживания // Учебное пособие – 2-е изд. – URSS, 2012. – 304 с.

3. Жидкова Л.А., Моисеева С.П. Исследование числа занятых приборов в системе $MPP|M|\infty$ с повторными обращениями // Вестник Томского государственного университета. – 2014. – С. 53-62.

ПРОВЕРКА ГИПОТЕЗЫ О НОРМАЛЬНОМ РАСПРЕДЕЛЕНИИ С ИСПОЛЬЗОВАНИЕМ НЕЙРОННЫХ СЕТЕЙ

К.Е. Дунаева

**Новосибирский государственный технический университет, г.
Новосибирск, dunaeva-1998@inbox.ru**

Научный руководитель: Чимитова Е.В., д.т.н., доцент

В данной работе рассматривается проблема проверки отклонения распределения от нормального закона. Разработан критерий для решения задачи проверки гипотезы о нормальности распределения, обобщающий несколько классических критериев согласия, основанный на нейросетевой модели. Разработан обучающий набор данных, проведены экспериментальные исследования работы критерия.

In this paper, the problem of verifying deviations from the normal distribution has been considered. The criterion for solving the task of testing the normal distribution hypothesis, generalizing several classical consent criteria, based on the neural network model, was developed. The training dataset was developed and experimental studies of the criterion was carried out.

Одно из основополагающих предположений многих классических методов математической статистики заключается в том, что наблюдаемые данные принадлежат нормальному закону распределения. Таким образом, проверка на принадлежность выборки нормальному закону распределения является одной из наиболее распространенных статистических процедур.

Проблему проверки нормальности можно сформулировать следующим образом.

Пусть имеется реализация выборки $x = [x_1, x_2, \dots, x_n]$ из распределения случайной величины X , требуется проверить гипотезу H_0 против конкурирующей гипотезы H_α :

- нулевая (H_0), что X следует нормальному распределению ($X \sim N$);
- альтернативная гипотеза (H_α) о том, что X не следует нормальному распределению ($X \not\sim N$).

Следует отметить, что проверяется сложная гипотеза о принадлежности выборки нормальному распределению.

В данной работе мы разработали новый обобщенный критерий проверки нормальности, основанный на искусственных нейронных сетях, с целью увеличения точности по сравнению со стандартными критериями нормальности. Наша исследовательская гипотеза заключалась в том, что нейронные сети могут эффективно обучаться правилам классификации из набора данных, состоящих из нормальных и ненормальных выборок.

Насколько нам известно, это четвертая попытка организовать процедуру проверки нормальности с помощью нейронных сетей, предыдущие три из которых - исследования Уилсона и Энгеля [1], Сигута [2] и Милоша Симича [3]. Уилсон и Энгель были первыми, предпринявшими попытки проверить нормальность с помощью нейронных сетей. Они тренировали нейронную сеть на наборе данных из выборок, взятых из стандартного нормального и выбранных ненормальных распределений, каждая выборка состояла ровно из 30 элементов. Перед тренировкой сети, Уилсон и Энгель сопоставили каждой выборке вектор из 16 статистик (которые не указаны в документе). Сигут применил аналогичный подход. Он обучил нейронную сеть на наборе данных, состоящем из множества нормальных и ненормальных выборок, взятых из большого количества различных распределений с использованием системы распределений Джонсона [4, 5]. В исследовании Сигута, выборки могут содержать любое количество элементов, что требует этапа предобработки для преобразования выборок в векторы одного размера. Сигут сопоставляет каждую выборку с вектором нескольких статистик: асимметрия, эксцесс, W-статистика теста Шапиро-Уилка [6], статистика теста Линь-Мудхолькара [7], статистика энтропийного теста Васичека [8] и размер выборки. В работе Симича были представлены дескрипторы для обучения нейронной сети. Дескрипторы представляли собой набор характеристик, описывающих выборку. Они состояли из среднего

значения, стандартного отклонения, минимума, максимума, медианы и других описательных характеристик выборки. Идея, лежащая в основе подходов Уилсона и Энгеля, Сигута и Симича состоит в том, что нейронные сети могут использовать информацию, которую дают все эти различные статистики в выборке и превосходить тесты, основанные только на индивидуальной статистике.

Основной идеей нашего подхода является создание набора данных, состоящего из выборок различных размеров, взятых из различных нормальных и ненормальных распределений и обучении нейронной сети на этом наборе данных, как если бы это была любая другая проблема бинарной классификации. Классифицировать выборку $x = [x_1, x_2, \dots, x_n]$, сгенерированную методом обратных функций:

- в соответствии с нормальным распределением, присваивая ей метку 0; или
- в соответствии с ненормальным распределением, присваивая ей метку 1.

Кроме того, поскольку мы самостоятельно можем моделировать набор данных, мы не ограничены наличием данных, как в других проблемах машинного обучения, а можем использовать столько данных, сколько позволяют наши вычислительные возможности.

Тем не менее, нейронные сети требуют, чтобы все объекты, которые они классифицируют, имели одинаковый размер: это означает фиксированное число и порядок характеристик. Даже если наиболее очевидным способом кажется представление выборки по ее содержанию, т.е. с помощью массива, содержащего элементы выборки, такой подход не является приемлемым для алгоритмов машинного обучения, так как тогда потребуется обучение отдельного классификатора для каждого возможного значения размера выборки, что не является осуществимой стратегией. Поэтому требуется этап предварительной обработки для преобразования входной выборки, независимо от ее объема, к вектору с фиксированным числом измерений. Мы представим выборку с помощью достигнутых уровней значимости (p-value) нескольких статистических критериев, которые были выбраны заранее. Смысл этого метода заключается в том, что взятые в совокупности различные критерии рассматривают выборки со взаимодополняющих точек зрения, так что нейронная сеть, которая комбинирует критерии, может быть более точной, чем отдельные критерии.

Для проверки гипотезы нормальности будут моделироваться выборки статистик и p -value для 6 критериев: Колмогорова [9], Крамера-Мизеса-Смирнова [10, 11], Андерсона-Дарлинга [12, 13], Никулина-Рао-Робсона [14, 15], а также специальных критериев отклонения наблюдаемого распределения от нормального закона (критерий Эпса-Палли [16] и критерий проверки на симметричность и значение эксцесса, модифицированный D'Agostino [17]). Мы выбрали эти критерии, поскольку они чаще всего изучаются в ходе моделирования и являются наиболее мощными критериями проверки нормальности.

Сеть, которую мы спроектировали в этом исследовании, представляет собой простую сеть с прямой связью с одним скрытым слоем. Мы использовали ReLU как функцию активации и Adam в качестве оптимизатора. Выходом нейронной сети является значение статистики полученного нами обобщенного критерия проверки нормальности. Для принятия решения о принадлежности выборки нормальному распределению необходимо округлить значение полученной статистики.

Несмотря на то, что описанная выше архитектура не является сложной, оказалось, что она смогла успешно освоить эффективные правила классификации из данных.

Поскольку открытого набора данных по выборкам, взятых из различных распределений, пригодного для обучения не было доступно во время нашего исследования, мы создали собственный набор данных для построения и оценки нашей сети.

Набор состоит из 50000 выборок из нормального распределения 60 595 выборок из распределений, отличных от нормального. Выборки имеют размеры от 10 до 1000. Набор сбалансирован по размерам выборок. Нормальные образцы были взяты из стандартного нормального распределения. Ненормальные выборки были смоделированы из распределения Вейбулла и обобщенного нормального распределения. Каждое распределение было задано с помощью подбора его параметров, наиболее близких к стандартному нормальному распределению и последующему удалению от него.

В этом исследовании мы создали и обучили искусственную нейронную сеть, способную проверять гипотезу о нормальности распределения. Сеть доказала свою конкурентоспособность в отношении многих стандартных статистических тестов для проверки нормальности.

Точность разработанной сети была 86% независимо от объема выборок. Такие показатели как TPR и TNR достигали значений 0.93 и 0.77 соответственно. TPR отражает отношение количества истинных положительных результатов к общему количеству положительных наблюдений, а TNR отражает отношение количества истинных отрицательных результатов к общему количеству отрицательных наблюдений. Принимая FPR (False Positive Rate) в качестве абсциссы и TPR (True Positive Rate) в качестве ординаты, мы можем сделать рисунок, называемый ROC-кривой. Очевидно, что чем ближе к точке левого верхнего угла, тем лучше эффект. AUC (область под кривой) определяется как область под кривой ROC. Очевидно, что чем больше значение, тем лучше. AUC нашей сети был 0.85, что приближается к 1 как к теоретическому максимальному AUC, соответствующему совершенному классификатору. То есть сеть характеризуется более высокой степенью точности классификации по сравнению с многими критериями. Возможным объяснением является следующее: нейронная сеть изучает эффективные правила классификации на основе совокупности множества критериев, взаимодополняющих друг друга, в то время как критерии строятся по специальным математическим свойствам нормального распределения и ограничиваются для классификации выборок путем проверки только конкретных свойств. Но разработанная сеть все еще имеет погрешность и не идеальный TPR и TNR. Эта проблема может быть преодолена с помощью усложнения архитектуры сети и увеличения обучающего набора данных.

Литература:

1. Wilson P. R., Engel A. B. Testing for normality using neural networks // First International Symposium on Uncertainty Modeling and Analysis, 1990: Proceedings. – P. 700–704.
2. Sigut J., Pineiro J., Estevez J., Toledo P. A neural network approach to normality testing // Intelligent Data Analysis. – 2006. – Vol. 10, N 6. – P. 509–519.
3. Milos Simic Normality Testing with Neural Networks // Neural Computing and Applications. – 2020. – Vol. 33, N 23. – P. 16279–16313.
4. Johnson N. L. Bivariate distributions based on simple translation systems // Biometrika. – 1949. – 36(3/4). – P. 297–304.
5. Johnson N. L. Systems of frequency curves generated by methods of translation // Biometrika. – 1949. – 36(1/2). – P. 149–176.

6. Shapiro S. S., Wilk M. B. An analysis of variance test for normality (complete samples) // *Biometrika*. – 1965. – 52(3/4). – P. 591–611.
7. Lin C. C., Mudholkar G. S. A simple test for normality against asymmetric alternatives // *Biometrika*. – 1980. – 67(2). – P. 455–461.
8. Vasicek O. A test for normality based on sample entropy // *Journal of the Royal Statistical Society. Series B (Methodological)*. – 1976. – 38(1). – P. 54–59.
9. Kolmogoroff A. N. Sulla determinazione empirica di una legge di distribuzione. // *G. Ist. Ital. attuar.* – 1933. – Vol. 4. – № 1. – P. 83–91.
10. Cramér H. On the Composition of Elementary Errors // *Scandinavian Actuarial Journal*. – 1928. – 1 – P. 13–74.
11. von Mises R. E. *Wahrscheinlichkeit, Statistik und Wahrheit* // Julius Springer. – 1928.
12. Anderson T. W., Darling D. A. Asymptotic theory of certain “Goodness of fit” criteria based on stochastic processes // *AMS*. – 1952. – Vol. 23. – P. 193–212.
13. Anderson T. W., Darling D. A. A test of goodness of fit // *J. Amer. Statist. Assoc.* – 1954. – Vol. 29. – P. 765–769.
14. Никулин М. С. Критерий хи-квадрат для непрерывных распределений с параметрами сдвига и масштаба // *Теория вероятностей и ее применение*. – 1973. – Т. XVIII, №3. – С. 583-591.
15. Никулин М. С. О критерии хи-квадрат для непрерывных распределений // *Теория вероятностей и ее применение*. – 1973. – Т. XVIII, №3. – С. 675-676.
16. Epps T. W., Pulley L. B. A test for normality based on the empirical characteristic function // *Biometrika*. – 1983. – Vol. 70. – P. 723-726.
17. D’Agostino R. B. Transformation to normality of the null distribution of g_1 // *Biometrika*. – 1970. – Vol. 57. – P. 679–681.

СИСТЕМА МОНИТОРИНГА СТЕПНЫХ И ЛЕСНЫХ ПОЖАРОВ

А.В. Тырышкин, С. Жуань
Национальный исследовательский Томский политехнический
университет, г. Томск, syren@tpu.ru
Научный руководитель: Тырышкин А.В., к.т.н., доцент

Цель исследования – разработать систему мониторинга степных и лесных пожаров. В статье рассмотрены составляющие устройства, принцип действия устройства, и его преимущества. Научная новизна работы заключается в том, что без постоянного электропитания. В результате исследования доказано, на основе которого произведён расчёт необходимых частей и ведутся практические испытания по слиянию всех компонентов в единое устройство.

The purpose of the paper is to develop a monitoring system for steppe and forest fires. The article discusses the components of the device, the principle of operation of the device, and the advantage. The scientific novelty of the work lies in the fact that without a constant power supply. As a result of the paper, it is proved, on the basis of which the calculation of the necessary parts is made and practical tests are carried out to merge all the components into a single device.

С момента развития науки и техники, учёные стали уделять внимание охране природы, в частности, лесным пожарам. В 2010 году в России сначала на территории Центрального федерального округа, а затем и в других регионах возникла сложная пожарная обстановка. По статистике, площадь сгоревших лесов в Томской области с 2010 по 2018 год составила 23,7 тысячи гектаров. Это второй показатель по Сибирскому федеральному округу после Красноярского края, говорится в докладе Минприроды России. Наибольшие площади уничтоженных пожаром лесов отмечены в Красноярском крае (69,637 тыс. га), Томской области (23,736 тыс. га), Иркутской области (14,323 тыс. га), Республике Саха (Якутия) (10,819 тыс. га). В таблице 1 показана динамика лесных пожаров на территории СССР за период 1991-2019 гг. [1]

Как видно из таблицы 1, несмотря на отлаженную систему тушения лесных пожаров, их число в отдельные годы составляло от 30 до 40 тысяч, а площадь пройденных лесов в 1991 и 2019 годах составила 1,85 млн. га и 1,7 млн. га соответственно. В настоящее время учёные разработали достаточное количество средств для

тушения и предупреждения пожаров, однако эти средства не всегда являются эффективными и имеют свои преимущества и недостатки.

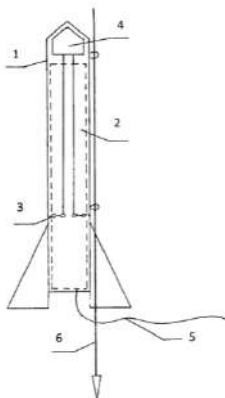


Рис.1 – Устройство сигнализации о лесном пожаре

Таблица 1 – динамика лесных пожаров на территории СССР за период 1991-2019 гг.

Годы	Число лесных пожаров	Площадь лесного фонда, пройденная пожарами, тыс. га.		
		общая	лесная	нелесная
1991	40169	1849,0	1459,9	389,1
2001	29285	284,7	213,1	71,6
2011	18189	238,2	170,2	68,0
2015	15201	698,2	492,7	205,5
2019	25345	1126,0	1384,7	310,5

Существует ряд технологий обнаружения лесных пожаров. Это и традиционное наблюдение с пожарных вышек, и облёт лесных массивов, как пилотируемыми аппаратами, так и беспилотниками, и различные технические средства типа пожарных извещателей.

В таблице 2 показаны разные виды устройств и оборудования для обнаружения и тушения лесных пожаров.

Как видно из таблицы 2, на сегодняшний день не существует надёжной системы оповещения о лесном пожаре, которая бы удовлетворяла существующим требованиям лесоохраны.

Для обнаружения лесных пожаров существуют множество решений, однако их достигать до реализации уже нужно подумать о расчете устройства.

Таблица 2 – виды устройств и оборудования для обнаружения и тушения лесных пожаров.

Виды средств охраны	Достоинства	Недостатки
Сотрудники лесоохраны	Индивидуальный подход к осмотру территории	Высокие трудозатраты Низкая производительность
Лесопожарные самолеты танкеры	Высокая мобильность	Высокая стоимость и ограниченность количества
Система мониторинга пожаротушения FSM-2500-EP	Реализация дальности управления и своевременное получение сигнализации.	Высокая стоимость и сложность
Система обнаружения лесного пожара. Ru 2617138 C1	Техническое обеспечение представляет собой раннее локальное обнаружение	Оператор радиопередачи имеет стесненность в радиусе 20 км.
Способ обнаружения пожара. Ru 2492899 C1	Обеспечение автономности системы, увеличение срока службы системы.	Невысокая надежность и автономность, большая вероятность ложных срабатываний

Предложен ряд технических решений, защищённых патентами на изобретения. Так, например, известно устройство для сигнализации о лесном пожаре [RU 2617138 C1], которое содержит датчиков, каждый из которых содержит акустический сенсор, первый температурный выключатель, первый выход которого соединен с источником питания, а второй выход связан с первым входом электропитания радиопередатчика, выход которого соединен с радиопередающей антенной. Первый выход второго температурного выключателя соединен с источником питания, а

второй выход подключен к входу таймера с управляемым ключом, выход которого связан с первым входом электропитания радиопередатчика, второй вход которого подключен к выходу усилителя, первый вход электропитания которого соединен со вторым выходом первого температурного выключателя, а второй вход связан с выходом акустического сенсора.

Недостатком устройства является наличие, как минимум, 2n соединительных проводов, что затрудняет практическое применение устройства в лесу, и снижает надежность устройства.

Литература:

1. Защищает древесину от возгорания, останавливает распространение пламени во время пожара, режим доступа: <http://mtnsk.mos.ru/Handlers/Files.ashx/Download?ID=3696>

2. Реактивные Двигатели для моделей, режим доступа: <http://real-rockets.ru/wp-content/uploads/2016/04/rd1-10.pdf>

3. Кабель терморезистивный в термостойкой изоляции предназначен для изготовления термоэлектрических преобразователей. Режим доступа: http://tpchel.ru/KTI-НА_KTI-НК

СОВРЕМЕННЫЕ СИСТЕМЫ УПРАВЛЕНИЯ ДАННЫМИ В НАУЧНЫХ ЭКСПЕРИМЕНТАХ

А.Ю. Кайда

**Национальный исследовательский Томский политехнический
университет, г. Томск, ayk13@tpu.ru**

Научный руководитель: Савельев А.О., к.т.н., доцент

Данная статья описывает историю развития систем управления научным экспериментом, их эволюцию и отражение в современных условиях управления экспериментами, в том числе проблематику экосистемы больших данных.

This article describes the evolution process of EMS and reflection in modern conditions of experiment management, including the problems of the big data ecosystem.

Проведение научных экспериментов в условиях оперирования большими данными непосредственно связано с вопросами организации вокруг эксперимента такой экосистемы, которая

обеспечивает взаимосвязь сложных систем между собой, как и интеграцию разнородных (гетерогенных) данных. Перед системами управления экспериментом, по мере развития технологий, ставятся новые вызовы. Вместе с тем, ученые отмечают проблемы разработки таких систем общего назначения.

Исходная информация при проведении научных экспериментов поступает в виде сырых данных. Основная задача науки о данных состоит в том, чтобы переработать поток таких данных в вид, пригодный для практической деятельности [1]. Разведочный анализ данных (Exploratory Data Analysis, EDA) – одно из ключевых понятий науки о данных, характерное на текущий момент для любой исследовательской задачи, появилось благодаря книге статистика Джона Тьюки «Анализ результатов наблюдений» [1,2], где автором подчеркивается важность разведочного анализа как первого шага исследований в различных областях. Современными примерами тому могут послужить исследование в области снижения рисков возникновения лесных пожаров и проектировании систем их предотвращения [3], в области заболеваний сердечно-сосудистой системы [4], в области высокопроизводительных вычислений [5] и т.д.

Система управления научным экспериментом (EMS) – система, обеспечивающая поддержку ученых на протяжении всех стадий проведения эксперимента, представляющая собой специальный интерфейс, связывающий проводящего эксперимент ученого и данные. EMS позволяет спланировать и упорядочить эксперимент, управлять данными и анализировать полученные результаты, иными словами – обеспечивать полный жизненный цикл данных [6].

Первые системы управления научным экспериментом появились до «эры больших данных». Одним из примеров такой системы является настольная («десктопная») система управления экспериментом Висконсинского университета, Zoo, созданная в 1996 г [7].

Авторами отмечается, что система Zoo была создана как настольная система управления средой эксперимента общего назначения (Desktop Experiment Management Environment, DEME), и для развертывания ее в конкретной лаборатории, а следовательно – превращения в полноценную настраиваемую систему управления экспериментом (Customized Desktop Experiment Management System, CDEMS) – требуется создание дополнительных модулей. Тем не менее, Zoo является важным образцом в истории развития систем управления экспериментами [8].

С момента описания «Zoo» условия проведения экспериментов значительно усложнились. В 2005 году Европейским биоинформатики (Кембридж) в коллаборации с крупнейшими вузами Великобритании была описана система для проведения исследований в области биоинформатики, Taverna. В архитектуре системы можно выделить два основных компонента: трехуровневую модель данных, описывающую взаимодействие ресурсов разных уровней абстракции (application data flow, execution data flow, invocation level), а также специальный фреймворк. Для авторов данной статьи ключевыми требованиями к EMS-системе в области биоинформатики были следующие:

1. Допущение использования любого ресурса.
2. Простая и быстрая разработка операционного потока.
3. Модель данных, ориентированная на потоки данных, которыми оперирует система.
4. Отказоустойчивость.
5. Поддержка эксперимента на всех этапах жизненного цикла данных [9].

В 2014 году учеными из университета Тулузы была описана система Exro, предназначавшаяся, в первую очередь, для балансировки вычислительной нагрузки при крупномасштабном моделировании в ходе проведения экспериментов в области электромагнитных полей. Система построена таким образом, что специальный описательный файл эксперимента интерпретируется двумя модулями и передается модулю-исполнителю, взаимодействующему с платформами, на которых производятся вычисления. Преимущество архитектуры данной системы в ее простоте и легковесности, несмотря на сложность самого эксперимента [10].

Другой системой, использующей единое описание эксперимента, является система OMF. Данное описание представлено в виде скрипта и включает в себя используемое программное обеспечение, рабочие нагрузки, ход эксперимента, а также оценку полученных данных. При этом скрипт описания может быть передан другому исследователю для воспроизведения эксперимента на другой рабочей станции [11].

Ученые Мадридского университета при проведении исследований в области гидрологии с использованием облачных сервисов, системы Apache Spark и парадигмы высокопроизводительных вычислений, отмечают, что по состоянию на 2016 г. возникла потребность в системах автоматизации

крупномасштабных экспериментов во многих научных областях, таких как экономика, геология, метеорология, ядерная физика, биология и социология. Экспериментам необходимы не только инструменты управления и анализа данных, но также инструменты визуализации и симуляции [12]. Авторы приводят в пример систему CyberShake для изучения сейсмической активности [13].

Авторами рассмотренных научных статей отмечается необходимость наличия систем управления метаданными экспериментов. В ходе проведения аналитического обзора было установлено, что первые системы управления средой эксперимента, существенно отличающиеся по уровню сложности в сравнении с современными компонентами BDE (Big Data Environment, или экосистема больших данных), появились еще в 90-е годы прошлого столетия (в частности, система Zoo). По мере развития технологий распределенных вычислений и появления понятия больших данных, в типовую экосистему современного научного эксперимента добавились новые элементы, не учитываемые в ранее описываемых системах.

Можно выделить следующие типовые проблемы современного научного эксперимента в области больших данных:

1. Гетерогенность данных.
2. Гетерогенность вычислительной инфраструктуры.
3. Стремительный рост данных.
4. Наличие объектов стохастически накапливаемой информации в сыром, непредобработанном виде, – озер данных.
5. Необходимость масштабирования.

Отдельно были выделены следующие требования к EMS-системе:

1. Отказоустойчивость.
2. Поддержка работы с озерами данных, т.е. ориентация на организацию процесса обработки данных.
3. Возможность подключения условно любого источника данных, т.е. гибкость организации процесса обработки данных.
4. Возможность поддержки эксперимента в течение всего жизненного цикла данных.
5. Отсутствие специальной грид-системы. Как правило, грид-системы используются для экспериментов класса мегасайенс, к которым нельзя отнести текущие эксперименты лаборатории.

Опираясь на жизненный цикл данных, можно выделить следующие функции, которые должна обеспечивать система:

1. Проектирование эксперимента.
2. Развертывание эксперимента.
3. Поэтапное проведение эксперимента (запуск, остановка).
4. Повторное воспроизведение эксперимента.
5. Анализ результатов и управление данными [14].

В ходе работы был проведен обзор литературы по системам поддержки экспериментов, по результатам которого были выделены общие для большинства систем проблемы и специфичные черты систем, а также задачи исследователей вне зависимости от отрасли. На сегодняшний день EMS-системы остаются крайне востребованными и имеют широкую область применения. Основной задачей является проектирование и реализации основных компонентов EMS-системы (generic EMS). На текущий момент ведется работа в экспериментах лаборатории обработки неструктурированной информации ИШИТР ТПУ. В текущей работе учитывается опыт работы в решении задачи построения базы знаний научного эксперимента ДКВ для эксперимента ATLAS в CERN.

Исследование выполнено при финансовой поддержке ГЗ «Наука», в рамках проекта FSWW-2020-0014

Литература:

1. П.Брюс, Э.Брюс. Практическая статистика для специалистов Data Science: Пер. с англ. – СПб.: БХВ-Петербург, 2019. – С. 19-21.
2. J.W. Tukey. Exploratory Data Analysis. – Pearson, –1977, pp. 1-7.
3. G. Lima, J. Dantas V. Barbosa, V.E. Beal, M. A. Moret S. Gonçalves, B. Machado, J. Gerber, B. S. Lazarus, Exploratory analysis of fire statistical data and prospective study applied to security and protection systems. International Journal of Disaster Risk Reduction, Vol. 61, 2021, p.102308.
4. R. Indrakumari, T. Poongodi, S.R. Jena. Heart Disease Prediction using Exploratory Data Analysis. Procedia Computer Science, Vol. 173, 2020, pp.130–139.
5. Y. Daqing, L. Wuji, W. Chase, S.V. Nageswara, K. Rajkumar, Exploratory analysis and performance prediction of big data transfer in High-performance Networks. Engineering Applications of Artificial Intelligence, Vol. 102, 2021, p.104285.

6. Y. Ioannidis, M. Livny, E. Haber, R. Miller, and O. Tsatalos, J. Wiener, Desktop Experiment Management. Technical Report. Stanford InfoLab. Publication Note: Appeared in IEEE Data Engineering Bulletin, 16(1), March 1993.
7. Y. Ioannidis, M. Livny, S. Gupta, and N. Ponnekanti. ZOO: A desktop experiment management environment. Proc. 22nd International VLDB Conference, 1996, pp. 274–285.
8. Y. Ioannidis and M. Livny. MOOSE: Modeling objects in a simulation environment. Information Processing, Vol. 89, 2019, pp. 821–826.
9. T. Oinn, M. Greenwood, M. Addis, M.N. Alpdemir, J. Ferris, K. Glover, C. Wroe, Taverna: lessons in creating a workflow environment for the life sciences. Concurrency and Computation: Practice and Experience, 18(10), 2006, pp.1067–1100.
10. C. Ruiz, M. Alenxandru, O. Richard, T. Monteil, H. Aubert, Platform calibration for load balancing of large simulations: TLM case, Proceedings of the 14th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing, 2014, pp. 465–472.
11. C. Dwertmann, E. Mesut, G. Jourjon, M. Ott, T. Rakotoarivelo, I. Seskar, Mobile experiments made easy with OMF/orbit, Proceedings of Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications, 2009.
12. S. Caino-Lores, A. Lapin, J. Carretero, P. Kropf, Applying big data paradigms to a large scale scientific workflow: Lessons learned and future directions. Future Generation Computer Systems. Vol.110, 2020, pp.440–452.
13. S. Callaghan, P. Maechling, P. Small, K. Milner, G. Juve, T.H. Jordan, E. Deelman, G. Mehta, K. Vahi, D. Gunter, K. Beattie, C. Brooks, Metrics for heterogeneous scientific workflows: A case study of an earthquake science application. International Journal of High Performance Computing Applications. Appl. Vol. 25(3), 2011, pp. 274–285.
14. T. Buchert, C. Ruiz, L. Nussbaum, O. Richard. A survey of general-purpose experiment management tools for distributed systems. Future Generation Computer Systems, Vol. 45, 2015, pp.1–12.

БЫСТРЫЕ АЛГОРИТМЫ РАЗЛОЖЕНИЯ ФУНКЦИЙ В РЯД ПО СФЕРИЧЕСКИМ ГАРМОНИКАМ

Д.С. Клементьев¹, А.В. Терехов^{1,2}

¹Новосибирский государственный технический университет,

²Институт вычислительной математики и математической
геофизики СО РАН, г. Новосибирск, a.terexov@corp.nstu.ru

Научный руководитель: Терехов А.В., д.ф.-м.н., профессор

Математическое моделирование климата, задачи прогноза погоды, вычислительной геофизики, астрономии и биологии – это неполный перечень приложений сферических функций. В рамках исследования рассмотрен простой и эффективный способ сжатия матриц для сферических функций на основе стандартной процедуры FFT. Это позволяет экономично умножить вектор на исходную матрицу преобразования в тригонометрическом базисе, что снижает оценку вычислительной трудоемкости метода. Исследование выполнено при финансовой поддержке РФФИ и Новосибирской области в рамках научного проекта № 20-41-540003

Mathematical modeling of climate, weather prediction problems, computational geophysics, astronomy and biology - this is an incomplete list of applications of spherical functions. The study considers a simple and efficient way to compress matrices for spherical functions based on the standard FFT procedure. This allows to economically multiply the vector by the original transformation matrix in the new trigonometric basis, which reduces the estimate of computational labor intensity of the method. The research was funded by RFBR and Novosibirsk region, project number 20-41-540003

Проблема построения эффективного численного метода разложения функции в ряд по сферическим гармоникам и обратно – сложная и актуальная задача, которая может быть сформулирована в виде:

$$f(\theta, \phi) = \sum_{m=0}^N \exp(im\phi) \sum_{l=|m|}^N f_l^m P_l^m(\cos(\theta)), \quad (1)$$

$$f_l^m = \int_0^\pi \int_0^{2\pi} \exp(-im\phi) P_l^m(\cos(\theta)) f(\theta, \phi) \sin(\theta) d\phi d\theta, \quad (2)$$

где θ, ϕ – стандартные сферические координаты на двумерной

сфере $S^2 \in \mathbb{R}^3$, $P_l^m(x)$ –присоединённые функции Лежандра степени l и порядка m , а f_l^m –коэффициенты разложения. Основная трудность при построении экономичных методов вычисления значений двойного ряда (1) состоит в расчёте $f_m(\theta)$ по формуле

$$f_m(\theta) = \sum_{l=|m|}^N f_l^m P_l^m(\cos(\theta)), \quad f(\theta, \phi) = \sum_{m=0}^N f_m(\theta) \exp(im\phi),$$

тогда как суммирование по индексу m может быть выполнено посредством FFT алгоритма. Вычисление двойного интеграла (2) производится в обратном порядке

$$f_l^m = \sum_{j=1}^{N_\theta} w_j f_m(\theta_j) P_l^m(\cos(\theta_j)), \quad f_m(\theta) = \int_0^{2\pi} f(\theta, \phi) \exp(-im\phi) d\phi.$$

Как прямое, так и обратное преобразования по вычислительной сложности эквивалентны и требуют $O(N^3)$ операций. Для вычисления $f_m(\theta)$ можно использовать метод экстра-компонент, однако с ростом m степень сжатия матрицы преобразования для функций $P_l^m(x)$ будет сокращаться (Рисунок 1). Для решения этой проблемы рассмотрим несколько подходов. Первый подход состоит в том, чтобы сначала разложить функцию в ряды вида $f(x) = \sum_{l=0}^N f_l^1 P_l^1(x)$, $f(x) = \sum_{l=0}^N f_l^2 P_l^2(x)$, посредством метода экстра-компонент, а затем с помощью быстрого алгоритма для умножения квази-разделяемых матриц перейти от p^1, p^2 к p^m за $O(N \log N) \log(1/\varepsilon)$ арифметических действий [1]. Предлагаемое усовершенствование состоит в том, что метод экстра-компонент позволит выполнить быстрое преобразование для любого набора узлов $x \in [-1, 1]$, тогда как в работе [1] разлагаемая функция предварительно разлагается в ряд по полиномам Чебышева. Второй подход состоит в том, чтобы коэффициенты f_l^m с порядком $m < m_0 < N$ вычислять посредством метода экстра-компонент, а для всех последующих $m_0 \leq m \leq N$ с помощью алгоритмов [2,3], которые не позволяют эффективно выполнить сжатие матрицы

преобразования для первых значений m . Метод экстра-компонент отличается от существующих простой программной реализацией и высокой эффективностью. На подготовительном этапе, выполняемом однократно, исходная матрица посредством Фурье преобразования приводится к матрице с разреженным портретом. Далее на вычислительном этапе матрично-векторное произведение может быть вычислено достаточно быстро в силу игнорирования относительно близких к нулю элементов сжатой матрицы. Однако наиболее перспективным является подход на основе двумерного сжатия, которое разработано для реализации быстрого преобразования Лагерра [4]. Из Рисунков 2 и 3 видно, что с ростом m степень сжатия матрицы преобразования увеличивается, кроме значений m близких к N , т.е. когда порядок умножаемой матрицы становится незначительным. Интересным является тот факт, что по сравнению с методом на основе локального косинус преобразования [2], сжатая матрица имеет компактный портрет, что позволяет производить ее умножение в спектре очень эффективно, так как не требуется специальных форматов данных для ее хранения. Таким образом, принципиальная возможность применения метода экстра-компонент для сферических гармоник показана.

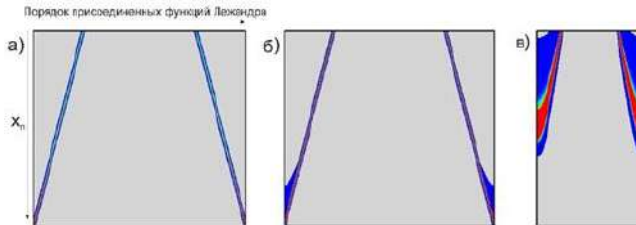


Рис. 1 – Амплитуда элементов сжатой матрицы FW_{NC} для присоединённых функций Лежандра $P_l^m(x)$ для а) $m=1$, б) $m=10$, в) $m=1000$

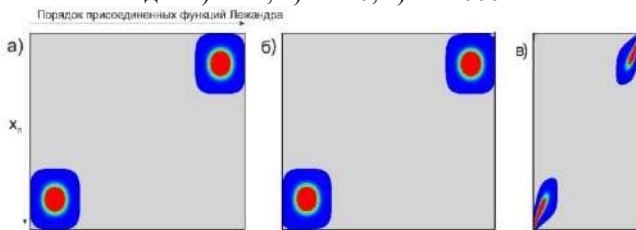


Рис. 1 – Амплитуда элементов сжатой матрицы $FW_{NC}CW_MF^*$ для присоединённых функций Лежандра $P_l^m(x)$ для а) $m=1$, б) $m=10$, в) $m=1000$

В работе рассмотрен эмпирический подход для построения быстрых алгоритмов вычисления матрично-векторных произведений, связанных со сферическими функциями. Расчёты показали, что метод экстра-компонент является весьма перспективным для решения довольно сложной проблемы о разложении функции в ряд по сферическим функциям. Для решения дифференциальных уравнений на современных суперкомпьютерах многообещающей выглядит комбинация рассмотренных в статье быстрых методов и параллельный алгоритм дихотомии [5,6], который можно трактовать как быстрый метод умножения матрицы обратной к блочно-трёхдиагональной матрице. Один из ключевых факторов, влияющих на производительность – это векторизация вычислений, которая не всегда может быть эффективно использована в рамках существующих алгоритмов. Для современных микропроцессоров векторные инструкции требуют, чтобы данные в оперативной памяти были непрерывны и одновременно выровнены по определённым значениям адресов. Структура рассматриваемых в исследовании алгоритмов позволяет удовлетворить этим ограничениям для векторных вычислений, тем самым гарантируется высокий уровень производительности для широкого диапазона размерности преобразований.

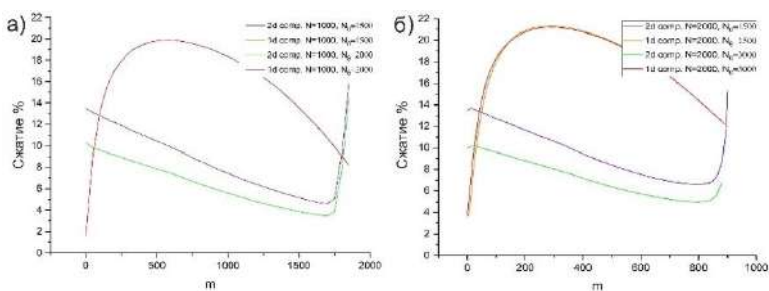


Рис. 2 – Зависимость степени одномерного и двумерного сжатия матрицы преобразования для присоединённых функций Лежандра $P_l^m(x)$ от порядка m для а) $0 < m < 1000$, $m < l < 1000$ и б) $0 < m < 2000$, $m < l < 2000$

Как следствие, если в теоретическом плане оценки числа операций метода экстра-компонент не лучше оценок для существующих подходов, то реальная производительность может быть даже выше, чем производительность высокооптимизированных библиотечных процедур.

Литература:

1. Rokhlin V. , Tygert M. Fast algorithms for spherical harmonic expansions // SIAM Journal on Scientific Computing. – 2006. – Vol. 27, N 6. – P. 1903–1928.
2. Mohlenkamp M. J. A fast transform for spherical harmonics // Journal of Fourier Analysis and Applications. – 1999. – Vol. 5, N 2. – P. 159–184.
3. Tygert M. Fast algorithms for spherical harmonic expansions, ii // Journal of Computational Physics. – 2008. – Vol. 227, N 8. – P. 4260–4279.
4. Terekhov A. V. Generating the Laguerre expansion coefficients by solving a one-dimensional transport equation // Numerical Algorithms. – 2021. – doi:10.1007/s11075-021-01115-8.
5. Terekhov A.V. A highly scalable parallel algorithm for solving Toeplitz tridiagonal systems of linear equations // Journal of Parallel and Distributed Computing. – 2015. – Vol. 87. – P. 102–108.
6. Terekhov A V. Parallel dichotomy algorithm for solving tridiagonal system of linear equations with multiple right-hand sides // Parallel Computing. – 2008. – Vol. 36, N 8. – P. 423–438.

**КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ ВЛИЯНИЯ
РАДИАЦИОННОГО ДАВЛЕНИЯ СОЛНЕЧНОГО
ИЗЛУЧЕНИЯ НА КОРПУС КОСМИЧЕСКОГО АППАРАТА**

И.А. Кондратьев

**Новосибирский государственный технический университет,
г. Новосибирск, igor36igor36@gmail.com**

Научный руководитель: Черникова О.С., к.т.н., доцент

Разработано программно-математическое обеспечение расчета силы давления радиационного потока на космический аппарат, в основе которого лежит моделирование траектории движения фотонов.

Software for calculating the pressure force of the solar radiation on the spacecraft has been developed. The algorithms are based on modeling the trajectory of photons.

Учет влияния силы радиационного давления Солнца (РДС) на элементы конструкции навигационного спутника является одной из основных проблем при построении высокоточной модели движения космического аппарата (КА).

В настоящее время построению математической модели влияния РДС на корпус КА посвящено достаточное количество работ [1-3].

В работе предлагается с помощью компьютерного моделирования движения потока фотонов имитировать влияния на корпус КА радиационного давления солнечного излучения и на основе расчета траектории их движения оценивать силу РДС на различные элементы аппарата.

Рассмотрим компьютерную трехмерную модель КА в связанной с ним системе координат $X_c Y_c Z_c$ (рис. 1). Начало системы координат совпадает с центром масс КА. Ось OX_c направлена на геометрический центр Земли, а в плоскости $OX_c Y_c$, называемой плоскостью «Солнце-Объект-Земля» (СОЗ), расположен вектор направления на Солнце.

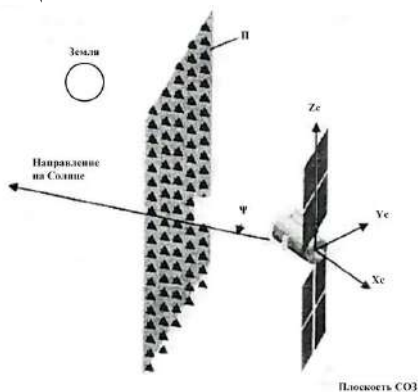


Рис. 1 – Трехмерная модель КА

Будем считать, что при взаимодействии с поверхностью КА радиационный поток Солнца разделяется на несколько составляющих (см. рис. 2).

Общий световой поток будем находить по формуле:

$$f(x, y, z) = f_T + f_{-\tau} + f_3 + f_d + f_{\Gamma} \quad (1)$$

Составляющие общего светового потока (1) определим по следующим формулам [4]:

$f_{\Gamma} = n(n, \tau)(1 - \alpha\rho - \chi)\tau$ – обусловленный падающим потоком, Составляющие общего светового потока (1) определим по следующим формулам [4]:

$f_{\Gamma} = n(n, \tau)(1 - \alpha\rho - \chi)\tau$ – обусловленный падающим потоком,

$f_{-\tau} = h\alpha\beta(n, \tau)\tau$ – обратно отражённым потоком,

$f_{\text{д}} = \frac{2}{3}h\alpha(1 - \rho - \beta)(n, \tau)n$ – диффузно отраженным потоком,

$f_{\text{Т}} = z_n \frac{2}{3}h(1 - \alpha - \chi)(1 - \eta)(n, \tau)n$ – несимметричным переизлученным потоком в ИК диапазоне,

$f_{\text{З}} = h\alpha\rho(n, \tau)^2n$ – зеркально отраженным потоком,

где τ – единичный вектор направления падающего светового потока, n – единичный вектор нормали к освещаемой поверхности. (n, τ) – косинус угла между векторами n и τ , E – мощность светового потока, приходящего на единицу поверхности, нормальной к этому потоку, c – скорость света, $h = E/c$.

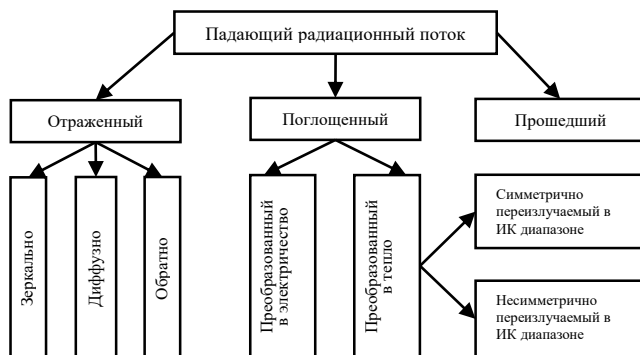


Рис. 2 – Схема разделения светового потока

Назовем равномерно распределенную случайную величину фотоном. Движение фотона будем представлять в виде луча, выходящего из заданной точки на Солнце и имеющего направление распространения перпендикулярно плоскости СОЗ. Будем считать, что при взаимодействии с КА такой фотон делится на два: $q_{1з}$ и $q_{1д}$. $q_{1з}$ с зеркально-отраженной энергией движется по траектории, обусловленной законом зеркального отражения. $q_{1д}$ с диффузно отражённой энергией – в направлении, определяемом принятым законом диффузного отражения. В компьютерной модели направление его распространения S является случайным, с углом от

0 до 180 градусов по отношению к взаимодействующей плоскости и определяется по формуле:

$$S = \left(X_l, Y_l, \frac{(X_l N_x + Y_l N_y)}{N_z} \right) + Z_l N,$$

где $X_l, Y_l, Z_l \in R, X_l, Y_l \in [-1, 1], Z_l \in [0, 1], N = (N_x, N_y, N_z)$ - вектор нормали.

Отметим, что при моделировании потока фотонов необходимо учитывать многократное переотражение лучей в точках пересечения с поверхностью и перераспределение энергии по поверхности КА. Так, например, некоторые элементы корпуса КА допускают экранно-вакуумную тепловую изоляцию, наличие которой влияет на процесс отражения лучей. В процессе построения модели следует проследить траектории двух образовавшихся фотонов, каждый из которых при наличии пересечений с поверхностью в точках (x_2, y_2, z_2) и (x_3, y_3, z_3) разделится ещё на два. Получим четыре фотона $q_{23}, q_{24}, q_{33}, q_{34}$ соответственно. Продолжаем этот процесс до тех пор, пока энергия переотражённых фотонов не станет меньше некоторого определенного уровня. При каждом соударении фотона вычисляется вектор-функция силы РДС по формуле (1) с учетом уменьшения уровня энергии. Тогда результирующая функция распределения силы, создаваемой начальным фотоном, выглядит следующим образом:

$$\begin{aligned} \hat{f}(U, V)dUdV = & f^{(0)}(x_1, y_1, z_1) + f^{(1)}(x_2, y_2, z_2) + f^{(1)}(x_3, y_3, z_3) \\ & + f^{(2)}(x_4, y_4, z_4) + f^{(2)}(x_5, y_5, z_5) \\ & + f^{(3)}(x_6, y_6, z_6) + f^{(3)}(x_7, y_7, z_7) + \dots, \end{aligned}$$

где $f^{(j)}(x_i, y_i, z_i)$ – сила, созданная фотоном при одном ударе о поверхность в точке с координатами (x_i, y_i, z_i) , а j – индекс точки, которая порождает два фотона, отраженных зеркально и диффузно.

Таким образом, задача вычисления силы РДС на КА сводится к нахождению интеграла

$$F(\psi) = \int_{\Pi} \hat{f}(U, V)dUdV$$

Разработанное программное обеспечение позволяет визуализировать движение потока фотонов, имитировать влияния

РД солнечного излучения на корпус КА и предоставляет пользователю:

- осуществлять настройку плотности потока фотонов,
- выбирать интересующую деталь КА,
- изменять размер деталей КА и/или добавлять материал экранно-вакуумной тепловой изоляции.
- отображать величину силы РДС в режиме реального времени.

Литература:

1. Arnold D. CODE's new solar radiation pressure model for GNSS orbit determination // Journal of Geodesy. – 2015. – Vol. 1, № 89. – P. 775–791.
2. Springer T.A., Beutler G. and Rothacher M. A new solar radiation pressure model for GPS satellites // GPS Solutions. – 1999. – Vol. 2, № 3. – P. 50–62.
3. Duan B., Hugentobler U., Hofacker M., Selmke I. Improving solar radiation pressure modeling for GLONASS satellites // Journal of Geodesy. – 2020. – Vol. 1, № 94. – P. 14.
4. Гречкосеев А.К., Марарескул Т.А. Построение модели сил радиационного давления излучения Солнца, влияющего на движение центра масс космического аппарата // Известия вузов. Приборостроение. – 2004. – № 4. – С. 10–14.

НАСТРОЙКА НЕЙРОСЕТИ ДЛЯ ПРЕДСКАЗАНИЯ ВЫСШЕЙ ТЕПЛОТЫ СГОРАНИЯ БИОМАССЫ

А.Е. Крохина

**Новосибирский государственный университет,
Институт химии твердого тела и механохимии, г. Новосибирск
Научные руководители: Матвеева А.Г., к.х.н.,
Бычков А.Л., д.х.н.**

Разработан алгоритм настройки искусственной нейронной сети. Выявлена ключевая роль количества и качества данных для работы искусственной нейронной сети для предсказания высшей теплоты сгорания биомассы.

An algorithm for setting up an artificial neural network has been created. The key role and quality of data for the operation of an artificial neural network for predicting the highest heating value of biomass has been identified.

В последнее время наблюдается тенденция к использованию возобновляемого топлива, одним из классов которого являются отходы производства растительного сырья. Действительно, биомасса представляет перспективную альтернативу ископаемому топливу, так как сочетает в себе низкую стоимость, доступность и экологичность. Использование растительного сырья значительно снижает количество вредных выбросов, таких как SO_2 и NO , за счет низкого содержания серы и азота в биомассе, кроме того, растительное сырье обладает CO_2 – нейтральной характеристикой. Все это позволяет повысить экологическую и энергетическую безопасность. [1]

Актуальной проблемой является определение оптимальных по содержанию энергии источников биотоплива. Поэтому совершенствуются методы анализа, позволяющие определять теплофизические свойства исходного сырья и получаемого из него топлива.

Наиболее точным является метод определения теплоты при адиабатическом горении. Также часто используют приблизительную оценку теплоты сгорания на основе данных элементного и технического анализа. На сегодняшний день известно более 150 эмпирических корреляционных уравнений [1-7], которые позволяют рассчитать высшую теплоту сгорания (ВТС) различных видов лигноцеллюлозного сырья. При этом оба этих метода не универсальны – первый применим не ко всем видам сырья, второй требует трудоемких вычислений и подбора формул для каждого вида сырья по-отдельности [7, 8].

В последние годы в связи с развитием математических методов возник интерес к применению самообучающихся искусственных нейронных сетей (ИНС) в работе с большими массивами данных. Опубликованы работы, в которых показаны возможности ИНС в предсказании ВТС на основании данных технического и элементного анализа сырья [9-12]. Проводятся сравнения результатов предсказания классическими корреляционными уравнениями и ИНС, анализируется эффективность логики, сжигания и пиролиза смеси угля и биомассы [13-15].

Однако, ИНС нуждается в предварительном обучении и настройке гиперпараметров: требуется провести много итераций обучения – настройки, каждый раз оценивая результат обучения на тестовой части выборки и меняя гиперпараметры. На сегодняшний день отсутствует четко прописанный алгоритм такой итерационной настройки. Исследователь сам решает, какие гиперпараметры и в какой последовательности нужно менять, чтобы получить

оптимальный результат, действуя «методом проб и ошибок». При этом, в отличие от пионерских работ в данном направлении [16], авторы статей про прикладное использование ИНС [17] зачастую не приводят ни найденные оптимальные значения гиперпараметров, ни обученную с оптимальными гиперпараметрами нейросеть, что делает результаты их работы невоспроизводимыми. Кроме того, часть опубликованных работ не учитывают важный методологический аспект – величину и равномерность обучающей выборки. Часты случаи, когда для обучения ИНС используется недостаточное количество (около 100, редко более 200-300) исходных данных, слабо отличающихся друг от друга [18-23]. При таком подходе качество работы алгоритма ожидаемо высокое, но применимость алгоритма, обученного на равномерной и узкоспециализированной выборке, для предсказания ВТС других видов биомассы (например, более лигнифицированной или с большей зольностью) находится под вопросом.

Целью данной работы является разработка алгоритма настройки искусственной нейронной сети для предсказания высшей теплоты сгорания биомассы при анализе большого и неоднородного массива данных и сравнение результатов предсказания ИНС с результатами эмпирических формул.

В работе использовалась выборка данных, собранных из открытой базы данных Phyllis2 [24] и опубликованных обзорных статей [8, 25-27]. Каждый элемент (вид биомассы) выборки имеет данные о ВТС, данные элементного анализа (содержание углерода, водорода и азота), данные технического анализа (зольность, содержание летучих веществ и связанного углерода). Все данные были представлены в пересчёте на сухое вещество. Размер датасета – 720 образцов. Статистическая обработка не проводилась.

Для работы с ИНС использовалась библиотека `scikit-learn`: `MLPRegressor` [[sklearn.neural_network.MLPRegressor — scikit-learn 0.24.2 documentation](#)], а также встроенные алгоритмы предобработки данных [[sklearn.preprocessing.StandardScaler — scikit-learn 0.24.2 documentation](#)].

В качестве критерия качества предсказаний ИНС был выбран R^2 -критерий, «доля объясненной выборки» (y_i – истинное значение ВТС сырья i -ого образца, $y_{pred\ i}$ – предсказание ИНС для того же образца, y_{mean} – среднее значение ВТС по всей выборке):

$$R^2 = 1 - \frac{\sum(y_i - y_{pred\ i})^2}{\sum(y_i - y_{mean})^2} \quad (1)$$

В ходе работы был проведен анализ принципа работы ИНС и ранжирование гиперпараметров по их важности. В результате было сделано предположение, что отдельные гиперпараметры можно считать независимыми. На основе сформулированной гипотезы мы провели настройку гиперпараметров как последовательный процесс выбора оптимальных значений отдельных гиперпараметров в порядке убывания их значимости. Предлагаемый нами алгоритм настройки позволяет снизить временные затраты со степенной зависимости от числа допустимых значений каждого гиперпараметра до зависимости в виде произведения на это число.

Для итоговой оценки результатов настроенной ИНС были использованы два метода: k-folds cross-validation – состоит в определении величины R^2 при нескольких вариантах разбиения исходной выборки на обучающую и тестовую часть. Метод позволяет определить не среднее значение обобщающей способности и стабильность предсказаний [28-30]. В нашей задаче этот подход показывает результат $R^2 = 0.880 \pm 0.025$. Альтернативный способ - расчет предсказания ИНС по всей использованной выборке [21, 22]. В нашей задаче этот метод дает $R^2 = 0.884$, то есть оба метода дают одинаковые (в пределах ошибки) оценки результатов работы ИНС.

Для сравнения качества предсказания MLPRegressor'a с существующими эмпирическими моделями были выбраны уравнения 2 и 3 (C – содержание углерода, N – содержание азота, H – содержание водорода) из работы [8], поскольку для них показана наибольшая применимость для разрозненных выборок:

$$Q = 0.4373 \cdot C - 1.6701 \quad (2)$$

$$Q = 0.00355 \cdot C^2 - 0.232 \cdot C - 2.230 \cdot H + 0.0512 \cdot C \cdot H + 0.131 \cdot N + 20.600 \quad (3)$$

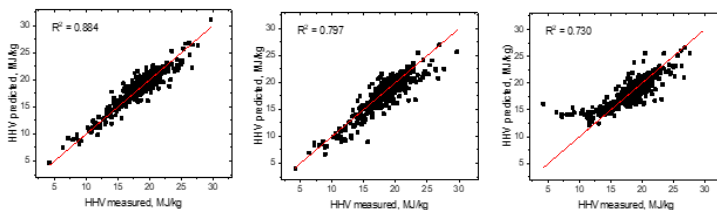


Рис. 1 – Сравнение ИНС с эмпирическими уравнениями: а) предсказание ИНС, б) расчет по уравнению 10, с) расчет по уравнению 11. Красной линией показана прямая НТС

В результате работы мы выяснили, что наиболее важную роль в улучшении качества предсказания играет предварительная обработка данных (а именно исключение зависимых переменных и шумящих значений, центрирование выборки) и размер обучающей выборки.

Работа поддержана грантом РФФИ 18-29-24028

Литература:

1. Pang C. H., Lester E., Wu T. Influence of lignocellulose and plant cell walls on biomass char morphology and combustion reactivity // *Biomass and Bioenergy*. – 2018. – Т. 119. – С. 480-491.

2. Kim J.-H., Jung S., Lin K.-Y.A., Rinklebe J., Kwon E.E. Comparative study on carbon dioxide-cofed catalytic pyrolysis of grass and woody biomass // *Bioresource Technology*. 2021; 323. <https://doi:10.1016/j.biortech.2020.124633>.

3. Yin, C.-Y. Prediction of higher heating values of biomass from proximate and ultimate analyses. *Fuel*. 2011; 90(3); 1128–1132. <https://doi:10.1016/j.fuel.2010.11.031>.

4. Vargas-Moreno J.M., Callejon-Ferre A.J., Perez-Alonso J., Velazquez-Marti B. A review of the mathematical models for predicting the heating value of biomass. *Renewable and Sustainable Energy Reviews*. 2012; 16; 3065-3083. <https://doi:10.1016/j.rser.2012.02.054>.

5. Qian, C., Li, Q., Zhang, Z., Wang, X., Hu, J., & Cao, W. Prediction of higher heating values of biochar from proximate and ultimate analysis. *Fuel*. 2020;265; 116925. <https://doi:10.1016/j.fuel.2019.116925>.

6. Górnicki, K., Kaleta, A., & Winiczenko, R. Prediction of higher heating value of oat grain and straw biomass. *E3S Web of Conferences*. 2020; 154; 01003. <https://doi:10.1051/e3sconf/202015401003>.

7. Maksimuk, Y., Antonava, Z., Krouk, V., Korsakova, A., & Kursevich, V. Prediction of higher heating value based on elemental composition for lignin and other fuels. *Fuel*. 2019, 116727. <https://doi:10.1016/j.fuel.2019.116727>.

8. Bychkov A.L., Denkin A.I., Tikhova V.D., Lomovsky O.I. Prediction of higher heating values of plant biomass from ultimate analysis data // *Journal of Thermal Analysis and Calorimetry*. 2017; 130; 1399-1405. <https://doi:10.1007/s10973-017-6350-0>.

9. Xing, J., Luo, K., Wang, H., Gao, Z., & Fan, J. A comprehensive study on estimating higher heating value of biomass from proximate and ultimate analysis with machine learning approaches. *Energy*. 2019; 116077. <https://doi:10.1016/j.energy.2019.116077>.

10. Obafemi, O., Stephen, A., Ajayi, O., & Nkosinathi, M. A survey of Artificial Neural Network-based Prediction Models for Thermal Properties of Biomass. *Procedia Manufacturing*. 2019; 33; 184–191. <https://doi:10.1016/j.promfg.2019.04.103>.
11. Estiati, I., Freire, F. B., Freire, J. T., Aguado, R., & Olazar, M. Fitting performance of artificial neural networks and empirical correlations to estimate higher heating values of biomass. *Fuel*. 2016; 180; 377–383. <https://doi:10.1016/j.fuel.2016.04.051>.
12. Uzun, H., Yıldız, Z., Goldfarb, J. L., & Ceylan, S. Improved prediction of higher heating value of biomass using an artificial neural network model based on proximate analysis. *Bioresource Technology*. 2017; 234; 122–130. <https://doi:10.1016/j.biortech.2017.03.015>.
13. Cao, H., Xin, Y., & Yuan, Q. Prediction of biochar yield from cattle manure pyrolysis via least squares support vector machine intelligent approach. *Bioresource Technology*. 2016; 202; 158–164. <https://doi:10.1016/j.biortech.2015.12.024>.
14. Ozonoh, M., Oboirien, B. O., & Daramola, M. O. Optimization of process variables during torrefaction of coal/biomass/waste tyre blends: Application of artificial neural network & response surface methodology. *Biomass and Bioenergy*. 2020; 143; 105808. <https://doi:10.1016/j.biombioe.2020.105808>.
15. Goettsch, D., Castillo-Villar, K. K., & Aranguren, M. Machine-Learning Methods to Select Potential Depot Locations for the Supply Chain of Biomass Co-Firing. *Energies*. 2020; 13(24); 6554. <https://doi:10.3390/en13246554>.
16. Akkaya, E., & Demir, A. Predicting the Heating Value of Municipal Solid Waste-based Materials: An Artificial Neural Network Model. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*. 2010; 32(19); 1777–1783. <https://doi:10.1080/15567030902882950>.
17. L.K. Abidoye, F.M. Mahdi Novel linear and nonlinear equations for the higher heating values of municipal solid wastes and the implications of carbon to energy ratios. *Journal of Energy Technologies and Policy*. 2014; 4 (5); 14-27.
18. Uzun, H., Yıldız, Z., Goldfarb, J. L., & Ceylan, S. Improved prediction of higher heating value of biomass using an artificial neural network model based on proximate analysis. *Bioresource Technology*. 2017; 234; 122–130. <https://doi:10.1016/j.biortech.2017.03.015>.
19. Li, H., Xu, Q., Xiao, K. et al. Predicting the higher heating value of syngas pyrolyzed from sewage sludge using an artificial neural

network. *Environ Sci Pollut Res.* 2020; 27; 785–797. <https://doi:10.1007/s11356-019-06885-2>.

20. Obafemi O. Olatunji, Stephen Akinlabi, Nkosinathi Madushele, Paul A. Adedeji, Ishola Felix. Multilayer perceptron artificial neural network for the prediction of heating value of municipal solid waste. *AIMS Energy.* 2019; 7(6); 944-956. <https://doi:10.3934/energy.2019.6.944>.

21. Cao, H., Xin, Y., & Yuan, Q. Prediction of biochar yield from cattle manure pyrolysis via least squares support vector machine intelligent approach. *Bioresource Technology.* 2016; 202; 158–164. <https://doi:10.1016/j.biortech.2015.12.024>.

22. Dashti, A., Noushabadi, A. S., Raji, M., Razmi, A., Ceylan, S., & Mohammadi, A. H. Estimation of biomass higher heating value (HHV) based on the proximate analysis: Smart modeling and correlation. *Fuel.* 2019; 257; 115931. <https://doi:10.1016/j.fuel.2019.115931>.

23. Elmaz, F., Büyükcakır, B., Yücel, Ö., & Mutlu, A. Y. Classification of solid fuels with machine learning. *Fuel.* 2020; 266; 117066. <https://doi:10.1016/j.fuel.2020.117066>.

24. Phyllis2[Internet], database for (treated) biomass, algae, feedstocks for biogas production and biochar. TNO Biobased and Circular Technologies. Available from: <https://phyllis.nl/>.

25. Yin, C.-Y. Prediction of higher heating values of biomass from proximate and ultimate analyses. *Fuel.* 2011; 90(3); 1128–1132. <https://doi:10.1016/j.fuel.2010.11.031>.

26. Parikh, J., Channiwala, S. A., & Ghosal, G. K. A correlation for calculating HHV from proximate analysis of solid fuels. *Fuel.* 2005; 84(5); 487-494. <https://doi:10.1016/j.fuel.2004.10.010>.

27. Krishnan, R., Hauchhum, L., Gupta, R., & Pattanayak, S. Prediction of equations for higher heating values of biomass using proximate and ultimate analysis. In 2018 2nd International Conference on Power, Energy and Environment: Towards Smart Technology (ICEPE). 2018; 1-5. <https://doi:10.1109/epetsg.2018.8658984>.

28. Kohavi, R. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Ijcai.* 1995; 14(2); 1137-1145.

29. Rodriguez, J. D., Perez, A., & Lozano, J. A. Sensitivity analysis of k-fold cross validation in prediction error estimation. *IEEE transactions on pattern analysis and machine intelligence* 2009; 32(3); 569-575. <https://doi.org/10.1109/TPAMI.2009.187>.

30. Bengio, Y., & Grandvalet, Y. No unbiased estimator of the variance of k-fold cross-validation. *Journal of machine learning research* 5. 2004; 1089-1105.

**ПРЕДЛОЖЕНИЕ АЛГОРИТМИЧЕСКОГО
ПРОЕКТИРОВАНИЯ ФАСАДА НА ОСНОВЕ ВВОДНОГО
ВАРИАТИВНОГО ТВОРЧЕСКОГО СОСТАВЛЯЮЩЕГО
И ГЕОМЕТРИЧЕСКОЙ СТРУКТУРЫ**

Е.С. Логачев, Ю.А. Калпакова

**Новосибирский государственный технический университет,
г. Новосибирск,**

**Новосибирский государственный архитектурно-строительный
университет (Сибстрин), г Новосибирск
e.logachev@edu.sibstrin.ru**

**Научный руководитель: Бурило Н.А., ст. преподаватель каф.
АРГС**

В данной статье предложен практический алгоритм для разработки структурных декоративных навесных фасадов с применением геометрических основ и любых творческих двухцветных зарисовок.

This article offers a practical algorithm for the development of structural decorative curtain facades using geometric foundations and any creative two-color sketches.

При работе с навесными фасадами нередко у архитектора возникает вопрос о параметризации некоторых элементов фасада, но определиться с принципом, при котором в процесс можно внедрить и легкую руку художника, и геометрически правильную привязку форм с математической точки зрения, довольно сложно. Именно этот принцип нам бы и хотелось осветить в данной статье.

Цель: предложить алгоритм, который включает в себя математическое построение форм с использованием творческих тематических зарисовок для построения фасада.

Задачи:

1. Выбрать геометрическую светопрозрачную основу;
2. Разработать и внедрить в алгоритм блок, отвечающий за творческую подложку к основной геометрии фасада;
3. Продумать к данным блокам инвариантность по габаритам и масштабным характеристикам.

При выборе основного принципа геометрического построения фасада учитывались следующие критерии:

1. Структурность, тектоника;

2. Конструктивно выгодная система, которая будет иметь узловые точки соединения;
3. Структура, состоящая из центров (множество точек, аттракторов). Данный критерий и позволит завуалировать в повторяющейся геометрии художественный смысл [1].

Исходя из выше приведенных критериев, была избрана диаграмма Вороного - случайного множества точек на плоскости фасада [2].

Для построения данной диаграммы было использовано программное обеспечение Rhinoceros и плагин Grasshopper. На плоскости фасада вариативно задано множество точек, которые и будут основой всего. (Рисунок 1)

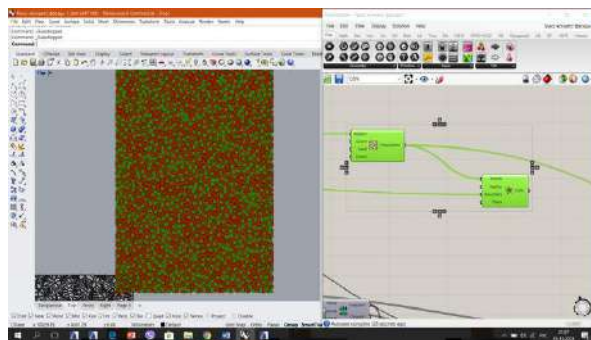


Рис.1 – Множество точек и нод воногоі. Демонстрация множества точек на плоскости и построения структуры геометрической основы диаграммы Вороного

Далее был разработан блок, сформированный на основе зарисовки от руки. Необходимое условие для данного алгоритма то, что рисунок необходимо отредактировать до сочетания чисто черного и чисто белого цвета. Тогда мы сможем применить алгоритм, который построен на дереве данных. Это дерево состоит из двух значений «0» и «1». После чего рисунок с этими данными проецируется на множество точек (Рисунок 2), более плотное, чем основа на рисунке 1 (Рисунок 1).

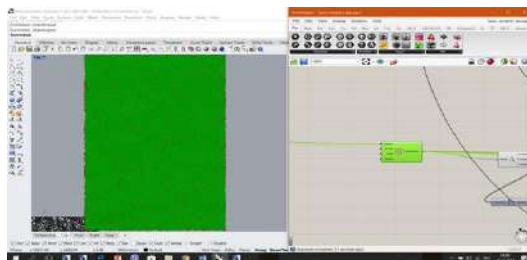


Рис.2 – Множество точек с большей плотностью. Демонстрация множества точек с плотностью в 8 раз превышающее предыдущее множество точек

Далее необходимо спроецировать по масштабу подготовленный рисунок с деревом данных и подключить разделяющий параметр, т.к. в рисунке есть переходящие градиенты на границах между белым и черным, чтобы избрать выделение на черном фоне (соответствие с «0») зададим сравнение с диапазоном от 0 до 0,6. Впоследствии точки из множества, попадающие на черный фон, именуются как 0, а благодаря сравнению значений можно выделить данные точки и отделить их от лишних (Рисунок 3)

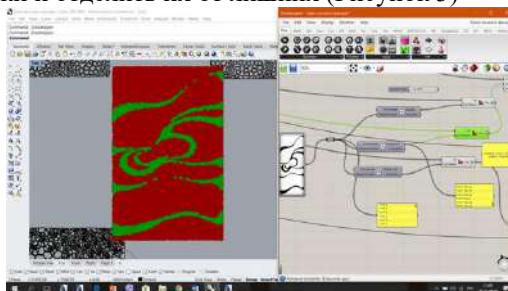


Рис.3 – Выделение необходимого множество точек. Демонстрация структурирования множества точек исходя из первоначальной зарисовки для фасада

Следующим действием будет структурирование данных. Необходимо соединить 2 дерева: список 1-го множества точек (Рисунок 1) и 2-го получившегося множества точек (Рисунок 3). На основе этого объединения уже строится конечная диаграмма Вороного (Рисунок 4) [2,3], которая впоследствии будет рассчитана на нагрузки и реализована в проектных и рабочих документациях.

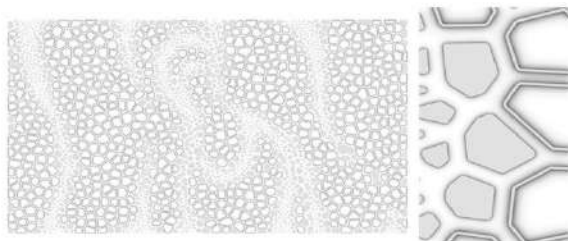


Рис.4 – Конечный результат совмещения двух множеств точек на основе творческого рисунка и геометрического алгоритма диаграммы Вороного

В заключении хотелось бы отметить, что на данный момент удалось завязать параметр на совмещении двух структур одного геометрического типа. Благодаря этому алгоритму можно разработать множество навесных светопрозрачных фасадов различных габаритов, плотностей сеток и тематических образов. После обработки данных удалось вывести 3д-модели, которые впоследствии проектировщики могут проверить на несущую способность и подобрать необходимое сечение. Из данного параметра можно вывести раскладку панелей с обмерами для рабочих чертежей и последующего монтажа. В перспективе планируется завязать параметр и на третью плоскость редактирования (по направлению нормали фасада), что даст более выраженный и динамичный результат [4]. Также планируется завязать параметр и другими структурами, к примеру, с триангуляцией Делоне. Но самым главным плюсом этого алгоритма является инвариантность образов и замыслов архитектора.

Данное практическое исследование выполнено в рамках работы над объектом, в настоящее время находящимся на стадии архитектурного проектирования, который в дальнейшем будет реализован на территории Новосибирской области.

Литература:

1. *Препарата, Ф.* Вычислительная геометрия: введение / Ф. Препарата, М. Шеймос; Перевод с англ. С. А. Вичеса, М. М. Комарова; Под ред. Ю. М. Баяковского. - М. : Мир, 1989. - 478 с. Ф.

Препарата, М. Шеймос. Вычислительная геометрия: Введение. — М.: Мир, 1989. Стр. 295.

2. *G.F. Voronoi.* Nouvelles applications des paramètres continus à la théorie de formes quadratiques (фр.) // Journal für die reine und angewandte Mathematik. — 1908. — Vol. 134. — P. 198—287.

3. *Александров А. Д., Вернер А. Л., Рыжик В. И.* Стереометрия. Геометрия в пространстве: Учеб. пособие для уч. ст. кл. и абитуриентов. — Висагинас, Alfa, 1998.— 576 с.

4. *Логачев, Е. С.* Внедрение энергоэффективных систем динамические малые архитектурные объекты на базе альтернативных источников энергии / Е. С. Логачев, Н. А. Бурило // Интеллектуальный потенциал Сибири : материалы 28-ой Региональной научной студенческой конференции: в 3 частях, Новосибирск, 13–22 мая 2020 года / Под. ред. Соколовой Д.О.. – Новосибирск: Новосибирский государственный технический университет, 2020. – С. 556-557.

МЕТОДЫ ОРГАНИЗАЦИИ ХРАНЕНИЯ БОЛЬШИХ ОБЪЁМОВ ДАННЫХ

А.В. Лузьинин

**Научный руководитель: к.т.н. Н.Д. Ганелина
Новосибирский государственный технический университет,
г. Новосибирск, 261-96@mail.ru**

Работа посвящена исследованию методов хранения больших объёмов данных, позволяющих быстро производить операции поиска, удаления и обновления в структуре больших данных. Также в данной работе представлен способ оптимизации хранимых данных и результаты применения данного способа для оптимизации базы данных.

The work is devoted to research methods for storing large amounts of data that allows you to provide quick search, delete and update operations in the big data structure. Also, this work presents the method to optimize stored data and the results of implementing this method for database optimization.

Введение. В настоящее время одной из актуальных проблем, требующих решения, является проблема хранения больших данных

и работа с ними. Традиционно для хранения больших объёмов данных используется реляционная база данных, использующая язык структурированных запросов (Structured Query Language). Такая организация данных позволяет относительно легко и быстро оперировать огромными объёмами данных. Однако существуют задачи, для которых организация данных в реляционной базе данных не может быть оптимальной: в случае, если она используется для хранения однотипных данных, таких как параметры устройств. Среди этих данных может присутствовать большое количество дубликатов, хотя каждая структурная единица этих данных будет уникальной и предназначена для конкретного устройства. Таким образом, база данных раздувается наличием дубликатов, и с течением времени становится всё тяжелее управлять этим объёмом данных, а также растут затраты на их хранение.

Основная часть. В данном случае исследуемая система представляет собой реляционную базу данных, содержащую в себе среди прочего данные о том, какие параметры должны быть применены к конечному устройству, которое запрашивает с сервера необходимое для своей работы приложение и параметры в виде модели данных (XML).

Проблема существующей организации структуры данных состоит в том, что текущая организация хранения данных порождает огромное количество фактических дубликатов данных, хранящихся в системе. Например, существует 2 уникальных значения параметра в системе: “On” и “Off”. Количество устройств, использующих данный параметр равно 1000, что порождает 1000 экземпляров данных значений в базе данных. Таким образом, вместо того, чтобы хранить 2 уникальных значения параметра, каждый параметр хранится как отдельная сущность. Также существуют шаблоны параметров устройств, которые позволяют снизить количество хранимых данных, однако при изменении параметра одного конкретного устройства требуется хранить отдельный экземпляр той же сущности. Очевидно, что решение неоптимально и влечёт за собой высокие затраты на обслуживание, управление и хранение этих данных.

Данную структуру можно оптимизировать, проведя сжатие данных до экземпляров уникальных значений, а также создав способ организации связи и редактирования хранимых данных. Это решение можно реализовать с использованием как SQL, так и NoSQL баз данных. Поскольку структура данных является иерархической, было принято решение использовать

графоориентированную СУБД для их хранения. С этой целью была проведена оптимизация хранимых данных, а также осуществлён их перенос из реляционной базы данных в графоориентированную, а также создана система для манипуляции этими данными.

Заключение. Проведя упомянутые выше преобразования с базой данных, удалось снизить объём хранимых данных примерно на 85% (с 4 гигабайт до 600 мегабайт), сохранив базовую логику работы системы и снизив затраты на её обслуживание.

Можно сделать вывод о том, что преобразования, представленные в данном докладе, могут быть применены для систем с подобной организацией данных.

Литература:

1. Gyssens M., Paredaens J., Van den Bussche J. A Graph-Oriented Object Database Model // IEEE Transactions on Knowledge and Data Engineering. 1994. Vol. 6. P. 572-586.

2. Alabdullah B., Beloff, N., White M. Rise of Big Data – Issues and Challenges // 2018 21st Saudi Computer Society National Computer Conference (NCC) Location: Riyadh, Saudi Arabia, 2018. Vol. 1. P. 1-6

АНАЛИЗ ДАННЫХ ОТКЛИКА РЕЗИСТИВНОГО СЕНСОРА НА ВОЗДЕЙСТВИЕ КИСЛОРОДА

С.П. Новиков

**Южный федеральный университет, г. Таганрог,
n_serg7@mail.ru**

Научный руководитель: Плуготаренко Н.К., к.т.н., доцент

Ставится задача рассмотреть фазы анализа данных сигналов от резистивных газовых датчиков на основе диоксида циркония и кремния, применяемых для оценки концентрации кислорода. Показано изменение нормированного сопротивления отклика датчика при воздействии кислородом и проведен анализ первой производной кривой отклика. Описана методика обработки сигналов, снижающая влияние шума и дрейфа сигнала. При этом используется экстремум первой производной отклика датчика, предложенный в качестве калибровочной зависимости.

The task is to consider the phases of data analysis of signals from resistive gas sensors based on zirconium dioxide and silicon used to

assess oxygen concentration. The change in the normalized resistance of the sensor response when exposed to oxygen is shown and the analysis of the first derivative of the response curve is carried out. A signal processing technique that reduces the influence of noise and signal drift is described. In this case, the extremum of the first derivative of the sensor response, proposed as a calibration dependence, is used.

Введение. Полупроводниковые сенсоры используются в качестве датчиков различных газов. При большом количестве положительных свойств, таких как: чувствительность, быстродействие, миниатюрность, невысокая стоимость, проблемой остается низкая селективность. Повышение селективности полупроводниковых сенсоров является актуальным применительно ко многим задачам газового анализа, некоторые способы решения этой проблемы найдены [1,2].

Постановка задачи. Использование методов обработки сигналов может обеспечить большую избирательность газового датчика. Авторы [3,4] использовали частотно-временной и переходный анализ для динамической обработки сигналов. Ключевые характеристики сигнала могут быть рассмотрены и проанализированы для идентификации или количественной оценки обнаруженного газа с помощью этих методов.

Анализ данных отклика датчика позволяет решить проблему дрейфа сигнала, возникающего из-за длительной работы или внешнего воздействия [5] и может быть использован для уменьшения погрешности измерения и повышения стабильности и избирательности датчика [6].

Проблема шума сигнала должна быть решена при обработке сигнала датчика [7], при этом могут использоваться разные методы фильтрации. В дополнение к широко используемой зависимости абсолютного изменения сопротивления от концентрации газа были предприняты попытки использовать первую производную отклика газового датчика.

Газочувствительные пленки $\text{SiO}_2\text{-ZrO}_2$ были получены по золь-гель технологии из спиртового раствора на основе тетраэтоксисилана.

Экспериментальная часть. Исходный сигнал – большой массив данных о сопротивлении в зависимости от времени, поступающий от датчика, который накапливается во времени. При этом регистрируется время начала воздействия газа на датчик, а также время завершения воздействия. Данные, полученные в ходе

измерений, состоят из суммы полезного сигнала, случайного шума и ошибок, полученных в ходе измерений.

Также в сигнале присутствуют шумы, которые возникают в электронных системах записи, когда происходит изменение условий окружающей среды и т.д. Для повышения воспроизводимости результатов анализа, системой обработки выполняется предварительное сглаживание исходных данных.

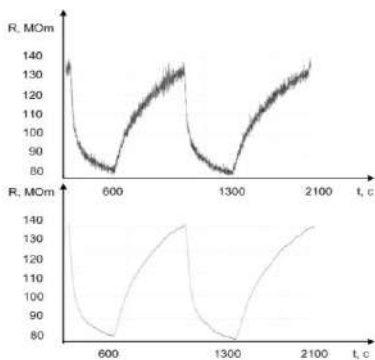


Рис.1 – отображение оригинального и отфильтрованного сигнала

Далее определяются локальные экстремумы, соответствующие точкам минимума и максимума при отдельном измерении. Данные нормализуются для более глубокой обработки. Отклик газового датчика рассчитывался по формуле:

$$S = \frac{R_t}{R_0},$$

где R_0 - сопротивление пленки до воздействия газа, а R_t - сопротивление пленки при воздействии газа в момент времени t .

В дальнейшем рассчитываются параметры динамики отклика сопротивления датчика, а именно скорость реакции датчика на газ как первая производная чувствительности по времени (dS/dt). Была проведена серия экспериментов с различными концентрациями газа.

Результаты. В качестве рабочей температуры выбрана температура 250 градусов цельсия. На рисунке 2 показаны графики параметров отклика для начального участка под воздействием кислорода после нормализации и приведения к единому источнику.

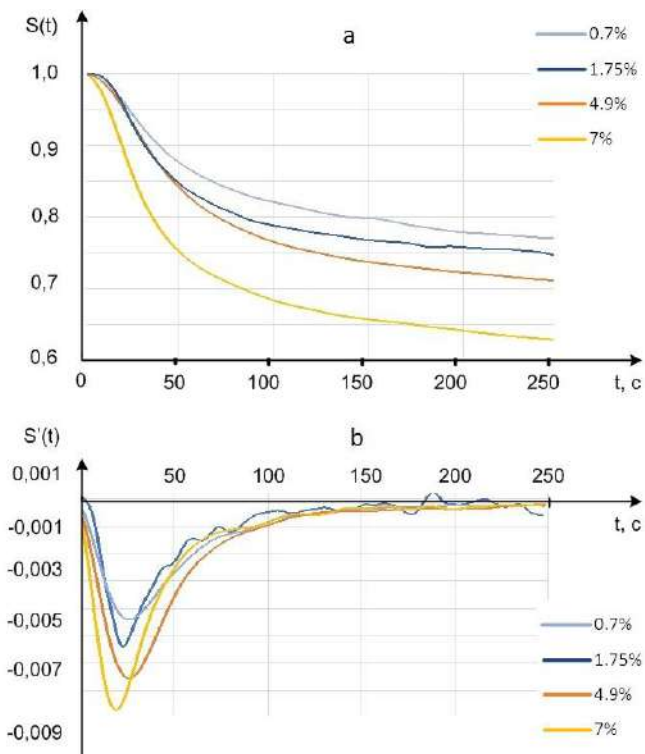


Рис.2 – Отображение реакции адсорбции $S(t)$ (a); первой производной $S'(t)$ (b) при различных концентрациях кислорода

В работе [8] было показано, что экстремумы первой и второй производных $S(t)$ зависят от концентрации газа, поэтому их можно использовать для построения калибровочных зависимостей. На рисунке 3 показаны калибровочные кривые, построенные для абсолютных отклонений $S(t)$, а также для экстремумов $S'(t)$ для исследуемого датчика.

Можно отметить, что коэффициент достоверности аппроксимации линейной зависимости для первой производной максимален.

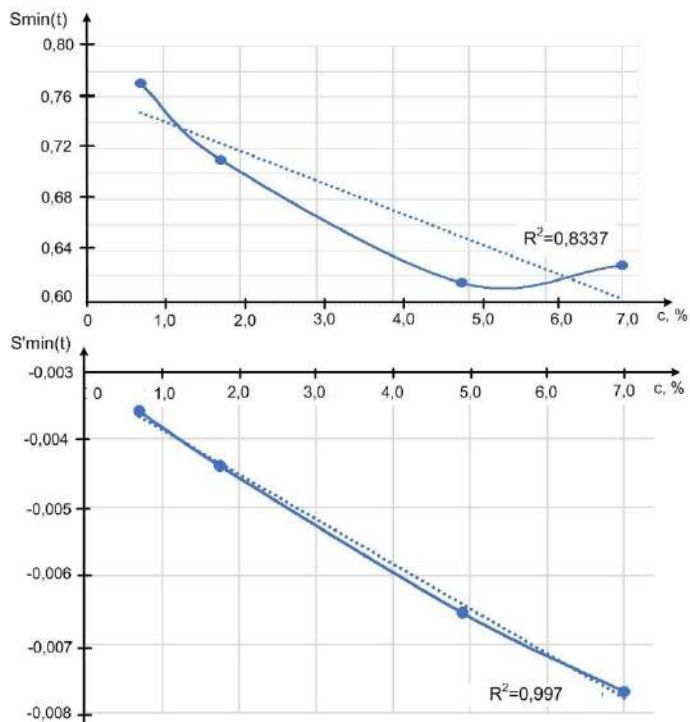


Рис.3 – Калибровочные кривые

Выводы. Таким образом, параметр экстремум производной отклика газового датчика $S'(t)$, показал высокую степень линейаризации при построении калибровочных кривых.

Предлагаемая схема обработки сигналов датчиков и построение калибровочных кривых позволяют не дожидаться наступления установившегося режима при подаче газа для оценки концентрации газа и сократить время детектирования. В дальнейшем оценка нескольких параметров динамики отклика датчика может быть способом оценки не только концентрации газа, но и распознавания самого газа.

Литература:

1. Sauter D., Weimar U., Noetzel G., Mitrovics J., Göpel W. Ibid.,2000, v. 69, p. 1—9.
2. Обвинцева Л.А., Губанова Д.П. Журнал аналитической химии, 2004, № 8, т. 59, с. 876.

3. Guangfen, W.; Wen, A.; Fajin, G.; Zhenan, T.; Jun, Y. The Hilbert-Huang Transform and Its Application in Processing Dynamic Signals of Gas Sensors. In Proceedings of the 2009 International Conference on Information Engineering and Computer Science, Wuhan, China, 19–20 December 2009; pp. 1–4.

4. A. Wen, Y. Chun-Ying and W. Guang-Fen, "Dynamic signal processing for gas sensors based on Hilbert-Huang Transform", Proc. of the 3rd IEEE Int. Conf. on computer Science and Information Technology, pp. 241-245, 2010.

5. Self-Repairing classification algorithms for chemical sensor array Gabriele Magnaa, Corrado Di Nataleb, Eugenio Martinellia Sensors & Actuators: B. Chemical 297 (2019) 126721.

6. Jakub Sorocki and Artur Rydosz, A Prototype of a Portable Gas Analyzer for Exhaled Acetone Detection. Appl. Sci. 2019, 9(13), 2605; DOI: 10.3390/app9132605.

7. Peter F. Satterthwaite, Sebastian Eberleb, Stefan Nedelcub, Cosmin Romanb, Christofer Hierold Transient and steady-state readout of nanowire gas sensors in the presence of low-frequency noise Sensors & Actuators: B. Chemical 297 (2019).

8. Du H, Xie G, Su Y, Tai H, Du X, Yu H, Zhang Q. A New Model and Its Application for the Dynamic Response of RGO Resistive Gas Sensor. Sensors. 2019; 19(4):889. <https://doi.org/10.3390/s19040889>.

АНАЛИЗ ВРЕМЕНИ РАБОТЫ АЛГОРИТМА СС-ABOD И МОДИФИЦИРОВАННОГО МЕТОДА ТОРГЕРСОНА

Н.С. Олейник

Новосибирский государственный технический университет,

г. Новосибирск, olejnik.2015@stud.nstu.ru

Научный руководитель:

к.т.н., доцент кафедры МиС В.Ю. Щеколдин

Аннотация. При анализе многомерных данных при помощи статистических методов часто возникают вычислительные сложности, связанные с существенным увеличением времени работы алгоритмов по мере роста размерности данных. В связи с этим проводится анализ и сравнение временных затрат при работе с данными больших размерностей алгоритма СС-ABOD и

модифицированного метода Торгерсона. Оценены уравнения для расчёта времени работы программы для заданного объёма выборки.

Annotation. When analyzing multidimensional data using statistical methods, computational difficulties often arise due to a significant increase in the running time of algorithms as the dimension of the data grows. In this regard, the analysis and comparison of the time spent when working with data of large dimensions of the CC-ABOD algorithm and the modified Thorgerston method is carried out. Equations for calculating the running time of the program for a given sample size are estimated.

В различных приложениях технического, экономического, социального характера приходится иметь дело с большими объемами данных, поскольку исследуемые объекты и процессы, как правило, имеют очень сложную природу, а их функционирование и эволюция напрямую связаны с сохранением и анализом больших объемов информации с различным количеством факторов. В то же время при выборе определенных методов исследования важную роль играет возможность получения правильных и хорошо объясненных результатов, в частности, позволяющих строить различные графические интерпретации, удобные для восприятия как специализированными специалистами, так и людьми, не обладающими глубокими знаниями в области прикладной статистики, но использующими ее подходы в своей деятельности.

Сами проблемы «больших данных» и данных больших размерностей начали обсуждаться ещё с 2008 г. [1, 2] Чаще всего их связывают с необходимостью анализа неструктурированных данных больших объёмов. Одной из основных сложностей при работе с многомерными данными является адекватное их восприятие. Если число факторов равно двум, то мы получаем простую плоскую проекцию, для трёх измерений возникает необходимость добавления третьей оси и включение пространственного воображения.

1. Постановка задачи

С точки зрения решения вопроса о построении качественной визуализации многомерных данных, основную задачу можно сформулировать следующим образом. Необходимо разместить исходные объекты в определенном метрическом пространстве, по возможности двумерном, при выполнении условия сохранения их взаимного расположения: похожие элементы должны располагаться в этом пространстве близко друг к другу, а отличающиеся – далеко.

При этом под «схожестью» объектов будем понимать величину, обратную выбранной в метрическом пространстве мере расстояния между ними.

Пусть имеется симметричная матрица $D = \|D_{ij}\|_{n \times n}$ различий D_{ij} между объектами, расположенными в n -мерном пространстве. Необходимо построить пространство меньшей размерности $r < n$, в котором матрица различий $d = \|d_{ij}\|_{n \times n}$ между анализируемыми объектами была бы в смысле некоторого критерия близка к исходной матрице различий.

В качестве критерия, определяющего свойства визуализации, будем использовать оценку классификатора наблюдений, строящегося на основе ABOD-подхода. Основная идея этого метода, применяемого для выявления выбросов, состоит в том, что при анализе геометрии расположения исходных данных может быть получена более полная информация, которая бы описывала их внутреннюю структуру. Вычислительная схема ABOD-подхода будет представлена ниже.

Для оценки влияния параметров задачи на скорость вычислений требуется отметить, что тестирование будет проводиться в системе с 32 ГБ памяти DDR4 и процессором i7-8700K с тактовой частотой 3.7 ГГц. Под параметрами будем иметь в виду объём выборки и размерность данных. Основная задача состоит в том, чтобы исследовать временные затраты работы алгоритма CC-ABOD (Cumulative Curves for Angle Based Outlier Detection) [4, 5] и модифицированного метода Торгерсона [3], используемого для снижения размерности решаемой задачи. Соответствующие

2. Описание исследования

Для получения временных затрат на работу алгоритма воспользуемся встроенной функцией языка C#, которая позволяет проводить замеры различных участков программного кода.

```
Stopwatch stopwatch = new Stopwatch();  
stopwatch.Start();  
stopwatch.Stop();
```

Для обеспечения статистической корректности, проводится серия экспериментов с числом репликаций $R = 10$ с изменением объёма N от 100 до 1000 с шагом 100 и с изменением количества факторов от 5 до 10 с шагом 1 при фиксированном объёме. Полученные результаты усредняются и по вычисленным показателям формируются дальнейшие выводы и рекомендации.

Для данной реализации в разработанный программный комплекс MDPV (Multidimensional Data Processing and Visualization), который позволяет проводить анализа данных на наличие выбросов, а также имеет функционал визуализации многомерных данных, был усовершенствован для проведения описанного тестирования.

3. Обсуждение результатов

В таблице 1 представлены временные затраты анализируемых алгоритмов в зависимости от объёма выборки и числа экзогенных факторов. По усреднённым по факторам и логарифмированным значениям из таблицы 1 были построены адекватные модели регрессии оценки времени общей работы, построения проекций и работы алгоритма классификатора СС-АВОД, представленные на рисунке 1.

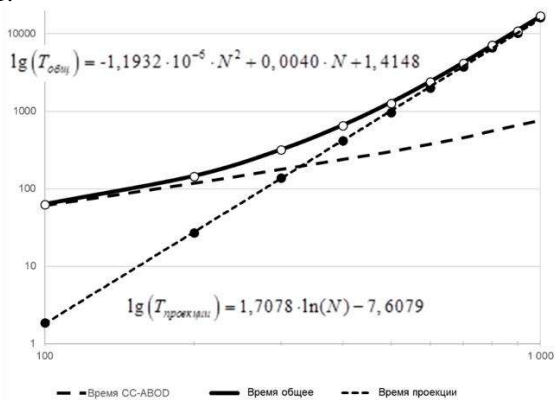


Рисунок 3 – Графическая интерпретация регрессионных зависимостей моделей оценки времени работы алгоритмов

Таблица 1 – Результаты замеров времени работы тестируемых алгоритмов от объёма выборки и количества факторов

Объём	Факторы	Время проекции (с)	Время общее (с)	Объём	Факторы	Время проекции (с)	Время общее (с)
100	5	1,729	60,210	600	5	1983,084	2357,128
100	6	1,903	64,537	600	6	1988,763	2363,017
100	7	1,918	63,438	600	7	1975,963	2349,950
100	8	1,877	62,966	600	8	1990,857	2365,555
100	9	1,878	62,525	600	9	2193,750	2586,335
100	10	1,900	62,886	600	10	1979,628	2354,245
200	5	29,136	150,939	700	5	3695,375	4149,059
200	6	29,452	155,278	700	6	3698,169	4151,898
200	7	28,914	150,390	700	7	3700,259	4153,607
200	8	24,868	138,015	700	8	3715,027	4169,989
200	9	24,715	137,718	700	9	3857,422	4317,008
200	10	25,049	138,520	700	10	3753,563	4209,160
300	5	128,477	302,293	800	5	6760,678	7327,818
300	6	145,340	327,839	800	6	6497,836	7051,189
300	7	147,671	330,613	800	7	7007,232	7580,105
300	8	150,555	336,196	800	8	6514,515	7066,331
300	9	130,926	305,264	800	9	6507,308	7058,377
300	10	125,647	298,113	800	10	6461,717	7012,344
400	5	395,656	631,523	900	5	10304,543	10955,702
400	6	450,630	699,372	900	6	10182,419	10827,669
400	7	448,238	694,771	900	7	9967,762	10604,191
400	8	396,226	631,522	900	8	10008,415	10646,442
400	9	394,794	630,167	900	9	10396,335	11062,641
400	10	400,129	636,972	900	10	10346,979	11062,614
500	5	963,201	1264,801	1000	5	10304,543	10955,702
500	6	963,433	1265,022	1000	6	10182,419	10827,669
500	7	982,900	1288,716	1000	7	9967,766	10604,191
500	8	964,825	1266,949	1000	8	10008,415	10646,442
500	9	965,505	1267,331	1000	9	10396,335	11062,641
500	10	956,061	1256,557	1000	10	10346,979	11062,614

Выводы:

Анализ данных показал, что число экзогенных факторов не оказывает влияния на общее время работы алгоритма, и поэтому достаточно изучать лишь влияния объёма выборки.

Рисунок 1 наглядно демонстрирует, что по мере увеличения объёма данных, доля времени работы метода СС-ABOD (классификатора) становится всё меньше, тогда как время на построение проекции (визуализатора) существенно растёт, достигая при больших объёмах выборки (1000 и более) 95% от общего времени работы.

Исходя из полученных результатов можно предположить, что разрабатываемый алгоритм может быть существенно улучшен только в разделе построения проекций в использование возможностей распараллеливания вычислений на языке C#, так и разработка дополнительных модификаций для сокращения числа вычислений.

Литература:

1. Черняк Л. Большие данные – новая теория и практика // Открытые системы. СУБД – 2011. – №10. – С.18-25.
2. Jacobs A. The pathologies of big data // Communications of the ACM. – 2009. – V. 52. – №. 8. – p.36-44.
3. Torgerson W. S. Theory and Methods of Scaling. – N.Y.: Wiley, 1958. – 245 p.
4. Олейник Н. С., Щёколдин В. Ю. Разработка модификации метода главных проекций Торгерсона с применением анализа кумулятивных кривых в задаче выявления выбросов в данных больших размерностей // Вычислительные технологии. - 2020. – Т. 25, № 3. – С. 119–129.
5. Kriegel H., Schubert M., Zimek A. Angle-Based Outlier Detection in High-dimensional Data. // Proc. of the 14th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, Las Vegas, NV, 2008. – P. 444-452.

ОПТИМАЛЬНОЕ ПЛАНИРОВАНИЕ ЭКСПЕРИМЕНТА НА НАДЕЖНОСТЬ СВЕТОИЗЛУЧАЮЩИХ ДИОДОВ

Е.А. Осинцева, Е.В. Чимитова

Новосибирский государственный технический университет,
г. Новосибирск, osinceva.j@gmail.com

Построение оптимального плана эксперимента на примере данных об исследовании светодиодов (LED) на основе винеровской деградационной модели. В основе оптимального планирования эксперимента лежит определение оптимальных величин нагрузки, количества объектов, соответствующих данным величинам, и моментов времени измерения деградационного показателя.

Construction of an optimal design using the example of LED degradation analysis based on Wiener degradation model. An algorithm for constructing an optimal design, which includes determining optimal stress levels, number of tested devices and time moments for measuring the degradation index, has been developed.

Важным этапом разработки новых высокотехнологических изделий является проведение испытаний на надежность. При проведении таких испытаний обычно проводят ускоренные эксперименты, в которых исследуемые объекты подвергаются повышенным нагрузкам. Однако даже в таких условиях данных об отказах может быть недостаточно. В этом случае привлекают дополнительную информацию о деградации исследуемых изделий. В данной работе рассматривается широко используемая в приложениях винеровская деградационная модель с учетом объясняющих переменных (ковариат) [1]. Поскольку время, отведенное на проведение эксперимента всегда ограничено, для повышения точности оценок параметров модели и показателей надежности целесообразно использовать оптимальное планирование эксперимента [2,3]. Предложенный в данных работах алгоритм основан на оптимизации функционала от информационной матрицы Фишера при заданных ограничениях на величину нагрузки, длительность проведения эксперимента и минимальный интервал между моментами времени измерения показателя деградации.

Светодиоды имеют более высокую яркость и меньшее энергопотребление, чем традиционные источники света. В [4] была зафиксирована деградация 24 светодиодов при двух уровнях электрического тока: 35 мА и 40 мА, при нормальном уровне 25 мА.

Интенсивность света уменьшалась на протяжении всего эксперимента, который длился в течение 250 часов.

Целью данной работы является построение оптимального плана эксперимента для проведения испытаний на надежность светодиодов, которое включает в себя определение оптимальных величин нагрузки, количества объектов, соответствующих данным величинам, и моментов времени измерения деградационного показателя.

Определим оптимальный план эксперимента:

$$\xi = \left\{ \begin{matrix} x_{(1)} & \dots & x_{(q)} \\ \omega_1 & \dots & \omega_q \end{matrix}, t_1 \dots t_k \right\},$$

где $x_{(1)}, \dots, x_{(q)}$ представляю собой совокупность уникальных точек;

$$\omega_i = \frac{n_i}{n}, \quad \omega_i \geq 0, \quad \sum_{i=1}^q \omega_i = 1, \quad n_i - \text{количество объектов,}$$

исследованных под различными нагрузками $x_{(i)}$, $i = \overline{1, q}$, $q \leq n$,

$\sum_{i=1}^q n_i = n$, и $t_1 \dots t_k$ - моменты времени измерения деградационного

показателя. Все объекты выборки разбиты на группы, соответствующие разным значениям вектора ковариаты.

Первым этапом необходимо на основе результатов предварительных испытаний построить вероятностную модель надежности, в качестве которой рассмотрим винеровскую деградационную модель [1].

Случайный процесс, характеризующий процесс деградации исследуемых изделий, называется деградационным винеровским процессом, если данный процесс центрирован, с независимыми приращениями и данные приращения подчиняются нормальному закону распределения:

$$f(u, \theta_1, \theta_2) = \frac{1}{\theta_2 \sqrt{2\pi}} \exp\left(-\frac{(u - \theta_1)^2}{2\theta_2^2}\right),$$

где $\theta_1 = \mu(\rho(t + \Delta t) - \rho(t))$ - параметр сдвига, $\theta_2 = \sigma\sqrt{\rho(t + \Delta t) - \rho(t)}$ - параметр масштаба, $\sigma > 0$. В качестве функции тренда рассмотрим степенную функцию вида $\rho(t) = \left(\frac{t}{\exp(\beta \cdot x)}\right)^y$, где x - скалярная ковариата (в рассматриваемой задаче это величина электрического тока).

Оценки максимального правдоподобия неизвестных параметров $\hat{\sigma} = 0.45$, $\hat{\mu} = 0.1$, $\hat{\gamma} = 0.48$, $\hat{\beta} = -0.18$.

На рисунках 1 и 2 показаны графики функций тренда и значения индекса деградации при силе тока 35 и 40 мА.

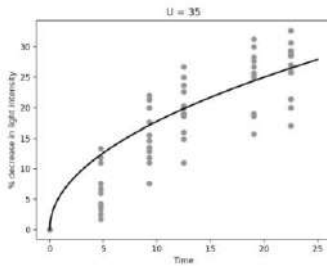


Рис. 1. – деградация светодиодов при силе тока 35

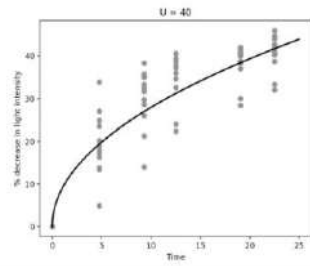


Рис. 2. – деградация светодиодов при силе тока 40

На основе полученной модели предлагается определить план эксперимента по предложенному в [2] алгоритму, позволяющему повысить точность оценок параметров модели в результате решения оптимизационной задачи:

$$\begin{cases} M(I(\xi)) = -\det(I(\xi)) \rightarrow \min, \\ 30 \leq x_{(i)} \leq 40, i = \overline{1, q}, \\ t_0 = 0, t_k \leq 250, t_{j+1} - t_j > 0, j = \overline{0, k}, \end{cases}$$

где $q = 2$ и $k = 5$.

Результаты построения D – оптимального плана эксперимента представлены в таблице 1.

Таблица 1. D – Оптимальный план эксперимента

Начальный план эксперимента	$\xi^0 = \left\{ \begin{matrix} 35 & 40 & 50 & 100 & 150 & 200 & 250 \\ 0,5 & 0,5 & & & & & \end{matrix} \right\}$ $\det(I(\xi^0)) = 5,45e - 20.$
Оптимальный план эксперимента	$\xi^* = \left\{ \begin{matrix} 30 & 40 & 100 & 111 & 115 & 129 & 148 \\ 0,5 & 0,5 & & & & & \end{matrix} \right\}$ $\det(I(\xi^*)) = 3,06e - 16$

Целью построения оптимального плана эксперимента является повышение точности оценок параметров модели. Проанализируем статистические свойства оценок параметров модели методом Монте-Карло. В данной работе с помощью средств компьютерного моделирования было сгенерировано $N = 100000$ выборок в соответствии с начальным и оптимальным планами эксперимента.

В таблице 2 приведены значения определителя оценки ковариационной матрицы

$$\left\| \frac{1}{N} \sum_{l=1}^N (\hat{\theta}_l^i - \bar{\theta}^i) (\hat{\theta}_l^j - \bar{\theta}^j) \right\|_{4 \times 4},$$

где $(\theta^1, \theta^2, \theta^3, \theta^4) = (\sigma, \mu, \gamma, \beta)$, $\hat{\theta}_l^i$ - оценка максимального правдоподобия параметра θ^i , полученная из l -ой выборки и $\bar{\theta}^i = \frac{1}{N} \sum_{l=1}^N \hat{\theta}_l^i$.

Таблица 2 – Определитель оценочной ковариационной матриц

Начальный план эксперимента	Оптимальный план эксперимента
1,75e-08	1,73e-09

Как видно из таблицы 2, определитель оценочной ковариационной матрицы уменьшился в десять раз для оптимального плана, что подтверждает, что оценки параметров для оптимального плана действительно более точны.

Заключение:

С использованием предложенного алгоритма получен оптимальный план эксперимента для проведения испытаний на надежность светоизлучающих диодов. Методами статистического моделирования показано, что оценки максимального правдоподобия параметров винеровской деградационной модели на порядок выше при использовании оптимального плана эксперимента по сравнению с использованием начального плана.

Литература:

1. Chimitova E.V., Chetvertakova E.S., Sergeeva S.A., Osinceva E.A. A comparative analysis of the Wiener, Gamma and Inverse gaussian degradation models. [Proc. of the intern. workshop “Applied methods of statistical analysis. Nonparametric methods in cybernetics and system analysis”]. Krasnoyarsk. 2017. pp. 160–167.
2. Осинцева Е.А., Чимитова Е.В. Построение оптимальных планов эксперимента на основе винеровской деградационной модели. Обработка информации и математическое моделирование [материалы Рос. науч.-техн. Конф.], Новосибирск, 2018, с. 75-85.
3. Осинцева Е.А., Чимитова Е.В. Информационная матрица Фишера для винеровской деградационной модели с учетом объясняющих переменных. Обработка информации и математическое моделирование [материалы Рос. науч.-техн. Конф.], Новосибирск, 2019, с. 92-99.
4. Chaluvadi V. Accelerated Life Testing of Electronic Revenue Meters. Master Thesis, Clemson University, Clemson, SC, USA, 2008.

ДВУХЭТАПНАЯ ПРОЦЕДУРА ПОСТРОЕНИЯ ПРОГНОЗА РАСХОЖДЕНИЯ ШКАЛ ВРЕМЕНИ НА ОСНОВЕ ПРИМЕНЕНИЯ НЕЙРОННОЙ ЭХО-СЕТИ

К.Е. Перевалов, А.А. Богданов, О.С. Черникова
Новосибирский государственный технический университет,
г. Новосибирск, kiperee@gmail.com

В работе приводятся результаты прогнозирования расхождения шкал времени на основе нейронной эхо-сети.

The paper presents the results of forecasting the divergence of time scales based on echo state networks.

Часто при решении различных научно-технических задач в области космической навигации не требуется приводить шкалы времени в состояние синхронных физическими методами, достаточно определить с заданной точностью значение разности между числовыми выражениями положения любого события. Одним из актуальных вопросов при решении этой задачи является определение математической модели, описывающей расхождение шкал времени (РШВ) и использование ее для прогнозирования на заданный интервал с необходимой точностью.

В настоящее время используются различные подходы для описания РШВ [1,2]. Наиболее часто используемый метод решения этой задачи – это экстраполяция степенным полиномом. Для выполнения всех требований, предъявляемых к перспективным космическим аппаратам (КА) необходимо усовершенствовать полиномиальные модели, описывающие расхождение бортовых шкал времени КА относительно шкалы времени системы, либо искать новые подходы к математическому описанию РШВ.

В настоящей работе приводятся результаты исследования применимости нейронной эхо-сети (Echo State Networks, ESN) для описания временного ряда значений РШВ космических аппаратов относительно системной шкалы времени и прогнозирования на короткие интервалы длительностью до двух часов. Приводится оценка точности построенного прогноза на различных интервалах в зависимости от выбора длительности мерного интервала, на котором представлены измерительные данные для построения модели.

Предложим следующую двухэтапную процедуру прогнозирования РШВ:

Этап 1. Применение фильтра Ходрика-Прескотта для выявления трендовой и шумовой составляющих.

Среди распространенных методов статистической фильтрации временных рядов исследователи выделяют фильтр Ходрика-Прескотта [3]. При использовании данного фильтра априорно предполагается, что временной ряд имеет структуру, допускающую его нестационарность, наличие тренда и присутствие циклических изменений. Тренд во временном ряду может изменяться, но за это вводится некоторый штрафной параметр, и тогда ряд изменяет свой средний вектор движения, получаясь при этом гладким в соответствии с выбранным значением штрафа. Считается, что общая структура временного ряда допускает представление:

$$y(t) = g(t) + c(t) + \varepsilon(t),$$

где временные ряды $g(t), c(t), \varepsilon(t)$ определяют трендовую, циклическую и шумовую компоненты соответственно. Применение фильтра Ходрика-Прескотта позволяет представить ряд в виде:

$$y(t) = g(t) + v(t), \quad v(t) = c(t) + \varepsilon(t).$$

Для выделения трендовой составляющей решается следующая оптимизационная задача:

$$\min_{g(t)} \left(\sum_{t=1}^T (y(t) - g(t))^2 + \lambda \sum_{t=2}^{T-1} (g(t+1) - 2g(t) + g(t-1))^2 \right). \quad (1)$$

В формуле (1) второе слагаемое представляет собой колебания тренда с заданным штрафом $\lambda > 0$ за такие изменения. При решении задачи (1) возникает проблема выбора значения параметра λ . В работе [3] авторы считают, что выбор λ является хотя и произвольным, но напрямую зависит амплитуды сглаживания ряда для конкретной решаемой задачи. На практике рекомендуются следующие статистически обоснованные значения параметра в зависимости от интервала наблюдения значений временного ряда: 14400 – для месячных, 1600 – для квартальных и 100 – для годовых временных рядов. Отметим, что при $\lambda \rightarrow 0, g(t) \rightarrow y(t)$, а при $\lambda \rightarrow \infty, g(t)$ приближается к линейно й по времени функции.

Этап 2. Прогнозирование остаточной и шумовой составляющих с помощью нейронной эхо-сети.

К современным инструментам для прогнозирования временных рядов можно отнести нейронные эхо-сети [4-6]. ESN являются разновидностью рекуррентных нейронных сетей, где связи между нейронами скрытого слоя фиксированы и распределены случайным образом. В ESN каждый входной узел получает нелинейный сигнал. Обновляются веса только тех нейронов скрытого слоя, которые соединены с выходным слоем, что существенно сокращает длительность обучения.

Исключив из тренда скользящее среднее, сформируем остаточную компоненту $s(t)$. Компоненты $s(t)$ и $v(t)$ независимо предсказываются с помощью ESN.

В качестве характеристик точности модели на интервале прогноза использованы: среднеквадратическое отклонение, значение погрешности по уровню вероятности 0.95.

Для апробации модели использовались апостериорные данные о расхождении шкал времени трех КА системы ГЛОНАСС относительно системной шкалы времени, предоставленные службой [7].

Численные результаты оценки точности построенного двухэтапной процедурой прогноза приведены в таблицах 1-3. Графики зависимости измерительных данных РШВ y_i и прогнозных значений РШВ на 12-часовой интервал y_i^{pr} для R01-R03 представим на рис.1-3.

Таблица 1 – Погрешность модели РШВ на интервалах прогноза 30, 60 и 120 минут для спутника R01

Длительность мерного интервала, часы	Среднеквадратическое отклонение / погрешность по уровню вероятности 0.95, нс		
	30 минут	60 минут	120 минут
6	0.89/1.54	1.25/1.7	1.33/1.89
12	0.21/0.42	0.21/0.42	0.25/0.54
24	0.26/0.56	0.43/0.69	0.55/1.15

Таблица 2 – Погрешность модели РШВ на интервалах прогноза 30, 60 и 120 минут для спутника R02

Длительность мерного интервала, часы	Среднеквадратическое отклонение / погрешность по уровню вероятности 0.95, нс		
	30 минут	60 минут	120 минут
6	0.31/0.69	0.64/1.42	0.63/1.38
12	0.62/1.17	1.48/2.83	1.63/2.79
24	0.61/1.61	0.66/1.43	1.34/2.64

Таблица 3 – Погрешность модели РШВ на интервалах прогноза 30, 60 и 120 минут для спутника R03

Длительность мерного интервала, часы	Среднеквадратическое отклонение / погрешность по уровню вероятности 0.95, нс		
	30 минут	60 минут	120 минут
6	0.19/0.42	0.57/1.14	1.16/1.90
12	0.20/0.39	0.33/0.63	0.36/0.71
24	0.42/0.73	0.51/0.81	0.60/1.04

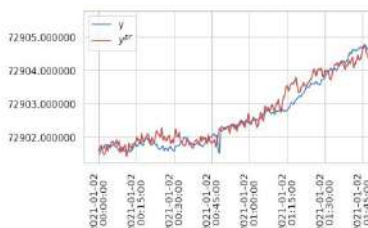


Рис.1 – График зависимости y_i и y_i^{pr} для R01

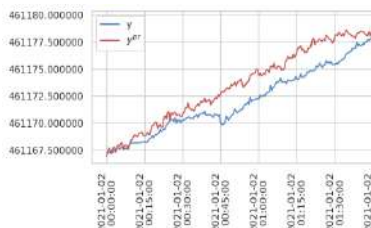


Рис.2 – График зависимости y_i и y_i^{pr} для R02

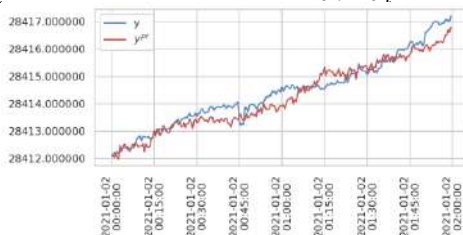


Рис.3 – График зависимости y_i и y_i^{pr} для R01

Анализ полученных результатов оценки точности предложенной двухэтапной процедуры построения прогноза РШВ для спутников R01-R03 позволяет сформулировать следующие выводы:

1. мерный интервал длительностью 12 часов обеспечивает наименьшую погрешность модели для рядов РШВ для R01 и R03 по сравнению с интервалами 6 и 24 часа и сопоставимую погрешность для R02 при 24-часовом интервале;

2. СКО модельных значений РШВ не превышает 0.3 нс. для R01-при прогнозе до 120 минут (на 12-часовом мерном интервале), для R02- при прогнозе до 30 минут (на 6-часовом мерном интервале), для R03- при прогнозе до 60 минут (на 12-часовом мерном интервале).

Проведение дальнейших исследований будет связано с изучением статистических характеристик точности построенной модели прогноза для всех спутников системы ГЛОНАСС на различных интервалах наблюдения.

Литература:

1. Forecasting at scale / Taylor S.J, Letham B. // PeerJ Preprints. – 5:e3190v2. – 2017. – <https://doi.org/10.7287/peerj.preprints.3190v2>.

2. Real-time clock offset prediction with an improved model / Guan Wen Huang, Qin Zhang, Guo Chang Xu // GPS Solutions. – 2014. – iss.18. – p. 95–104. – doi:10.1007/s10291-013-0313-0.

3. Business cycles: an empirical investigation / Hodrick R., Prescott E., Postwar U. // Journal of money, credit and banking. – 1997. – vol. 29. – №1, p. 1-16.

4. Reinforcement Learning with Echo State Networks / Szita I., Gyenes V., Lőrincz A. // International Conference on Artificial Neural Networks. 2006. pp. 830–839.

5. Harnessing nonlinearity: predicting chaotic systems and saving energy in wireless communication / Jaeger H, Haas H. // Science. 2004 Apr 2;304(5667):78-80. doi: 10.1126/science.1091277.

6. A review of designs and applications of echo state networks / C. Sun, M. Song, S. Hong, and H. Li // arXiv preprint arXiv:2012.02974, 2020.

7. Система высокоточного определения эфемерид и временных поправок - <http://www.glonass-svoevp.ru>.

ИССЛЕДОВАНИЕ МОЩНОСТИ КРИТЕРИЯ АССОЦИАЦИЙ, ПОСТРОЕННОГО С ПОМОЩЬЮ НЕЙРОННОЙ СЕТИ

Д.В. Поверин, С.И. Кушмуратов, С.Н. Постовалов
Новосибирский государственный технический университет,
г. Новосибирск, foxlandg@gmail.com

В данной работе предлагается подход построения оптимального критерия ассоциаций, с помощью технологий машинного обучения. Так же приведено краткое описание уже существующих критериев ассоциаций и проведения мета-анализа. Для реализации программного обеспечения использовался язык программирования Python. Для машинного обучения применялись библиотеки TensorFlow и Keras. Для получения обучающей выборки были использованы критерии хи-квадрат, Кокрена-Армитеджа, MAX3 и MERT. Объем обучающей и контрольной выборок в сумме составил 567 тысяч экземпляров. В результате, был построен оптимальный критерий ассоциации, имеющий точность 86,62%, относительно данных контрольной выборки. По правилу Вальда и Сэвиджа новый критерий ассоциаций оказался самым мощным, среди рассмотренных базовых критериев.

In this research proposed an approach for constructing an optimal association test using machine learning technologies. Also provided a brief description of the already existing associations tests and meta-analysis. To implement the software, the Python programming language was used. For machine learning, the TensorFlow and Keras libraries were used. Chi-square, Cochran-Armitage, MAX3 and MERT tests were used to obtain the training dataset. The size of the training and test datasets was 567 000 elements. As a result, the optimal association criterion was constructed, having an accuracy of 86.62%, relative to the control sample data. According to the rule of Wald and Savage, the new test of associations is the most powerful among the considered basic tests.

Полногеномный анализ ассоциаций (ПГАА) в наши дни является одним из важных и перспективных направлений исследований в биоинформатике, которые направлены на выявление связи между геномом человека и наблюдаемыми заболеваниями. В основе данного метода лежит проверка гипотезы независимости между генотипом индивидуума и наличием заболевания (фенотипом).

В полногеномном анализе ассоциаций изучают связь геномных мутаций с любыми фенотипическими (внешними) признаками, но особый интерес представляют трудно диагностируемые и редкие заболевания, например, различные раковые опухоли, болезнь Паркинсона, сердечнососудистые заболевания и пр. Одной из самых главных проблем для выявления зависимостей является недостаток объёма рабочей выборки испытуемых, ввиду низкой распространённости наблюдаемого заболевания в исследуемой популяции.

На данный момент существует множество критериев ассоциаций, такие как классический критерий хи-квадрат, критерий отношения правдоподобий с ограничением (CLRT), критерии максимизации минимальной эффективности (MERT), критерий тренда Кокрена-Армитеджа (САТТ), для разных генетических моделей наследования. Позже, появились робастные критерии, основанные на САТТ, которые имеют хорошую мощность, в условиях неопределённости генетической модели, такие как критерий МАХЗ, критерий выбора и исключения генетической модели (GMS и GME). Существуют, также, подходы обобщения результатов независимых экспериментов ПГАА, основанные на комбинировании p -value [1,3,4]. Такие подходы, при определённых условиях, дают очень хороший прирост мощности, в результате чего, уменьшается частота ложных срабатываний, при относительно малых значениях вероятности ошибки первого рода [6]. Но каков будет прирост мощности, если разработать интеллектуальную систему выбора оптимального критерия ассоциаций, в условиях неопределённости генетической модели? В данной работе, предлагается подход, построения оптимального критерия ассоциаций, с помощью технологий машинного обучения.

В рамках данной работы был получен новый критерий ассоциаций, с помощью нейронной сети [2,5] и был произведён анализ его мощности.

Объём обучающей и контрольной выборок в сумме составил 567 тысяч экземпляров.

Далее обучающая выборка подавалась на вход многослойной нейронной сети. Полученная в результате нейронная сеть может рассматриваться как новый статистический критерий, который позволяет предсказывать верную гипотезу.

Нейронной сети на вход подавались p -value критериев тренда САТТ, критерия хи-квадрат, критерия MERT, и критерия МАХЗ. Выбор данных критериев, для построения нейронной сети

обусловлен тем, что для данных критериев известны предельные законы распределений их статистик, и что в свою очередь позволяет легко оценить достигаемые уровни значимости. Был проведён анализ различных архитектур нейронной сети. Рассматривались разное число слоёв сети, разное число нейронов и различные функции активации.

Наибольшую точность на контрольной выборке удалось получить при использовании нейронной сети из 13 слоёв, с числом нейронов, указанных в таблице 1.

Таблица 1 - Оптимальная архитектура нейронной сети

Слой, №	1	2	3	4	5	6	7	8	9	10	11	12	13
Число нейронов	32	64	32	64	64	64	32	64	128	32	64	16	1

На первых 12 слоях применялась функция активации *relu*, на последнем слое применялась функция *sigmoid*. Полученная точность модели на контрольной выборке составила 86,62%. На обучающей выборке, точность модели составила 87,75%.

Был произведён анализ мощности нового критерия, на 9 различных конкурирующих гипотезах, приведённых в таблице 2.

Таблица 2 – Конкурирующие гипотезы ассоциаций

МОI (вид наследования)	p (частота аллели)	ψ_1	ψ_2	$\rho_{кл.}$
R	0,1	1,00	1,35	0,00102
R	0,5	1,00	1,0755	0,00101
R	0,9	1,00	1.0850	0,00100
A	0,1	1.0765	2.1530	0,00100
A	0,5	1.0470	2.0940	0,00105
A	0,9	1.0785	2.1570	0,00100
D	0,1	1.0850	1.0850	0,00105
D	0,5	1.0800	1.0800	0,00109
D	0,9	1.4090	1.4090	0,00100

По правилам Вальда (рис. 1) и Сэвиджа (рис. 2) лучшим по мощности является новый критерий, построенный с помощью нейронной сети (ANN), в сравнении с базовыми критериями ассоциаций.

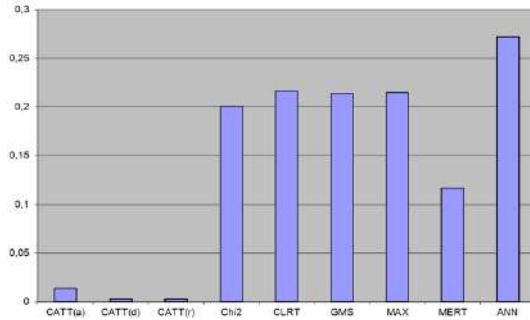


Рис. 1 – Сравнение мощности критериев по правилу Вальда

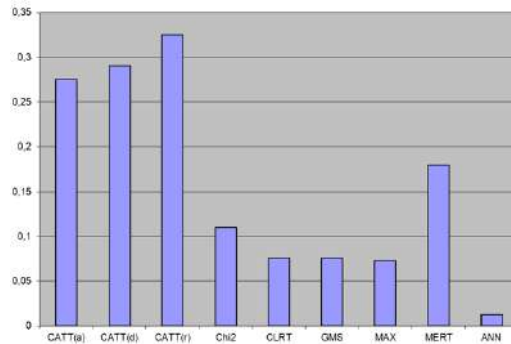


Рис. 2 – Сравнение мощности критериев по правилу Сэвиджа

Литература:

1. A meta-analysis of genome-wide association studies identifies 17 new Parkinson's disease risk loci / Chang D, Nalls M.A., Hallgrímsson I.B., Hunkapiller J., van der Brug M., Cai F.; International Parkinson's Disease Genomics Consortium; 23andMe Research Team, Kerchner G.A., Ayalon G., Bingol B., Sheng M., Hinds D., Behrens T.W., Singleton A.B., Bhangale T.R., Graham R.R. // *Nat Genet.* 2017. – Vol. 49(10). P. 1511-1516. doi: 10.1038/ng.3955.
2. Bishop, C. M. (2006), *Pattern Recognition and Machine Learning*, Springer, ISBN 978-0-387-31073-2
3. Evangelos E., Ioannidis J.P.A. Meta-analysis methods for genome-wide association studies and beyond // *Nature Reviews Genetics* – 2013. – Vol. 14. – P. 379-389.
4. Chen, Z. A new statistical approach to combining p-values using gamma distribution and its application to genome-wide association study / Z. Chen, W. Yang, Q. Liu // *BMC Bioinformatics.* – 2014. – DOI: 10.1186/1471-2105-15-S17-S3.

5. Hopfield, J. J. (1982). "Neural networks and physical systems with emergent collective computational abilities". Proc. Natl. Acad. Sci. U.S.A. 79 (8): 2554–2558. Bibcode:1982PNAS...79.2554H. doi:10.1073/pnas.79.8.2554. PMC 346238. PMID 6953413

6. Поверин Д. В. Исследование вероятности обнаружения новых геномных ассоциаций при комбинировании результатов полногеномного анализа ассоциаций = Estimating of probability discovery the new association by combining the results of genome-wide association study / Д. В. Поверин, С. Н. Постовалов // Обработка информации и математическое моделирование : материалы Рос. науч.-техн. конф., [Новосибирск, 25–26 апр. 2019 г.]. – Новосибирск : Изд-во СибГУТИ, 2019. – С. 100–113. - ISBN 978-5-91434-048-0.

ПРИМЕНЕНИЕ АДДИТИВНОЙ РЕГРЕССИОННОЙ МОДЕЛИ ДЛЯ ПРОГНОЗИРОВАНИЯ РАСХОЖДЕНИЯ ШКАЛ ВРЕМЕНИ

**А.С. Чирихина¹, И.Г. Данченко¹,
О.С. Черникова¹, Т.А. Марарескул²**

**¹Новосибирский государственный технический университет,
г. Новосибирск, chirixina.2018@stud.nstu.ru**

**²АО «ИСС» имени академика М. Ф. Решетнёва,
г. Железногорск**

В работе приводятся результаты прогнозирования расхождения шкал времени на основе нелинейной аддитивной регрессионной модели.

The paper presents the results of forecasting the divergence of time scales based on a nonlinear additive regression model.

Синхронизация разнесенных в пространстве стандартов частоты является одной из важнейших задач для различных научно-технических приложений и особую актуальность приобретает для космических и навигационных систем. При этом зачастую не требуется приводить реализуемые ими шкалы времени в состояние синхронных физическими методами, достаточно осуществить их привязку, то есть определить с заданной точностью значение разности между числовыми выражениями положения любого события в них. Одним из аспектов решения этой задачи является

построение математической модели, описывающей случайный процесс расхождения шкал времени (РШВ) и обеспечивающей заданную точность прогнозирования на некотором интервале.

В настоящее время для построения модели РШВ применяются различные подходы [1-3], такие как экстраполяция степенным полиномом, основанная на применении метода наименьших квадратов; экстраполяция степенным полиномом с экспоненциальным сглаживанием; построение моделей авторегрессии и интегрированного скользящего среднего; прогнозирование на основе фильтра Калмана.

Данная работа проводилась с целью исследования применимости аддитивной регрессионной модели (Generalized Additive Models, GAM) для описания временного ряда значений РШВ космических аппаратов относительно системной шкалы времени и прогнозирования на короткие интервалы длительностью до одного часа. Задачами исследования являлись выбор длительности мерного интервала, на котором представлены измерительные данные для построения модели, оценка точности модели на этом интервале и на интервале прогноза.

Аддитивная регрессионная модель может быть записана в следующем виде [1]:

$$y(t) = g(t) + s(t) + h(t) + e(t), \quad (1)$$

где

– $g(t)$ - кусочно-линейная функция или логистическая функция роста, моделирующие непериодические изменения значений РШВ;

–

$$g(t) = (k + a^T(t)\delta)t + m + a^T(t)\gamma$$

или

$$g(t) = \frac{C(t)}{1 + \exp(-(k + a^T(t)\delta)(t - (m + a^T(t)\gamma)))},$$

k - скорость роста, m - параметр смещения, δ - изменение (корректировка) скорости роста, $a(t) = \begin{cases} 1, t \in S, \\ 0, t \notin S, \end{cases}$ S – множество

точек, в которых происходит изменение тренда, γ - функция, обеспечивающая непрерывность функции $g(t)$ путем регулирования параметра m . Правильная регулировка в точке j

$$\text{вычисляется как: } \gamma_j = (s_j - m - \sum_{l < j} \gamma_l) \left(1 - \frac{k + \sum_{l < j} \delta_l}{k + \sum_{l \leq j} \delta_l} \right);$$

- $s(t)$ - периодические колебания, описываемые рядом Фурье:

$$s(t) = \sum_{n=1}^N \left(a_n \cos\left(\frac{2\pi n t}{P}\right) + b_n \sin\left(\frac{2\pi n t}{P}\right) \right),$$

P – регулярный период временного ряда;

- $h(t)$ – функция-индикатор, отражающая аномальные значения временного ряда. Она имеет вид:

$$h(t) = Z(t)q, q \in \text{Norm}(0, v^2),$$

где v – параметр сглаживания, $Z(t) = \begin{cases} 1, t \in L, \\ 0, t \notin L, \end{cases}$, L - набор

точек, задающийся пользователем, в которых ожидаются аномальные значения ряда (например, праздники);

- $e(t)$ — нормально распределенные случайные возмущения.

В качестве характеристик точности модели на мерном интервале и на интервале прогноза использованы:

- среднеквадратическое отклонение,
- значение погрешности по уровню вероятности 0.95.

Для апробации модели использовались апостериорные данные о расхождении шкал времени трех космических аппаратов системы ГЛОНАСС относительно системной шкалы времени, предоставленные службой [4].

Численные результаты оценки точности модели, полученные с использованием пакета Prophet, приведены в таблицах 1-4.

Таблица 1 – Погрешность модели РШВ на интервале определения

Длительность мерного интервала, часы	Среднеквадратическое отклонение / погрешность по уровню вероятности 0.95, нс		
	Спутник R01	Спутник R02	Спутник R03
12	0.17 / 0.35	0.30 / 0.58	0.19 / 0.46
24	0.21 / 0.45	0.43 / 0.88	0.21 / 0.44
48	0.27 / 0.53	0.71 / 1.44	0.32 / 0.67

Таблица 2 – Погрешность модели РШВ на интервалах прогноза 10, 30 и 60 минут для спутника R01

Длительность мерного интервала, часы	Среднеквадратическое отклонение / погрешность по уровню вероятности 0.95, нс		
	10 минут	30 минут	60 минут
12	0.26 / 0.40	0.29 / 0.51	0.49 / 0.82
24	0.52 / 0.71	0.67 / 0.98	0.95 / 1.40
48	0.47 / 0.64	0.58 / 0.88	0.85 / 1.28

Таблица 3 – Погрешность модели РШВ на интервалах прогноза 10, 30 и 60 минут для спутника R02

Длительность мерного интервала, часы	Среднеквадратическое отклонение / погрешность по уровню вероятности 0.95, нс		
	10 минут	30 минут	60 минут
12	0.21 / 0.37	0.49 / 1.03	0.58 / 1.00
24	0.77 / 1.05	0.69 / 0.90	0.73 / 0.98
48	2.41 / 2.74	2.42 / 2.67	2.65 / 3.10

Таблица 4 – Погрешность модели РШВ на интервалах прогноза 10, 30 и 60 минут для спутника R03

Длительность мерного интервала, часы	Среднеквадратическое отклонение / погрешность по уровню вероятности 0.95, нс		
	10 минут	30 минут	60 минут
12	0.68 / 0.77	0.80 / 1.09	0.80 / 1.08
24	0.77 / 0.87	0.90 / 1.22	0.93 / 1.22
48	0.16 / 0.24	0.18 / 0.30	0.25 / 0.52

Анализ полученных результатов оценки точности предложенной аддитивной регрессионной математической модели РШВ для трех спутников позволяет сформулировать следующие выводы:

1. мерный интервал длительностью 12 часов обеспечивает наименьшую погрешность модели для всех трех рядов РШВ по сравнению с интервалами 24 и 48 часов;

2. на выбранном мерном интервале СКО модельных значений РШВ не превышает 0.3 нс, погрешность по уровню вероятности 0.95 не превышает 0.6 нс для трех спутников;

3. при выбранном мерном интервале рассматриваемая модель позволяет осуществлять прогноз РШВ с субнаносекундной

точностью, при этом:

– СКО прогнозных значений РШВ на интервале 10 минут не превышает 0.7 нс, в среднем по трем спутникам составляет 0.4 нс;

– при выбранном мерном интервале СКО прогнозных значений РШВ на интервале 30 минут не превышает 0.8 нс, в среднем по трем спутникам составляет 0.5 нс;

– при выбранном мерном интервале СКО прогнозных значений РШВ на интервале 60 минут не превышает 0.8 нс, в среднем по трем спутникам составляет 0.6 нс.

Исходя из изложенного, можно считать целесообразным проведение дальнейших исследований рассмотренной математической модели для построения эмпирической зависимости погрешности прогноза РШВ от длительности мерного интервала, оценки эффективности в условиях наличия аномальных значений и возможных потерь измерительных данных, а также получения статистических характеристик ее точности для всех спутников системы ГЛОНАСС на длительных интервалах наблюдения.

Литература:

1. Forecasting at scale / Taylor S.J, Letham B.// PeerJ Preprints. - 5:e3190v2.- 2017. - <https://doi.org/10.7287/peerj.preprints.3190v2>.

2. Real-time clock offset prediction with an improved model/ Guan Wen Huang, Qin Zhang, Guo Chang Xu //GPS Solutions. -2014. – iss.18. – P. 95–104. – doi:10.1007/s10291-013-0313-0.

3. Скобелин А. А., Баушев С. В. Метод и алгоритм определения погрешности прогнозирования расхождения шкал времени // Изв. вузов. Приборостроение. - 2019. - Т. 62. - № 4. - С. 301—311.

4. Система высокоточного определения эфемерид и временных поправок - <http://www.glonass-svoevp.ru>.

ПРИМЕНЕНИЕ НЕЙРОСЕТЕВОЙ МНОГОФАКТОРНОЙ НЕЛИНЕЙНОЙ МОДЕЛИ ДЛЯ ПРОГНОЗИРОВАНИЯ СПРОСА НА АВИАТОПЛИВО В УСЛОВИЯХ НЕСТАБИЛЬНОСТИ

Д.И. Чурквейдзе

Московский авиационный институт (национальный исследовательский университет), г. Москва,
churkveidzedi@mail.ru

Научный руководитель: Хромова О.М., к.т.н., доцент

В работе описывается и анализируется нейросетевая многофакторная модель для прогнозирования спроса на авиакеросин. Устанавливаются закономерности между архитектурами и их эффективностями. Результаты могут быть использованы для оценки потребности на авиатопливо в нестабильных экономических условиях.

The paper describes and analyzes a neural network multi-factor model for forecasting the demand for jet fuel. Patterns are established between architecture and their efficiency. The results can be used to estimate the demand for aviation fuel in unstable economic conditions.

Введение. Среди авиакомпаний распространена практика хеджирования (страхования от возможных убытков). Для эффективного управления денежными потоками необходимо выбирать удачные моменты для заключения контрактов. Затраты на авиатопливо это самые большие эксплуатационные расходы, которые составляет до 30% от всех затрат авиакомпании, и поэтому они напрямую влияют на прибыльность [1]. Ставится задача о возможности прогнозирования спроса на авиакеросин. Применяя нейросетевое моделирование можно решить поставленную задачу.

В качестве статистических данных была найдена информация о спросе на авиакеросин в рублях за тонну в период с 2014 по 2021 год от Федерального агентства воздушного транспорта, также было найдено 1030 факторов (индексы потребительских цен), которые влияют на объект прогнозирования с корреляцией от 0,5% до 75%, данные взяты у Федеральной службы государственной статистики [2-3].

Слабая корреляция факторов с объектом прогнозирования также имеет значение наряду с факторами, обладающими сильной

корреляцией, и поэтому необходимо обращаться к максимально возможному числу индексов цен [4].

Оптимальная настройка сети. Для построения среднесрочных прогнозов рассмотрены следующие архитектуры нейронных сетей в таблице 1.

Таблица 1 – Архитектуры

Номер	1	2	3
Архитектура	Входной слой; Скрытый слой (64 нейрона); Выходной слой.	Входной слой; Скрытый слой 1 (64 нейрона); Скрытый слой 2 (32 нейрона); Выходной слой.	Входной слой; Скрытый слой 1 (64 нейрона); Скрытый слой 2 (32 нейрона); Скрытый слой 3 (16 нейронов); Выходной слой.

На вход нейронной сети подаются значения каждого фактора [5]. Рассматриваются три ширины окна 800, 1600 и 2400 дней, а также следующие наборы прогнозных лагов 120, 180 и 365 дней.

Метод масштабируемых сопряженных градиентов (SCG) выбран для обучения нейросети [6]. При применении среды Matlab этот алгоритм оказывается наиболее точным среди других для данной модели и имеющегося объема данных.

В качестве функции активации используется тангенциально-сигмоидная функция. Применение скрытых слоев в описанной модели приводит в некоторых случаях к ускорению обучения.

Вариации параметров и обзор результатов. Средняя абсолютная ошибка (MAE, Mean Absolute Error) и среднеквадратичная ошибка (RMSE, Root Mean Square Error) использованы в качестве метрик. Соотношение тестового к обучающему набору составляет 20% к 80%.

Для различных параметров сети на персональном компьютере были проведены тесты, а результаты занесены в таблицу 2.

При переходе от архитектуры типа I к типу II наблюдается значительное падение точности. Ошибка RMSE архитектуры II больше на 42%, а ошибка MAE также больше на 42% по сравнению с первой архитектурой. При этом длительность обучения сокращается на 22%.

Таблица 2 – Результаты для разных архитектур и окон

Размер окна, дни	Лаг, дни	Конфигурация архитектуры	RMSE, тыс. руб.	MAE, тыс. руб.	Время обучения, с
800	120	I	3,5334	3,2047	10,2660
		II	4,4700	4,1471	9,4429
		III	4,6973	3,9866	7,6157
	180	I	2,5994	1,9938	10,0802
		II	5,2022	4,2393	9,6299
		III	11,0600	9,9362	6,7901
	365	I	10,0326	7,7437	8,8809
		II	10,0820	8,6992	7,0990
		III	11,3876	10,3449	12,3718
1600	120	I	2,9590	2,5296	17,2648
		II	5,6373	4,7879	11,8831
		III	5,6692	4,3482	25,8887
	180	I	2,3568	1,8082	8,9042
		II	3,1227	2,3012	9,2603
		III	4,2177	2,8464	27,8521
	365	I	8,9626	7,4906	16,0804
		II	9,6562	7,8467	12,5170
		III	6,5008	5,7844	10,6022
2400	120	I	1,1530	0,9721	24,4097
		II	2,0784	1,4679	15,5130
		III	2,8026	2,4027	26,4159
	180	I	4,3258	3,2515	18,0788
		II	6,0276	4,5243	17,6657
		III	6,2476	5,0813	16,5497
	365	I	11,2292	9,0028	16,0442
		II	11,5795	10,2169	12,4125
		III	12,5485	10,7590	19,6490

Если добавить ещё один скрытый слой, как в третьей архитектуре, то обучение замедляется, а точность сопоставима с архитектурой II. Ошибка RMSE архитектуры III больше на 20%, ошибка MAE больше на 24% по сравнению со второй архитектурой. Обучение замедляется на 16%.

Графики прогнозов (рисунок 1) на 120 шагов для окна 2400 дней и 365 шагов для окна в 1600 дней с архитектурой I имеют наибольшую точность.

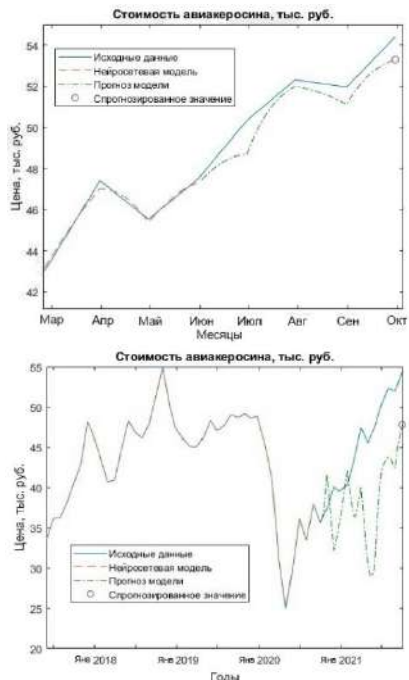


Рис.1 – слева изображён прогноз на 4 месяца, справа на 12 месяцев

Среднесрочный прогноз относительно дня n строится (по окну от k до $n-k$) из факторов, где k соответствует длине прогноза (120, 180 или 365).

Заключение. Таким образом, опробована нейросетевая многофакторная нелинейная модель изменения цен на авиатопливо с применением Matlab.

Несмотря на низкую коррелированность некоторых индексов цен с авиакеросином, они вносят полезный вклад в нейросетевую модель, поэтому в ней присутствует максимально возможное число факторов.

Было установлено, что есть тенденция к росту скорости обучения для двух слоёв. Однако дальнейшее увеличение числа слоёв, приводит к снижению точности.

Наличие локальных трендов и выбросов приводят к увеличению погрешности (см. 2020 год). Убирать выбросы нельзя, так как при их

отсутствии модель не будет готова к будущим сильным изменениям экономической ситуации, что приведёт к увеличению погрешности.

Рассмотренная модель и выводы могут быть использованы для прогнозирования стоимости авиакеросина.

Литература:

1. Из чего состоит керосиновая диета для самолетов // [Электронный ресурс] URL: <https://www.vedomosti.ru/opinion/articles/2019/10/15/813726-kerosinovaya-dieta> (дата обращения: 13.11.2021).

2. Цены на авиаГСМ // [Электронный ресурс] URL: <https://favt.gov.ru/deyatelnost-ajeroporty-i-ajerodromy-ceny-na-aviagsm/> (дата обращения: 11.10.2021).

3. Цены // [Электронный ресурс] URL: <https://rosstat.gov.ru/price> (дата обращения: 11.10.2021).

4. Губин Е. И. Методика подготовки больших данных для прогнозного анализа // Наука и бизнес: пути развития. – 2020. – №. 3. – С. 33-35.

5. Тухбатуллин Т. И., Курбанов Б. Реализация нейронных сетей для прогнозирования в среде MATLAB // Цифровизация образования: теоретические и прикладные исследования современной науки. – 2021. – С. 31-34.

6. Ильичов Д. Э., Головушкин А. А., Ионкин Д. А. Методы обучения нейронных сетей: градиентный спуск, метод Ньютона, сопряженный градиент // Наука и инновации в XXI веке: актуальные вопросы, открытия. – 2021. – С. 55.

**Секция ЧИСЛЕННОЕ МОДЕЛИРОВАНИЕ ФИЗИЧЕСКИХ
ПРОЦЕССОВ В ТЕХНОЛОГИЯХ И ПРИРОДНЫХ ЯВЛЕНИЯХ**

**ИССЛЕДОВАНИЕ ВЫЧИСЛИТЕЛЬНОЙ ЭФФЕКТИВНОСТИ И
ТОЧНОСТИ МОДЕЛИРОВАНИЯ ПРОЦЕССА ФИЛЬТРАЦИИ В
ЗАВИСИМОСТИ ОТ ВЫБОРА ПОРОГОВОГО ЗНАЧЕНИЯ
ШАГА ПО ВРЕМЕНИ**

**Г.Ж. Алтынбекова, А.С. Овчинникова, М.Г. Персова
Новосибирский государственный технический университет,
г. Новосибирск, ovchinnikova.2014@stud.nstu.ru
Научный руководитель: Персова М.Г., д.т.н., профессор**

Рассматривается подход с неявным расчетом давления методом конечных элементов и явным переносом фаз между ячейками сетки на каждом временном шаге, процедура группирования элементов. Исследуется влияние выбора шага по времени.

An approach with an implicit calculation of pressure by the finite element method and an explicit phase transference between grid cells at each time step, and an elements grouping procedure are considered. The influence of the choice of the time step is investigated.

При разработке нефтегазовых месторождений важную роль играет 3D моделирование процессов многофазной фильтрации в пористых средах. Повышение эффективности управления процессом разработки достигается за счет, например, построения гидродинамических моделей месторождений путем решения обратных задач [1,2] и синтеза оптимального управления разработкой [3,4]. В связи с этим требуется разработка эффективных вычислительных схем для моделирования разработки нефтяных месторождений.

В данной работе для моделирования процесса многофазной фильтрации при решении задач нефтедобычи используется подход с неявным расчетом давления методом конечных элементов и явным переносом фаз между ячейками сетки на каждом временном шаге [2]. Основной идеей данного подхода является вычисление по полученному полю давления потоков фаз $Q_{\Gamma_i}^m$, перетекающих через грани Γ_i конечных элементов Ω_e . Новый фазовый состав в элементах Ω_e рассчитывается по втекающим и вытекающим за временной интервал Δt объемам фаз $V_{\Gamma_i}^m$, вычисленным по формуле

$$V_{\Gamma_i}^m = Q_{\Gamma_i}^m \cdot \Delta t.$$

Обсудим выбор временного шага Δt . Прежде всего, он непосредственно влияет на точность аппроксимации по времени. Кроме того, его предельное (максимальное) значение зависит от объема m -й фазы в ячейке Ω_e , из которой эта фаза вытекает:

$$\Delta t \leq \frac{(\text{mes}(\Omega_e) \Phi(S_{\Omega_e}^m - S_{\Omega_e}^{m, \text{res}}) + \Delta V_{\Omega_e}^m)}{\sum_{i \in I_{\Omega_e}^{\text{out}, m}} |Q_{\Gamma_i}^m|}, \quad \forall \Omega_e, \quad \forall m, \quad (1)$$

где $\text{mes}(\Omega_e)$ – объем ячейки Ω_e , $S_{\Omega_e}^m$ – насыщенность m -й фазы, $S_{\Omega_e}^{m, \text{res}}$ – остаточная насыщенность m -й фазы, $\Delta V_{\Omega_e}^m$ – дефицит/профицит объема m -й фазы, $I_{\Omega_e}^{\text{out}, m}$ – множество номеров i граней Γ_i , через которые m -й фаза вытекает из конечного элемента Ω_e .

Таким образом, временной шаг Δt зависит от величины $S_{\Omega_e}^m$ (которая в процессе моделирования может получиться довольно малой) и поэтому может стать недопустимо маленьким (даже несмотря на то, что в случае уменьшения $S_{\Omega_e}^m$ величины $Q_{\Gamma_i}^m$ в знаменателе правой части неравенства (1), как правило, тоже резко уменьшаются из-за уменьшения фазовой проницаемости фазы).

Чтобы сократить вычислительные затраты, связанные с резким уменьшением шага по времени, предлагается использовать процедуру «подмены» фаз. Для этого вводится пороговое значение Δt^{thr} , которое определяет минимально допустимое значение шага по времени Δt . Если существуют фазы $\{m_{\Omega_e}^{\text{thr}}\}$, насыщенность которых приводит к уменьшению временного шага ниже Δt^{thr} , то предлагается уменьшить вытекающие объемы этих фаз таким образом, чтобы за выбранный шаг $\Delta t \geq \Delta t^{\text{thr}}$ весь имеющийся в ячейке Ω_e объем этих фаз вытек через ее грани. При этом если направления потоков фаз $\{m_{\Omega_e}^{\text{thr}}\}$ и потоков смеси совпадают, то для выполнения закона сохранения фильтрующей смеси увеличиваются объемы остальных фаз, вытекающих с Ω_e , в противном случае уменьшаются объемы фаз, втекающих в элемент через эту грань. За счет выполнения такой процедуры получают сбалансированные значения объемов фаз $V_{\Gamma_i}^m$, перетекающих через грань Γ_i за временной интервал $\Delta t \geq \Delta t^{\text{thr}}$.

Таким образом, временной шаг Δt должен удовлетворять только критерию достаточности объема смеси

$$\Delta t \leq \Delta t^{crit} \leq \frac{\left(\text{mes}(\Omega_e) \Phi + \sum_m \Delta V_{\Omega_e}^m \right)}{\sum_{i \in I_{\Omega_e}^{out}} |Q_{\Gamma_i}|}, \quad \forall \Omega_e, \quad (2)$$

где $I_{\Omega_e}^{out}$ – множество номеров i граней Γ_i , через которые поток смеси Q_{Γ_i} вытекает из Ω_e . Очевидно, что пороговое значение Δt^{thr} не должно превышать Δt^{crit} из формулы (2).

Для сокращения вычислительных затрат, вызванных сильным уменьшением шага Δt , также используется процедура группирования элементов. Данная процедура позволяет зафиксировать глобальный шаг по времени Δt^{main} , и распределить ячейки конечноэлементной сетки по группам, в каждой из которых может быть использован допустимый и наиболее близкий к требуемому шаг по времени.

Численные эксперименты проведем на модельной задаче, имитирующей процесс добычи нефти на поздних этапах разработки месторождения, когда насыщенность нефти в коллекторе близка к ее значению остаточной насыщенности.

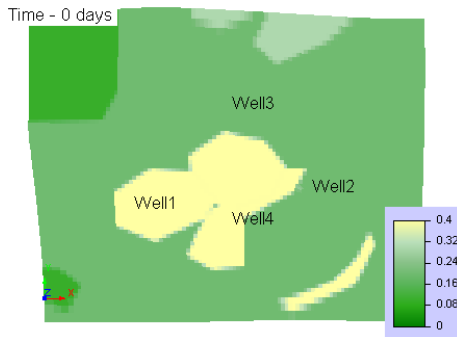


Рис.1 – Насыщенность нефти в начальный момент времени

Стартовая нефтенасыщенность – 0.4, остаточная насыщенность нефти – 0.313. Модель включает в себя одну добывающую (Well1) и три нагнетательных скважины (Well2, Well3 и Well4). Количество КЭ 19556. Режимы работы скважин следующие: мощность Well1 постоянна и равна

63 м³/сут; мощности Well2, Well3 и Well4 равны 13.333 м³/сут, первые 625 суток они нагнетают смесь воды с полимером, доля которого равна 0.001, после 625-го дня и до 2000-го дня нагнетается вода. Время жизни «месторождения» – 2000 суток. Глобальный шаг по времени Δt^{main} равен 30 суткам.

В таблице 1 приведены вычислительные затраты расчетов с различными значениями Δt^{thr} .

Таблица 1 – Сравнение вычислительных затрат для различных значений Δt^{thr}

№	Δt^{thr}	Количество элементов с выполненной «подменной» фаз	Время расчета, с	Отклонение значения накопленного отбора нефти, %
1	без группирования	165	644.6	0.07
2	30 суток	315	13.7	2.68
3	15 суток	298	13.6	1.33
4	11 часа	178	17.7	0.11
5	42 минуты	38	50.5	0.03
6	без «подмены» фаз	0	71.8	0

Как видно из таблицы результаты расчетов с уменьшением Δt^{thr} практически не изменяются. Отклонение значения накопленного отбора нефти в момент времени 400 суток для расчетов с $\Delta t^{thr} = 1$ и без «подмены» фаз составили 1.33 %. При этом с уменьшением Δt^{thr} снижаются вычислительные затраты. Так временные затраты для расчета без применения «подмены» фаз составили – 71.8 с, а для $\Delta t^{thr} = 1$ – 13.6 с. Таким образом, с использованием рассмотренного подхода вычислительные затраты уменьшились в 5.3 раза.

Литература:

1. Comprehensive framework for gradient-based optimization in closed-loop reservoir management / V. Bukshtynov, O. Volkov, L.J. Durlofsky, K. Aziz // Computational Geosciences. – 2015. – Vol. 19(4). – P. 877–897. – doi: 10.1007/s10596-015-9496-5.

2. The design of high-viscosity oil reservoir model based on the inverse problem solution / M.G. Persova, Yu.G. Soloveichik, D.V. Vagin, A.M. Grif, D.S. Kiselev, I.I. Patrushev, A.V. Nasybullin & B.G. Ganiev // J. Pet. Sci. Eng. – 2021. – Vol. 199, art. 108245. – doi: 10.1016/j.petrol.2020.108245.

3. Optimization of injection scheme to maximizing cumulative oil steam ratio based on improved artificial bee colony algorithm / H.M. Ni, Y.J. Liu & Y.C. Fan // J. Pet. Sci. Eng. – 2019. – Vol. 173. – P. 371–380. – doi: 10.1016/j.petrol.2018.10.032.

4. Persova M.G., Soloveichik Yu.G., Vagin D.V., Grif A.M., Patrushev I.I. & Ovchinnikova A.S. Oil production optimization based on the finite-element simulation of the multi-phase flow in porous media and inverse problem solution // GeoBaikal 2020, oct. 2020: Conference proceedings. – Irkutsk, 2020. – P. 1–6.

ЧИСЛЕННОЕ ИССЛЕДОВАНИЕ ТЕЧЕНИЯ ВЯЗКОЙ ЖИДКОСТИ В МИКРОКАНАЛЕ С ТРЕУГОЛЬНЫМИ ЛОВУШКАМИ

А.З. Булатова, О.А. Солнышкина, Н.Б. Фаткуллина

**Центр Микро- и наномасштабной динамики дисперсных систем,
Башкирский государственный университет, г. Уфа,
bulatova29@yandex.ru**

Научный руководитель: Солнышкина О.А., к.ф.-м.н.

Настоящее исследование посвящено изучению гидродинамических потоков в микроканале с треугольными ловушками. В основе численного подхода лежит метод граничных элементов (МГЭ). В работе приведены результаты триангуляции канала рассматриваемой геометрии, а также получена картина течения и поля скоростей.

The present study focuses on the hydrodynamic flows in a microchannel with triangular traps. The numerical approach is based on the boundary element method (BEM). The results of triangulation of the channel of the considered geometry, the flow pattern, and velocity field are given in the paper.

В последние годы основные достижения науки связаны с уменьшением размеров различных устройств и улучшением их технических характеристик. Изучение и конструирование микрофлюидных устройств являются одним из наиболее актуальных тем [1]. Микрофлюидные устройства находят широкое применение во многих областях науки и техники. Например, системы микроканалов разнообразной формы и размеров нужны при проведении экспериментальных исследований, микрожидкостные устройства используются в качестве охлаждающей системы в

высокопроизводительных микросхемах, когда по микроканалам прокачивается охлаждающая жидкость, микрореакторы применяются для лучшего смешивания реагентов, а биочипы для проведения экспресс-тестирования в медицине. Таким образом, комплексное изучение микрофлюидных систем позволило найти множество прикладных приложений этих устройств.

На течение флюидов на микроуровне существенное влияние оказывают форма и размер микроканалов. Поэтому изучение особенностей течения жидкости в каналах со сложной геометрией важно как с точки зрения практического применения, так и при выводе теоретических закономерностей подобных процессов, поскольку эти эффекты могут оказывать значительное воздействие на характер движения дисперсных включений в жидкости.

В настоящей работе проведено численное исследование гидродинамических потоков микроканале с треугольными ловушками. Рассматривается медленное периодическое течение жидкости, поэтому силы вязкости значительно сильнее сил инерции. Таким образом, установившееся течение жидкости описывается уравнениями Стокса. Для течения в канале задается условие прилипания на стенках, которое дополняется периодичностью течения на входном и выходном сечении микроканала.

В основе численного подхода лежит метод граничных элементов (МГЭ), который очень эффективен при изучении динамики систем в сложных областях, поскольку все вычисления связаны только с границами рассматриваемых объектов [2].

В рамках исследования была получена качественная триангуляция микроканала с ловушками, имеющих форму равностороннего треугольника. Общее количество расчетных узлов $N_{\Delta} = 6236$. На рис. 1 показана триангуляция канала рассматриваемой формы.

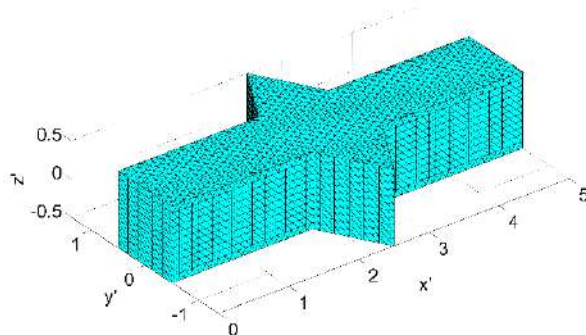


Рис. 1 – триангуляция микроканала

с треугольными ловушками ($N_{\Delta} = 6236$)

Далее проводился расчет установившегося течения вязкой жидкости в канале рассматриваемой геометрии при заданном постоянном перепаде давления. Были получены картины течения и продольные компоненты скорости течения жидкости (рис. 2).

Как видно из рисунка, в узкой части канала линии тока параллельны и направлены вдоль оси канала, а в области треугольных ловушек они образуют небольшие вихри. Компонента U_x поля скорости показывает, что максимальная скорость находится ближе к оси канала, но в области расширения канала скорость падает почти до минимума.

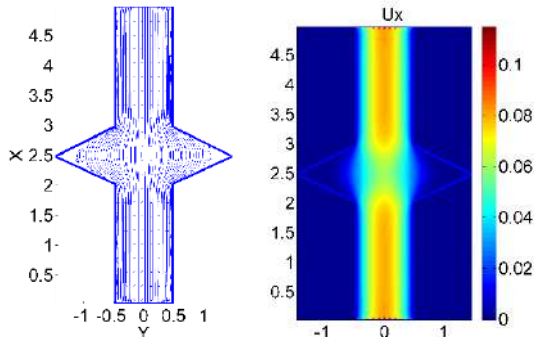


Рис. 2 – Линии тока и компонента U_x поля скорости потока

Работа выполнена при поддержке гранта РФФ 21-79-10212.

Литература:

1. Foudeh A., Didar T., Veres T. et al. Microfluidic Designs and Techniques Using Lab-on-a-Chip Devices for Pathogen Detection for Point-of-Care Diagnostics // *Lab on a Chip*. – 2012. – V. 12. – pp 3249–66.
2. Pozrikidis C. Boundary integral and singularity methods for linearized viscous flow // Cambridge (Cambridge University Press). – 1992 – p. 259.

ПРИМЕНЕНИЕ МЕТОДА БАЛАНСИРОВКИ ДЛЯ УТОЧНЕНИЯ ЧИСЛЕННЫХ ПОТОКОВ НА СКВАЖИНАХ ПРИ МОДЕЛИРОВАНИИ МЕСТОРОЖДЕНИЙ

А.М. Гриф

Новосибирский государственный технический университет,
г. Новосибирск alexgrif@inbox.ru

В работе представлена математическая модель для гидродинамического моделирования месторождений. Приводится результат применения метода балансировки для уточнения численных потоков на скважинах.

The paper presents a mathematical model for reservoir simulation. The result of the balancing method is presented to clarify the numerical flows in the wells.

Введение. При моделировании месторождений полезных ископаемых довольно часто возникают ситуации, когда скважины работают при некотором фиксированном давлении, и, соответственно, объем смеси является неизвестным. В этом случае качество аппроксимации потоков главным образом зависит от точности расчета поля давления. При этом, для уточненного расчета давления в основном используется стандартная p - h -технология, которая влечет увеличение вычислительных затрат [1]. В данной работе для улучшения аппроксимации потоков смеси используется специальный метод балансировки потоков, который основан на минимизации функционала баланса смеси с регуляризацией [2].

Математическая модель. Процесс многофазной фильтрации несжимаемой неперемешиваемой изотермической жидкости в расчетной области Ω описывается с помощью системы дифференциальных уравнений

$$-\operatorname{div} \left(\sum_{m=1}^M \frac{\kappa^m}{\mu^m} \mathbf{K} \left(\operatorname{grad} (P + P_c^m) + (0, 0, \rho^m g)^T \right) \right) = \sum_{m=1}^M f^{m,\Omega} \quad (1)$$

с краевыми условиями

$$P|_{\Gamma_1} = P^\Gamma, \quad (2)$$

$$\sum_{m=1}^M \frac{\kappa^m}{\mu^m} \mathbf{K} \left(\operatorname{grad} (P + P_c^m) + (0, 0, \rho^m g)^T \right) \Big|_{\Gamma_2} \cdot \vec{n} = \sum_{m=1}^M f^{m,\Gamma}, \quad (3)$$

где Γ^1 – границы расчетной области Ω с заданным давлением P^Γ , а Γ^2 – зоны перфорации скважин, через которые осуществляется отбор или нагнетание.

Для решения задачи (1)-(3) применяется вычислительная схема с неявной аппроксимацией поля давления, явным пересчетом насыщенностей и специальная процедура для балансировки потоков [2].

Для балансировки перетекающих через каждый конечный элемент объемов смеси вводятся корректирующие слагаемые δQ_i , обеспечивающие равенство объемов втекающей и вытекающей смеси. Они могут быть найдены в результате минимизации функционала

$$\sum_{e=1}^{N^e} \beta_e \left(\sum_{i \in I_{\Omega_e}} (Sg_{\Gamma_i}^{\Omega_e} \cdot |Q_{\Gamma_i}| + Sg_{\Gamma_i}^{\Omega_e} \cdot \delta Q_i) \right)^2 + \sum_{i=1}^{N^f} \alpha_i (\delta Q_i)^2. \quad (4)$$

где e – номер конечного элемента, N^e и N^f – число конечных элементов и граней соответственно, I_{Ω_e} – множество номеров граней Γ_i конечного элемента Ω_e , а β_e , α_i – параметры регуляризации.

В результате минимизации функционала (4) получается СЛАУ

$$(\mathbf{B} + \mathbf{A})\mathbf{q} = \mathbf{d}.$$

где \mathbf{q} – это вектор с корректирующими добавками, \mathbf{A} – диагональная матрица, а компоненты матрицы \mathbf{B} и вектора правой части \mathbf{d} вычисляются с помощью соотношений

$$B_{ij} = \begin{cases} \sum_{k \in I_{\Gamma_i}} \tilde{\beta}_k, & i = j, \\ \tilde{\beta}_e \cdot Sg_{\Gamma_i}^{\Omega_e} \cdot Sg_{\Gamma_j}^{\Omega_e}, & i, j \in I_{\Omega_e}, i \neq j, e = 1 \dots N^e, \\ 0, & \text{иначе,} \end{cases}$$

$$d_i = - \sum_{e=1}^{N^e} \left(\tilde{\beta}_e Sg_{\Gamma_i}^{\Omega_e} \cdot \left(\sum_{j \in I_{\Omega_e}} Sg_{\Gamma_j}^{\Omega_e} |Q_{\Gamma_j}| \right) \right).$$

Вычислительный эксперимент. Расчет будем проводить на модели месторождения, состоящего из одного слоя и включающего четыре

скважины [3]. Одна скважина является добывающей с отбором смеси в 40 м³/сут. Три скважины – нагнетательные, между которыми равномерно распределен объем закачиваемой смеси. Суммарный объем нагнетания соответствует объему добычи.

На рисунке 1 представлено распределение нефтенасыщенности в коллекторе на начало моделирования (вид в плане).

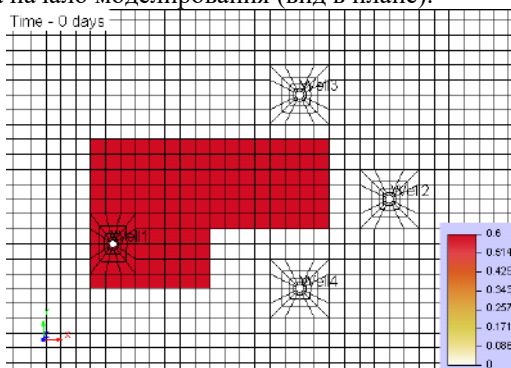


Рисунок 1 – Распределение нефтенасыщенности на начало моделирования

Для оценки качества уточнения численных потоков с помощью процедуры балансировки оставим «открытыми» для балансирования границы на группе скважин. Так, на добывающей скважине Well1 и двух нагнетательных скважинах Well3 и Well4 будут искажаться соответствующие корректирующие добавки (4). В этом случае качество уточнения потоков будет определяться по близости скорректированного объема смеси к истинному.

В таблице 1 представлены истинные, численные и скорректированные значения объемов для исследуемых скважин.

Таблица 1 – Истинные, численные и скорректированные значения объемов для исследуемых скважин

Отбор/нагнетание	Well1	Well3	Well4
Истинный	40	-13.3(3)	-13.3(3)
Численный	29.74	-9.46	-9.46
Балансировка потоков	38.43	-12.87	-12.23

Из таблицы видно, что применение метода балансировки позволило существенно уменьшить погрешность в значениях отборов относительно

численных значений. Для скважин Well1 и Well3 погрешность составила менее 4%, для скважины Well4 – менее 9%.

Заключение. Представленный в работе метод балансировки позволяет моделировать месторождения с фиксированным давлением скважин, при этом для моделирования могут быть использованы конечноэлементные сетки с довольно крупными ячейками.

В результате вычислительного эксперимента было показано, что применение метода балансировки позволило существенно уменьшить погрешность в значениях отборов относительно численных значений.

Литература:

1. Соловейчик, Ю.Г. Метод конечных элементов для скалярных и вектор-ных задач / Ю.Г. Соловейчик, М.Э. Рояк, М.Г. Персова. – Новосибирск : НГТУ, 2007. – 869 с.
2. Persova, M.G., Soloveichik, Yu.G., Vagin, D.V., Grif, A.M., Kiselev, D.S., Patrushev, I.I., Nasybullin, A.V. & Ganiev, B.G. (2021) The design of high-viscosity oil reservoir model based on the inverse problem solution. J. Pet. Sci. Eng. vol. 199. art. 108245.
3. Persova, M.G., Soloveichik, Yu.G., Grif, A.M. & Patrushev, I.I. (2018) Flow Balancing in FEM Modelling of Multi-Phase Flow in Porous Media. Novosibirsk: 2018 XIV Int. Scientific-Technical Conference on Actual Problems of Electronics Instrument Engineering (APEIE). pp. 205–211.

РЕШЕНИЕ ДВУМЕРНОЙ ОБРАТНОЙ ЗАДАЧИ МАГНИТОРАЗВЕДКИ С ИСПОЛЬЗОВАНИЕМ ГЕОМЕТРИЧЕСКОЙ ПАРАМЕТРИЗАЦИИ

М.С. Дьячкова, А.М. Гриф

**Новосибирский государственный технический университет,
г. Новосибирск, dyach_m@mail.ru**

В работе описан метод решения обратной задачи магниторазведки с параметризацией тел, заданных многоугольными контурами в плоскости XZ. Рассмотрен случай задания аномальных объектов несколькими прямоугольными контурами.

The paper describes a method for solving the inverse problem of magnetic prospecting with the parameterization of bodies specified by polygonal contours in the XZ plane. The case of specifying anomalous objects by several rectangular contours is considered.

Введение.

В настоящее время помимо необходимости достаточно точного описания геометрии реального аномального объекта, важным критерием использования конкретного метода инверсии данных магниторазведки является способность к уменьшению области эквивалентности решений.

Наиболее геологически обоснованным является геометрический метод решения обратной задачи. Данный метод рассматривался для работы с объектами-призмами с многоугольниками (в частности, прямоугольниками) в основании в плоскости Oxz , в том числе, в областях с несколькими аномальными включениями.

Математическая модель. Решение обратной задачи с геометрической параметризацией представляет собой вектор параметров \vec{u} , состоящий из координат точек – вершин многоугольного контура-основания найденного объекта, а также значений трехкомпонентного вектора его намагниченности \vec{p} . Задача является нелинейной и может быть решена с помощью метода Гаусса-Ньютона [1]. Входными данными являются значения вектора магнитной индукции \vec{B} в точках, установленных в области приемников, а также некоторое обязательное начальное приближение вектора \vec{u} .

Задача инверсии сводится к минимизации функционала [2]:

$$\Phi(\vec{u}) = \sum_i^N (B_{zi}(\vec{u}) - B_{zi}^*)^2 + \sum_j^M \alpha_j (u_j - \bar{u}_j)^2, \quad (1)$$

где $B_{zi}(\vec{u})$, B_{zi}^* – z-компоненты B , полученного при текущем \vec{u} и истинного B соответственно, N – количество приемников, M – размерность вектора параметров.

Для оценки качества работы геометрического метода используется относительный функционал

$$\Phi_{rel i} = \Phi_i / \Phi_0, \quad (2)$$

где Φ_i – функционал от текущего приближения на i -той итерации цикла, Φ_0 – функционал от начального приближения объекта, вычисленные согласно формуле (1).

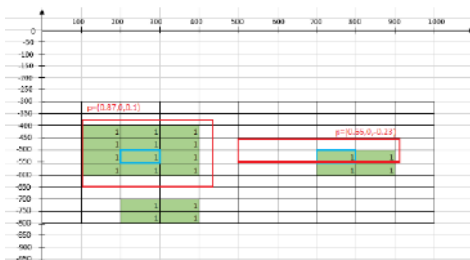
Использование в функционале регуляризирующей добавки с вектором параметров \vec{a} позволяет ограничить величину изменений каждого из параметров, являющуюся непосредственным решением СЛАУ на каждой итерации цикла обратной задачи. Ограничения могут задаваться с учетом предыдущих значений вектора \vec{u} , например, следующим образом: для параметров, отвечающих за положение, изменение может составлять не более 50% от приблизительной средней величины текущей конфигурации аномалии, для параметров вектора намагниченности – не более значения модуля текущего \vec{p} .

Результаты исследования. Для нескольких заданных аномальных объектов было проведено сравнение методов геометрической инверсии

(с различными начальными приближениями) и ячеистой инверсии. Истинная конфигурация аномалий представляет собой трех прямоугольных объекта с намагниченностью (1,0,0), обозначенные на рисунках зеленым цветом. Задача была решена при различных начальных приближениях, представленных двумя и тремя прямоугольными контурами (на рисунках обозначены синим цветом). Координаты полученного решения представлены на рисунках контурами красного цвета.

Результаты геометрической инверсии приведены на рисунках 1 и 2.

При использовании начального приближения из двух контуров (рисунок 1(а)) полученная конфигурация и значения векторов намагниченности не соответствуют истинным, несмотря на достаточно низкое значение относительного функционала (2) – $6.94803e-05$, что может быть интерпретировано как попадание в локальный минимум в области решений обратной задачи. При использовании трех контуров полученное решение близко к истинному, значение функционала – $8.57119e-09$ (б). Таким образом, для корректной работы метода в случае поиска нескольких объектов необходимо начальное приближение с количеством контуров, позволяющим описать истинную геометрию аномалии достаточно точно.



а



б

Рис.1 – Решение обратной задачи для трех прямоугольных контуров с геометрической параметризацией с начальным приближением двух контуров (а) и трех контуров (б)

Для истинной модели, представленной на рисунке 1, обратная задача была также решена с использованием ячеистой параметризации [3]. Результаты приведены на рисунке 2.

Значение соответствующего функционала при использовании ячеистой инверсии составило $1.58869\text{e-}14$. Результаты ячеистой инверсии плохо оценивают аномальные объекты, расположенные друг под другом на разной глубине. Однако данные результаты могут быть использованы как начальное приближение контуров для геометрической инверсии.

Заключение. Метод геометрической инверсии позволяет получать более точное положение аномальных объектов и значения их векторов намагничённости, чем метод ячеистой инверсии. Однако для случая с несколькими намагничёнными объектами необходимо корректное начальное приближение, полученное предварительной оценкой области аномалий, например, с помощью ячеистой инверсии.

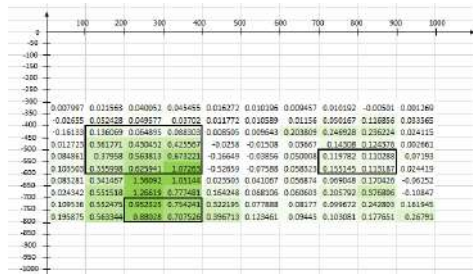


Рис.2 – Решение обратной задачи для трех прямоугольных контуров с ячеистой параметризацией

Литература.

1. Алгоритмы решения обратных задач для дифференциальных уравнений в частных производных: методические указания к лабораторным работам по курсам "Современные проблемы прикладной математики и информатики" и "Современные компьютерные технологии" для магистрантов ФПМИ, направление 010400 / Новосиб. гос. техн. ун-т; [сост.: М. Г. Персова, Ю. Г. Соловейчик, П. А. Домников]
2. Вагин, Д.В. Оценка параметров в обратных задачах: учебное пособие / Д.В. Вагин. – Новосибирск: Изд-во НГТУ, 2019. – 48 с.
3. Персова М.Г. Современные компьютерные технологии: конспект лекций / М.Г. Персова, Ю.Г. Соловейчик, П.А. Домников. – Новосибирск: Изд-во НГТУ, 2014. - 80 с.

МЕТОД ЛИУ-СТЕФФАНА ДЛЯ РЕШЕНИЯ ОДНОМЕРНЫХ ЗАДАЧ ГАЗОВОЙ ДИНАМИКИ

А.И. Кантарбаева

Томский государственный университет,

г. Томск, arukantar@gmail.com

Научный руководитель: Моисеева К. М., к.ф.-м.н., доцент

Рассматривается применение метода Лиу-Стеффана для моделирования нестационарного течения сжимаемого газа. Представлены численные решения одномерных задач газовой динамики, описанных уравнениями Эйлера. При моделировании сверхзвуковых течений выбранным методом наблюдаются разрывы в численных решениях. Полученные результаты сравнивались с работой других авторов.

The Liou-Steffan method for modeling the flow of an unsteady compressible gas is considered. Numerical solutions of one-dimensional problems of gas dynamics described by the Euler equations are presented. When simulating supersonic flows by the chosen method, discontinuities in the numerical solutions are observed. The results obtained were compared with the work of other authors.

Для описания течения нестационарного сжимаемого газа используются уравнения Эйлера, вычисление которых сопряжено с некоторыми особенностями. Одной из таких особенностей является возникновение разрывов решения. При этом решение задач газовой динамики нередко требует больших вычислительных затрат.

Противопоточные схемы часто применяются для решения уравнений Эйлера. Они относительно просты в реализации и обладают хорошей точностью при разрешении разрывов. Целью данной работы являлось исследование и тестирование метода Лиу-Стеффана или AUSM (Advection Upstream Splitting Method) [1].

Математическая постановка задачи включает в себя систему уравнений Эйлера и уравнение состояния идеального газа. Рассматривалась область размером L . В середине области расположена перегородка. Значения плотности скорости и давления газа слева и справа от перегородки определяются начальными условиями. В начальный момент времени перегородка моментально убирается, что приводит к возникновению распада произвольного разрыва.

Основной идеей метода Лиу-Стеффана является расщепление вектора потока \mathbf{F} на две составляющие: конвективный вектор $\mathbf{F}^{(c)}$ и вектор давления $\mathbf{F}^{(p)}$:

$$\mathbf{F}^{(c)} = \begin{pmatrix} \rho u \\ \rho u^2 \\ \rho u H \end{pmatrix} = M \begin{pmatrix} \rho c \\ \rho c u \\ \rho c H \end{pmatrix} = M \hat{\mathbf{F}}^{(c)}, \quad \mathbf{F}^{(p)} = \begin{pmatrix} 0 \\ p \\ 0 \end{pmatrix}$$

Вектор потока на границах ячеек определяется как сумма конвективного вектора и вектора давления. Конвективная составляющая вектора потока рассчитывается по противопоточной схеме:

$$\hat{\mathbf{F}}_{i+1/2}^{(c)} = \begin{cases} \hat{\mathbf{F}}_i^{(c)}, & \text{если } M_{i+1/2} \geq 0 \\ \hat{\mathbf{F}}_{i+1}^{(c)}, & \text{если } M_{i+1/2} < 0 \end{cases}$$

Число Маха и вектор давления на границах ячеек определялись как сумма неотрицательной и неположительной составляющих. Вектор давления определяется также как сумма неотрицательной и неположительной составляющих.

Были решены тестовые задачи (Таблица 1) [2]. Расчет проводился при следующих значениях параметров: $c_p = 1005 \text{ Дж/кг/К}$, $\gamma = 1.4$, $L = 1 \text{ м}$. Перегородка расположена в точке $x_0 = 0.5 \text{ м}$. Значение числа Куранта задавалось равным 0.1, шаг по пространству - $h = 10^{-2} \text{ м}$, шаг по времени - $\Delta t = 10^{-4}$.

Таблица 1 – Начальные условия для тестовых задач

Тест	Параметры слева			Параметры справа			Время
	ρ	u	p	ρ	u	p	
1	1	0	1	0,125	0	0,1	0,15
2	0,445	0,698	3,528	0,5	0	0,571	0,15
3	1	0	0,5	0,5	0	0,5	0,42
4	1	0,5	0,5	0,5	0,5	0,5	0,43
5	1	0	1	0,02	0	0,02	0,15
6	3,857	0,920	10,333	1	3,55	1	0,09

На рисунках 1-3 приведены результаты решения задачи Сода (тест 1) и сравнения результатов с [2]. В задаче Сода течение является

дозвуковым. Задача Лакса (тест 2), также моделирует дозвуковое течение. Однако, интенсивность волны разрежения меньше чем в первом тесте

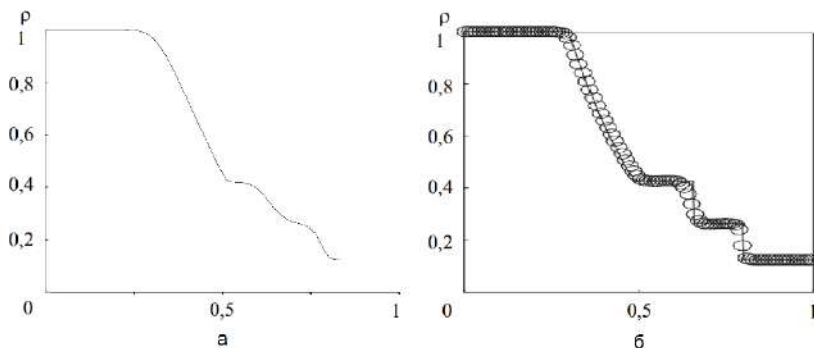


Рис. 1 – Профиль плотности в решении задачи Сода:
а – полученное решение, б – решение из [2]

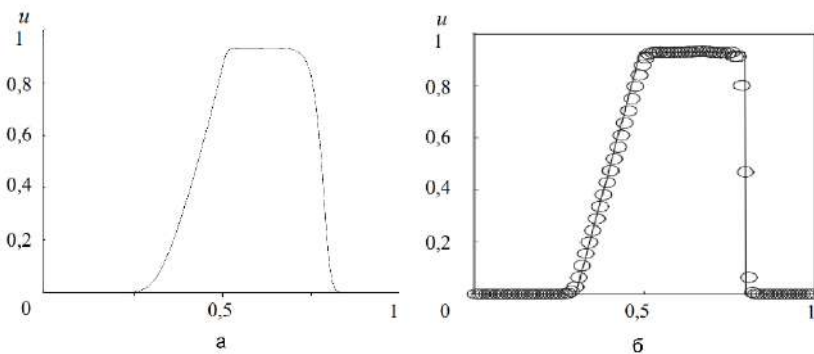


Рис. 2 – Профиль скорости в решении задачи Сода:
а – полученное решение, б – решение из [2]

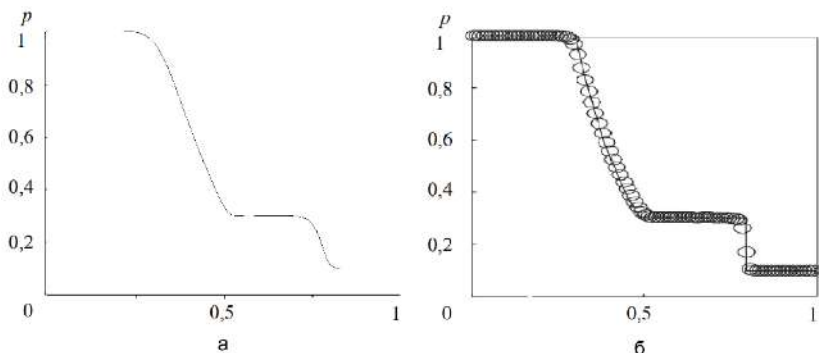


Рис. 3 – Профиль давления в решении задачи Сода:
 а – полученное решение, б – решение из [2]

Тест 3 соответствует задаче о неподвижном контактном разрыве. Полученное решение показывает, что контактный разрыв не перемещается и не распадается. Значения плотности и давления в тесте 4 аналогичны значениям в предыдущем тесте, но в этом случае газ обладает скоростью. В результате, можно наблюдать медленно перемещающийся контактный разрыв.

В задаче о сверхзвуковой ударной трубе (тест 5) образуется сверхзвуковое течение. Также, сверхзвуковое течение наблюдалось в тесте 6, где число Маха равнялось 3. В ходе решения данных задач наблюдались разрывы решения.

Написан и протестирован алгоритм для решения уравнений Эйлера. Результаты, полученные в тестах 1, 2, 5 и 6, хорошо соответствуют результатам из [2]. Однако, в тестах 3 и 4, полученные результаты претерпевают разрыв. Следовательно, для моделирования сверхзвуковых течений необходимо применение схем с более высоким порядком точности.

Работа выполнена при финансовой поддержке гранта Президента МК-421.2020.8

Литература:

1. *Liou M.-S., Steffen C. J.* A new flux splitting scheme // *Journal of computational physics.* – 1993. – N 107. – P. 23–29.
2. *Булат П.В., Волков К.Н.* Одномерные задачи высокой разрешающей способности // *Научно-технический вестник информационных технологий, механики и оптики.* – 2015. – Т.15, №4. – С. 731–740.

РАЗРАБОТКА КРОССПЛАТФОРМЕННОЙ ПРОГРАММЫ ДЛЯ РАСЧЕТОВ СВЕРХЗВУКОВЫХ ТЕЧЕНИЙ

Д.О. Ларкин

Национальный исследовательский Томский государственный университет, г. Томск, larkin935@gmail.com

Научный руководитель: Кагенов А.М. к.ф-м.н., доцент

Разработана кроссплатформенная GUI программа для расчета одномерных задач газовой динамики с использованием IDE «QT Creator» на языке программирования C++. Алгоритм программы основан на решении уравнений Эйлера по схеме С.К. Годунова. Определение потоков через грани ячеек выполняется из точного решения задачи Римана. Проведены тестовые расчёты.

A cross-platform GUI program has been developed for calculating one-dimensional problems of gas dynamics using the IDE "QT Creator" in the C++ programming language. The algorithm of the program is based on the solution of Euler equations according to the scheme of S.K. Godunov. The determination of flows through the cell faces is performed from the exact solution of the Riemann problem. Test calculations were carried out

Введение. С развитием вычислительных систем требуется разработка универсального программного обеспечения для моделирования физических процессов. Такими системами являются компьютеры на различных архитектурах процессоров отличных от x86, в том числе и отечественных процессорах (Байкал и Эльбрус), таких как ARM, MIPS и т.д. При этом в зависимости от архитектуры процессора используются операционные системы различных дистрибутивов Linux, например, Astra Linux. Поэтому целью работы является разработка кроссплатформенной программы для расчётов одномерных задач газовой динамики с использованием схемы С.К. Годунова

Методика расчёта. Рассматривается труба постоянного сечения. В точке $x=x_0$ расположена мембрана, разделяющая два газа находящихся при разных давлениях, скоростях и плотностях. В момент времени $t=0$ перегородка мгновенно убирается и произвольный разрыв распадается на несколько разрывов. В зависимости от начальных условий образуется та или иная конфигурация распада [1]. Для описания течения используется система уравнений Эйлера. Для замыкания системы уравнений используется уравнение состояния идеального газа.

Для решения системы уравнений применяется метод С.К. Годунова первого порядка точности [2]. Алгоритм нахождения параметров газа на

границах ячеек реализуется с помощью точного решения задачи Римана о распаде произвольного разрыва. Программная реализация выполнена на языке программирования «C++», пользовательский интерфейс реализован с использованием IDE «QT Creator» [3]. Реализована возможность построения графиков с применением внешней библиотеки «QcustomPlot».

Результаты. Проведены тестовые расчёты. Варианты расчетов приведены в таблице 1 [4]. Вариант 1 является модифицированной задачей Сода, вариант 2 – состоит из двух симметричных волн разряжения и контактной волны нулевой скорости, вариант 3 – состоит из сильной ударной волны, контактного разрыва и левой волны разряжения. На рисунках (1-3) показано сравнение результатов расчетов (пунктирная линия) с эталонными решениями (сплошная линия) рассматриваемых задач в различные моменты времени, где изображены графики давления, плотности, скорости и внутренней энергии. Во всех расчетах использовались следующие параметры: показатель адиабаты $k=1.4$, число Куранта $C=0.9$, число разбиений $N = 100$.

Получено, что реализованный алгоритм и программа расчета дают достоверные результаты расчетов для рассмотренных тестовых задач. Показаны хорошие результаты на контактных разрывах и сильных ударных волнах.

Заключение. Разработана кроссплатформенная программа для расчета одномерных задач газовой динамики в IDE «QT Creator» на языке программирования C++. Алгоритм программы основан на решении уравнений Эйлера по схеме С.К. Годунова. Проведены тестовые расчёты для модифицированной задачи Сода, задачи состоящей из двух симметричных волн разряжения и контактной волны нулевой скорости и задачи о сильной ударной волне.

Таблица 1 – Тестовые задачи

Вариант	Левое состояние			Правое состояние			x_0
	ρ	u	p	ρ	u	p	
1	1.0	0.75	1.0	0.125	0.0	0.1	0.3
2	1.0	-2.0	0.4	1.0	2.0	0.4	0.5
3	1.0	0	1000.0	1.0	0.0	0.01	0.5

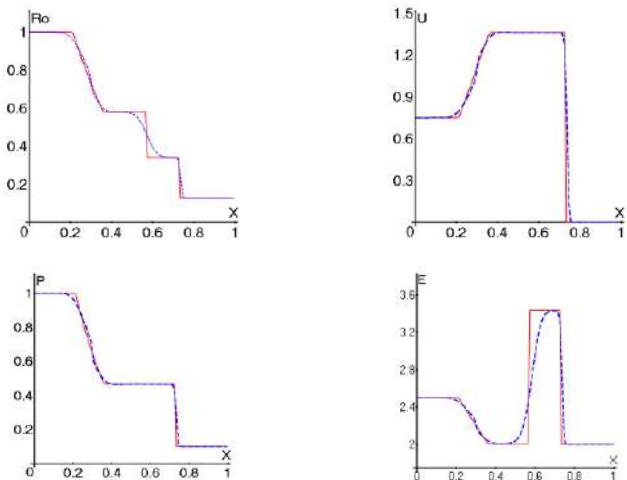


Рис. 1 – Вариант 1 в момент времени $t=0,2$

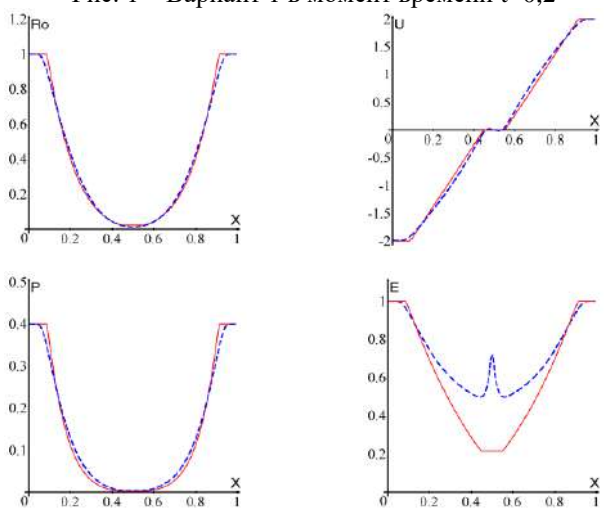


Рис. 2 – Вариант 2 в момент времени $t=0.15$

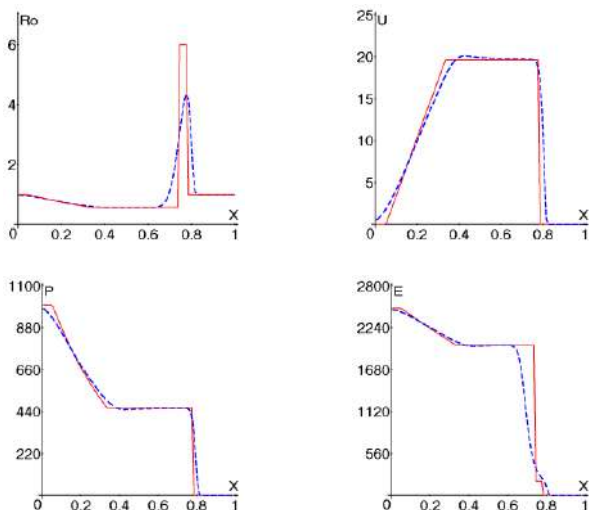


Рис. 3 – Вариант 3 в момент времени $t=0.012$

Литература:

1. Куликовский, А.Г. Математические вопросы численного решения гиперболических систем уравнений / А.Г. Куликовский, Н.В. Погорелов, А.Ю. Семенов. – М.: ФИЗМАТЛИТ, 2001. – 608 с.
2. Численное решение многомерных задач газовой динамики / С.К. Годунов, А.В. Забродин, М.Я. Иванов и др.; под ред. С.К. Годунова. М.: Наука, 1976. – 400 с.
3. Шлее, М. Qt 5.10. Профессиональное программирование на C++ / М. Шлее. – СПб.: БХВ-Петербург, 2018. – 1072 с.
4. Toro E.F. Riemann solvers and numerical methods for fluid dynamics. Third Edition – SpringerVerlag, Berlin Heidelberg, 2009, 724 p. DOI: 10.1007/b79761

ПОСТРОЕНИЕ КОНЕЧНОЭЛЕМЕНТНЫХ СЕТОК С ГОРИЗОНТАЛЬНЫМИ СКВАЖИНАМИ ПРИ МОДЕЛИРОВАНИИ ПРОЦЕССОВ МНОГОФАЗНОЙ ФИЛЬТРАЦИИ

Д.А. Леонович, Ю.Г. Соловейчик

Новосибирский государственный технический университет,
г. Новосибирск, leonovich.d.a@yandex.ru

Научный руководитель: д.т.н., профессор Ю.Г. Соловейчик

В работе рассматриваются два подхода к построению конечноэлементных сеток при моделировании разработки нефтяных месторождений с использованием горизонтальных скважин.

The paper discusses two approaches to the construction of finite element meshes for modeling the oil reservoirs where horizontal wells are used.

В настоящее время в связи с исследованием и применением методов увеличения нефтеотдачи расширяется область применения технологии горизонтальных скважин при разработке углеводородных месторождений. Так, например, горизонтальные скважины получили широкое распространение при разработке месторождений высоковязкой и сверхвязкой нефти с использованием тепловых методов повышения нефтеотдачи: закачка горячего пара или воды, технология парогравитационного дренажа (или SAGD) и другие методы [2]. В связи с этим становится актуальным моделирование процессов нефтедобычи с использованием технологии горизонтальных скважин.

Для моделирования многофазной фильтрации в пористых средах используется метод конечных элементов [1]. Для этого с использованием закона Дарси и закона сохранения масс составляется краевая задача для расчета давления. На каждом временном шаге на основе найденного давления вычисляются потоки фаз, перетекающие через грани ячеек конечноэлементной сетки, и определяется новый фазовый состав в конечных элементах. Более подробное описание метода численного решения данной задачи представлено в работах [3, 4].

В данной работе рассматриваются два подхода к построению конечноэлементных сеток с горизонтальными скважинами. Первый из них основан на явной встройке скважины в сетку, т.е. на удалении элементов из базовой сетки на месте скважины и построении новых шестигранных ячеек, описывающих скважину. При этом в область самой скважины элементы сетки не встраиваются.

Алгоритм построения конечноэлементной сетки заключается в следующем. Задаются параметры скважины: радиус, начало и конец скважины, границы зон перфорации. Также задаются следующие параметры сетки: коэффициент разрядки (для сгущения сетки вблизи скважины), количество дроблений между скважиной и базовой сеткой, коэффициент, определяющий во сколько раз область встройки скважины должна быть больше радиуса скважины. В базовой сетке находятся все элементы, попадающие в область встройки скважины, и удаляются из сетки. После чего строятся новые элементы: от каждого узла по краю вырезанной области строится прямая до центральной оси скважины. На этих прямых откладываются узлы для разбиения встраиваемой области на более мелкие ячейки. После чего эти узлы последовательно обходятся и объединяются в конечные элементы сетки. Параметры пористости среды (проницаемость, пористость и насыщенности фаз) присваиваются новому элементу в зависимости от того, на какой удаленный элемент из базовой сетки приходится центр нового конечного элемента. При этом в область встраивается вся скважина, часть граней которой образует заданные зоны перфорации. Пример построения приведен на рисунке 1. Слева представлена сетка со встроенной скважиной, разрезанная в плоскости XZ; красным прямоугольником выделена область с зоной перфорации. На правом рисунке вид на скважину в плоскости YZ; прямоугольником выделена область встройки скважины.

Во втором способе встройки горизонтальных скважин в конечноэлементную сетку их зоны перфорации задаются с помощью маленьких (в соответствии с радиусом скважин) ячеек. При этом чтобы избежать избыточного дробления всей сетки, используется технология несогласованных сеток. Отметим также, что при численном решении задачи фильтрации элементы, приходящиеся непосредственно на «внутреннюю» часть скважин, считаются непроницаемым, т.е. их фазовое состояние не пересчитывается на каждом временном шаге.

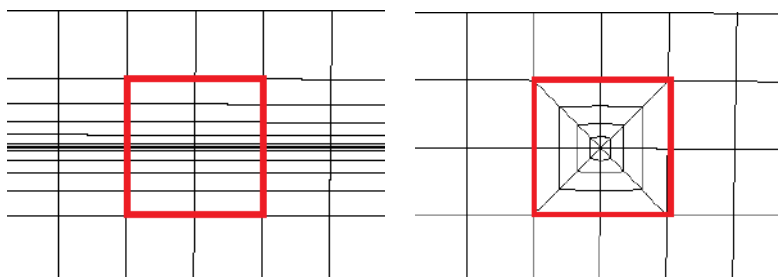


Рис. 1 – Пример построения конечноэлементной сетки с помощью явной встройки скважины

На рисунке 2 представлено изображение построенной сетки в области зоны перфорации. Красным квадратом выделена область встройки скважины. Слева показана скважина в плоскости XZ, справа – в плоскости YZ.

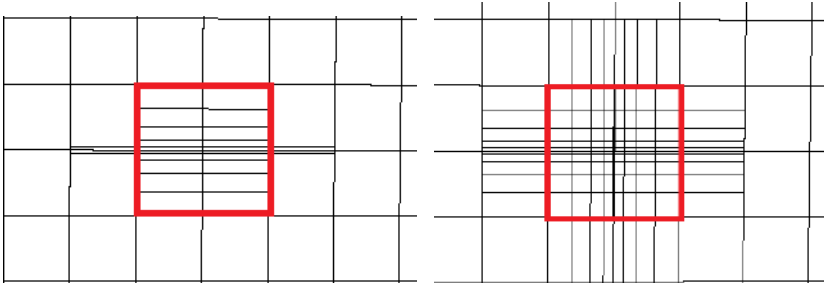


Рис. 2 – Пример построения конечноэлементной сетки с помощью неявной встройки скважины

Основные отличия второго подхода от явной встройки скважины заключаются в том, что дополнительные ячейки добавляются только для зон перфорации, а остальная часть скважины не встраивается в сетку (из предположения, что радиус скважины слишком мал), и в том, что в поперечном сечении скважина квадратная, а не многоугольная, так как используются только конечные элементы в виде параллелепипедов.

Литература:

1. Соловейчик Ю. Г., Рояк М. Э., Персова М. Г. Метод конечных элементов для решения скалярных и векторных задач. — Новосибирск: Изд-во НГТУ, 2007. — 896 с.
2. Mokheimer, E., Hamdy, M., Abubakar, Z., Shakeel, M. R., Habib, M. A., & Mahmoud, M. A comprehensive review of thermal enhanced oil recovery: Techniques evaluation // *Journal of Energy Resources Technology*, 2019, T. 141, №. 3.
3. Persova M.G. et al. Numerical Modeling of Multi-Phase Flow for Various Junctions of Water and Oil Saturated Layers in 3-D Porous Media // 2018 14th International Scientific-Technical Conference on Actual Problems of Electronic Instrument Engineering, APEIE 2018 - Proceedings. Institute of Electrical and Electronics Engineers Inc., 2018. P. 212–215.
4. Persova M.G. et al. Flow Balancing in FEM Modelling of Multi-Phase Flow in Porous Media // 2018 14th International Scientific-Technical Conference on Actual Problems of Electronic Instrument Engineering, APEIE 2018 - Proceedings. IEEE, 2018. P. 205–211.

РАЗВИТИЕ МЕТОДА СТОХАСТИЧЕСКОГО МОЛЕКУЛЯРНОГО МОДЕЛИРОВАНИЯ ПРОЦЕССОВ ПЕРЕНОСА РАЗРЕЖЕННЫХ ГАЗОВ В НАНОКАНАЛАХ

Д.Н. Любимов

Новосибирский государственный архитектурно – строительный
университет, г. Новосибирск, danilch@mail.ru

Научный руководитель: Рудяк В.Я., д-р физ.-мат. наук, проф.

Предлагается новый алгоритм для расчета коэффициентов переноса разреженного газа в стесненных условиях. В его основе лежит стохастическое моделирование фазовых траекторий рассматриваемой молекулярной системы. Коэффициенты переноса вычисляются с помощью формул Грина–Кубо. Взаимодействие молекул газа со стенкой описывается зеркальным, диффузным или зеркально-диффузным законами. Работоспособность алгоритма продемонстрирована на расчете теплопроводности разреженных газов в наноканалах при нормальных условиях. Варьируется высота канала. Показано, что теплопроводность в наноканале существенно анизотропна.

A new algorithm is proposed for calculating the transport coefficients of a rarefied gas in confined conditions. It is based on stochastic modeling of the phase trajectories of the considered molecular system. The transfer coefficients are calculated using the Green – Kubo formulas. The interaction of gas molecules with the wall is described by mirror, diffuse, or mirror-diffuse laws. The performance of the algorithm is demonstrated by calculating the thermal conductivity of a rarefied gas in nanochannels under normal conditions. Gas pressure and channel height vary. It is shown that the thermal conductivity in the nanochannel is substantially anisotropic.

В последние несколько десятилетий в производство активно внедряется микросистемная техника (МСТ) и соответствующие технологии. В ряде случаев МСТ является единственно возможной, например, при создании компьютеров нового поколения, в различных биомедицинских приложениях. Во многих устройствах МСТ имеют место течения жидкостей или газов. Сложность экспериментального исследования течений флюидов в стесненных условиях (в микро- и наноканалах, порах и т.п.) вполне очевидна. Как правило, здесь можно получить лишь некоторую интегральную информацию. Не удастся, в частности, измерить коэффициенты переноса, хотя их знание совершенно необходимо для описания соответствующих течений. В такой ситуации естественной альтернативой экспериментальному

изучению коэффициентов переноса является их моделирование. Поскольку коэффициенты переноса имеют молекулярную природу, то и их моделирование требует использования молекулярного уровня описания. Типичным методом такого сорта является метод молекулярной динамики (см., например, [1]). Однако этот метод неприменим для моделирования коэффициентов переноса разреженных газов в наноканалах, так как требует огромного количества молекул (характерный размер канала должен быть много больше длины свободного пробега молекул газа). Ранее в работах [2-4] разработан метод стохастического молекулярного моделирования (СММ) для моделирования коэффициентов переноса разреженного газа в объеме. Целью данной работы является обобщение этого метода для моделирования процессов переноса разреженного газа в стесненных условиях. Применимость метода продемонстрирована на моделировании теплопроводности молекул разреженных газов в наноканалах.

Ячейка моделирования создавалась в виде прямоугольного параллелепипеда с квадратным сечением. Вдоль оси ячейки использовались периодические граничные условия. Взаимодействие молекул газа со стенками канала описывается зеркальным или зеркально-диффузным законами. В первом случае скорость молекулы вдоль поверхности не меняется, а нормальная изменяет знак. При зеркально-диффузном отражении вводится коэффициент аккомодации θ , определяющий долю молекул, взаимодействующих со стенкой диффузно, оставшаяся часть $(1 - \theta)$ молекул отражается зеркально. Скорость диффузно отраженной молекулы разыгрывается по функции распределения Максвелла по скоростям при заданной температуре. Взаимодействие молекул в данной работе описывается потенциалом Леннарда-Джонса.

Коэффициенты переноса разреженного газа вычисляются с помощью формул Грина–Кубо, которые определяют их как интегралы по времени от двухвременных корреляционных функций соответствующих динамических переменных [5]. Вычисление корреляционных функций требует в свою очередь определения фазовых переменных системы (координат и импульсов всех N молекул моделируемого газа). Поскольку корреляционные функции вычисляются по равновесной функции распределения, то необходимо смоделировать эволюцию молекул равновесного газа.

Первым шагом построения алгоритма является создание начального состояния. Для этого в соответствии с заданной массовой плотностью газа ρ молекулы распределяются равномерно по ячейке моделирования. Затем в соответствии с температурой газа T разыгрываются скорости

молекул \mathbf{v}_i согласно распределению Максвелла. Суммарный импульс всех N молекул равен нулю. После того, как сформировано начальное состояние системы, формируется набор координат и импульсов всех молекул газа в последующие моменты времени. Время расчета фазовой траектории t_s разбивается на интервалы длительностью $\tau_i = \sigma/v_{max}$ где σ – эффективный диаметр молекул, а v_{max} – максимальная из скоростей молекул в данный момент времени. Пусть в момент времени t состояние молекул характеризуется скоростями $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_N$ и координатами $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N$. Для получения значений динамических переменных в момент времени $(t + \tau_1)$ их динамика расщепляется по процессам. Сначала определяется положение молекулы i , она сдвигается в конфигурационном пространстве по закону: $\mathbf{r}'_i = \mathbf{r}_i + \mathbf{v}_i \tau_1$. Если молекула достигает стенки, то ее скорость изменяется в соответствии с законом отражения.

Если взаимодействие со стенкой не произошло, тогда проверяется возможность ее столкновения с другой молекулой. На основе кинетической теории вычисляется вероятность соударения за время τ_1 молекулы с любой другой: $P_{c1} = 4\tau_1 \rho \sigma^2 \sqrt{\pi RT/m}$ [6]. Если молекула сталкивается с какой-либо другой молекулой, тогда их скорости изменятся в соответствии с законами сохранения импульса и энергии. Молекулы, которые не сталкивались за время τ_1 , своих скоростей не меняют. Эта процедура последовательно выполняется для всех N молекул. В результате формируется полный список новых фазовых переменных системы в момент времени $(t + \tau_1)$. Затем эта процедура повторяется для всех последующих шагов τ_i пока не закончится заданное время расчета τ_s . Таким образом, получаем полный набор динамических переменных всех молекул моделируемой системы в последовательные моменты времени.

Коэффициент теплопроводности определяется следующей флуктуационно-диссипационной теоремой

$$\lambda = \frac{k}{3VT^2} \int_0^{\tau_p} \langle \mathbf{j}(0) \cdot \mathbf{j}(t) \rangle dt, \quad (1)$$

$$\mathbf{j}(t) = \sum_i (\mathbf{v}_i(t) - \mathbf{v}_c(t)) e_i(t), \quad \mathbf{v}_c(t) = \frac{1}{N} \sum_i \mathbf{v}_i(t).$$

Здесь V – объем системы, k – постоянная Больцмана, e_i – кинетическая энергия молекулы ($m\mathbf{v}_i^2/2$), а \mathbf{v}_c – скорость центра масс системы. Ненулевые флуктуации центра масс системы возникают из-за взаимодействия молекул со стенками канала.

В данной работе моделировался коэффициент теплопроводности аргона, криптона, ксенона, кислорода и метана при температуре 300 К и

атмосферном давлении. Рассчитывались продольный, поперечный и полный коэффициенты теплопроводности. Коэффициент теплопроводности вдоль достаточно длинного канала должен совпадать с соответствующим значением в объеме. Установлено, что это действительно так. Таким образом, метод СММ адекватно моделирует теплопроводность газа. Однако теплопроводность газов в наноканалах имеет важные особенности. На рис. 1 в качестве примера представлена зависимость от высоты канала продольного (вдоль оси OY) значения коэффициента теплопроводности аргона (рис. 1а) и поперечного (вдоль оси OX, рис. 1б). Здесь черные круги обозначают значения коэффициента теплопроводности с учетом флуктуаций центра масс, а квадраты – без этого учета. Пунктирная линия на рис. 1а обозначает экспериментальное значение коэффициента теплопроводности в объеме [7].

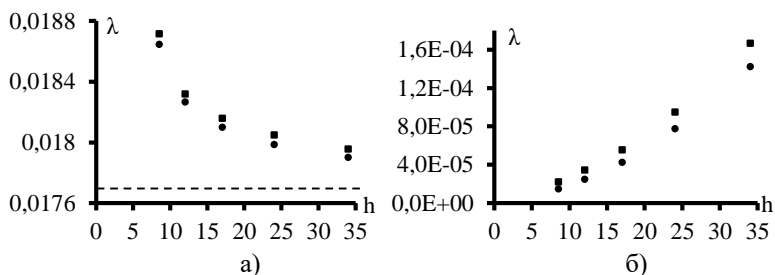


Рис. 1 – Зависимость продольного а) и поперечного б) коэффициентов теплопроводности λ , Вт/(м·К), аргона от высоты h , нм, наноканала

Сопоставление рис. 1а и 1б показывает, что теплопроводность газа в наноканале ниже соответствующего значения в объеме и она существенно анизотропна. С ростом высоты канала теплопроводность вдоль канала стремится к теплопроводности в объеме при тех же условиях. Анизотропия существенно зависит от закона взаимодействия газа со стенкой и снижается с ростом высоты канала. Так в канале с зеркальными стенками при высоте канала 12 нм отношение продольного коэффициента теплопроводности к поперечному около 2700, а в канале высотой 1000 нм – всего 2.3. Тем не менее эта анизотропия имеет место, как видно не только в нано-, но и в микроканалах.

Работа выполнена при частичном финансировании Российского фонда фундаментальных исследований (грант № 20-01-00041) и Мегагранта Минобрнауки РФ (соглашение № 075-15-2021- 575).

Литература:

1. Rapaport, D.C., The art of molecular dynamics simulation, Cambridge (UK): Cambridge University Press, 2004.
2. Rudyak V.Ya., Lezhnev E.V. Стохастический метод моделирования коэффициентов переноса разреженного газа // Матем. моделирование. 2017. Т. 29. № 3. С. 113–122.
3. Rudyak V.Ya., Lezhnev E.V. Stochastic algorithm for simulating gas transport coefficients. J. Comp. Physics. 2018. Vol. 355. P. 95–103.
4. Рудяк В.Я., Лежнев Е.В., Любимов Д.Н. Имитационное моделирование коэффициентов переноса разреженных газов и наногазовзвесей // Вестн. Томск. гос. ун-та. Матем. и мех. 2019. № 59. С. 105–117.
5. Рудяк В.Я. Статистическая аэрогидромеханика гомогенных и гетерогенных сред. Т. 2. Гидромеханика. Новосибирск: НГАСУ, 2005. 468 с.
6. Chapman S., Cowling T.G. The Mathematical Theory of Non-Uniform Gases. Cambridge: Cambridge University Press, 1990. 423 p.
7. Григорьев И.С., Мейлихова Е.З. Физические величины. – М.: Энергоатомиздат, 1991. – 1234 с.

РАСЧЕТ ПОЛЯ ДАВЛЕНИЯ В ЗАДАЧАХ МНОГОФАЗНОЙ ФИЛЬТРАЦИИ СЖИМАЕМЫХ ФАЗ

А.С. Овчинникова

**Новосибирский государственный технический университет,
г. Новосибирск, ovchinnikova.2014@stud.nstu.ru**

Рассматривается подход к моделированию многофазной фильтрации смеси сжимаемых фаз. Верификация предложенного метода проведена на тестовой задаче из сравнительного проекта SPE-1.

An approach to modeling multiphase flow of compressible phases is considered. The verification of the proposed method was carried out on a test problem from the comparative project SPE-1.

В данной работе рассматривается проблема моделирования газонефтяных месторождений. Отличительными чертами такой задачи являются высокая степень зависимости плотности газа от давления и процесс выделения/поглощения газа в нефти, что также приводит к значительному изменению объема нефти в коллекторе. Для адекватного

моделирования течения смеси флюидов в таких условиях необходимо учитывать эти особенности.

Для численного моделирования процессов многофазной фильтрации решается следующая краевая задача

$$-\operatorname{div}\left(\sum_{m=1}^{NP}\frac{\kappa^m}{\mu^m}\mathbf{K}\left(\operatorname{grad}P+(0,0,\rho^m g)^T\right)\right)=f^\Omega(P) \quad (1)$$

с краевыми условиями, определенными на границе $\partial\Omega=\Gamma_1\cup\Gamma_2$ расчетной области Ω ,

$$P|_{\Gamma^1}=P^\Gamma, \quad (2)$$

$$\sum_{m=1}^{NP}\frac{\kappa^m}{\mu^m}\mathbf{K}\left(\operatorname{grad}P+(0,0,\rho^m g)^T\right)\Big|_{\Gamma^2}\cdot\vec{n}=\sum_{m=1}^{NP}f^{m,\Gamma}(P). \quad (3)$$

где \mathbf{K} – тензор структурной проницаемости, m – номер фазы, NP – число фаз, κ^m, μ^m, ρ^m – коэффициент относительной фазовой проницаемости, вязкость и плотность фазы m , P – давление. g – ускорение свободного падения, z – координата на вертикальной оси, P^Γ – давление, заданное на границах Γ^1 , а $f^\Gamma(P)=\sum_{m=1}^{NP}f^{m,\Gamma}(P)$ – поток смеси сжимаемых фаз, заданный на границах Γ^2 .

Для решения уравнения (1) функция $f^\Omega(P)$, характеризующая изменение объема смеси $\Delta V(P)$ в расчетной области Ω за временной интервал Δt , линеаризуется в окрестности давления P_0

$$f^\Omega(P)\approx\frac{\Delta V(P_0)}{\operatorname{mes}(\Omega)\cdot\Delta t}+\alpha^{\Delta V}(P_0-P),$$

где $\operatorname{mes}(\Omega)$ – объем расчетной области Ω , $\alpha^{\Delta V}$ – коэффициент, характеризующий изменение объема смеси в зависимости от изменения давления.

Коэффициент $\alpha^{\Delta V}$ рассчитывается, исходя из значений объемов смеси V при давлении P_0 и при давлении $\tilde{P} = P_0 + \Delta P$ (ΔP – некоторое приращение к P_0) по формуле

$$\alpha^{\Delta V}(P_0) = \frac{V(P_0) - V(\tilde{P})}{\Delta P \cdot \text{mes}(\Omega) \cdot \Delta t}.$$

Таким образом, расчет давления на очередном шаге Δt представляет собой итерационный процесс решения нелинейного уравнения (1), алгоритм которого показан на рисунке 1.

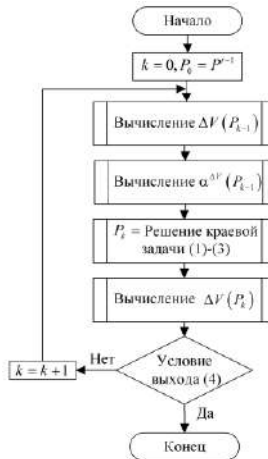


Рис.1 – Алгоритм решения нелинейного уравнения для давления

Критерием окончания итерационного процесса является выполнение двух условий:

$$\left| \frac{\Delta V(P_k) - \Delta V(P_{k-1})}{\Delta V(P_k)} \right| < \varepsilon_{\Delta V}, \quad \left| \frac{\sum_h P_k^h - \sum_h P_{k-1}^h}{\sum_h P_k^h} \right| < \varepsilon_P, \quad (4)$$

где P_k^h – давление на активной зоне перфорации с номером h .

Краевая задача (1)–(3) решается с использованием метода конечных элементов [1,2]. Полученное распределение давления используется для

вычисления численных потоков смеси, перетекающих через грани конечных элементов. Для выполнения закона сохранения фильтрующейся смеси используется метод балансировки потоков [3], корректирующий численные потоки. Сбалансированные потоки смеси делятся на потоки фаз, которые используются для расчета нового распределения насыщенности [4,5].

Верификация рассмотренного подхода проведена на задаче из сравнительного проекта SPE-1 [6]. Модель представляет собой куб с размерностью 10000x10000x100 фут³. В углах модели расположены две скважины: нагнетательная и добывающая. Нагнетательная скважина используется для закачки 10^8 SCF/D газа, для добывающей скважины определен отбор нефти – $2 \cdot 10^4$ STB/D.

На рисунке 2 продемонстрируем полученные результаты на примере графиков дебита нефти и газового фактора отбираемой нефти.

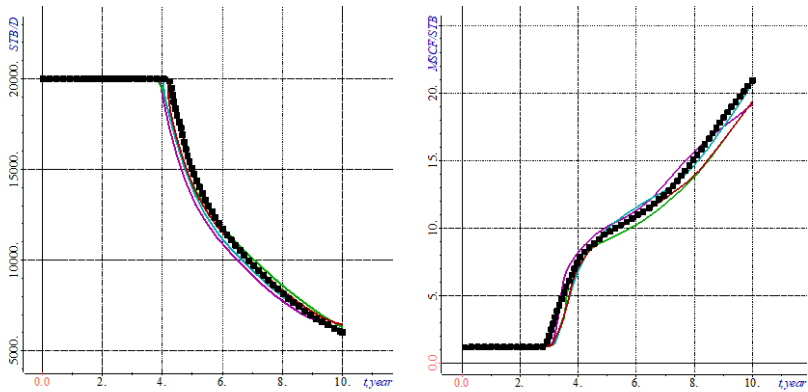


Рис.2 – Мгновенный отбор нефти (слева) и газовый фактор (справа):
 кривая с метками – расчеты, полученные с использованием рассматриваемого подхода; кривые без меток – расчеты участников проекта SPE-1

Как видно из представленных на рисунке 2 результатов, графики, полученные с помощью предложенного подхода, лежат либо между графиками участников проекта SPE-1, либо достаточно близко к ним.

Отметим также что, количество итераций, необходимое для решения нелинейного уравнения (1), на каждом временном шаге составляет 1–2 итерации.

Литература:

1. Finite element solution to 3-D airborne time-domain electromagnetic problems in complex geological media using non-conforming hexahedral meshes / M.G. Persova, Yu.G. Soloveichik, D.V. Vagin, D.S. Kiselev, Yu.I. Koshkina // *Journal of Applied Geophysics*. – 2020. – Vol. 172. – Art. 103911. – doi: 10.1016/j.jappgeo.2019.103911.
2. Finite-element solution to multidimensional multisource electromagnetic problems in the frequency domain using non-conforming meshes / Yu.G. Soloveichik, M.G. Persova, P.A. Domnikov, Yu.I. Koshkina, D.V. Vagin // *Geophysical Journal International*. – 2018. – Vol. 212, iss. 3. – P. 2159–2193. – doi: 10.1093/gji/ggx530.
3. Persova M.G., Soloveichik Yu.G., Grif A.M., Patrushev I.I. Flow balancing in FEM modelling of multi-phase flow in porous media // *Actual problems of electronic instrument engineering, AREIE–2018*, Novosibirsk, Russia, 2–6 oct. 2018: Conference proceedings. – Novosibirsk, 2018. – P. 205–211.
4. The design of high-viscosity oil reservoir model based on the inverse problem solution / M.G. Persova, Yu.G. Soloveichik, D.V. Vagin, A.M. Grif, D.S. Kiselev, I.I. Patrushev, A.V. Nasybullin, B.G. Ganiev // *Journal of Petroleum Science and Engineering*. – 2021. – Vol. 199. – Art. 108245. – doi: 10.1016/j.petrol.2020.108245.
5. Persova M.G., Soloveichik Yu.G., Vagin D.V., Grif A.M., Patrushev I.I., Ovchinnikova A.S. Oil production optimization based on the finite-element simulation of the multi-phase flow in porous media and inverse problem solution // *GeoBaikal 2020*, oct. 2020: Conference proceedings. – 2020. – Vol. 2020. – P. 1–6.
6. Comparison of Solutions to a Three-Dimensional Black-Oil Reservoir Simulation Problem / A.S. Odeh // *Journal of Petroleum Technology*. – 1981. – Vol. 33. – P. 13–25. – doi: 10.2118/9723-pa.

СТАБИЛИЗИРОВАННЫЕ КОНЕЧНОЭЛЕМЕНТНЫЕ СХЕМЫ ДЛЯ РЕШЕНИЯ ЗАДАЧ КОНВЕКЦИИ-ДИФФУЗИИ

А.Д. Пастернак

Новосибирский государственный технический университет,
г. Новосибирск, Pasternak.2015@stud.nstu.ru
Научный руководитель: Шурина Э.П., д.т.н., профессор

Разработаны, реализованы и исследованы вычислительные схемы на базе стабилизированного метода SUPG для решения стационарных задач конвекции-диффузии-реакции.

Developed, implemented and investigated computational schemes based on the stabilized SUPG method for solving stationary convection-diffusion-reaction problems.

Многие прикладные задачи, описывающие природные явления, являются сингулярно-возмущенными и содержат в уравнении доминирующий конвективный член, из-за чего решение, полученное с помощью метода Галеркина, может содержать в себе большие численные колебания и выбросы по всей области исследования, а также полностью не соответствовать истинному решению. Таким образом за последние 30 лет были разработаны стабилизирующие методы конечных элементов. Одним из таких методов является streamline upwind/PetrovGalerkin(SUPG). Основная идея этого метода заключается в добавлении стабилизирующего члена в вариационную постановку, который позволяет сохранить точность и увеличить устойчивость схемы.

Стационарная задача конвекции-диффузии-реакции заключается в нахождении функции $u(x)$, для которой верно:

$$-div(-\lambda grad(u)) + a grad(u) + \sigma u = f, x \in \Omega$$

с краевыми условиями:

$$u = g, x \in \partial\Omega,$$

где λ - коэффициент диффузии, $a(a_1(x), a_2(x))$ - вектор скорости, σ - коэффициент реакции источника, $f = (x)$ - функция источника.

Введем функциональное пространство Соболева $H^1(\Omega)$,

$$H^1 = \{u \in L^2(\Omega) : \nabla u \in (L^2(\Omega))\}$$

Скалярное произведение для элементов пространства $L^2(\Omega)$ вычисляются по формуле:

$$(u, v) = \int u(x)v(x)dx$$

Норма для пространства $H^1(\Omega)$, $L^2(\Omega)$:

$$\|u\|_{L^2(\Omega)} = \sqrt{(u, u)}, \|u\|_{H^1(\Omega)}^2 = \|u\|_{L^2(\Omega)}^2 + \|\nabla u\|_{L^2(\Omega)}^2$$

Воспользовавшись формулой Грина интегрирования по частям, в итоге получим:

$$Lu = f, x \in \Omega$$

$$u = g, x \in d\Omega$$

Тогда будем искать решение $u \in H_0^1, f \in (L^2\Omega)$, удовлетворяющее следующему равенству:

$$\lambda \int_{\Omega} \nabla u \nabla v d\Omega + \int_{\Omega} (a \nabla u) v d\Omega + \sigma \int_{\Omega} u v d\Omega = \int_{\Omega} f v d\Omega, \forall v \in H_0^1(\Omega)$$

Любой конечно-элементный метод подразумевает разбиение области Ω множеством ячеек. В данной работе рассматривались двумерные задачи и разбиение производилось на треугольники. Пример расчетной области (Рисунок 1) представлен ниже. Конечный элемент – треугольник (Рисунок 2). В качестве базиса использовались линейные базисные функции

Множество стабилизированных методов можно определить следующим образом [5,7]:

$$b(u, v) + \sum_T \tau (Lu - f, \bar{L}v) = (f, v)$$

где τ - скалярный параметр, называющийся стабилизированным коэффициентом. При $\tau = 0$ получим классический метод конечных элементов.

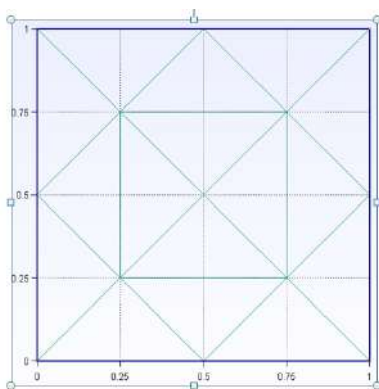


Рис. 1 – Пример расчетной области

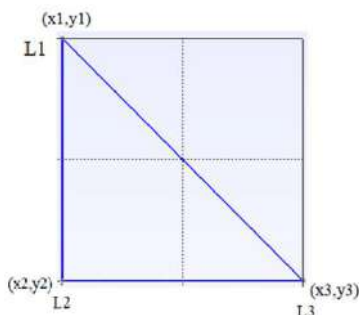


Рис. 2 – Конечный элемент

Стабилизированный оператор для метода SUPG имеет следующий вид:

$$\bar{L}v = a\nabla v + \sigma v$$

Разработанная вычислительная схема была верифицирована на классе модельных задач, имеющих аналитическое решение. Применимость данной схемы была исследована на задаче о изменении концентрации вещества в химическом реакторе.

Литература:

1. Hundsdorfer W., Verwer J. Numerical Solution of Time-Dependent Advection-Diffusion- Reaction Equations. Springer Berlin Heidelberg, 2003. 471 с.

2. Verwer J.G., Blom J.G., Hundsdorfer W. An implicit-explicit approach for atmospheric transport-chemistry problems // *Appl. Numer. Math.* 1996. T. 20, № 1–2. С. 191–209.
3. McRae G.J., Goodin W.R., Seinfeld J.H. Numerical solution of the atmospheric diffusion equation for chemically reacting flows // *J. Comput. Phys.* Academic Press, 1982. Т. 45, № 1. С. 1–42.
4. Google Scholar [Электронный ресурс]. URL: <https://scholar.google.com/> (дата обращения: 10.06.2021).
5. Franca L.P., Hauke G., Masud A. Revisiting stabilized finite element methods for the advective-diffusive equation // *Comput. Methods Appl. Mech. Eng.* 2006. Т. 195, № 13–16. С. 1560–1572.
6. Shishkin G.I. Roos, H.-G.; Stynes, M.; Tobiska, L.: *Numerical Methods for Singularly Perturbed Differential Equations. Convection-Diffusion and Flow Problems.* Berlin etc., Springer-Verlag 1996. XVI, 348 pp., DM 148,00. ISBN 3-540-60718-8 (Springer Series in Computation // ZAMM - J. Appl. Math. Mech. / Zeitschrift für Angew. Math. und Mech. 1997. Т. 77, № 5. С. 403–403.
7. Hughes T.J.R. A simple scheme for developing ‘upwind’ finite elements // *Int. J. Numer. Methods Eng.* 1978. Т. 12, № 9. С. 1359–1365.
8. Hughes T.J.R., Franca L.P., Hulbert G.M. A new finite element formulation for computational fluid dynamics: VIII. The galerkin/least-squares method for advective-diffusive equations // *Comput. Methods Appl. Mech. Eng.* North-Holland, 1989. Т. 73, № 2. С. 173–189.
9. Reed W.H., Hill T.R. Triangular mesh methods for the neutron transport equation // National topical meeting on mathematical models and computational techniques for analysis of nuclear systems, Ann Arbor, Michigan, USA, 8 Apr 1973. New Mexico: Los Alamos Scientific Laboratory, 1973. 23 с.
10. Oden J.T., Babuška I., Baumann C.E. A Discontinuous hp Finite Element Method for Diffusion Problems // *J. Comput. Phys.* Academic Press, 1998. Т. 146, № 2. С. 491–519.
11. Arnold D.N. An Interior Penalty Finite Element Method with Discontinuous Elements // *SIAM J. Numer. Anal. Society for Industrial and Applied Mathematics*, 1982. Т. 19, № 4. С. 742–760.
12. Abouchakra R. Hierarchical Tetrahedral Elements using Orthogonal Polynomials // *Proceedings of 1996 Canadian Conference on Electrical and Computer Engineering.* Calgary, Alberta, Canada, Canada: IEEE, 1996. С. 525–528.
13. Díez A., Ortega A., Garcia de la Torre J. Brownian dynamics simulation of analytical ultracentrifugation experiments // *BMC Biophys.* 2011. Т. 4, № 1. С. 6.

14. Cockburn B., Shu C.-W. The Runge–Kutta Discontinuous Galerkin Method for Conservation Laws V: Multidimensional Systems // J. Comput. Phys., Academic Press, 1998. Т. 141, № 2. С. 199–224.

**ГРУППИРОВАНИЕ КОНЕЧНЫХ ЭЛЕМЕНТОВ
СО СГЛАЖИВАНИЕМ ПЕРЕНОСА ФАЗ ПРИ
МОДЕЛИРОВАНИИ НЕСТАЦИОНАРНЫХ МНОГОФАЗНЫХ
ПОТОКОВ В ПОРИСТЫХ СРЕДАХ**

И.И. Патрушев

**Новосибирский государственный технический университет,
г. Новосибирск, patrushev.ii@ami.nstu.ru**

Работа посвящена численному методу моделирования разработки нефтяных месторождений. Рассматривается модификация процедуры группирования конечных элементов, позволяющая сглаживать процесс переноса объемов фаз между ячейками конечноэлементной сетки. Численные эксперименты проведены на модели реального месторождения.

The work is devoted to a numerical method for modeling the development of oil fields. A modification of the finite element grouping procedure is considered, which makes it possible to smooth the transfer of phase volumes between the cells of the finite element mesh. Numerical experiments were carried out on a model of a real oil field.

В настоящее время активно развиваются методы моделирования разработки нефтяных месторождений [1]. В данной работе используется метод моделирования многофазной фильтрации в пористых средах, в основе которого лежит использование метода конечных элементов [2], в котором для аппроксимации давления используются скалярные базисные функции из гильбертова пространства H^1 функций, имеющих суммируемые с квадратом производные. Данный метод модифицирован таким образом, чтобы массы перетекающих фаз сохранялись в численном решении. Для этого используется специальная процедура балансировки потоков смеси через грани конечных элементов [3]. Для дополнительного существенного сокращения вычислительных затрат используется процедура группирования [4], позволяющая пересчитывать состояния разномасштабных ячеек со своим допустимым шагом по

времени. В данной работе рассматривается модификация процедуры группирования.

Пусть даны потоки $Q_{\Gamma_i}^m$ фазы m через грани ячеек Γ_i . Тогда объемы фазы m , которые за время Δt перетекают через грани Γ_i , можно вычислить по формуле

$$V_{\Gamma_i}^m = Q_{\Gamma_i}^m \cdot \Delta t. \quad (1)$$

Новые насыщенности фаз $\tilde{S}_{\Omega_e}^m$ в ячейках рассчитываются по следующей формуле

$$\tilde{S}_{\Omega_e}^m = \left(S_{\Omega_e}^m \text{mes}(\Omega_e) \Phi + \sum_{i \in I_{\Omega_e}^{\text{in},m}} V_{\Gamma_i}^m - \sum_{i \in I_{\Omega_e}^{\text{out},m}} V_{\Gamma_i}^m \right) / (\text{mes}(\Omega_e) \Phi), \quad (2)$$

где $\text{mes}(\Omega_e)$ – объем ячейки Ω_e ; $I_{\Omega_e}^{\text{out},m}$, $I_{\Omega_e}^{\text{in},m}$ – множества номеров граней элемента Ω_e , через которые m -я фаза вытекает из Ω_e и, соответственно, втекает в Ω_e .

Очевидно, что для корректности пересчета новых насыщенностей необходимо, чтобы вытекающий из ячейки объем фазы не превышал имеющийся в ячейке объем подвижной фазы. Таким образом, максимальный допустимый шаг $\Delta t_{m,\Omega_e}$ для каждой фазы в каждой ячейке можно вычислить следующим образом:

$$\Delta t_{m,\Omega_e}^{\text{max}} = \left((S_{\Omega_e}^m - S_{\Omega_e}^{m,\text{res}}) \text{mes}(\Omega_e) \Phi + \Delta V_{\Omega_e}^m \right) / \sum_{i \in I_{\Omega_e}^{\text{out},m}} |Q_{\Gamma_i}^m|, \quad (3)$$

где $S_{\Omega_e}^{m,\text{res}}$ – остаточная насыщенность фазы m .

Для ускорения расчета новых насыщенностей используется процедура группирования ячеек сетки и специальный алгоритм расчета перетоков фаз. Подробнее этот алгоритм описан в [4]. Суть процедуры группирования заключается в том, что все множество ячеек сетки разбивается на группы G_g ($G_g = \{\Omega_e, e \in I_g\}$, где g – номер группы), для которых выбирается шаг по времени $\Delta t_{G_g} = \Delta t^{\text{main}} / 2^{g-1}$, а номер группы для ячеек выбирается по следующему условию:

$$\min \left\{ g : \Delta t_{G_g} \leq \min_m \Delta t_{m, \Omega_e}^{\max} \right\}. \quad (4)$$

Стоит отметить, что в данном случае шаг Δt^{main} может быть выбран достаточно произвольно и довольно крупным (например, 30 суток), поскольку насыщенности во всех ячейках будут пересчитываться с шагом, не превышающим максимально возможный.

Поскольку описанная процедура группирования не учитывает взаимное расположение ячеек в сетке, то могут возникать ситуации, когда соседние ячейки будут находиться в группах, номера которых могут сильно отличаться. Например, если соседствующие ячейки Ω_e и Ω_k попали в первую и в шестую группу, соответственно, то это означает, что за один пересчет насыщенностей в ячейке Ω_e , насыщенности в ячейке Ω_k будут пересчитаны 32 раза.

Сглаживание в процедуре группирования накладывает ограничение на разницу номеров группы соседствующих ячеек. Самое сильное сглаживание будет в случае, если номер группы ячейки будет отличаться от номера группы ее соседей не более, чем на 1. Очевидно, что для выполнения сглаживания часть ячеек будет принудительно перемещено в группы с меньшими номерами, что приведет к увеличению количества пересчета в них насыщенностей фаз и, следовательно, может значительно увеличить вычислительные затраты.

Заметное влияние сглаживания в процедуре группирования можно увидеть при моделировании поздних стадий разработки месторождения, когда в области остается малое количество подвижной фазы нефти. Покажем это на модели реального месторождения (р. Татарстан), полученного в результате автоадаптации скважинных данных [5]. На рисунке 1 показано распределение насыщенности нефти через 40 лет после начала разработки.

На рисунке 2 показаны графики среднесуточного объема добываемой нефти на добывающей скважине, обозначенной стрелкой на рисунке 1, красным цветом – расчет без сглаживания, черным цветом – с выполнением сглаживания (разница групп на соседний ячейках не более 1). В таблице 1 приведено распределение количества конечных элементов по группам на первом временном шаге для расчетов со сглаживанием и без сглаживания. Время расчета без сглаживания составило 23 минуты, а расчета со сглаживанием – 81 минуту.

В заключении необходимо отметить, что при моделировании месторождения с достаточным количеством подвижной фазы нефти

процедура сглаживания не повышает качество аппроксимации многофазного потока, при значительном росте вычислительных затрат.

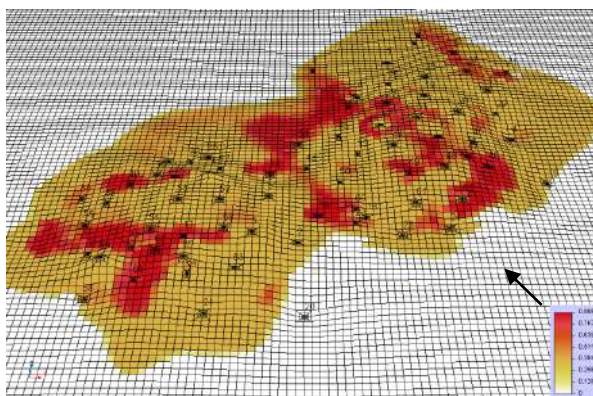


Рис. 1 – Распределение поля насыщенности нефти через 40 лет после начала разработки месторождения

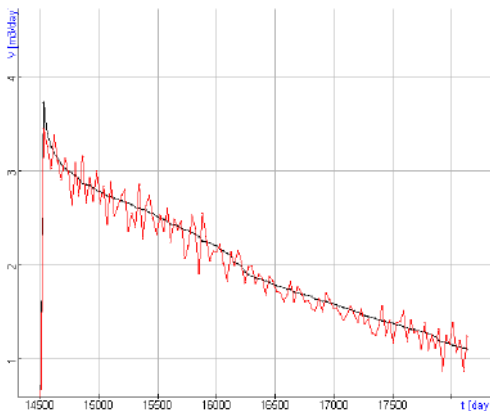


Рис. 2 – Среднесуточный дебит нефти на контрольной скважине (красный – без сглаживания, черный – со сглаживанием)

Таблица 1 – Распределение количества ячеек сетки по группам

№ группы	0	1	2	3	4	5	6	7	8	9	10
Без сглаживания	130095	4850	4196	2851	1859	1234	891	636	516	417	338
Со сглаживанием	75653	5496	6846	8464	10033	11170	11701	9980	6022	2180	338

Литература:

1. Doyle B., Riviere B., Sekachev M. A multinumerics scheme for incompressible two-phase flow // Comput. Methods Appl. Mech. Eng. Elsevier B.V., 2020. Vol. 370. P. 113213.
2. Соловейчик Ю. Г., Рояк М. Э., Персова М. Г. Метод конечных элементов для решения скалярных и векторных задач. — Новосибирск: Изд-во НГТУ, 2007. — 896 с.
3. Marina G. Persova, Yuri G. Soloveichik, Alexander M. Grif, Ilya I. Patrushev Flow balancing in FEM modelling of multi-phase flow in porous media // Actual problems of electronic instrument engineering (AREIE–2018): междунар. науч.-техн. конф., Новосибирск 2-6 окт. 2018 г.:– Новосибирск : Изд-во НГТУ, 2018. – С. 205-211.
4. Persova M.G., Soloveichik Yu.G., Patrushev I.I., Ovchinnikova A.S. Application of the Finite Element Grouping Procedure to improve the efficiency of unsteady multiphase flow simulation in high-heterogeneous 3D porous media // Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie vychislitel'naja tehnika i informatika, 2021, №57 [В печати]
5. Persova M.G. et al. The design of high-viscosity oil reservoir model based on the inverse problem solution // J. Pet. Sci. Eng. Elsevier B.V., 2021. Vol. 199. P. 108245.

ОПРЕДЕЛЕНИЕ ЭФФЕКТИВНОГО КОЭФФИЦИЕНТА ТЕПЛОПРОВОДНОСТИ ГЕТЕРОГЕННЫХ СРЕД

Е.Е. Патрушева

Новосибирский государственный технический университет,

г. Новосибирск, elizaveta.patrusheva@gmail.com

Научный руководитель: Иткина Н.Б., к.т.н., доцент

В рамках данной работы рассматривается определение эффективного коэффициента теплопроводности гетерогенной среды с контрастными микровключениями. Вычислительная схема строится на базе метода виртуальных элементов и верифицируется на классе модельных задач.

The article considers the determination of the effective thermal conductivity coefficient of a heterogeneous medium with contrast micro inclusions. The computational scheme is based on the virtual element method and verified on the model problems class.

Многие современные прикладные задачи связаны с материалами, имеющими гетерогенную структуру, характеризующуюся наличием разномасштабных включений сложной формы с различающимися физическими характеристиками. Теория гомогенизации позволяет использовать представленные гетерогенной среды как гомогенной с некоторым коэффициентом, который называется эффективным. Данная характеристика может быть использована при исследовании процесса теплопереноса в нативных средах, а также при разработке композитных материалов с заданными свойствами. Для определения эффективного коэффициента используются две группы методов: явные аналитические и численные.

Аналитические методы основаны на предположении о периодичности или квазипериодичности функции теплопроводности. Для сред, заполненных периодически или случайно расположенными сферическими включениями применимы методы определения эффективного коэффициента теплопроводности по формулам Максвелла, Рэлея, Бруггемана, Мередита и др. [1-3]. Стоит отметить, что в этих формулах используется относительно грубая модель среды, что влияет на точность вычисления эффективного коэффициента. Кроме того, значительным недостатком данных методов являются ограничения на форму и расположение микровключений.

Численная гомогенизация позволяет реализовать другой подход к решению задачи определения эффективного коэффициента. Можно выделить следующие классы методов численной гомогенизации:

1) методы, основанные на теории осреднения дифференциальных операторов [4]. Данные методы, в основном, используются для задач с равномерным распределением микровключений;

2) методы численной гомогенизации, разработанные на основе конечно-элементных или конечно-разностных дискретизаций дифференциальных операторов. Данные методы являются достаточно эффективными, однако для учета гетерогенной структуры с высокой концентрацией включений необходимо достаточно подробное разбиение расчетной области, что приводит к резкому росту размерности дискретного аналога;

3) многомасштабные методы [5]. Данные методы позволяют разработать вычислительную схему с высокой степенью параллелизма, что обеспечивает сокращение времени решения дискретной задачи без потери точности определения эффективного коэффициента [6].

Метод виртуальных элементов (VEM) [7] является модификацией многомасштабных методов. Общая идея данного метода состоит в следующем: расчетная область разбивается на виртуальные элементы

(макроэлементы), на каждом макроэлементе решается серия подзадач на конечных элементах для построения многомасштабных неполиномиальных функций формы, которые формируют базис глобального пространства. Функции формы достаточно точно отражают структуру расчетной области, что позволяет с высокой степенью точности учесть геометрические неоднородности области моделирования, не увеличивая при этом размерность дискретного аналога. Для обеспечения устойчивости постановки вводятся специальные проектирующие операторы и стабилизаторы. При этом макроэлементы могут иметь произвольную форму, что позволяет гибко работать с расчетной областью в зависимости от локализации включений.

Численные методы определения эффективного коэффициента теплопроводности основаны на решении обратных коэффициентных задач. Задача теплопроводности решается с учётом всех неоднородностей среды, а затем, при тех же начальных и граничных условиях, вычисляются такие параметры гомогенной среды, при которых могут быть получены те же потоки тепла, что и для гетерогенной среды. Поскольку обратная задача является некорректной, необходимо использовать специальные алгоритмы регуляризации.

В рамках данной работы рассматривается определение эффективного коэффициента гетерогенной среды с микровключениями на базе метода виртуальных элементов. Рассматривается стационарная задача теплопроводности в области $\Omega \subset R^2$ с границей $\partial\Omega$

$$-\nabla \cdot (\lambda \nabla u) = f \text{ в } \Omega, \quad (1)$$

$$u = g \text{ на } \partial\Omega, \quad (2)$$

где λ – теплопроводность (Вт/м·К).

Необходимо определить эффективный коэффициент теплопроводности неоднородной среды.

Вариационная формулировка VEM для задачи теплопроводности. Вариационная постановка задачи (1)-(2) в форме Галеркина формулируется следующим образом: найти $u \in H^1(\Omega)$ такую, что

$$A(u, v) := \int_{\Omega} \lambda \nabla u \cdot \nabla v d\Omega = \int_{\Omega} f v d\Omega := (f, v) \quad \forall v \in H_0^1(\Omega). \quad (3)$$

Введем разбиение расчетной области на виртуальные элементы P и определим пространство $V_h \subset H^1(\Omega)$. Получим аппроксимацию постановки (3) на V_h : найти $u_h \in V_h \subset H^1(\Omega)$ такую, что

$$A_h(u_h, v_h) := \int_{\Omega} \lambda \nabla u_h \cdot \nabla v_h dV = \int_{\Omega} f v_h dV := (f, v_h) \quad \forall v_h \in V_{h,0}.$$

Для удовлетворения условий теоремы о сходимости представим $A_{h,P}$ в виде суммы двух компонент [7]:

$$A_{h,P} = A_{h,P}^0 + A_{h,P}^1,$$

$$A_{h,P}^0 = (\Pi_m^{\nabla} u_h, v_h) = A_P(\Pi_m^{\nabla} u_h, \Pi_m^{\nabla} v_h),$$

$$A_{h,P}^1 = ((I - \Pi_m^{\nabla}) u_h, v_h) = S_P((I - \Pi_m^{\nabla}) u_h, (I - \Pi_m^{\nabla}) v_h),$$

где $\Pi_m^{\nabla} : V_{h,P} \rightarrow P_1(P)$, $\nabla(\Pi_m^{\nabla} v_h) = \overline{\nabla v_h} = \frac{1}{|P|} \int_P \nabla v_h dV$,

$S_P(u_h, v_h) = \mu \int_{\partial P} [|u_h|] \times [|v_h|] ds^0$, $[|q|] = q^R \overline{n_1} - q^L \overline{n_1}$, $\overline{n_k}$ – внешняя нормаль к границе k -го элемента.

Для построения функций формы явно зададим их на ∂P , предполагая, что их след является линейным полиномом. Внутри P зададим функции как решение уравнения в частных производных.

Вычисление эффективного коэффициента теплопроводности. Для определения параметров модели теплопроводности в однородной среде сформулируем обратную задачу [8]

$$\mathbf{A}\Theta = \mathbf{F}, \tag{4}$$

где $\mathbf{A} : X \rightarrow Y$ – оператор обратной задачи, $\Theta \in X$ – вектор неизвестных параметров, в нашем случае $\Theta = \{\lambda\}$, $\mathbf{F} \in Y$ – модель наблюдения, X и Y – гильбертовы пространства.

Получить приближённое решение задачи (4) можно путём минимизации сглаживающего функционала А.Н. Тихонова [8]

$$J_{\alpha}(\Theta) = \|\mathbf{A}\Theta - \mathbf{F}\|^2 + \alpha \|\Theta\|^2, \tag{5}$$

где α – параметр регуляризации, согласованный с величиной ошибки [9].

Измеряемыми данными является температура u в исследуемой гетерогенной среде. Оператор \mathbf{A} является оператором решения прямой задачи (1)–(2). Поскольку VEM позволяет определить решение прямой задачи в пространстве $H^1(\Omega)$, то $Y = H^1(\Omega)$.

Для минимизации функционала (5) используется метод сопряжённых градиентов Флетчера-Ривса [10]. Градиентные итерационные методы чувствительны к выбору начального приближения $\Theta^{(0)}$, и в окрестности точного решения может наблюдаться осцилляция градиента функционала, поэтому предлагается использовать метод Ньютона для локальной минимизации функционала с начальным приближением метода Флетчера-Ривса по достижению им критерия останова [10].

Программно-алгоритмическая реализация была верифицирована на классе модельных задач. Были проведены вычислительные эксперименты по исследованию влияния размеров, формы, контрастности и расположения микровключений на величину эффективного коэффициента теплопроводности, а также сравнение результатов применения аналитических формул и результатов численного моделирования.

Литература:

1. Rayleigh Lord. On the Influence of Obstacles Arranged in Rectangular Order upon the Properties of a Medium // *Phil. Mag.* – 1892. – Vol. 34, iss. 211. – P. 481–502.
2. Bruggeman D.A.G. Dielektrizitätskonstanten und Leitfähigkeiten der Mischkorperausotropen Substanzen // *Ann. Phys.* – 1935. – Vol. 416, iss. 7. – P. 636–664.
3. Meredith R.E., Tobias C.W. Conductivity of Emulsions // *J. Electroch. Soc.* – 1961. – Vol. 103. – P. 286–290.
4. Жиков В.В., Козлов С.М., Олейник О.А. Усреднение дифференциальных операторов. – М.: Издательская фирма "Физико-математическая литература", 1993. – 464 с.
5. Efendiev Y.R., Hou T.Y. Multiscale finite element methods. Theory and application. – Springer. 2009. – 239 p.
6. Эпов М.И., Шурина Э.П., Артемьев М.К. Численная гомогенизация многомасштабных гетерогенных сред // *Геофизический журнал НАН Украины.* – 2012. – Т. 34, №4. – С. 16-21.
7. Basic principles of Virtual Element Methods / Beirão da Veiga L., Brezzi F., Cangiani A., Manzini G., Marini L.D., Russo A. // *Mathematical Models And Methods in Applied Sciences.* – 2013. – Vol. 23, iss. 1. – P. 199–214.

8. Тихонов А.Н. О некорректных задачах линейной алгебры и устойчивом методе их решения // ДАН СССР. – 1965. – Т. 163, № 3. – С. 591-594.

9. Тихонов А.Н., Гончарский А.В., Степанов В.В., Ягола А.Г. Численные методы решения некорректных задач. – М.:Наука, 1990. – 232 с.

10. Бойко О.А., Зеркаль С.М., Иткина Н.Б. Применение методов планирования эксперимента при решении обратных коэффициентных задач теплопереноса. – Новосибирск, 2004. – 20с. – (Препринт / РАН Сиб. Отд. Ин-т математики, №125).

ОЦЕНКА ЭФФЕКТИВНОСТИ ОПЕРЕЖАЮЩЕГО ЗАКИСЛЕНИЯ РУДНОЙ ЗАЛЕЖИ ПРИ ОТРАБОТКЕ БЛОКА МЕТОДОМ СКВАЖИННОГО ПОДЗЕМНОГО ВЫЩЕЛАЧИВАНИЯ

К.Е. Попова

**Северский технологический институт Национального
исследовательского ядерного университета «МИФИ», г. Северск,
popova.k5enia@yandex.ru**

**Научный руководитель: Носков М.Д., доктор физ.-мат. наук,
профессор**

Рассмотрено применение опережающего закисления рудной залежи для повышения эффективности добычи урана методом скважинного подземного выщелачивания. Представлены результаты математического моделирования извлечения урана при рядной схеме вскрытия эксплуатационного блока. Проведен сравнительный анализ геотехнологических показателей отработки блока при различных режимах закисления.

The application of advanced acidification of an ore deposit to increase the efficiency of uranium mining by in-situ leaching was considered. The results of mathematical modeling of uranium extraction with a row-by-row scheme of the operational unit opening are presented. A comparative analysis of geotechnological indicators of the operational unit mining under different acidification modes has been carried out.

В настоящее время широко применяемым способом добычи урана является скважинное подземное выщелачивание (СПВ), которое

характеризуется технологичностью, экономичностью и экологичностью процесса по отношению к другим способам. СПВ - метод разработки месторождений полезных ископаемых, при котором происходит воздействие на залежь на месте ее залегания с целью перевода полезных компонентов в раствор и последующее их извлечения. Данный способ эффективен при разработке бедных и глубокозалегających месторождений, которые характеризуются сложными гидрогеологическими и горно-технологическими условиями [1].

Разрабатываемые в России способом скважинного подземного выщелачивания месторождения урана характеризуются сложными горно-геологическими условиями из-за чего не всегда удается добиться наивысших показателей отработки эксплуатационного блока. С целью повышения эффективности добычи урана применяют различные схемы вскрытия блока: изменение расположения и количества скважин, режимов откачных и нагнетательных скважин или изменение концентрации кислоты и окислителя в выщелачивающих растворах и т.д. Одним из вариантов повышения эффективности является опережающее закисление рудной залежи перед началом отработки эксплуатационного блока. Положительный эффект в этом случае достигается благодаря более быстрому изменению геохимической обстановки в контуре эксплуатационного блока и растворению минералов урана.

Исследования проводились методом математического моделирования процесса СПВ с использованием специализированного программного обеспечения «Курс», разработанного в СТИ НИЯУ МИФИ [2, 3]. Программа «Курс» позволяет создавать цифровые модели эксплуатационных блоков, включающие в себя модели геологической среды и технологических объектов, а также проводить имитационное моделирование процесса выщелачивания урана. Работа программы основана на математической модели, описывающей физико-химические процессы, происходящие в продуктивном горизонте при сернокислотном выщелачивании урана. Геотехнологические расчеты проводятся с учетом гидрологических и геологических особенностей строения продуктивного горизонта, режимов работы технологических скважин и составов нагнетаемых растворов.

В данной работе рассматривалась отработка эксплуатационного блока рядной схемой со следующими параметрами: расстояние между нагнетательными скважинами 30 м, расстояние между откачными скважинами 30 м, расстояние между рядами откачных и закачных скважин 40 м. Геотехнологические параметры эксплуатационного блока: ГРМ 609,6 (тыс.т), эффективная мощность 11,5 (м), площадь 28,8 (тыс.м²), средняя исходная продуктивность 9,2 кг/м², средний

метропроцент 0,5 (м*%). Начальное распределение продуктивности и расположение скважин в эксплуатационном блоке представлены на рисунке 1.

В ходе исследования рассматривалась эффективность опережающего закисления рудной залежи при отработке блока методом скважинного подземного выщелачивания. В первом варианте отработка блока проводилась без опережающего закисления. В последующих расчетах закисление проводилось с различной продолжительностью - от 1 до 4 месяцев. На стадии опережающего закисления нагнетательные скважины не работали, а все откачные работали в нагнетательном режиме. Дебит нагнетания - 1 м³/ч. На стадии активного выщелачивания к работе подключались нагнетательные скважины, а откачные переключались в свой основной режим. Дебиты нагнетательных и откачных скважин распределялись равномерно, средний дебит откачной скважины составлял 2 м³/ч, а в целом по блоку 24 м³/ч. Во всех вариантах отработки средняя концентрация кислоты в выщелачивающих растворах составляла 12,5 г/л.

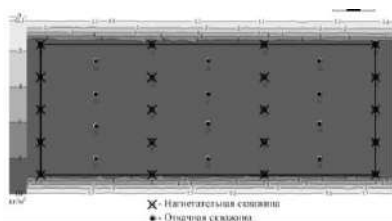


Рис.1 – Распределение продуктивности и расположение скважин в эксплуатационном блоке

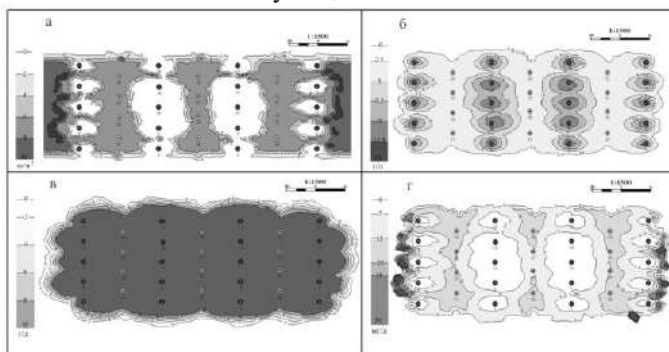


Рис.2 – Карты на момент 80% извлечения урана: распределение продуктивности (а), серной кислоты (б), SO₄²⁻ (в), урана (г) в продуктивном горизонте при отработке эксплуатационного блока без опережающего закисления

Моделирование осуществлялось до момента достижения 80% извлечения урана из контура эксплуатационного блока. Результаты референтного расчета (без опережающего закисления) представлены на рисунке 2.

Геотехнологические показатели отработки рудного тела представлены в таблице 1.

Таблица 1 – Геотехнологические показатели отработки при извлечении 80% урана из эксплуатационного блока

Показатель	Референтный расчёт	Время опережающего закисления, мес.			
		1	2	3	4
Время отработки, лет	12	12,4	12	11,6	11,4
Ж/Т, м ³ /т	4,1	4,3	4,1	3,9	3,8
Расход кислоты, тыс.т	23,7	24,4	23,5	22,6	22,2
Кислотоемкость, кг/т	38,9	40,1	38,6	37,1	36,5
Удельный расход кислоты, кг/кг	111,8	115,2	110,8	106,6	104,8
Средняя концентрация U в ПР, мг/л	84,1	81,7	85,2	89,0	90,9
Средняя концентрация кислоты в ПР, г/л	3,1	3,1	3,1	3,2	3,2

На основе данных, приведенных в таблице 1, можно сделать вывод, что опережающее закисление рудной залежи приводит к сокращению времени отработки эксплуатационного блока, уменьшению отношения Ж/Т (масса рабочего раствора, приходящегося на единицу выщелачиваемой горнорудной массы) и снижению удельного расхода кислоты.

Таким образом, результаты математического моделирования показывают, что применение опережающего закисления рудной залежи при отработке блока методом скважинного подземного выщелачивания приводит к повышению эффективности добычи урана.

Литература:

1. Акимова И.Д., Бабкин А.С., Воронцова О.М. и др. Геотехнология урана (русский опыт): монография / Под ред. И.Н. Солодова, Е.Н. Камнева. – М.: «КДУ», «Университетская книга», 2017. – 576 с.

2. Программный комплекс для управления разработкой месторождения полезных ископаемых методом скважинного подземного выщелачивания / А.Д. Истомин, М.Д. Носков, А.Г. Кеслер, С.Н. Носкова, А.А. Чеглоков // Горный информационно-аналитический бюллетень. – 2011. - № 8. – С. 376–381.

3. Применение математического моделирования для решения геотехнологических и экологических задач при добыче урана способом подземного выщелачивания / М.Д. Носков, М.В. Гуцул, А.Д. Истомин, А.Г. Кеслер, С.Н. Носкова // Горный информационно-аналитический бюллетень. – 2012. - № 7. – С. 361–366.

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ГОРЕНИЯ УГЛЕ-МЕТАНО-ВОЗДУШНОЙ СМЕСИ В ДВУХМЕРНОЙ ПОСТАНОВКЕ

Е.И. Рожкова

Научный руководитель: Моисеева К. М., к. ф.-м. н.

Томский государственный университет,

г.Томск, katushar2801@mail.ru

Работа посвящена численному решению задачи горения метано-воздушной смеси с частицами угольной пыли. Математическая постановка основывалась на подходах динамики двухфазных реагирующих сред. Решение задачи выполнено с использованием метода Ван Лира. Из решения задачи определялась видимая скорость горения в зависимости от радиуса частиц и массовой концентрации угольной пыли.

The work is devoted to the numerical solution of the problem of combustion of a methane-air mixture with coal dust particles. The mathematical formulation was based on the approaches of the dynamics of two-phase reacting media. The problem was solved using the Van Leer method. From the solution of the problem, the apparent burning rate was determined depending on the radius of the particles and the mass concentration of coal dust.

Воспламенение и горение газовых смесей с низким содержанием горючего является одной из актуальных проблем физики горения и взрыва. Метан и угольная пыль в смеси с воздухом могут образовывать взрывчатые системы. Так, воспламенение метана и угольной пыли в

горном производстве становится причиной аварий, которые в большинстве случаев носят характер катастроф [1].

Целью настоящего исследования является исследование влияния угольной пыли на видимую скорость горения угле-метано-воздушной смеси.

Рассмотрен плоский канал с источником зажигания. Канал заполнен метано-воздушной смесью, с равномерно распределёнными частицы угольной пыли. Содержание метана в смеси составляет 7%, массовая концентрация и радиус частиц угольной пыли варьировались. Между горючим и кислородом протекает химическая реакция, на поверхности частиц протекает гетерогенная реакция первого порядка по кислороду.

Математическая постановка задачи описывается уравнениями сохранения массы, импульса, записываемого для компонент скорости, и энергии для газа, уравнениями сохранения импульса и энергии для частиц, уравнением баланса массы частиц, уравнениями баланса горючего компонента и окислителя в смеси, уравнением счетной концентрации частиц и уравнением состояния газа. На правой границе реализован выход в атмосферные условия, сверху- свободное вытекание.

Численно задача решалась с помощью программы на языке Fortran, основанной на методе Ван Лира. Шаг по пространству составил 10^{-5} м, шаг по времени определялся из условия устойчивости Куранта [2]. Результаты были получены для четверти канала с размерами 0,05 м по осям x и y , источник зажигания находился в левой нижней части канала. Проверка методики решения системы уравнений и программы ЭВМ на достоверность численного решения осуществлялось путем проверки законов сохранения.

В работе показано влияние размеров частиц и массы угольной пыли на видимую скорость горения и выгорание частиц. Изменение массы угольной пыли в пределах 0,05 г слабо влияет на видимую скорость горения. При увеличении радиуса частиц, видимая скорость горения уменьшается, на рисунках 1-2 показаны распределения температур при разных радиусах частиц угольной пыли. В случае, когда радиус частиц составляет 1 мкм, видимая скорость горения примерно 3,95 м/с, а при увеличении радиуса частиц вдвое видимая скорость горения уменьшилась до 3,79 м/с.

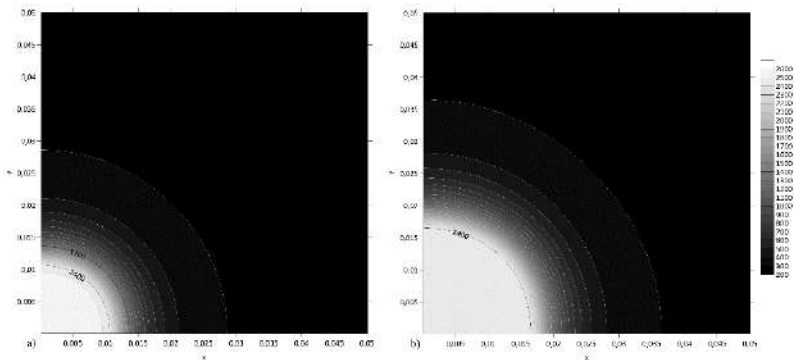


Рис. 1- Распределение температуры в моменты времени а) $0.3 \cdot 10^{-2} \text{ с}$, б) $0.45 \cdot 10^{-2} \text{ с}$ при радиусе частиц 1 мкм , массе угольной пыли 0.1 г .

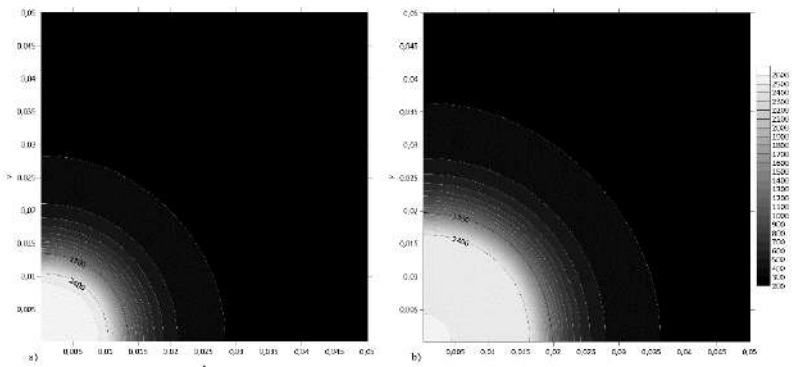


Рис. 2- Распределение температуры в моменты времени а) $0.3 \cdot 10^{-2} \text{ с}$, б) $0.45 \cdot 10^{-2} \text{ с}$ при радиусе частиц 2 мкм , массе угольной пыли 0.1 г .

При увеличении радиуса частицы быстрее выгорают (рис. 3), а с увеличением массы угольной пыли частицы выгорают ближе к источнику зажигания (рис. 4).

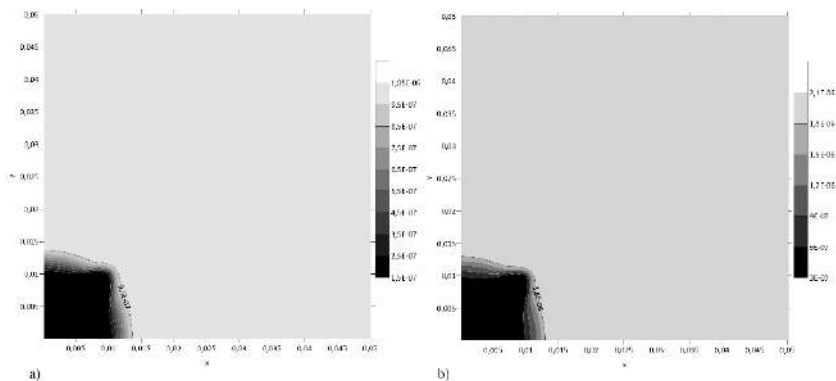


Рис. 3- Радиус частиц в момент времени $0.3 \cdot 10^{-2}$ с, масса угольной пыли 0,05 г, начальный радиус частиц а) 1 мкм, б) 2 мкм

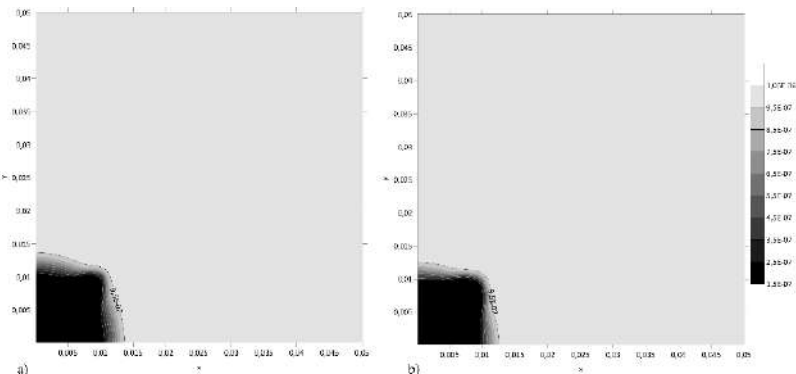


Рис. 4- Радиус частиц в момент времени $0.3 \cdot 10^{-2}$ с, начальный радиус частиц 1 мкм, масса угольной пыли а) 0,05 г, б) 0,1 г

Выводы.

Разработана физико-математическая модель горения метано-воздушной смеси с примесью частиц угольной пыли в канале с источником зажигания. Получено, что при изменении массы угольной пыли в пределах 0,05 г почти не влияет на видимую скорость горения, при увеличении радиуса частиц, видимая скорость горения уменьшается. Также показана зависимость скорости выгорания частиц от массы угольной пыли и радиуса частиц.

Работа выполнена при поддержке Российского научного фонда (грант № 21-71-10034).

Литература:

1. Шевцов Н.Р. Взрывозащита горных выработок, Донецк. 2002.
2. Годунов С.К., Забродин А.В., Иванов М.Я., Крайко А.Н., Прокопов Г.П. Численное решение многомерных задач газовой динамики, Москва: Наука, 1976.
3. Миньков Л.Л., Шрагер Э.Р. Основные подходы к численному решению одномерных уравнений газовой динамики// Учебное пособие / Томск, Издательство Общество с ограниченной ответственностью "СТТ". 2016. 136 с.

ИССЛЕДОВАНИЕ ОДНОКОМПОНЕНТНОЙ СИСТЕМЫ ЛЕННАРД-ДЖОНСА ПРИ НИЗКИХ СКОРОСТЯХ ОХЛАЖДЕНИЯ МЕТОДАМИ МОЛЕКУЛЯРНОЙ ДИНАМИКИ И РЕГРЕССИОННОГО АНАЛИЗА

А.Р. Фархутдинов

**Казанский (Приволжский) федеральный университет,
г. Казань, farhutdinovalbert@yandex.ru**

Научный руководитель: Мокшин А.В., д.ф.-м.н., профессор

Разработаны алгоритмы получения функций радиального распределения частиц в случае охлаждения системы Леннард-Джонса при низких скоростях охлаждения методами регрессионного анализа.

Radial distribution function calculation algorithms in the case of the Lennard-Jones system cooling at low rates are developed using regression analysis methods.

Скорость охлаждения γ , используемая при моделировании процессов охлаждения системы классическими методами молекулярной динамики, обычно составляет $\sim 10^{12}$ К/с. Однако, максимальные экспериментально измеренные скорости охлаждения γ не превышают значений $\sim 10^8$ К/с. Таким образом, классические методы оказываются неэффективными при моделировании системы в условиях эксперимента. Целью данной работы является разработка методики, позволяющей предсказать характеристики системы в будущие моменты времени на основе

результатов относительно небольшого числа итераций моделирования процесса охлаждения при условиях, близких к реальным.

Рассматривается система из $N = 1372$ частиц, взаимодействие между которыми описывается потенциалом Леннард-Джонса [1]:

$$U(r_{ij}) = 4\varepsilon \left[\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^6 \right] \quad (1)$$

Здесь r_{ij} – расстояние между частицами i и j . Температура системы меняется в диапазоне от $1.3\varepsilon/k_B$ до $10^{-4}\varepsilon/k_B$, где k_B – постоянная Больцмана. Плотность системы $\rho \approx 0.75\sigma^{-3}$ поддерживается постоянной. Интегрирование уравнений движения осуществляется с помощью скоростного алгоритма Верле [2] с временным шагом $\Delta t = 0.005\tau = 0.005\sigma(m/\varepsilon)^{1/2}$, где m – масса частицы. Скорость охлаждения γ принимает значение $\gamma = 4 \cdot 10^{-4} \varepsilon/k_B\tau \sim 10^{10}$ К/с.

Для некоторых температур из указанного диапазона вычислялась функция радиального распределения частиц [3]:

$$g(\vec{r}) = \frac{2V}{N^2} \left\langle \sum_{i=1}^{N-1} \sum_{j>i}^N \delta(\vec{r} - \vec{r}_{ij}) \right\rangle \quad (2)$$

Полученные результаты использовались в качестве тренировочных данных при обучении различных регрессионных моделей.

Обученная модель позволяет получить функцию радиального распределения частиц для заданной температуры, не прибегая к моделированию классическими методами молекулярной динамики.

Работа была поддержана грантом Фонда развития теоретической физики и математики «БАЗИС» (проект № 20-1-2-38).

Литература:

1. Jones J.E. On the Determination of Molecular Fields. I. From the Variation of the Viscosity of a Gas with Temperature // Proc. R. Soc. Lond. A. – 1924. – Vol. 106. – P. 441-462. – doi: 10.1098/rspa.1924.0081.

2. A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters / W.C. Swope, H.C. Andersen, P.H. Berens, K.R. Wilson // J. Chem. Phys. – 1982. – V. 76, iss. 1 – P. 637-649. – doi: 10.1063/1.442716.

3. Rapaport D.C. The Art of Molecular Dynamics Simulation – Cambridge: Cambridge University Press, 2004. – 549 p.

3D МОДЕЛИРОВАНИЕ ГИДРОДИНАМИЧЕСКИХ ПОТОКОВ В СТРУКТУРНОЙ МОДЕЛИ ДВОЙНОЙ ПОРИСТОСТИ

Н.Б. Фаткуллина, О.А. Солнышкина, А.З. Булатова
Центр Микро- и наномасштабной динамики дисперсных систем,
Башкирский государственный университет, г. Уфа,
nazgulbay1999@gmail.com
Научный руководитель: Солнышкина О.А., к.ф.-м.н.

Рассматривается задача о течении вязкой несжимаемой жидкости в микроканале со сложной геометрией, представляющей собой структурную модель двойной пористости. Проведено 3D моделирование гидродинамических потоков с использованием ускоренного метода граничных элементов, а также исследовано влияние геометрических параметров канала на картину течения флюида. Полученные результаты могут быть использованы для верификации моделей, описывающих течения в средах с двойной пористостью.

The problem of viscous incompressible fluid flow in a microchannel with complex geometry, which is a structural model of double porosity, is considered. 3D simulation of hydrodynamic flows using the accelerated boundary element method has been carried out, and the influence of the geometric parameters of the channel on the flow pattern has been investigated. The obtained results can be used to verify the models describing flows in media with double porosity.

На данный момент более 40% мировых запасов нефти связано с карбонатными коллекторами. Особенностью таких коллекторов является зональная неоднородность по фильтрационно-емкостным и упруго-механическим свойствам. Модель двойной пористости, рассматриваемая в данной работе, учитывает неоднородности таких типов. В данной модели предполагается, что пласт состоит из двух компонентов: матрицы, являющейся пористой средой, и твердых включений. Твердые включения, рассматриваемые в настоящей модели, имеют существенную пористость, но аномально низкую проницаемость [1]. На течение жидкости в пористой среде (например, фильтрации флюида в трещиновато-пористой породе) влияет большое количество параметров, в связи с этим изучение процессов в таких средах является сложной задачей, требующей детального изучения гидродинамических потоков в микромасштабе. Это связано с тем, что макроэффекты, возникающие при фильтрации нефти в пласте, во многом определяются особенностями динамики жидкости на микроуровне в поровых каналах. Моделирование фильтрации флюидов в пласте имеет практическое значение для

разработки технологий увеличения нефтеотдачи. Также проведение численного моделирования данных процессов целесообразно с экономической точки зрения. Использование таких моделей позволяет уменьшить число лабораторных и полевых экспериментов, оптимизировать существующие механизмы и стратегии добычи, а также оценить возможные риски.

Целью данной работы является применение цифрового модуля для численного исследования особенностей гидродинамических потоков в микромасштабной модели двойной пористости. Одним из важнейших факторов, который следует учитывать при моделировании течения жидкости в пористых средах, является геометрия области, в которой происходит фильтрация. В литературе применяются два основных подхода к описанию пористой среды – капиллярная и структурная. Капиллярная представлена в виде сети микроканалов, а структурная состоит из множества микроразмерных твердых включений (например, недеформируемых столбиков). В данном исследовании рассматривается последняя модель пористой среды.

Рассматриваются гидродинамические потоки вокруг недеформируемых элементов внутри плоского закрытого микроканала, в совокупности представляющих собой модель двойной пористости. Такие потоки возникают при медленном течении вязкой несжимаемой жидкости под действием заданного постоянного перепада давления при малых числах Рейнольдса. Для описания таких течений используются стационарные уравнения Стокса. На поверхности недеформируемых элементов микроканала задается условие прилипания. Расчетная область представлена в виде массива твердых цилиндрических элементов, расположенных в плоском закрытом прямоугольном микроканале поперек потока. Для формирования второго масштаба пористости при равном количестве столбиков менялось их пространственное распределение – выделялись группы по 9 элементов, варьировалось расстояние между этими группами. Таким образом, общая пористость микроканала оставалась неизменной и равнялась 0.8.

Задача решалась с использованием ускоренного метода граничных элементов. Данный метод применим к задачам в областях со сложной геометрией, а также к трехмерным задачам в бесконечных областях. Поверхности канала и столбиков покрывались треугольной сеткой. Поскольку поставленная задача обладает высокой вычислительной сложностью, стандартный метод граничных элементов был ускорен с помощью усовершенствованного масштабируемого алгоритма (быстрый метод мультиполей) и с использованием гетерогенной вычислительной архитектуры (многоядерные процессоры и графические процессоры). Это позволяет осуществлять непосредственное крупномасштабное

трехмерное моделирование на персональной рабочей станции. Используемый численный подход более подробно описан в работе [2].

На микроуровне проведено численное исследование особенностей гидродинамических потоков в микроканалах с недеформируемыми цилиндрическими элементами, моделирующими структурные особенности пористых сред. Получены картины течения для различных конфигураций расположения цилиндрических элементов: для различных соотношений ширины поровых каналов k от 1 до 2.5. Получены картины течения, гистограммы распределения скорости, а также продольная и поперечная компоненты скорости течения жидкости в канале с двумя масштабами пористости. Численное изучение течения жидкости в модели двойной пористости показало, что при увеличении соотношения гидравлического диаметра широких поровых каналов k узким распределение скоростей меняется, и выделяются два характерных максимума скорости.

В рамках данной работы проведено исследование влияния параметров микромоделей двойной пористости на структуру потока жидкости на основе эффективного численного подхода. Полученные результаты могут быть использованы для верификации моделей, описывающих течения в средах с двойной пористостью. Подобные модели используются для описания фильтрации в горных породах, технологических процессов изготовления волокнистых композитов и других прикладных задач. Кроме того, изучение течений в подобных структурах является важной задачей, связанной с повышением эффективности теплообменников для микроэлектронных устройств. Применяемый численный подход с некоторыми модификациями может быть применен при объяснении физических эффектов, возникающих при фильтрации многофазных сред в моделях пористой среды, таких как закупоривание поровых каналов и увеличение подвижности дисперсных систем в сложных пористых структурах.

Исследование проведено при финансовой поддержке гранта РНФ 21-79-10212.

Литература:

1. Лобковский Л. И., Рамазанов М. М. К теории фильтрации в среде с двойной пористостью // Доклады Академии наук. – 2019. – Т. 484. – №. 3. – С. 348-351.
2. Abramova (Solnyshkina) O.A., Pityuk Y.A., Gumerov N.A., Akhatov I.S. Three-dimensional simulation of Stokes flow around a rigid structure using FMM/GPU accelerated BEM // Communications in Computer and Information Science. – 2019. – Vol. 965. – P. 427-438.

МОДЕЛИРОВАНИЕ ТЕХНОЛОГИЧЕСКИХ ПАРАМЕТРОВ ИНДУКЦИОННОЙ ПОВЕРХНОСТНОЙ ЗАКАЛКИ С УЧЕТОМ ОБОБЩЕНИЙ ЭКСПЕРИМЕНТАЛЬНЫХ ИССЛЕДОВАНИЙ

Ю.В. Шанин, А.А. Хоршев, А.С. Бондарь
Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина),
г. Санкт-Петербург, yurashanin@yandex.ru
Научный руководитель: Чмиленко Ф.В., д.т.н., профессор

В статье приводятся результаты моделирования технологии индукционной поверхностной закалки низколегированных углеродистых сталей. Для исследования технологических параметров обобщены экспериментальные данные о процессах аустенизации и гомогенизации. В результате численного моделирования рассчитаны режимы индукционной поверхностной закалки для углеродистой стали Ст45.

The article presents the results of modeling the technology of induction surface hardening of low-alloy carbon steels. To study the technological parameters, experimental data on the processes of austenization and homogenization are summarized. As a result of numerical simulation, induction surface hardening modes for carbon steel St45 are calculated.

В работе предлагается усовершенствовать подход А.Е. Слухоцкого к расчету мощности для поверхностной индукционной закалки, предложенный в [1, 2]. Для этого электротепловой расчет выполняется численными методами, а также учитывается влияние скоростного нагрева на изменение температур аустенизации и гомогенизации [3].

Система дифференциальных уравнений электромагнитного поля и теплопроводности в одномерной постановке для нагрева в продольном магнитном поле имеет вид [4]:

$$\text{Модель 1: } \frac{1}{x'} \frac{\partial}{\partial x} \left(\rho x' \frac{\partial H}{\partial x} \right) = j \omega \mu_0 \dot{H} \quad (1)$$

$$\text{Модель 2: } \frac{1}{x'} \frac{\partial}{\partial x} \left(\rho x' \frac{\partial H}{\partial x} \right) = \mu \mu_0 \frac{\partial H}{\partial t} \quad (2)$$

$$\gamma C \frac{\partial T}{\partial t} - \frac{1}{x'} \frac{\partial}{\partial x} \left(\lambda x' \frac{\partial T}{\partial x} \right) = q \quad (3)$$

где ω – круговая частота; q – удельная объемная мощность внутренних источников теплоты; l – показатель геометрии системы, $l=0$ – плоская, $l=1$ – цилиндрическая; x – пространственная координата, $x \in [x_n; x_e]$, остальные обозначения общепринятые.

При сильных нелинейностях при нагреве стали до точки Кюри, следует использовать решение во временной области (формула (2)).

В ряде работ, опубликованных Н.В. Зиминим и Ф.В. Безменовым [3], [5], [6] представлены результаты экспериментальных исследований по определению температур аустенизации и гомогенизации.

В настоящей работе была проведена экспертная оценка экспериментальных данных [5], которым были присвоены весовые коэффициенты, а затем построены формулы квадратичной регрессии для низколегированных сталей с содержанием углерода от 0,2 до 0,8 % при скорости нагрева от 0,1 до 1000 °C/сек для размера зерна исходной структуры от 1 до 100 мкм (табл. 1).

Общая формула нелинейной регрессии имеет следующий вид:

$$T = a + b \cdot x + c \cdot y + d \cdot z + e \cdot x \cdot y + f \cdot x \cdot z + g \cdot y \cdot z + h \cdot x^2 + i \cdot y^2 + j \cdot z^2 + k \cdot x^2 \cdot y^2 + l \cdot x^2 \cdot z^2 + m \cdot y^2 \cdot z^2 \quad (4)$$

где T – температура аустенизации или гомогенизации, °C; x – процентное содержание углерода в стали, %; y – десятичный логарифм размера зерна, мкм; z – десятичный логарифм скорости нагрева, °C/с; $a, b, c, d, e, f, g, h, i, j, k, l, m$ – коэффициенты нелинейной регрессии.

Для определения предельной допустимой твердости произведена аппроксимация экспериментальных данных о росте зерна аустенита углеродистых сталей от температуры (от 800 до 1100 °C) и скорости нагрева (0,1 до 1000 °C/сек) [6]:

$$G = a + b \cdot T + c \cdot v \quad (5)$$

где G – десятичный логарифм размера зерна, мкм; T – температура аустенизации, °C; v – десятичный логарифм скорости нагрева, °C/с; $a = -1.64862$, $b = 0.003$, $c = -0.17618$ – коэффициенты регрессии.

После достижения температуры аустенизации дальнейший рост зерна предполагается из соотношения 7% на каждые 10 градусов нагрева [6].

Твердость после охлаждения для стали марки Ст45 рассчитывается на основе экспериментальных данных [7] следующим образом:

$$\text{При } D \in [0; 20]: H(D, y) = -0.2 \cdot D + 3 \cdot y + 58 \quad (6)$$

$$\text{При } D \in [20; 40]: H(D, y) = -0.05 \cdot D + 3 \cdot y + 55 \quad (7)$$

где H – твердость по Роквеллу (HRC); D – размер зерна аустенита, мкм; y – десятичный логарифм скорости нагрева; °C/с.

Таблица 1 – Коэффициенты нелинейной регрессии

Коэффициенты	Температура, T	
	Аустенизация	Гомогенизация
A	923.005	925.424
B	-269.067	-265.810
C	17.148	19.427
D	4.337	2.113
E	-7.506	0.437
F	8.030	11.000
G	11.177	19.851
H	50.445	43.544
I	2.841	8.229
J	2.076	1.967
K	-4.208	-13.491
L	-2.100	-2.644
M	-0.134	0.444

Авторами был разработан итерационный алгоритм управления численной электротепловой моделью, который на основании скорости нагрева поверхности детали находит температуру гомогенизации и подбирает необходимую удельную мощность. На глубине закалки, как и в работах [1, 2], температура в конце нагрева 750 °C. На рис. 1 приведены результаты расчетов с использованием разработанного алгоритма для закалки цилиндрических деталей.

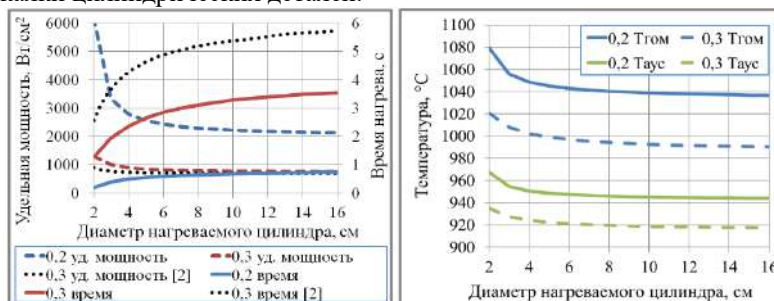


Рис. 1 – а): Результаты расчетов для закалки на частоте тока 10000 Гц стали Ст45 с размером исходного зерна 100 мкм на глубине закалки 0,2 и 0,3 см, б): Значения температур гомогенизации и аустенизации на поверхности детали для разных режимов нагрева

На рис. 1-б) приведена температура гомогенизации на поверхности детали, которая определяется на основании средней скорости нагрева поверхности после критической точки A_{c1} по формулам (4). Эта температура значительно выше значения $900\text{ }^{\circ}\text{C}$, которое использовалось в работах [1,2]. Таким образом, результаты исследования авторов отличаются от работы [2] не только в связи с использованием более точной модели, но и в связи с тем, что рассчитывались другие режимы индукционной поверхностной закалки, учитывающие фазовые превращения.

На рис. 2-а) представлены зависимости предельно достижимой твердости по Роквеллу на поверхности цилиндрической детали после индукционного нагрева на частотах 6000 и 10000 Гц при глубине закалки 0,2 и 0,3 см. Оптимальная частота тока 6000 Гц для глубины закалки 0,3 см позволяет получить большее значение твердости.

Рис. 2-б) с одной стороны показывает влияние исходной структуры на параметры термообработки, а с другой, демонстрирует влияние точности выбранной электротепловой модели на конечный результат.

В ходе выполнения работы была разработана комплексная численная модель для определения технологических параметров высокочастотной поверхностной закалки низколегированных углеродистых сталей, в которой учтено влияние температуры, скорости нагрева и исходного состояния структуры углеродистых сталей на процессы образования в них аустенита. Данная модель позволяет значительно повысить точность моделирования технологических параметров, что дает возможность более надежно гарантировать получение необходимой микроструктуры на требуемой глубине закалки.

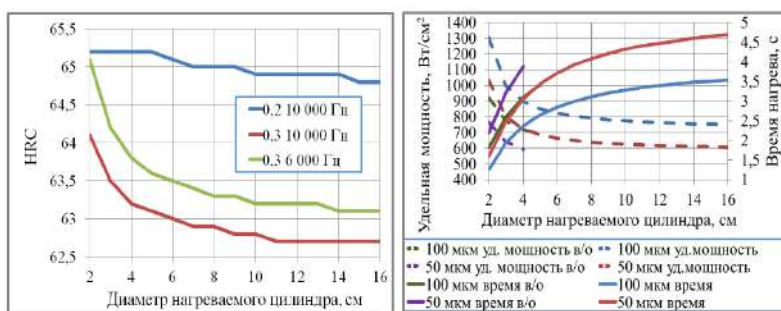


Рис. 2 – а): Зависимости предельно достижимой твердости по Роквеллу для глубины закалки 0,2 и 0,3 см, б): влияние исходной структуры и точности электротепловой модели на параметры термообработки

Литература:

1. *А.Е. Слухоцкий, С.Е. Рыскин.* Индукторы для индукционного нагрева. – Л.: Энергия, 1974. – 264 с.
2. *А.Е. Слухоцкий.* Индукторы. Под ред. А. Н. Шамова — 5-е изд., перераб. и доп. – Л.: Машиностроение. Ленингр. отд-ние, 1989. – 69 с.
3. *Н.В. Зимин.* О влиянии температуры, скорости нагрева и исходного состояния структуры углеродистых сталей на процессы образования в них аустенита // Металлообработка. – 2006, – №1 (31). – С. 41-47.
4. *В.Б. Демидович, Ф.В. Чмиленко.* Численные методы в теории индукционного нагрева. СПб.: Технолит, 2008. – 220 с.
5. *Ф.В. Безменов, Н.В. Зимин.* Математические аналоги экспериментальных зависимостей процессов аустенизации и гомогенизации от некоторых характеристик сталей и технологических параметров термообработки // Металлообработка. – 2006. – № 4 (34). – С. 30-36.
6. *Ф.В. Безменов, Н.В. Зимин.* Математические аналоги экспериментальных зависимостей результатов термической обработки сталей от исходного состояния структуры, размеров аустенитного зерна и параметров нагрева и охлаждения // Металлообработка. – 2006. – № 5-6 (35-36). – С. 59-63.
7. *Н.В. Зимин.* Влияние интенсификации охлаждения после различных видов нагрева на прокаливаемость и закаливаемость углеродистых нелегированных сталей // Металлообработка. – 2006. – № 2 (32). – С. 36-42.

**Секция ТЕХНОЛОГИИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА,
МАШИННОГО ОБУЧЕНИЯ И ПРЕДСКАЗАТЕЛЬНОГО
МОДЕЛИРОВАНИЯ**

**МОДЕЛЬ ОЦЕНКИ ТЕХНИЧЕСКОГО СОСТОЯНИЯ И
ПРОГНОЗИРОВАНИЯ ОТКАЗОВ ЭЛЕКТРОНАСОСНЫХ
АГРЕГАТОВ В УСЛОВИЯХ СКВАЖИННОГО ПОДЗЕМНОГО
ВЫЩЕЛАЧИВАНИЯ (СПВ)**

Н.Р. Адонин

Северский технологический институт НИЯУ «МИФИ»,

г. Северск, D273ANR@edu.ssti.ru

Научный руководитель: Щипков А.А., к.т.н., доцент

Работа описывает алгоритмы синтеза и функционирования адаптивной прогнозной модели для повышения срока службы и предсказания отказов электронасосных агрегатов на основе углубленного анализа параметров непрерывного мониторинга добычного полигона СПВ с совместным использованием методов экспертных оценок и машинного обучения.

The paper describes algorithms for the synthesis and functioning of an adaptive predictive model for increasing the service life and predicting failures of electric pumping units, based on an analysis of the parameters of continuous monitoring of the mining site with the joint use of expert assessment methods and machine learning.

Самым экологичным и безопасным способом добычи урана на сегодняшний день является скважинное подземное выщелачивание (СПВ). При использовании данной технологии рудное тело остается под землей, куда посредством системы нагнетательных скважин подается выщелачивающий раствор, извлекаемый впоследствии через откачные скважины. [1] Подъем продуктивного раствора на поверхность осуществляется скважинными электронасосными агрегатами (ЭНА), состоящими из центробежного насоса и погружного электродвигателя. Будучи важными звеньями технологической цепочки добычи урана методом СПВ, ЭНА вместе с тем являют существенную статью расходов, поскольку число эксплуатируемых откачных скважин только на одном месторождении исчисляется десятками или сотнями. Между тем характер наработок ЭНА на отказ на различных уранодобывающих предприятиях СПВ, как следует из суждений специалистов, разнится от соответствия номинальным значениям до полного несоответствия, что актуализирует проблемы увеличения срока их службы и прогнозирования отказов.

Очевидно, что предсказание выходов из строя на основе номинальной наработки на отказ для проблемных предприятий неэффективно, однако и оценка среднего срока службы не спасает положения ввиду большого разброса значений. Что касается продления службы ЭНА, то, как показывает практика, базовая автоматизация добычного полигона по ряду причин не справляется с задачей на должном уровне. Таким образом, назрела необходимость в более эффективной прогнозной модели.

Настоящая работа описывает алгоритмы синтеза и функционирования адаптивной модели оценки технического состояния и прогнозирования отказов ЭНА на основе анализа параметров непрерывного мониторинга добычного комплекса. Предполагается, что автоматизация добычного полигона, наподобие описанной в [2], уже проведена: контролируется расход, динамический уровень в откачных скважинах; ЭНА оборудованы современными станциями управления.

Различную частоту отказов ЭНА в зависимости от места эксплуатации мы связываем с комплексом гидрогеологических условий, для которого типичны те или иные неблагоприятные режимы работы, ведущие к ускоренному износу. Средства защиты системы автоматического управления ЭНА позволяют предотвратить бурное развитие ряда событий, ведущих к внезапному выходу из строя, однако возникает проблема адекватного определения параметров защит, к тому же факторы длительного действия при этом не учитываются, что вкуче может даже провоцировать новые. Таким образом, при сохранении положения о развитии отказов по модели накопления повреждений [3] считаем необходимым дополнить ее новыми факторами сверх времени активной работы.

Анализ вышедших из строя ЭНА на действующем предприятии СПВ показал, что причины отказов сводятся к проблемам отдельных относительно простых узлов, природа износа которых представляется известной. Поэтому *в качестве технического состояния ЭНА предлагается рассматривать совокупность износов наименее отказоустойчивых узлов, полагая их независимыми*: это первый шаг на пути синтеза прогнозной модели. Дальнейшая задача состоит в эвристическом выражении обобщенной функции износа через параметры мониторинга, покуда математическая модель, описывающая физику износа узла, либо не найдена для данного типа агрегата, либо применить ее в промышленных условиях невозможно из-за нехватки данных. Решать эту задачу предлагается путем синтеза методов машинного обучения и экспертных оценок.

Так, для каждого ранее выбранного узла ЭНА проводится экспертиза причин его износа и в рамках этих причин выявляется *поле независимых факторов, влияющих на износ по принципу суперпозиции*. Затем на основе

имеющегося набора параметров мониторинга по каждому фактору определяются возможные *интегральные показатели износа* с группировкой прямо пропорциональных между собой: полагается, что значение показателя растет по мере действия определенного фактора и накапливается со временем. При составлении показателей эксперты должны стремиться к линейности *частных функций износа* с одной стороны и разнообразию способов их расчета с другой. В самом деле, *представление о линейности функции износа при поиске ее конкретного выражения (обучении модели) позволяет уйти от СДУ в частных производных и необходимости оперировать полным набором архивных значений интегральных показателей в пользу решения СЛАУ с привлечением только конечных значений интегральных показателей*. Вместе с тем действительная функция износа может оказаться сугубо нелинейной, а *одновременный учет нескольких составимых не пропорциональных параметров способен сгладить недостатки допущения, улучшив приближение реальной функции*.

В качестве наиболее критичных узлов ЭНА нами выделены изоляция обмотки статора двигателя и подшипник, соединяющий двигатель с насосом. В табл. 1 представлены соответствующие причины, факторы и примерный вид интегральных показателей износа.

Таблица 1 – Факторы и примерные показатели износа

Фактор износа	Интегральный показатель износа
<i>Перегрев (изоляция обмотки статора)</i>	
Превышение тока	{Интеграл квадрата превышения тока, время работы с повышенным током}
Частые пуски	Число событий частых пусков
Плохое охлаждение	Интеграл превышения P/Q
<i>Повышенные вибрации (подшипник)</i>	
Механический резонанс	Интеграл превышения коэфф. резонансных пульсаций тока, число резонансных частот
Нерезонансные вибрации	Интеграл превышения коэфф. нерезонансных пульсаций тока
<i>Старение (общее)</i>	
Работа	{Наработка, число оборотов вала двигателя}, деградация средн. КПД, cos φ, коэфф. норм. пульсаций
Запуск	Число пусков, обратное суммарное время в разгоне
Авария по напряжению	Число аварий по напряжению

Для обучения модели износа необходима выборка из N ЭНА, вышедших из строя по данному узлу с наработкой более $\frac{1}{4}$ от номинальной. На основе накопленных за срок службы данных мониторинга для каждого ЭНА вычисляется итоговое значение всех интегральных показателей износа X_i ($i \in [1; n]$), при этом показатели, нулевые у всех ЭНА, отбрасываются. Далее формируются все возможные наборы из $n_0 < N$ оставшихся интегральных показателей, чтобы в одном наборе не было сверх одного показателя из пропорциональной группы, а число сочетаний из n_0 элементов по N ($C_N^{n_0}$) было достаточным для адекватной оценки среднего. Затем при заданном наборе показателей для каждого ЭНА составляются уравнения износа вида

$$1 = \sum_{i=1}^{n_0} k_i X_i,$$

где k_i – неизвестный коэффициент линейной функции износа i -го интегрального показателя. Уравнения поочередно группируют по n_0 шт. в системы так, чтобы для всякого интегрального показателя имелось хотя бы одно ненулевое значение, и разрешаются относительно коэффициентов. В результате получается до $C_N^{n_0}$ наборов $\{k_i\}$, из совокупности которых для соответственных коэффициентов вычисляются матожидания (МО) $\{\bar{k}_i\}$ и средние модули отклонения $\{\overline{\Delta k}_i\}$ от МО. Средние коэффициенты подставляют во все уравнения износа, получая набор $\{I_j = \sum_i k_i X_{i,j}\}$ из N расчетных износов для испытуемой группы интегральных показателей. Для расчетных износов находят сумму модулей отклонений от исходной единицы $\Delta = \sum_{j=1}^N |I_j - 1| / N$, что и будет основным параметром экспертной оценки применимости данного набора показателей износа.

После того как описанная процедура повторяется для всех наборов показателей и всех критичных узлов ЭНА, данные машинного обучения передаются экспертам в виде таблиц по узлам с упорядочением по возрастанию Δ : в них значения $\{\bar{k}_i\}$ и $\{\overline{\Delta k}_i\}$ взвешиваются по $\{\bar{X}_i\}$, что позволяет судить о вкладе каждого коэффициента в оценку износа и сравнивать поведение интегральных показателей в различных наборах. Из лучших наборов показателей эксперты выбирают наиболее удовлетворяющий представлениям об износе конкретного узла: он и будет использоваться в модели наряду с соответствующими коэффициентами.

Для прогнозирования отказов – по расчетным значениям износов узлов ЭНА из обучающей выборки строятся функции плотности распределения вероятностей, а для работающих ЭНА – функции износа узлов от времени: последние экстраполируются, в результате чего, рассматривая отказ по каждому узлу независимо, можно получить МО времени выхода из строя. Вновь отказавшие ЭНА ведут к повторному обучению модели и автокоррекции расчетных коэффициентов. При достижении некоторого количества отказавших по данному узлу ЭНА скользящим окном предлагается учитывать лишь последние. Так реализуется дообучение и адаптивность прогнозной модели. В свою очередь, состоявшийся ранее отбор лучших показателей износа узлов с их весовыми коэффициентами позволяет обосновать наиболее благоприятный в текущих условиях режим работы ЭНА и тем самым продлить срок его службы.

Преимуществом разработанной модели является то, что окончательный набор входных параметров получается в результате ее обучения, какое можно начать уже при незначительном числе отказавших ЭНА благодаря их комбинаторным сочетаниям и особому механизму взаимодействия экспертных оценок и машинного обучения, призванному нивелировать издержки малых выборок. Наиболее полное раскрытие представленной модели видится в составе экспертной системы, описанной в [4].

Литература:

1. Геотехнология урана (российский опыт): монография / Под ред. Солодова И.Н., Камнева Е.Н. — М.: «КДУ», «Университетская книга», 2017. — 576 с.

2. Адонин Н.Р., Юров А.В. Автоматизированная система управления эксплуатационным блоком добычи урана методом СПВ // Физико-технические проблемы в науке, промышленности и медицине (ФТПНПМ-2019): сборник научных трудов Международной научно-практической конференции студентов, аспирантов и молодых ученых; Томский политехнический университет. — Томск: Изд-во Томского политехнического университета, 2019. — С. 203–204.

3. Сидоров В.А. Техническая диагностика механического оборудования: учебник. — М., Вологда: «Инфра-Инженерия», 2021. 256 с.: ил., табл.

4. Адонин Н.Р., Щипков А.А. Информационно-управляющая система для повышения эффективности работы погружных насосов. // Инновации в атомной отрасли: проблемы и решения: отраслевая научно-практическая конференция студентов, аспирантов и молодых ученых, посвященная 60-летию СТИ НИЯУ МИФИ в рамках научной сессии НИЯУ МИФИ, 9–13 декабря 2019 г.: материалы конференции: [тезисы

докладов] / Министерство науки и высшего образования РФ, Национальный исследовательский ядерный университет «МИФИ», Информационный центр по атомной энергии г. Томска, Северский технологический институт — филиал НИЯУ МИФИ (СТИ НИЯУ МИФИ); под ред. М.Д. Носкова. — Северск: Изд-во СТИ НИЯУ МИФИ, 2019. — С. 43 — Текст (визуальный): электронный.

ВНЕДРЕНИЕ ЦИФРОВОГО ПОМОЩНИКА В ВЫСШЕМ УЧЕБНОМ ЗАВЕДЕНИИ КАК СПОСОБ ПОВЫШЕНИЯ ЭФФЕКТИВНОСТИ РАБОТЫ СТРУКТУРНЫХ ПОДРАЗДЕЛЕНИЙ

Д.В. Аксенова

**Сибирский государственный университет путей сообщения,
г. Новосибирск, aksenovadv_v@mail.ru**

Научный руководитель: Уланов А.А., к.т.н., доцент

В статье рассматриваются проблемы моделирования структурных подразделений высшего учебного заведения в условиях появления современных средств коммуникаций - цифровых помощников. Представляются основные подходы к планированию и внедрению цифрового документооборота. Анализируются основные преимущества цифровой трансформации, и ее функционирование внутри образовательной организации.

The article discusses the problems of modeling structural divisions of a higher educational institution in the context of the emergence of modern means of communication - digital assistants. The main approaches to the planning and implementation of digital document collection are presented. The main advantages of digital transformation and its functioning within the educational organization are analyzed.

В современном программировании технология искусственного интеллекта с каждым годом стремительно развивается и применяется повсеместно уже почти с десятков лет. Создание данной системы дало значительное количество интеллектуальных продуктов, которые активно используются в повседневной жизни.

Одним из ключевых признаков цифровой трансформации является появление нового класса систем, облегчающих человеку манипулирова-

ние артефактами цифровой среды. Наиболее яркими представителями таких систем являются персональные цифровые ассистенты [2].

Виртуальный цифровой помощник – веб-сервис и (или) приложение для смартфонов и ПК, фактически исполняющий роль личного секретаря при пользователе.

В условиях современного информационного общества эффективное управление информационными потоками — важнейшая задача руководителя образовательной системы.

Регулирование огромного массива данных, циркулирующих в образовательном процессе (документы, отчеты, новости, поручения, и т. д.), требует создания системы управления корпоративным контентом, включающей в себя новейшие цифровые сервисы и ресурсы.

В настоящее время хранение документов осуществляется на электронных носителях, но так как информационный обмен происходит с помощью компьютерной сети, то специалисту, отвечающему за управление информационными потоками, необходимо владеть знаниями в электронных ресурсах [2].

К современным инструментам управления корпоративным контентом, активно используемых в образовательных организациях города, можно отнести две группы: уже знакомый и удобный формат оффлайна для педагогов, и не всегда удобный и доступный формат онлайн инструментов (таблица 1).

Таблица 1 – Инструменты управления корпоративным контентом

<i>Оффлайн</i>	<i>Онлайн</i>
<i>Информационные стенды</i>	<i>Официальный сайт организации</i>
<i>Печатная продукция (брошюры, журналы, газеты)</i>	<i>Страницы в социальных сетях, чаты в мессенджерах</i>
<i>Раздаточный материал (памятки, флаеры, рабочие тетради)</i>	<i>Облачные технологии, электронные преподавательские</i>

Обычно, для управления корпоративным контентом, образовательные организации используют общедоступные цифровые сервисы такие как: электронная почта, чаты в мессенджерах (WhatsApp, Viber), социальные сети (ВКонтакте, Facebook), облачные инструменты (Google, Yandex).

В период пандемии COVID-19 многие образовательные организации оказались неготовыми к дистанционным формам передачи информации, что сильно повлияло на качество образовательных услуг[5].

Несмотря на повсеместный переход к цифровым инструментам, повышению цифровой компетентности сотрудников, административный корпус образовательных организаций по-прежнему испытывает значительные перегрузки в связи с интенсивностью документооборота и коммуникаций между участниками образовательного процесса.

Управление невероятно быстро растущими объемами информации и извлечение из них необходимых знаний является одной из наиболее сложных задач, стоящих перед современными организациями. Управление неструктурированной информацией («контент»), к которой относят документы, электронные таблицы, презентации, электронные письма, аудио- и видеофайлы, отсканированные документы и другие нормативно-локальные акты, является менее развитой сферой деятельности. Дело осложняется тем, что не весь контент является электронным — многие организации все еще утопают под горами бумажных документов, которые не все еще переведены в электронный вид.

Разработка персонального цифрового помощника в сфере образования позволит с помощью корпоративного банка данных упростить работу структурным подразделениям организации.

Создание персонального цифрового помощника предусматривает:

- информирование о расписании дня;
- предоставление необходимых документов;
- функцию создания заявки-сообщения вышестоящему начальству;
- сообщение актуальной информации о работе корпоративного университета, принятых решений на собраниях сотрудников;
- формирование новостной ленты, привязанной к официальному сайту образовательной организации;
- предоставление контактной информации о внутренних административных кабинетах [1, 3].

Большой плюс данного «хранилища» заключается в быстром и неограниченном доступе к необходимой информации, простоте использования и автоматическом сохранении данных на телефоне без использования браузера.

Для административного состава — это инновационная форма оптимизации документооборота и управления корпоративным контентом, который содержит в себе актуальную и единственно верную информацию для сотрудников [4].

Для преподавательского состава – это универсальный помощник в организации деятельности и коммуникации с управляющим составом.

Внедрение цифрового обмена документами позволит полностью исключить ручной ввод целого ряда информации, поступающей в управленческий отдел от контрагентов и различных подразделений крупных

учреждений. Фиксация времени поступления первичных документов позволит строго отслеживать график документооборота, контролировать сроки отражения операций, что существенно повысит эффективность внутреннего контроля и достоверность отчетности[5].

Переход на цифровой документооборот является одной из важных составляющих цифровой экономики. Данный подход не только позволяет не только обеспечить экономию ресурсов, но и прийти к автоматизации массовых операций.

Литература:

1. *Проватар А.И.* Особенности и проблемы виртуального общения с помощью чат-ботов // А.И. Проватар, К.А. Ключко. - URL: http://ict.informika.ru/ft/005773/2007_4_87-90.pdf

2. *Сазыкина Н.А., Морозов Б.Б.* Исследование современных средств для разработки веб-приложения виртуального цифрового помощника // Вестник науки. – 2019. - №9(8). – Т.4. – С. 38-39.

3. *Тихонова Л.П.* Моделирование функциональных возможностей цифрового ассистента и условий его реализации в вузовском учебном процессе // Вестник Череповецкого государственного университета. - 2019. - №4(91). – С. 215-225.

4. *Чиркова В. М.* Влияние информационно-коммуникационных технологий на сферу образования // Региональный вестник. – 2019 - №12(27). – С. 37-38.

5. *Ширшов Е.В., Ефимова Е.В.* Организация учебной деятельности в вузе на основе электронных, информационно-образовательных технологий. – Архангельск: Изд-во Арханг. гос. тех. ун-та, 2006. – 208 с.

НЕЙРОННЫЕ СЕТИ КАК ОСНОВА БУДУЩЕГО

И.Е. Бережнев, Ю.А. Сокольникова

Северо-Восточный федеральный университет имени

М.К. Аммосова, г. Якутск, vipkajklal@gmail.com

Научный руководитель: Сокольникова Ю.А., старший преподаватель кафедры Математической экономики и прикладной информатики Института математики и информатики

В данной работе будет дана теоретическая основа о нейронных сетях, их предпосылках, а также их потенциале. В практической ее части мы создадим собственную нейронную сеть для более яркого представления о данной технологии, а также рассмотрим имеющиеся

на данный момент алгоритмы и на примере открытой программной платформы Watson Studio проведем анализ данных и продемонстрируем пример того, что уже умеют алгоритмы в нейронных сетях.

In this work, a theoretical basis will be given about neural networks, their prerequisites, as well as their potential. In its practical part, we will create our own neural network for a brighter idea of this technology, and also consider the algorithms currently available and use the example of the open software platform Watson Studio to analyze data and demonstrate an example of what algorithms in neural networks can already do.

Сегодня нейронные сети ведут нас к кардинальным изменениям в бизнесе и повседневной жизни, выводя нас на новый уровень в области искусственного интеллекта. Имитируя работу взаимосвязанных клеток мозга, машины с поддержкой ИИ (включая смартфоны и компьютеры, которые мы используем ежедневно) теперь обучены учиться, распознавать закономерности и делать прогнозы в человеческой манере, а также решать проблемы в каждом секторе бизнеса.

Целью научной статьи является изучение алгоритмов обучения нейронных сетей. Для достижения данной цели были поставлены следующие задачи:

1. Ознакомление с имеющимися на данный момент нейросетями, а также рассмотрение перспектив данного направления.
2. Построение нейронной сети и её обучение с использованием полученных данных.
3. Тестирование полученной модели.
4. Оценка успешности классификации.

В качестве метода исследования был выбран высокоуровневый язык программирования Python, а также программная платформа IBM для анализа данных - Watson Studio.

У каждой нейронной сети имеется определенное количество алгоритмов, вычисляющих наилучший исход. В случае с Watson Studio, используется сразу несколько алгоритмов. В ходе эксперимента, мы вычислим какой из них лучше справляется с данными нашей БД, в которой расписаны сорта ириса.

И так, в нашей БД 150 данных для трёх классов (сортов ириса). Каждый класс имеет по 4 атрибута: длина чашелистика в см (sl), ширина чашелистика в см (sw), длина лепестка в см (pl), ширина лепестка в см (pw).

Классы: Iris Setosa, Iris Versicolour, Iris Virginica.

В Progress Map набор тренировочных данных был разбит между двумя алгоритмами кроссвалидации - "Logistic Regression" и "Gradient Boosting Classifier".

Каждый алгоритм проверял Pipeline на количество баллов кроссвалидации. Первая линия каждого алгоритма выводила результат без каких-либо улучшений, далее каждая новая линия была оптимизирована одним из улучшений. В этой схватке алгоритмов самым эффективным оказался алгоритм "Logistic Regression".

Анализируя полученные данные с помощью наилучшей модели, нейросеть выдала наименьший отчёт в виде 3 неполных чисел, которые являлись вероятностями того, какой класс это может быть.

Так же, в данной работе мы написали на Python одну из простейших нейронных сетей. В качестве примера была взята задача, которую можно найти практически во всех учебниках по нейронным сетям.

Таблица 1 - Простейшая задача нейронного программирования

	Ввод			Вывод
Пример 1	0	0	1	0
Пример 2	1	1	1	1
Пример 3	1	0	1	1
Пример 4	0	1	1	0
Пример 5	1	1	0	?

Программа получает на вход 3 числа, и должна рассчитать какое число будет на выходе. В качестве тренировочных данных имеется 4 примера с 3 входными данными и 1 выходом в каждом из них. Поняв закономерность, мы приходим к выводу, что в 5 примере вывод будет равен 1, то есть вводу. Поймет ли это нейросеть?

Сам подход к обучению нейросети подразумевает старт с неправильной позиции для нахождения правильной, при этом важен тот факт, что начальные синаптические веса не могут быть полностью одинаковыми, иначе в процессе обучения они не изменятся. А поскольку мы не знаем, какие веса идеальные и их нельзя делать одинаковыми, то случайные значения синаптических весов – наше решение. Именно веса синапсов выявляют взаимосвязь между входными данными и конечным результатом, который будет считаться решением нашей задачи. Сам же нейрон выявляет взаимосвязь с помощью двух простых действий.

Во-первых, мы производим умножение входных данных на их веса с последующим сложением получившихся чисел. Во-вторых, получившийся результат мы инициализируем в функцию активатор.

В написании кода программы (Рис.1) мы будем применять более подходящую для нас функцию активатор, а именно сигмоид (1), где z - скалярное умножение предыдущего скрытого слоя (или входов) и строки матрицы весов в дополнение к смещению.

$$\varphi(z) = \frac{1}{1+e^{-z}} \quad (1)$$

И так, после объявления функции активатора мы объявляем тренировочные данные. С этой целью мы создаем две переменные: `training_inputs` (она будет хранить в себе массив 4×3 с соответствующими входными данными) и `training_outputs` (она хранит в себе массив 1×4 , содержащий ожидаемые выходные данные). Так же в конце добавляем транспонирование чтобы содержание матрицы поменялось на 4×1 . Далее инициализируют веса, с помощью генератора случайных чисел. Но полученные синаптические веса пока не пригодны к использованию. Поэтому решено воспользоваться методом обратного распространения.

Верными значениями для нас являются $[0,1,1,0]$. Конечно, после использования функции активатора, такие значения получить невозможно, поскольку метод обратного распространения подразумевает многократное обучение нейронной сети, в ходе которого будет производиться регулирование весов. Поэтому мы повторим обучение 30 тысяч раз.

После запуска программы мы видим, что нейронная сеть научилась сама выявлять взаимосвязь между входными и выходными данными. Проверим, справится ли она с новой ситуацией. Передадим ей входные значения $[1,1,0]$. Как мы помним, во входных тренировочных данных такого значения нет. После запуска кода видим результат: наша нейронная сеть прекрасно справилась и поняла, что на выходе должна быть цифра 1 (Рис.2). Так мы получили простейшую нейронную сеть – перцептрон.

В заключении хотелось бы сказать: поскольку размер нейронных сетей увеличивается значительными темпами, возможности технологии для решения все более сложных задач становятся осуществимее. Включение этики ИИ, сочетания более крупных нейронных сетей, увеличивающейся вычислительной мощности, больших наборов данных и результатов десятилетий исследований— все это открывает нам в будущем захватывающую перспективу для применения искусственных нейронных сетей на благо общества.

```
import numpy as np

def sigmoid(x):
    return 1 / (1 + np.exp(-x))

# Инициализация весов и смещений
weights_hidden = np.random.randn(4, 3)
weights_output = np.random.randn(1, 3)
bias_hidden = np.zeros((1, 4))
bias_output = np.zeros((1, 3))

# Функция потерь
def loss(y_hat, y):
    return -np.sum(y * np.log(y_hat) + (1 - y) * np.log(1 - y_hat))

# Обучение
for epoch in range(1000):
    # Генерация входных данных
    x = np.random.randn(3, 1)
    # Прямой проход
    z_hidden = x.dot(weights_hidden) + bias_hidden
    a_hidden = sigmoid(z_hidden)
    z_output = a_hidden.dot(weights_output) + bias_output
    y_hat = sigmoid(z_output)
    # Обратный проход
    error_output = y_hat - y
    error_hidden = error_output.dot(weights_output)
    # Обновление весов и смещений
    weights_hidden -= lr * x.dot(error_hidden)
    bias_hidden -= lr * error_hidden
    weights_output -= lr * a_hidden.dot(error_output)
    bias_output -= lr * error_output
    # Вывод прогресса
    if epoch % 100 == 0:
        print(f'Эпоха: {epoch}, Потери: {loss(y_hat, y)}')
```

Рис.1 - Код программы

```
Случайные инициализирующие веса:
[[-0.16595999]
 [ 0.44064899]
 [-0.99977125]]
Веса после обучения:
[[10.94837682]
 [-0.20556845]
 [-5.26921262]]
Результат после обучения:
[[0.08512137]
 [0.99982147]
 [0.99659517]
 [0.08417383]]
Новая ситуация:
[0.9999784]
```

Рис.2 - Вывод программы

В результате написания статьи была разработана простейшая нейронная сеть и произведено ее обучение с использованием полученных данных. В ходе разработки нейронной сети была изучена предметная область, сформулированы цель и задачи работы. Был проведен обзор, демонстрация и тестирование имеющихся на данный момент нейронных сетей, а также рассмотрены перспективы данного направления. В результате чего было решено создать собственную сеть, которая способна из входных данных предугадать выходные данные.

Разработанную нейронную сеть можно рассматривать как удобное средство для практического понимания работы нейронных сетей.

Литература:

- 1. Васильев, А. Н. Принципы и техника нейросетевого моделирования / А.Н. Васильев, Д.А. Тархов. - Москва: Гостехиздат, 2014. - 218 с.
- 2. Васильев, А.Н. Тархов Д.А. Принципы и техника нейросетевого моделирования / А.Н. Васильев. - Москва: СИНТЕГ, 2015. - 513 с.

3. Галушкин, А. И. Нейрокомпьютеры. Учебное пособие / А.И. Галушкин. - М.: Альянс, 2014. - 528 с.
4. Гелиг, А. Х. Введение в математическую теорию обучаемых распознающих систем и нейронных сетей. Учебное пособие / А.Х. Гелиг, А.С. Матвеев. - М.: Издательство СПбГУ, 2014. - 224 с.
5. Круглов, В. В. Искусственные нейронные сети. Теория и практика / В.В. Круглов, В.В. Борисов. - М.: Горячая линия - Телеком, 2002. - 382 с.
6. Рамиля Латыпова Нейронные сети / Рамиля Латыпова. - М.: LAP Lambert Academic Publishing, 2012. - 418 с.
7. Редько, В.Г. Подходы к моделированию мышления / В.Г. Редько. - Москва: Машиностроение, 2016. - 736 с.
8. Редько, В.Г. Эволюция, нейронные сети, интеллект: Модели и концепции эволюционной кибернетики / В.Г. Редько. - Москва: СИНТЕГ, 2017. - 224 с.
9. Рутковская, Д. Нейронные сети, генетические алгоритмы и нечеткие системы: моногр. / Д. Рутковская, М. Пилиньский, Л. Рутковский. - М.: Горячая линия - Телеком, 2013. - 384 с.
10. Рышард, Тадеусевич Элементарное введение в технологию нейронных сетей с примерами программ / Тадеусевич Рышард. - М.: Горячая линия - Телеком, 2011. - 723 с.
11. Хайкин, Саймон Нейронные сети. Полный курс / Саймон Хайкин. - М.: Вильямс, 2016. - 562 с.

ОБУЧЕНИЕ НЕЙРОННОЙ СЕТИ ДЛЯ ПРОВЕРКИ ПРОСТЫХ ГИПОТЕЗ О ВИДЕ РАСПРЕДЕЛЕНИЯ

Д.А. Голованова

Новосибирский государственный технический университет,

г. Новосибирск, golovanova.2017@stud.nstu.ru

Научный руководитель: Чимитова Е.В., д.т.н., доцент

Объектом исследования являются простые гипотезы о виде распределения. Цель работы – разработка нового критерия согласия для проверки простых гипотез о виде распределения с более высоким значением мощности, чем любой из классических критериев согласия (Колмогорова, Крамера-Мизеса-Смирнова, Андерсона-Дарлинга и хи-квадрат Пирсона). Исследования полученного критерия проводились на двух парах близко конкурирующих гипотез. В результате предложен новый

критерий согласия, объединяющий широко известные классические критерии согласия.

The object of the research is simple statistical hypotheses about the form of the distribution. The aim is to develop a new goodness-of-fit test for testing simple hypotheses about the form of the distribution with a higher power than any of the classical goodness-of-fit tests (Kolmogorov, Cramer-Mises-Smirnov, Anderson-Darling, and Pearson's chi-square). The researches of the obtained goodness-of-fit test was investigated on two pairs of closely competing hypotheses. As a result, a new goodness-of-fit test was proposed that combines the well-known classical goodness-of-fit tests.

Идентификация закона распределения наблюдаемой случайной величины – задача выбора параметрической модели какого-либо закона распределения вероятностей, наилучшим образом соответствующей результатам экспериментальных наблюдений.

Пусть имеется выборка данных X_n . Необходимо проверить простую гипотезу H_0 , согласно которой функция распределения выборки F равна некоторой функции распределения F_0 :

$$H_0: F(x) = F_0(x, \theta),$$

где $\theta = (\theta_1, \theta_2, \dots, \theta_i)^T$ – известные значения параметров.

Проверка таких гипотез осуществляется с использованием критериев согласия: задается уровень значимости α и по выбранному критерию вычисляется значение статистики S_n – некоторой функции от данной выборки и теоретического закона распределения.

С результатом проверки, так как случайность выборки предполагает возможность ошибочного результата, связывают ошибки двух видов: 1-го рода, когда гипотеза H_0 отклоняется будучи верной, и 2-го рода, когда гипотеза H_0 принимается будучи неверной (справедливой оказывается некоторая конкурирующая гипотеза H_1).

Вывод о схожести вида распределения формируется на соответствии полученного значения статистики заданному уровню значимости. При этом узнать истинный закон распределения случайной величины невозможно, возможно лишь найти модель, наилучшим образом приближающую этот закон.

Использование разных критериев согласия может приводить к получению несогласованных результатов – мощность критериев (способность не допускать ошибку 2-го рода) существенно зависит от пары конкурирующих гипотез [1].

На практике широко применяются непараметрические критерии, которые не могут гарантировать точное распознавание распределения для выборки данных небольшого объема, но их достоинство в том, что предельные распределения их статистик при проверке простых гипотез не зависят от вида и параметров наблюдаемого закона распределения [2]. К таким критериям относятся критерий Колмогорова, Крамера-Мизеса-Смирнова, Андерсона-Дарлинга и хи-квадрат Пирсона.

Подход на основе нейронных сетей зарекомендовал себя как оптимальный по скорости и точности решения на небольших выборках данных по сравнению с классическими критериями [3]. Применение нейросетевых технологий представляется перспективным: за счет больших объемов данных для обучения и использования различных критериев согласия полученная модель может объединять в себе преимущества различных классических критериев, а значит, будет устойчивой относительно различных конкурирующих гипотез.

Нейронная сеть представляет собой систему соединенных и взаимодействующих между собой искусственных нейронов (узлов). Каждый из нейронов воздействует на входной сигнал и посылает обработанный сигнал на вход другому нейрону.

Способность обучаться – определять коэффициенты связей нейронов между друг другом – характерная черта, выделяющая нейронные сети перед другими алгоритмами. Обучение нейронной сети происходит постоянно, даже на данных, приобретенных через большой временной промежуток после внедрения нейронной сети [4].

В качестве архитектуры нейронной сети выбран многослойный перцептрон – класс сетей прямого распространения, состоящий как минимум из трех слоев: входного, скрытого, узлы которого не заметны со стороны входа или выхода, и выходного. Среди отличительных признаков многослойного перцептрона выделяют: наличие нелинейной функции активации в каждом нейроне; наличие одного или нескольких скрытых слоев; высокую степень связности нейронной сети [4].

Входной слой состоит из 11 нейронов, каждый из которых принимает на вход значение p -value, полученное в результате проверки одного из реализованных критериев согласия (Колмогорова, Крамера-Мизеса-Смирнова, Андерсона-Дарлинга и Хи-квадрат Пирсона для различного числа группирования).

Выходной слой представляет собой один нейрон – значение от 0 до 1, полученное с помощью сигмоидальной функции активации:

$$y_i = \frac{1}{1 + \exp(-v_i)},$$

представляющее собой вероятность принятия или отвержения рассматриваемой гипотезы.

В работе исследовались выборки из 92 непрерывных распределений с разными параметрами масштаба, сдвига и формы.

Предварительно для генерации выборки выбираются параметры: n – объем выборки (случайное целое число от 20 до 500 с шагом 20); loc – параметр сдвига, среднее значение случайной величины (случайное значение с плавающей запятой в промежутке $[-10, 10]$); scale – параметр масштаба – показатель рассеивания значения случайной величины от ее математического ожидания (случайное значение с плавающей запятой в промежутке $[0.5, 10]$) и параметры формы, заданные ограничениями конкретного вида распределения.

Конкурирующие гипотезы подбирались двумя способами. Из аналогичного распределения с близкими параметрами: дельта-изменения каждого параметра подбиралась из нормального распределения в пределах $[0, 0.2]$. Для дельты-изменения параметра масштаба учитывалось непопадание значения в окрестность нуля (от -0.02 до 0.02). А также из другого семейства распределения, методом подбора близких оптимальных значений. Набор данных состоит из половины принимающихся и половины отклоняющихся гипотез.

Далее полученные выборки данных, подчиняющиеся определенному закону распределения, подвергались проверке согласия с основной гипотезой по критериям: Колмогорова, Крамера-Мизеса-Смирнова, Андерсона-Дарлинга и хи-квадрат Пирсона с k -интервалами группирования ($k \in [3, 10]$).

На вход нейронной сети подаются значения p -value вместо статистики критерия, чтобы естественным образом отображать область значений статистики в интервале от 0 до 1. На выходном слое получается число – мощность построенного критерия о согласии вида распределения. Данное значение не учитывает интересующий уровень значимости, поэтому после следует оценить полученное значение мощности.

Для построения модели нейронной сети использовался оптимизатор, который автоматически настраивает набор гиперпараметров – переменных, управляющих процессом обучения нейронной сети. В множество поиска гиперпараметров входят число нейронов слоя, функция активации, скорость обучения, количество скрытых слоев и число эпох обучения. Для оптимизации используется метод Байесовской оптимизации.

В итоге, реализованная архитектура модели нейронной сети представляет собой многослойный перцептрон с одним скрытым слоем. Для данной модели были подобраны оптимальные гиперпараметры и достигнут определенный уровень точности (примерно 0.95 на валидационной выборке).

Данная нейронная сеть дает возможность использовать так называемый обобщенный критерий согласия, который минимизирует вероятность совершения ошибок 1-го или 2-го родов. То есть критерий с более высоким значением мощности, чем любой из одиннадцати представленных критериев в одиночку. В дальнейшем планируется исследование полученного критерия по мощности согласно различным уровням значимости.

Литература:

1. Лемешко Б.Ю., Лемешко С.Б., Постовалов С.Н. Сравнительный анализ мощности критериев согласия при близких конкурирующих гипотезах. I. Проверка простых гипотез // Сибирский журнал индустриальной математики. – 2008. – Т.11. – № 2(34). – С.96–111.
2. Лемешко Б.Ю. Непараметрические критерии согласия: Руководство по применению: Монография / Б.Ю. Лемешко. – М.: ИНФРА-М, 2014. – 163 с.
3. A Neural Network-Based Approach for Statistical Probability Distribution Recognition / С.-Т. Su, С.-J. Chou. – 2006. – doi: 10.1080/08982110600653598.
4. Хайкин С. Нейронные сети: полный курс, 2-е изд., испр.: Пер. с англ. /С. Хайкин.– М.:ООО И.Д. Вильямс, 2006. – 1104с.

ИСПОЛЬЗОВАНИЕ БОЛЬШИХ ДАННЫХ ДЛЯ ТЕЛЕМЕДИЦИНЫ В ОФТАЛЬМОЛОГИИ

Э.Р. Добров

**Новосибирский государственный университет
экономики и управления «НИНХ»,
г. Новосибирск, sib.2020@bk.ru**

Научный руководитель: Терещенко С.Н., к.т.н., доцент

В статье показаны типы аналитических задач, которые можно решать с применением передовых методов анализа больших данных для телемедицины в офтальмологии. Глубокое обучение в диагностировании заболеваний органов зрения.

The article shows the types of analytical tasks that can be solved using advanced methods of big data analysis for telemedicine in ophthalmology. Deep learning in the diagnosis of diseases of the organs of vision.

Издание Gartner в 2001 г. предложило основные характеристики больших данных: разнообразие, увеличивающаяся скорость прибавления и огромный объем. Сегодня к этим трем добавлены еще три характеристики: достоверность, изменчивость, ценность.

Типы задач медицинского анализа, которые можно решить с помощью интеллектуального анализа больших данных: описательный анализ, диагностический анализ, прогнозная аналитика, предписывающий анализ [1].

Закон о телемедицине вступил в силу с 1 января 2018 г. ФЗ от 29 июля 2017 г. N 242-ФЗ «О внесении изменений в отдельные законодательные акты Российской Федерации по вопросам применения информационных технологий в сфере охраны здоровья». (Федеральный закон Российской Федерации N 242-ФЗ от 29.07.2017 г. «О внесении изменений в отдельные законодательные акты Российской Федерации по вопросам применения информационных технологий в сфере охраны здоровья». Доступно по: consultant.ru/document/cons_doc_LAW_221184/ Ссылка активна на 18.11.2021).

1 февраля 2021 г. вступил в силу приказ Министерства здравоохранения России, согласно которому поликлиники и больницы получили возможность полностью отказаться от бумажных медицинских карт и перейти на электронный документооборот. (Приказ Министерства здравоохранения Российской Федерации от 07.09.2020 г. N 947н «Об утверждении Порядка организации системы документооборота в сфере охраны здоровья в части ведения медицинской документации в форме электронных документов» (Зарегистрировано в Минюсте России 12.01.2021 N 62054) Доступно по: consultant.ru/document/cons_doc_LAW_373853/ Ссылка активна на 18.11.2021).

Законодательство регламентирует вопрос информационной безопасности по риску утечки и обезличиванию личных персональных данных пациентов.

В рамках нацпроекта «Здравоохранение» с 2019 г. в России стартовала работа над федеральным проектом «Создание единого цифрового контура здравоохранения на основе единой государственной информационной системы здравоохранения (ЕГИСЗ)». («Паспорт национального проекта «Здравоохранение» (утв. президиумом Совета при Президенте РФ по стратегическому развитию и национальным проектам, протокол N 16 от 24.12.2018 г.) Доступно по: consultant.ru/document/cons_doc_LAW_319209/ Ссылка активна на 18.11.2021).

В литературе, посвященной большим данным в медицине, значительное место отводится обсуждению этапов работы с медицинской

информацией: сбору, обработке, хранению, анализу, моделированию и визуализации [2]. По прогнозам аналитического агентства Gartner, области использования ИИ в медицине классифицируются на уровни: проектирования, производственный, в части продвижения, оказания услуг [3]. За последние несколько лет по мнению ученых, рост эффективности демонстрирует применение глубокого обучения в задачах: компьютерного зрения, технологии обработки естественного языка, речевых технологий [4].

В литературе, посвященной решению задач распознавания изображений, значительное место отводится обсуждению методов, основанных на искусственных нейронных сетях. На примере офтальмологии анализ изображений подразделяется на две группы: методы классификаций изображений и методы семантической сегментации [5]. В настоящее время получены уникальные результаты по точности распознавания различных диагнозов глазных болезней с применением фундус-камеры. [1, 6]. В научных работах отмечено, что следующим шагом в цифровом обследовании пациентов будет косоглазие и онкология в офтальмологии [7].

Многие авторы уделяют внимание производительной работе алгоритмов, требуются обработанные наборы очищенных данных. [8]. Реализованные приложения согласно обзорам литературы устанавливают присутствие или отсутствие нарушений здоровья [6]. Вместе с тем остается открытым вопрос о изучении перспективных направлений лечения глазных заболеваний - это профилактика и ранняя диагностика больных людей [9, 10].

С февраля 2008 по июль 2017 г. было проведено дистантное скрининговое исследование зрения 28983 школьникам г. Новосибирска [11]. Сложность медицинской практики состоит в обследовании степени тяжести присутствующего заболевания у клиента. [12]. Изучение литературы в области телеофтальмологии показало, что реализация программных продуктов на основе метода глубокого обучения функционирует по примеру «черного ящика».

Исследователи компании Google резюмируют принципы интерпретируемости следующим образом: понимание того, что делают скрытые слои; понимание того, как активируются узлы; понимание того, как формируются концепции [13].

24 апреля 2018 г. вышел Приказ Министерства здравоохранения Российской Федерации N 186 «Об утверждении Концепции предиктивной, превентивной и персонализированной медицины». (Приказ Министерства здравоохранения Российской Федерации N 186 от 24.04.2018 г. «Об утверждении Концепции предиктивной, превентивной и персона-

лизированной медицины» Доступно по: consultant.ru/cons/cgi/online.cgi?req=doc&base=EXP&n=612584#07276999663002452 Ссылка активна на 18.11.2021).

Исходя из вышесказанного перспективным направлением для исследовательской работы является предоставление понятного, пошагового и развернутого метода решения поставленных задач. Многие исследователи отмечают, что в связи с ростом различных заболеваний глаз у населения требуется принять меры по увеличению штата мед учреждений с узкими специалистами. [14]. С развитием интернета телемедицина шагнула вперед, появилась возможность предоставления эффекта полного присутствия при первичном осмотре врача-диагноста [3]. Во время отсутствия в достаточном количестве персональных компьютеров в медицинских учреждениях на помощь приходят мобильные средства связи. [15]. Основная идея телепомощи офтальмологическим больным показана на схеме (рис. 1) [14].

Мобильное здравоохранение – отрасль телемедицинских технологий в рамках которых развиваются мобильные решения по всему миру.

Анализ данных на уровне мобильных приложений включает в себя облачные вычисления и периферийные вычисления.

Глубокое изучение предъявляет высокие требования к ресурсам для высокопроизводительных вычислений и существенно зависит от параллельных вычислений.

Описаны уже случаи что рынок мобильного здравоохранения представлен приложениями диагностики, справочниками-консультантами. [16, 17].



Рис. 1 – Схематическое устройство работы телемедицинского центра

Таким образом, исходя из проведённого обзора литературы, телемедицина характеризуется накоплением больших объемов данных медицинской информации. Кроме того, в последние годы наблюдается рост достижений в применении машинного обучения в распознавании изображений глазного дна для диагностирования пациентов.

Вместе с тем недостаточно исследованы мобильные приложения для постановки степени тяжести заболевания. Не до конца изучены объяснения основных принципов и алгоритмов действия методов глубокого машинного обучения. Предложения по решению обозначенных задач будут предприняты в следующих работах автора в рамках диссертационного исследования.

Литература:

1. Применение систем с искусственным интеллектом в диагностике офтальмологических заболеваний (обзор литературы) / Апрелева А.Е., Манкибаева Р.И., Манкибаев Б.С., Апрелева Е.В. // Вестник Башкирского государственного медицинского университета. - 2019. - №3. - С. 10-14.

2. Ophthalmic Diagnosis and Deep Learning - A Survey / Sourya Sengupta, Amitojdeep Singh, Henry A. Leopold, Vasudevan Lakshminarayanan [Electronic resource] // arXiv.org. 2018. - URL: arXiv:1812.07101v1 [cs.CV] (accessed 10.03.2021).

3. Tele-ophthalmology: Need of the hour / Mohita Sharma, Neha Jain, Sridhar Ranganathan, Naman Sharma, Santosh G Honavar, Namrata Sharma, Mahipal S Sachdev // Indian Journal of Ophthalmology. – 2020. – Vol. 68, N 7. – P. 1328–1338. - doi: 10.4103/ijo.IJO_1784_20.

4. Гусев А. В. Перспективы нейронных сетей и глубокого машинного обучения в создании решений для здравоохранения // Врач и информационные технологии. – 2017. - №3. – С. 92 – 105.

5. Grewal P.S., Oloumi F., Rubin U. Deep learning в офтальмологии: обзор // Новое в офтальмологии. – 2018. - № 4. – С. 36-40.

6. Applications of Artificial Intelligence in Ophthalmology: General Overview / Wei Lu, Yan Tong, Yue Yu, Yiqiao Xing, Changzheng Chen and Yin Shen // Hindawi. Journal of Ophthalmology – 2018. – N 6. – P. 1-15. – doi: 10.1155/2018/5278196 Article ID 5278196.

7. Методы машинного обучения в офтальмологии. Обзор литературы. / Гарри Д.Д., Саакян С.В., Хорошилова-Маслова И.П., Цыганков А.Ю., Никитин О.И., Тарасов Г.Ю. // Офтальмология. – 2020. - Т. 17 - № 1. – С. 20–31. – doi: 10.18008/1816-5095-2020-1-20-31.

8. Современный передовой уровень искусственного интеллекта для умной медицины / Колесниченко О.Ю., Мартынов А.В., Пулит В.В.,

Колесниченко Ю.Ю., Шакиров В.В., Варламов О.О., Минушкина Л.О., Сотник А.Ю., Жилина Т.Н., Дорофеев В.П., Смородин Г.Н., Жапаров М.К., Мазелис Л.С. // Ремедум. Журнал о рынке лекарств и медицинской технике. – 2019. - № 4. – С. 36 – 43. – doi: 10.21518/1561-5936-2019-04-36-43.

9. Терещенко А.В., Трифаненкова И.Г., Юдина Ю.А. Телемедицина в скрининге, диагностике и лечении активной ретинопатии недоношенных // Офтальмохирургия. – 2017. – №2. – С. 73-77. – doi: 10.25276/0235-4160-2017-2-73-77.

10. Julia E. Reid, Eric Eaton Artificial Intelligence for Pediatric Ophthalmology [Electronic resource] // arXiv.org. 2018. - URL: arXiv:1904.08796v1 [physics.med-ph] (accessed 10.03.2021).

11. Дистантный скрининг зрения школьников как практический метод применения телемедицины в офтальмологии: итоги и перспективы / Черных В.В., Плисов И.Л., Анциферова Н.Г., Пуцина В.Б. // Сибирский научный медицинский журнал. – 2018. – Т. 38. - №3. – С. 80-85. - doi: 10.15372/SSMJ201803012.

12. Muhammad Imran Razzak, Saeeda Naz and Ahmad Zaib Deep Learning for Medical Image Processing: Overview, Challenges and Future [Electronic resource] // arXiv.org. 2017. - URL: arXiv:1704.06825v1 [cs.CV] (accessed: 10.03.2021).

13. Lokman Balyen, Tunde Peto Promising Artificial Intelligence–Machine Learning–Deep Learning Algorithms in Ophthalmology // Asia-Pacific Academy of Ophthalmology. – 2019. – Vol. 8, N 3. – P. 264 – 272. - doi: 10.22608/APO.2018479.

14. Астахов Ю.С., Тургель В.А. Телемедицина в офтальмологии. Часть 1. «Общая телеофтальмология» // Офтальмологические ведомости. – 2020. – Т. 13. – № 1. – С. 43–52. – doi: 10.17816/OV19112.

15. Android mobile applications in eye care / Siddharth Karuppasamy Karthikeyan, Rajesh Thangarajan, Nagarajan Theruvathi, Krithica Srinivasan // Oman Journal of Ophthalmology. – 2019. – Vol. 12, N 2. – P. 73-77. - doi: 10.4103/ojo.OJO_226_2018.

16. Фокин С. Ю., Киричек Р. В. Обзор медицинских приложений, устройств и технологий связи интернета вещей // Информационные технологии и телекоммуникации. - 2016. - Т. 4. - № 4. – С. 67-80.

17. Smartphone applications in ophthalmology: A quantitative analysis / Lokeshwari Aruljyothi, Anuja Janakiraman, B Malligarjun, Balasundaram Manohar Babu // Indian Journal of Ophthalmology. – 2021. – Vol. 69, N 3. – P. 548 – 553. - doi: 10.4103/ij.o.IJO_1480_20.

ПРИМЕНЕНИЕ НЕЙРОННЫХ САМООРГАНИЗУЮЩИХСЯ СЕТЕЙ К ЗАДАЧЕ КЛАСТЕРИЗАЦИИ ДАННЫХ ДЛЯ ВЫЯВЛЕНИЯ ЗАКОНОМЕРНОСТЕЙ МЕЖДУ ЛИЧНОСТНЫМИ КАЧЕСТВАМИ И УСПЕВАЕМОСТЬЮ СТУДЕНТОВ

П. К. Ершов, Т. В. Авдеенко

Новосибирский Государственный Технический Университет,
г. Новосибирск, РМ162.Ershov@outlook.com

В работе рассматривается применение самоорганизующихся нейронных сетей в решении задачи кластеризации данных с целью выявления в психологических данных ключевых признаков.

The paper discusses the use of self-organizing neural networks in solving the problem of data clustering in order to identify key features in psychological data.

Для разработки цифровой среды современные университеты используют подход, основанный на актуальных педагогических и дидактических практиках, позволяющих учитывать индивидуальные психологические и интеллектуальные качества обучающихся. Для повышения качества обучающего процесса необходима персонализация образовательной среды. Целью исследования является выявления качества интеллекта и характера студента, которые влияют на его успеваемость. В дальнейшем, эти результаты можно будет использовать для разработки обучающих программ индивидуального характера. В качестве исследуемой выборки были использованы психологических данных студентов.

Для решения задачи будет использован один из вариантов нейронных сетей: самоорганизующиеся нейронные сети. Самоорганизующиеся нейронные сети имеют значительное преимущество перед другими типами нейронных сетей. Как понятно из названия, такие нейронные сети способны изменять свою структуру[1] в ходе процесса обучения, что позволяет им находить более оптимальные решения, чем сети с фиксированной структурой.

Сама сеть состоит из двух слоёв:

1. Первый слой получает на вход вектор данных и определяет топологию кластеров.
2. Второй слой получает на вход данные узлов-победителей первого слоя и определяет число кластеров.

Выборка состоит из психологических данных 68 студентов и данных об их успеваемости:

1. Тест Большая Пятёрка, показывающая пять признаков характера человека.
2. Краткий опросник Грей, показывающий два типа стимулов поведения.
3. Общий IQ тест из 9 признаков интеллекта.
4. Результаты успеваемости в гуманитарных, программистских, математических предметах и общая успеваемость в целом. Данные результаты не были использованы в кластеризации так как требовалось выявить влияние на успеваемость определённых признаков.

Было установлено, что в Большой пятёрке влияние на успеваемость оказывают только признаки Нейротизм и Добросовестность. В тоже время успеваемость изменяется же только для программистских предметов.

В Опроснике Грея влияние оказывает только BAS, система положительных стимулов. Подвержены влиянию результаты успеваемости среднего балла по всем предметам, программистские предметы и гуманитарные предметы.

Для IQ теста значимыми оказались только признак IQ3 – тест на поиск аналогий; IQ2 – тест на исключения слов; IQ9 – тест на память. Влияние оказывается на успеваемость по программистским и гуманитарным предметам.

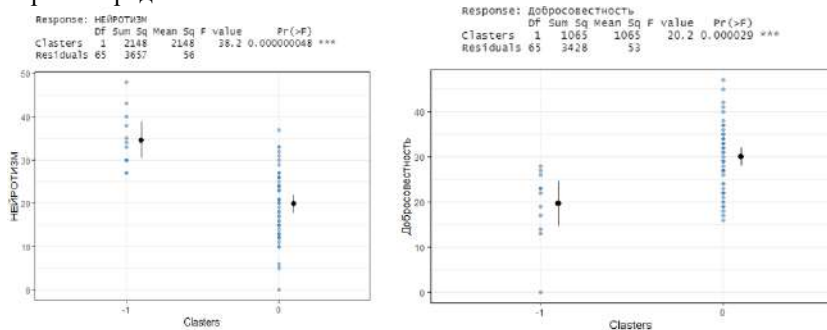


Рис. 1 – Результаты дисперсионного анализа для Большой Пятёрки

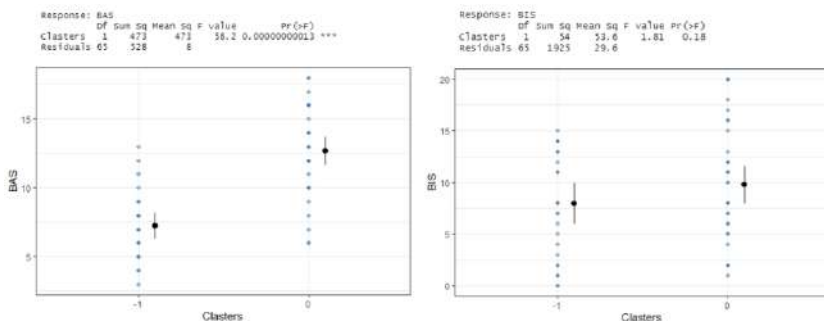


Рис. 2 - Результаты дисперсионного анализа для Теста Грея

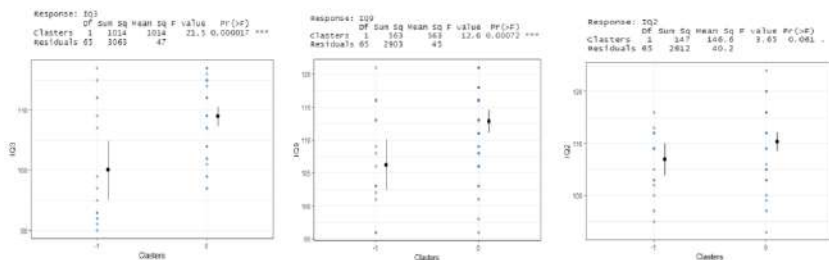


Рис. 3 - Результаты дисперсионного анализа для IQтеста

Таким образом, исходя из результатов, полученных в ходе исследований можно сделать следующие выводы, что основное влияние на успеваемость среди психологических факторов оказывают добросовестность, нейротизм, а так же положительная стимуляция. В тоже время, среди интеллектуальных качеств основное влияние показывают навыки поиска аналогий и исключения лишнего, а так же память.

Литература:

1. Сараев П.В. Нейросетевые методы искусственного интеллекта: учебное пособие / Сараев П.В. – Липецк: ЛГТУ, 2007. – 64 с.
2. Козин Н.Е., Фурсов В.А. Снижение размерности векторов признаков по критериям мультиколлинеарности // Компьютерная оптика, том 32. – 2008. – № 3. – С. 307-311.

ПРИМЕНЕНИЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ПРОГНОЗИРОВАНИЯ СТОИМОСТИ НЕДВИЖИМОСТИ

А.И. Павлова, А.А. Корж

Новосибирский государственный университет экономики и управления, г. Новосибирск, e-mail:annstab@mail.ru, ck1e.sty@gmail.com

В работе построена модель прогнозирования стоимости жилой недвижимости с применением машинных алгоритмов. Выполненный анализ методов: стохастический градиентный спуск (SGD), метод адаптивного градиента (Adagrad), метод адаптивного скользящего среднего градиентов (RMSprop), метод адаптивного шага обучения (Adadelata), метод Адама (Adam), метод RMSprop

The paper builds a model for predicting the value of residential real estate using machine algorithms. Methods analysis performed: Stochastic Gradient Descent (SGD), Adaptive Gradient Method (Adagrad), Adaptive Gradient Moving Average Method (RMSprop), Adaptive Learning Step Method (Adadelata), Adam Method (Adam), the RMSprop method.

Глубокие нейронные сети в настоящее время становятся одним из самых популярных методов машинного обучения. Способность к моделированию нелинейных процессов, работе с зашумленными данными и адаптивность дают возможности применять нейронные сети для решения широкого класса финансовых задач. Одной из актуальных проблем является прогнозирование стоимости жилой недвижимости. Применяемые для прогнозирования методы разнообразны.

Цель работы – анализ методов машинного обучения для прогнозирования стоимости жилой недвижимости.

В качестве исходных данных были использованы сведения о стоимости недвижимости, состоящий из 1460 строк и 10 столбцов. Набор данных содержит информацию о площади земельного участка, количестве спален, количество и качество ванных комнат, оценку общего качества жилья, оценку состояния жилой недвижимости, количество каминов, площади гаража, общее количество комнат.

Метод или методология проведения работы: в статье использованы методы машинного обучения обучения глубоких нейронных сетей: стохастический градиентный спуск (SGD), метод адаптивного градиента (Adagrad), метод адаптивного скользящего среднего градиентов (RMSprop), метод адаптивного шага обучения (Adadelata), метод Адама (Adam).

Модель реализована в виде глубокой нейронной сети с двумя скрытыми слоями. В качестве функций активации нейронов использованы сигмоидальная (Sigmoid) для выходного слоя и ReLu для промежуточных слоев. Обучение модели прогнозирования осуществилось на множестве данных, разделенных на три группы: обучающее (70% от общего числа примеров), тестовое (15%) и вариационное (15%).

Таблица 1 – Сравнение значений RMSE

	Обучение	Тестирование	Кросс-валидация
Adam	0.58	0.54	0.6
Adadelta	1.25	1.21	1.18
SGD	0.85	0.78	0.7
RMSprop	0.58	0.53	0.6
Adagrad	0.84	0.74	0.72

Таблица 2 – Результаты оценки точности нейронной сети

Результат обучения	Метод обучения сети				
	SGD	Adagrad	RMSprop	Adadelta	Adam
Общая оценка обучения	63%	85%	96%	90%	83%
Максимальная ошибка	0,65	1,97	1,94	1,65	1,97
Средняя абсолютная ошибка	0,47	0,79	0,66	1,00	0,65
Средняя квадратичная ошибка	0,22	0,72	0,66	1,25	0,65
Медианная абсолютная ошибка	0,47	0,97	0,94	0,62	0,97

Значение среднеквадратической ошибки составило 0,58 (для обучающей выборки), 0,53 (для тестовой выборки), 0,60 (для валидационной выборки) (таблица 1).

Для оценки точности использованы показатели: общая оценка обучения, максимальная ошибка, средняя абсолютная ошибка, средняя квадратичная ошибка и медианная абсолютная ошибка (таблица 2).

Закключение. В результате работы построена модель обучения нейронной сети для прогнозирования стоимости жилой недвижимости.

Анализ точности алгоритмов машинного обучения показал, что меньшие ошибки получены при использовании метода адаптивного скользящего среднего градиентов (RMSprop).

СРАВНЕНИЕ АЛГОРИТМОВ ПРОГНОЗА В ЗАДАЧЕ ОПРЕДЕЛЕНИЯ НЕРАВНОМЕРНОСТИ ВРАЩЕНИЯ ЗЕМЛИ

А.А. Кривецкий

**Новосибирский государственный технический университет,
г. Новосибирск, kriveczkij.2020@stud.nstu.ru
Научный руководитель: Карманов В.С., к.т.н.**

Рассмотрены векторный и рекуррентный варианты метода прогноза на основе спектрального сингулярного анализа. Реализованы алгоритмы, проведены сравнения точности прогноза.

The vector and recurrent versions of the forecast method based on spectral singular analysis are considered. Algorithms are implemented, comparisons of forecast accuracy are made.

Введение

Возникновение неравномерности вращения Земли обусловлено различными факторами. К наиболее значимым из них относятся: непостоянный момент импульса атмосферы, приливы и отливы мирового океана, диссипативные свойства земных недр [1]. Совокупность всех факторов дает сложную математическую модель процесса.

Параметры вращения Земли (ПВЗ) необходимы для расчета орбит космических аппаратов в режиме реального времени, также с помощью этих данных реализуется высокоточная навигация наземных объектов. К основным ПВЗ относятся координаты Северного полюса X_p, Y_p и поправка к координированному атомному времени dUT [2].

Одним из методов анализа временных рядов является метод SSA (Singular spectrum analysis), который заключается в выделении сингулярного базиса траекторного пространства временного ряда, преобразованного в многомерный ряд и представленного в виде траекторной матрицы [3, 4]. Цель данной работы состояла в сравнении алгоритмов прогнозирования ПВЗ, которые могут быть определены на основе метода SSA несколькими способами.

Постановка задачи

Для исследований выберем один из ПВЗ, а именно координату X_p . Рассмотрим временной ряд $F = (f_1, f_2, \dots, f_N)$ этого параметра, данные находятся в открытом доступе [5], шаг по времени – сутки, единица измерения – угловая секунда (секунда дуги). Требуется построить прогноз, т.е. продолжение ряда на M значений вперед.

Описание методов

Для сравнения алгоритмов прогноза, будем использовать два варианта метода: рекуррентный и векторный [4]. Оба алгоритма зависят от параметров L и r , где L – ширина временного окна для построения траекторной матрицы ряда, а $r < L$ – размерность подпространства значимых признаков, используемых при построении прогноза.

Рекуррентный прогноз $\tilde{F} = (\tilde{f}_{N+1}, \tilde{f}_{N+2}, \dots, \tilde{f}_{N+M})$ выполняется по формуле

$$\tilde{f}_i = \begin{cases} f_i, & i = \overline{1, N} \\ \sum_{j=1}^{L-1} a_j \tilde{f}_{i-L+j}, & i = \overline{N+1, N+M} \end{cases},$$

вектор коэффициентов $A = (a_1, a_2, \dots, a_{L-1})^T$ находится по формуле

$$A = \frac{1}{1 - \sum_{i=1}^r L(U_i)^2} \sum_{i=1}^r L(U_i) \cdot WL(U_i),$$

где U_i – левый сингулярный вектор траекторной матрицы ряда X , $L(U_i)$ – последняя координата вектора, $WL(U_i)$ – первые $L-1$ координат вектора.

Векторный прогноз производится по формулам

$$Z_i = \begin{cases} X_i & i = \overline{1, N-L+1} \\ P(Z_{i-1}) & i = \overline{N-L+2, N+M} \end{cases},$$

$$P(Z_i) = \begin{pmatrix} P^{(r)} \cdot WF(Z_i) \\ A^T \cdot WF(Z_i) \end{pmatrix},$$

$$P^{(r)} = [WL(U_1) \dots WL(U_r)][WL(U_1) \dots WL(U_r)]^T + (1 - \sum_{i=1}^r L(U_i)^2) AA^T,$$

где X_i – i -ый столбец тракторной матрицы временного ряда, Z_i – i -ый столбец прогнозной тракторной матрицы, P – матрица оператора формирования столбцов прогнозной тракторной матрицы, $P^{(r)}$ – матрица оператора проектирования на подпространство, заданного линейной оболочкой соответствующих левых сингулярных векторов без последней координаты, $WF(Z_i)$ – столбец прогнозной тракторной матрицы без первой координаты.

После формирования матрицы Z необходимо преобразовать её в вектор, выполнив диагональное усреднение по побочным диагоналям, при этом последние M значений и будут являться прогнозом исходного временного ряда.

Результаты расчетов

Для исследований был выбран набор данных, начальная дата 03.01.1994, конечная дата прогноза 24.04.2012. Объём обучающей выборки был выбран как $N = 3650$ значений (суток), по ней строился прогноз на $M = 100$ значений (суток) вперед.

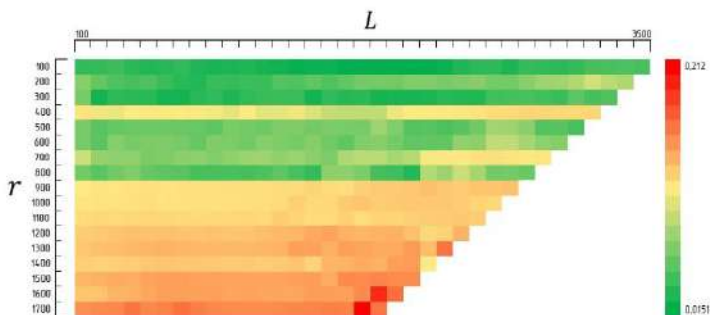


Рис. 1 – Тепловая карта средних по реализациям значений MAE, в зависимости от параметров алгоритма прогноза.

Далее рассчитывалась разность между прогнозным и известным фактическим значением ряда. После этого происходил сдвиг начальной даты на 1 сутки и процедура повторялась заново. В итоге получилось $R = 2937$ реализаций прогнозов на 100 суток и их сравнений с фактическими значениями.

На первом этапе с использованием дихотомического поиска были найдены значения оптимальных параметров L и r . Для оценки точности использовалась средняя абсолютная ошибка (MAE, Mean Absolute Error). Далее по множеству реализаций строилась тепловая карта зависимости средней ошибки от параметров L и r (Рис. 1). Пример прогнозирования по одной реализации приведен на Рис. 2.

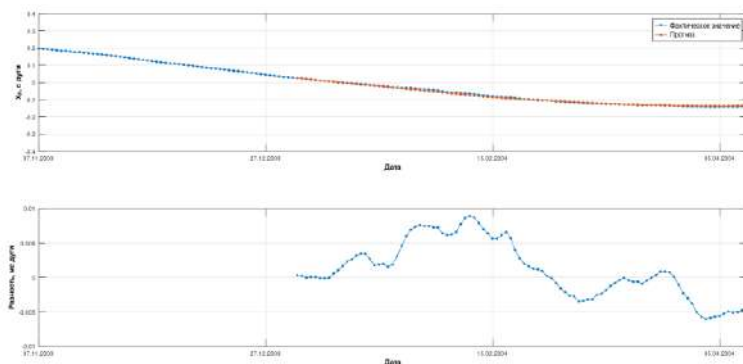


Рис. 2 – Пример реализации рекуррентного прогноза

После выбора параметров алгоритма качество прогнозирования оценивалось по абсолютной ошибке по уровню вероятности 0,95 (95%-квантиль), см. Табл. 1.

Таблица 1 – Показатель точности алгоритма, 95%-квантиль абсолютной погрешности для различных интервалов прогноза

Сутки прогноза	Рекуррентный алгоритм, мс дуги	Векторный алгоритм, мс дуги
10	7,5945	106,73
30	15,435	108,58
50	19,446	103,2
70	24,407	105,29
90	27,538	102,27

Анализ результатов расчетов показывает, что рекуррентный алгоритм показывает точность выше в несколько раз по сравнению с векторным.

Заключение

В результате исследования были рассмотрены два алгоритма прогнозирования временных рядов. Проведено исследование по влиянию параметров на точность, были определены наиболее подходящие парамет-

ры на примере задачи прогнозирования ПВЗ. Планируется проведение дальнейших исследований для сравнения с результатами других методов.

Литература:

1. Физика нестабильностей вращения Земли / Н.С. Сидоренков / Изд-во Физматлит; 2001: 376 с.

2. Современные и перспективные информационные ГНСС-технологии в задачах высокоточной навигации / В.А. Бартенев, А.К. Гречкосеев, Д.А.Козорез, М.Н. Красильщиков / Изд-во Физматлит; 2014: 275 с.

3. Модификация базового метода сингулярного спектрального анализа для повышения точности прогнозирования неравномерности вращения Земли / А. К. Гречкосеев, А. С. Толстиков, В. М. Тиссен, В. С. Карманов, А. И. Ваганова // Вычислительные технологии, 2020, том 25, № 3, с. 54–65.

4. Голяндина Н.Э. Метод “Гусеница”-SSA: прогноз временных рядов: Учеб. пособие. СПб.: Изд-во СПбГУ; 2004: 52 с.

5. Earth orientation data [Электронный ресурс]. Режим доступа: <https://www.iers.org/IERS/EN/DataProducts/EarthOrientationData/eop.html> (дата последнего обращения: 22.11.2021).

РАЗРАБОТКА СИСТЕМЫ РАСПОЗНАВАНИЯ ОБЪЕКТОВ С ПОМОЩЬЮ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ

И.М. Кузьмина

**Северо-Восточный федеральный университет имени
М.К. Аммосова, г. Якутск, irakuzmina199973@gmail.com
Научный руководитель: Иванова М.А., к.ф.-м.н., доцент**

Работа посвящена изучению сверточных нейронных сетей для разработки системы распознавания лица. Результаты работы могут найти применение в системах безопасности и видеоконтроля.

The work is devoted to the study of convolutional neural networks for the development of a face recognition system. The results of the work can find application in security and video control systems.

Система распознавания лиц используется во многих сферах, прежде всего в сфере безопасности. На данный момент уже существуют системы, которые позволяют найти людей, находящихся в перечне разыски-

ваемых. Компания «Apple» использует для собственных товаров в качестве PIN-кода систему распознавания лиц. Кроме того, есть приложения, которые позволяют находить людей, в точности похожих на человека, изображенного на фотографии.

Для такой системы присущи следующие требования:

- ✓ высокая скорость работы;
- ✓ высокая точность;
- ✓ масштабируемость.

Целью работы является реализация и обучение глубокой сверточной нейронной сети. Также проанализировать влияние параметров обучения, архитектуры нейронной сети, алгоритмов выбора изображений для обучения на точность определения.

Сверточные нейронные сети отлично подходят для работы с изображениями, имеют повышенную устойчивость к разным искажениям входных данных, но и способны достаточно эффективно описывать сложную функцию распределения объектов на изображениях.

Объединяя три архитектурные сети, сверточные нейронные сети обеспечивают инвариантность к изменению масштаба, повороту, сдвигу и пространственным искажениям:

- локальные рецепторные поля (обеспечивают локальную двумерную связность нейронов);
- общие синаптические коэффициенты (обеспечивают детектирование некоторых черт в любом месте изображения и уменьшают общее число весовых коэффициентов);
- иерархическая организация с пространственными подвыборками.

Поэтому в данной работе применялась сверточная нейронная сеть, которая основана на принципах неокогнитрона и дополнена по алгоритму обратного распространения ошибки.

Нейронная сеть в самой простой форме представляет собой перцептрон, который имеет входные и выходные данные и выполняет только самые простые функции.

Во многом, нейронная сеть похожа на человеческий мозг. Прежде чем распознать объект она должна обучиться на данных, которые мы предоставляем. Только обладая некоторыми данными и знаниями, мы можем идентифицировать что-то.

Для обучения нейронных сетей существуют специальные базы данных. Чтобы обучение прошло эффективно берем данные из внешней среды.

И на данный момент существуют множество алгоритмов для обучения, и их основной принцип состоит в минимизации эмпирических ошибок.

Из этих алгоритмов на основе исследования были выбраны следующие алгоритмы, основываясь на том, что они наиболее эффективные и востребованные:

- алгоритм обратного распространения ошибки (Backpropagation Algorithm);
- алгоритм Quick Propagation (QuickProp);
- алгоритм Resilent Propagation (RProp).

Далее после выбора алгоритмов идёт обучение разработанной нейронной сети. Критерий прекращения обучения будет служить изменение ошибки после 16 итераций. И на основе этого критерия делаем вывод, что наименьшую ошибку в обучении сверточной нейронной сети будет тот или иной алгоритм.

Следующий процесс- это тестирование алгоритма для обнаружения лиц. Для это взяли базу, которая состоит из 2000 изображений. Они были со сложными фонами и присутствовали также другие объекты.

Для более наглядного сравнения результатов, было решено проводить обучение нейронной сети несколькими алгоритмами, которые были описаны ранее.

Точность классификации вычислялась как отношение количества правильных классификаций к их общему количеству.

Тестирование выполнялось на следующие этапы:

1. Тестирование с различным уровнем освещенности.
2. Тестирование с различным уровнем отклонения лица на плоскости.
3. Тестирование с различным масштабом изображения лица.

При первом тестировании имело место быть зависимость точности распознавания от изменения интенсивности освещения. Т.е. какая будет точность, если менять уровень освещенности. И на основе этого выбрать средний диапазон.

Тестирование с различным уровнем отклонения лица на плоскости подразумевала в себе изучение зависимости точности распознавания лица от угла поворота изображения.

Третье тестирование основана на определении объекта, если взять изображение(лица) в различных масштабах.

Таким образом, при обучении сверточной нейронной сети применялись три алгоритма, которые выбирались по двум критериям- эффективность и точность. Для точности алгоритмов были учтены следующие

зависимости: от уровня освещенности, от уровня масштаба и от отклонения(угла) объекта.

Способность обучить нейронную сеть приводит к тому, что их можно оптимизировать и максимально увеличить функциональность.

Исследование нейронных сетей относится к самым перспективным областям в настоящее время, а в будущем они будут применяться практически везде. Они способны значительно облегчить труд, даже обезопасить человека.

Литература:

1. *Кухарев Г.А.* Поиск изображений лиц в больших базах данных // Мир измерений. – 2009. – № 4. – С. 22-30.

2. *Лентланд А.* Распознавание лиц для интеллектуальных сред // Открытые системы. – 2000. – № 3. – С.17-20.

3. *Кухарев Г.А.* Биометрические системы: Методы и средства идентификации личности человека. – СПб.: Политехника, 2001.– 240 с.

4. *Саймон Х.* Нейронные сети: полный курс, 2-е издание Пер. с англ. – М.: Издательский дом «Вильмс», 2006. – 1104с.

ПОКАДРОВОЕ ОПРЕДЕЛЕНИЕ ЭМОЦИЙ НА ВИДЕОЗАПИСИ ПРИ ПОМОЩИ МНОГОСЛОЙНЫХ НЕЙРОННЫХ СЕТЕЙ

В.В. Макаров

ФИЦ ИУ РАН, г. Москва, viktor.makarov@phystech.edu

Научный руководитель: Цурков В.В., д.ф.-м.н., профессор

Предложена архитектура многослойной нейронной сети для решения задачи определения эмоции человека на видеозаписи. Сформированы соответствующие признаки. Под эмоциями понимаются страх, радость, грусть и др. Для обучения используются известные и широко применяемые наборы данных. С помощью комбинирования в архитектуре удается доучить нейросеть и повысить точность ее окончательной версии.

The architecture of a multilayer neural network is proposed for solving the problem of determining human emotion on video. Corresponding signs have been formed. Emotions are understood as fear, joy, sadness, etc. Well-known and widely used datasets are used for training. Combination of architectures have been increased the accuracy of its final version.

Введение. Распознавание эмоций человека по мимике лица является важной научно-исследовательской проблемой, которая затрагивает множество дисциплин и областей [1-3]. Эта тематика актуальна в таких сферах, как медицина, психология [4] и безопасность. В данной работе рассматриваются подходы к распознаванию эмоций человека по визуальным признакам лица. Применяется глубокое обучение многослойных нейронных сетей.

1. Краткий обзор существующих подходов. Наиболее точно на сегодняшний день эмоции человека были описаны Полом Экманом в работе [5], где каждая эмоция была представлена при помощи кодирования лицевых движений, но данный подход сложно автоматизировать [6, 7]. Это связано с тем, что он содержит 46 основных категорий и более 50 дополнительных, а эмоции являются комбинациями таких групп. Поэтому для подготовки такие датасеты (пронумерованный набор изображений, фонограмм или видеозаписей с указанием исследуемых признаков каждого элемента) обрабатываются психологами вручную с учетом указанных обстоятельств.

Особый интерес представляет датасет Aff-Wild [8-10], который состоит из фрагментов видеороликов платформы YouTube, они не являются предзаписанными в видеостудии. Это значит, что условия записи материалов максимально приближены к действительности.

Существует два принципиально разных подхода в распознавании эмоций:

- с предварительным алгоритмическим извлечением визуальных признаков и последующей машинной классификацией [11],
- с использованием глубоких нейронных сетей без предварительного извлечения признаков.

Визуальные признаки могут быть извлечены при помощи выявления:

- геометрических объектов лица (брови, нос, рот, глаза и др.) [12, 13] методами дескриптора line edge map, сравнения направленности градиентов [14, 15], активной модели формы ASM [16], кривлет-преобразования [17], использования структурных моделей [18] и др.;
- текстурных особенностей методами фильтра Габора, дискретным вейвлет-преобразованием [19] и др.;
- глобальных и локальных объектов методами главных компонент [20], оптического потока [21], морфологическими преобразованиями [22] и др.

Но при условии наличия достаточного датасета наиболее высокой точности классификации удастся достичь именно при помощи автома-

тического выявления признаков и классификацией глубокими нейронными сетями [23-25].

2. Комбинирование архитектурой нейронных сетей и датасетов. В рамках проведения экспериментов были получены следующие результаты тестирования нейронной сети для видеофиксации.

- Датасет RAVDESS, точность – 69.35%.
- Датасет СК+: точность – 82.3%
- Датасет Aff-Wild: точность – 60.7%.

Здесь в каждой из трех групп вводятся неотрицательные коэффициенты (вероятности), сумма которых равна единице. Цель распознавания - нахождение максимального коэффициента.

Остановимся на результатах датасета Aff-Wild. Он является единственным из представленных, в котором материал для анализа содержит изображения в различных ракурсах. Из этого делается следующий вывод: в датасетах СК+ и RAVDESS наличие схожих кадров лиц неизбежно привело бы к ухудшению показателей результата.

На предыдущем этапе разработки была создана и обучена нейросеть, которая показывает хорошую точность в условиях, близким к идеальным: при правильном освещении, фоне, расстоянии от камеры до лица. Но при ухудшении условий точность результатов падает. Поэтому следующей задачей является создание нейросети, которая обучена уже на данных, приближенных к реальным.

Для обучения был выбран датасет Aff-Wild, который участвовал в предыдущем этапе разработки, но в качестве датасета результативной выборки. Сама организация процесса обучения выглядит следующим образом.

Обучаем первичную сеть на простой задаче, удаляя первый класс. При этом датасет сбалансирован, так как в исходной версии было много нейтральных кадров с нулевой результативностью. На этом этапе кадры обрабатываются по отдельности без последовательности. Точность на этом этапе составляет 72%.

Из обученной в предыдущем этапе нейросети убираем последние слои, дойдя до слоя укрупнения (max pool) с предыдущего блока слоев. Полученную нейросеть делаем первичной. После чего пропускаем через эту нейросеть датасет RAVDESS и сохраняем найденный промежуточный результат.

Таким образом получаем промежуточный датасет для обучения нейросетей по определению эмоций и их силы. В результате такой комбинации архитектур нейросети будут достаточно точны при пропуске через них материалов, приближенных к реальным условиям, потому

что первичная нейросеть уже умеет с ними работать, а студийные условия выступают лишь как частный случай.

Нейросети должны объединять в себе точность, скорость работы и простоту реализации. Работа с изображением все так же предполагает работу со сверточными нейронными сетями в силу их спецификации. Точность сети можно предварительно оценить, исходя из результатов теста cifar-10. В тесте убирается первый слой, а после через нейросеть пропускаются картинки размером 32x32, разбитые на 10 классов.

Особенность обеих архитектур состоит в том, что градиент ошибки, являющийся фактором обучения, не угасает от слоя к слою, а равномерно обучает все слои сети. Кроме того, обе архитектуры используют после каждого слоя (либо перед каждым слоем) нормализацию внутри партии. Это значит, что из исходных данных вычитается среднее, а отклонение делается равным единице. Этот процесс заметно стабилизирует и ускоряет обучение и заодно повышает точность, однако замедляет работу примерно на 30%.

В отличие от стандартной архитектуры, из модели удалены полностью связанные слои, образующие итоговый набор классов. Вместо них при достижении слоев размеров менее 5x5 происходит усреднение внутри слоев.

Контрольная выборка была взята из расширенного датасета RAVDESS, в которой три актера произносили одну фразу с разными эмоциональными оттенками. Итоговая нейросеть правильно классифицировала 91 файл из 100, тем самым можно сделать вывод об увеличении точности на 21% по сравнению с предварительными итогами.

Закключение. Последующее исследование будет связано с анализом и разработкой нейронных сетей для классификации эмоциональных состояний в фонограммах. Совместное применение аудио- и видео-систем увеличит точность распознавания.

Литература:

1. Александров А.А., Кирпичников А.П., Ляшева С.А., Шлеймович М.П. Анализ эмоционального состояния человека на изображении // Вестн. Технологического ун-та. 2019. Т. 22. № 8. С. 120–123.

2. Заболеева-Зотова А. В. Развитие системы автоматизированного определения эмоций и возможные сферы применения // Открытое образование. 2011. № 2. С. 59–62.

3. Люсин Д.В. Современные представления об эмоциональном интеллекте // Социальный интеллект: теория, измерение, исследования / Под ред. Д.В. Люсина, Д.В. Ушакова. М.: Изд-во Ин-та психологии РАН, 2004. С. 29–36.

4. Гранская Ю.В. Распознавание эмоций по выражению лица: Автореф. Дис. ... канд. Психологических наук по специальности 09.00.01. СПб., 1998.
5. Ekman P. Facial Action Coding System. Palo Alto, USA: Consulting Psychologist Press, 1978.
6. Kollias D., Zafeiriou, S. Expression, Affect, Action Unit Recognition: Aff-Wild2, Multi-Task Learning and ArcFace // arXiv preprint arXiv:1910.04855, 2019.
7. Kollias D. Face Behavior a la carte: Expressions, Affect and Action Units in a Single Network // arXiv preprint arXiv:1910.11111, 2019.
8. Kollias D. Analysing Affective Behavior in the First ABAW 2020 Competition // arXiv preprint arXiv:2001.11409, 2020.
9. Kollias D. Deep Affect Prediction in-the-wild: Aff-Wild Database and Challenge, Deep Architectures, and Beyond // International J. Computer Vision (IJCV). Berlin, Germany, 2019. №127. P. 907-929
10. Kollias D. Distribution Matching for Heterogeneous Multi-Task Learning: a Large-scale Face Study // arXiv preprint arXiv:2105.03790, 2020.
11. Abdulrahman M., Eleyan A. Facial Expression Recognition Using Support Vector Machines // Proc. 23rd Signal Processing and Communications Applications Conf. (SIU 2015). Malatya, Turkey, 2015. P. 276–279.
12. Tripathi A., Pandey S. Efficient Facial Expression Recognition System Based on Geometric Features Using Neural Network // Lecture Notes in Networks and Systems. 2018. V. 10. P. 181–190.
13. Hernandez-Matamoros A., Bonarini A., Escamilla-Hernandez E., Nakano-Miyatake M., Perez-Meana H. A Facial Expression Recognition with Automatic Segmentation of Face Regions // Communications in Computer and Information Science. 2015. V. 532. P. 529–540.
14. Jumani S.Z., Ali F., Gurirro S., Kandhro I.A., Khan A., Zaidi A. Facial Expression Recognition with Histogram of Oriented Gradients Using CNN // Indian J. Science and Technology. 2019. V. 12. N 24. P. 1–8.
15. Greche L., Es-Sbai N., Lavendelis E. Histogram of Oriented Gradient and Multi Layer Feed Forward Neural Network for Facial Expression Identification // Proc. Intern. Conf. on Control, Automation and Diagnosis (ICCAD 2017). Hammamet, Tunisia, 2017. P. 333–337.
16. Iqtait M., Mohamad F.S., Mamat M. Feature Extraction for Face Recognition Via Active Shape Model (ASM) and Active Appearance Model (AAM) // IOP Conf. Series: Materials Science and Engineering. Tangerang Selatan, Indonesia, 2018. V. 332. P. 1-8.

17. Candès E., Demanet L., Donoho D., Ying L. Fast Discrete Curvelet Transforms // *Multiscale Modeling & Simulation*. 2006. V. 5. №3. P. 861–899.

18. Себряков Г.Г., Визильтер Ю.В. Разработка методики построения специализированных экспертных систем для анализа цифровых изображений в задачах обнаружения и идентификации сложных структурных объектов // *Вестн. Компьютерных и информационных технологий*. 1997. № 3. С. 31.

19. Nigam S., Singh R., Misra A.K. Efficient Facial Expression Recognition Using Histogram of Oriented Gradients in Wavelet Domain // *Multimedia Tools and Applications*. 2018. V. 77. N 21. P. 28725–28747.

20. Varma S., Shinde M., Chavan S.S. Analysis of PCA and LDA Features for Facial Expression Recognition Using SVM and HMM Classifiers // *Techno-Societal 2018: Proc. 2nd Intern. Conf. on Advanced Technologies for Societal Applications*. Berlin, Germany, 2019. V. 1. P. 109–119.

21. Zhao J., Mao X., Zhang J. Learning Deep Facial Expression Features from Image and Optical Flow Sequences Using 3D CNN // *Visual Computer*. 2018. V. 34. №10. P. 1461–1475

22. Визильтер Ю.В., Выголов О.В., Желтов С.Ю., Князь В.В. Метрический подход к семантико-морфологическому сравнению изображений // *Вестн. Компьютерных и информационных технологий*. 2020. Т.17. № 5(191). С. 3-12.

23. Рюмина Е.В., Карпов А.А. Аналитический обзор методов распознавания эмоций по выражениям лица человека // *Научно-технический вестник информационных технологий, механики и оптики*. 2020. № 2. С. 163-176.

24. Talegaonkar I., Joshi K., Valunj S., Kohok R., Kulkarni A. Real Time Facial Expression Recognition Using Deep Learning // *Proc. Of Intern. Conf. on Communication and Information Processing (ICCIP)*. 2019 [Электронный ресурс]. URL: <https://ssrn.com/abstract=3421486>.

25. Визильтер Ю.В., Горбацевич В.С., Желтов С.Ю. Структурно-функциональный анализ и синтез глубоких конволюционных нейронных сетей // *Компьютерная оптика*. 2019. Т.43. №5. С.886-900

ПРОГНОЗ ОСТАТОЧНОГО ВРЕМЕНИ БЕЗОТКАЗНОЙ РАБОТЫ ТУРБОВЕНТИЛЯТОРНЫХ ДВИГАТЕЛЕЙ С ИСПОЛЬЗОВАНИЕМ РЕКУРРЕНТНЫХ НЕЙРОННЫХ СЕТЕЙ

Е.Р. Нестерова

Научный руководитель: д.т.н., доцент Е.В. Чимитова
Новосибирский государственный технический университет,
г. Новосибирск, nesterovaelenarus@mail@yandex.ru

Турбовентиляторный двигатель – важнейший компонент самолета. Чтобы обеспечить его надлежащее техническое обслуживание с целью повышения надежности системы, необходимо оценить остаточное время безотказной работы (Remaining Useful Life, RUL) двигателя. В данной работе реализована модель рекуррентной нейронной сети с долгой кратковременной памятью (Long short-term memory, LSTM) для прогнозирования RUL турбовентиляторных двигателей. В качестве входных данных используются измерения с датчиков за текущий временной промежуток и момент времени, на который рассчитывается прогноз RUL. Проведено исследование влияния ширины временного промежутка на точность получаемого прогноза.

The turbofan engine is an essential component of an aircraft. To ensure that it is properly maintained to improve the system reliability, the engine's Remaining Useful Life (RUL) must be assessed. In this paper, a long short-term memory (LSTM) recurrent neural network model is implemented for predicting the RUL of turbofan engines. Measurements from sensors for a time interval and a point in time are used as input data. The influence of the width of the time interval for the accuracy of the forecast obtained has been studied.

Введение

В области транспорта безопасность авиации является одной из проблем, которые больше всего беспокоят людей, поскольку в случае аварии — это может привести к катастрофическим последствиям и материальному ущербу. Турбовентиляторный двигатель – важный элемент 207вмолета, обеспечивающий его тягу. Прогнозирование остаточного времени работы для такого двигателя может повысить надежность системы и максимально продлить срок его службы.

Для прогнозирования остаточного времени работы турбовентиляторных двигателей можно использовать различные методы машинного обучения, такие как лес решений [1], метод опорных векторов [2], скры-

тые Марковские модели [3] и другие. Среди подходов, опирающихся на данные, нейронная сеть является одной из самых современных моделей для решения различных задач прогнозирования. В качестве исследуемой модели будет использоваться рекуррентная нейронная сеть, а именно, сеть с долгой краткосрочной памятью LSTM, которая способна запоминать предыдущую информацию о работе циклов эксплуатации.

Таким образом, целью данной работы является построение рекуррентной нейронной сети LSTM для прогнозирования остаточного времени безотказной работы турбовентиляторных двигателей.

1. Рекуррентная нейронная сеть LSTM

Рекуррентные нейронные сети – нейронные сети, в которых связи между элементами образуют направленную последовательность. Благодаря чему появляется возможность обрабатывать серии событий во времени или последовательные пространственные цепочки. Наиболее распространенной архитектурой рекуррентных нейронных сетей является LSTM-сеть. Запоминание информации на продолжительный срок – это одна из основных особенностей LSTM сетей, не требующая продолжительного обучения.

LSTM можно рассматривать как блок памяти, в котором есть три гейта (гейт забывания, входной гейт и выходной гейт) и состояние ячейки [4]. Определяются W_f , W_i , W_o и W_c как весовые матрицы гейта забывания, входного гейта, выходного гейта, состояния ячейки и соответствующие вектора смещения b_f , b_i , b_o и b_c .

Гейт забывания f_t принимает значения предыдущего выхода h_{t-1} и текущего входа x_t , они пропускаются через сигмоиду (σ).

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$$

Информация, поступающая в ячейку, контролируется входным гейтом.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$$

Выходная информация ячейки регулируется выходным гейтом.

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$$

После преобразования гиперболическим тангенсом \tanh , создается вектор новых значений-кандидатов, которые могут быть добавлены в ячейку состояния.

$$\begin{aligned}\tilde{C}_t &= \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \\ h_t &= o_t * \tanh(C_t)\end{aligned}$$

Значение старой ячейки состояния C_{t-1} обновляется до новой ячейки C_t .

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$

2. Подготовка данных

Для обучения нейронной сети использовались данные моделирования деградации турбовентиляторного двигателя, предоставленный NASA [5]. Набор данных C-MAPSS (Commercial Modular Aero-Propulsion System Simulation) состоит из четырех дополнительных наборов данных. Каждый набор состоит из 28 столбцов, которые включают номер двигателя, временной цикл, 3 признака, показывающие условия эксплуатации турбовентиляторных двигателей, 23 сенсора, показывающие измерения датчиков машин NASA. Для сравнения полученных результатов модели в наборе так же имеются данные с фактическим значением RUL двигателей. В данном исследовании будет использован только первый поднабор данных «FD001».

Так как некоторые признаки не изменяются во времени, был произведен отбор признаков, в результате которого были исключены 10 признаков.

Размер временного окна (TW) будет определять длину последовательности данных, которые будут отправлены в модель. Учитывая минимальный рабочий цикл, заданный тестовым набором данных, максимальный размер временного окна составляет 30 временных циклов.

Исходные данные нарезаются на последовательности равные ширине окна, значение целевой переменной будет определяться, как значение RUL для TW-го цикла. Следующая последовательность будет формироваться аналогично, после сдвига окна на цикл вперед.

3. Обучение модели

Модель прогнозирования принимает на вход последовательность из TW циклов работы двигателя. На выходе модели дается значение RUL для TW-го цикла конкретной последовательности. Для обучения модели был произведен подбор оптимальных гиперпараметров. Итоговая модель состоит из LSTM-слоя и 3-х Dense-слоев.

Используемый оптимизатор – Adam (adaptive moment estimation). Функция потерь MSE.

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

4. Обучение модели

Для оценивания точности прогнозирования RUL на основе полученной модели использовалась величина RMSE.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} = \sqrt{MSE}$$

На Рисунке 1 представлены отсортированные прогнозируемые результаты RUL для 100 экземпляров тестируемого двигателя с фактическими RUL, соответственно.

Влияние размера временного окна показано на рисунке 2. Из рисунка видно, что чем меньше временное окно, тем точнее результат прогнозирования, что можно объяснить тем, что в случае с уменьшением временного окна, объем обучающей выборки увеличивается.

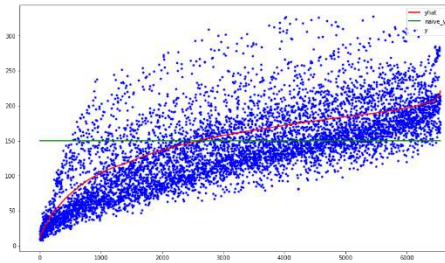


Рисунок 1 – предсказанные и фактические значения RUL

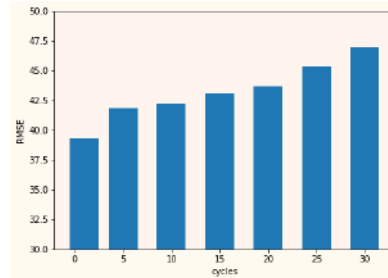


Рисунок 2 – зависимость точности модели от ширины временного окна

Заключение

Была реализована нейронная сеть для прогнозирования RUL турбовентиляторных двигателей. RMSE итоговой модели равна 39,33. В результате исследований влияния ширины временного окна на точность модели было показано, что с увеличением временного окна точность прогнозирования падает. В дальнейшем планируется провести сравнение нескольких методов прогнозирования, в частности, леса решений, градиентного бустинга и др.

Литература:

1. Xin Chen et al. Direct Remaining Useful Life Estimation Based on Random Forest Regression // 2020 Global Reliability and Prognostics and Health Management (PHM-Shanghai). 2021. – С. 7.
2. Racha Khelif et al. Direct Remaining Useful Life Estimation Based on Support Vector Regression // IEEE Transactions on Industrial Electronics. 2016. – Vol.64. – P. 2276 – 2285
3. Emmanuel Ramasso, Rafael Gouriveau. Remaining Useful Life Estimation by Classification of Predictions Based on a Neuro-Fuzzy System and Theory of Belief Functions // IEEE Transactions on Reliability. 2014 – Vol.63. – P. 555 - 566
4. Николенко С.И., Кадурын А.А., Архангельская Е.В. Глубокое обучение. Погружение в мир нейронных сетей. СПб.: Питер, 2018. – 480 с.
5. NASA [Электронный ресурс]. URL: <https://ti.arc.nasa.gov/tech/dash/groups/pscoe/prognostic-datarepository/#turbofan> (дата обращения: 08.10.2021).

**ПРЕДСКАЗАНИЕ АКТИВНОСТИ ХИМИЧЕСКОГО
СОЕДИНЕНИЯ С ПОМОЩЬЮ ГРАФОВОЙ НЕЙРОСЕТЕВОЙ
МОДЕЛИ**

Б.И. Пякилля

**Томский политехнический университет, г. Томск, morphism@tpu.ru
Научный руководитель: Гончаров В. И., д.т.н., профессор**

Методы машинного обучения, основанные на нейросетях, показали свою эффективность в задачах анализа большого количества данных и поиска нелинейных закономерностей. Для приложений в области химии особенно интересны задачи обучения с учителем, которые связаны с классификацией химических соединений. Примером такой задачи является задача биохимии, где приходится оценивать активность того или иного химического соединения для различных сигнальных путей ядерных рецепторов. Таким сигнальным путем, к примеру, является сигнальный путь Андрогенного рецептора (AR или NR3C4), рецептора стероидных гормонов. Данный рецептор влияет как на целостность мужского скелета, так и в целом на способность к репродукции организма. Однако, задача определения этой активности является сложной из-за вида биохимических процессов, которые определяют поведение рецептора в

связке с химической молекулой. Решением является использование большого количества размеченных данных и построение на них нейросетевой модели, которая обладает возможностью учета структуры химической молекулы в виде графа. Для такой цели подходит графовая сверточная сеть, оптимизации которой посвящена данная статья.

Machine learning methods, based on neural networks, showed efficiency in numerous problems of big data analysis and nonlinear functions search. The most interesting applications of supervised problems in chemistry are chemical compounds classification. For example, in biochemistry we must estimate an activity of different chemical compounds for a diverse set of nuclear receptor signal paths. As a target we can take Androgen receptor (AR or NR3C4), which is steroid hormone receptor. This receptor affects an integrity of a masculine body skeleton and reproductive ability. But the task of activity classification is complicated due to biochemical processes, which affect link between a chemical compound and the receptor. The solution lies in use of large amount of data and training of neural net model, which can recognize chemical compound structure based on its graph representation. For this task graph convolutional network is an appropriate choice and this paper directly describes the optimization of the network structure.

Нейросетевые методы давно вошли в арсенал основных инструментов специалистов различных областей естественных наук. Способность нейросетевых методов не только качественно интерполировать существующие экспериментальные зависимости, но и успешно экстраполировать на новые случаи заслуженно обратила на себя внимание ученых из областей, где особенно трудны повторные эксперименты. Трудность их проведения связана как с дороговизной процедуры, так и с невозможностью прямого наблюдения исследуемого процесса либо его сложностью. Примером области знаний, которая в последнее время стала активно использовать методы машинного обучения, является биохимия - наука, изучающая химию живых клеток и организмов.

Как и любую из сегодняшних наук, биохимию не обошли современные методы, связанные с информатикой; примером такого синтеза химии и компьютерных наук, является хемоинформатика. В задачи хемоинформатики входят компьютерные методы анализа почти любых химических процессов, таких, к примеру, как реакция различных рецепторов на химические соединения, которые используются в фармацевтической промышленности. При создании всевозможных лекарственных средств, важно учитывать их возможную токсичность, которая может выражаться в ненужной активации некоторых жизненно важных рецеп-

торов. Характерным примером такого рецептора является Андрогенный рецептор, необходимый для взаимодействия с тестостероном и дигидротестостероном. Сложность предсказания последствий применения того или иного соединения, а иногда и невозможность проведения лабораторного эксперимента (или его дороговизна) для этого рецептора приводят к тому, что все больше внимания уделяется компьютерным экспериментам с использованием методов машинного обучения. Данные методы, образующие специальный класс – обучение с учителем – позволяют экстраполировать поведение рецептора на те случаи, когда неизвестны последствия применения того или иного химического соединения. В статье рассматривается метод, показавший свою эффективность в данном типе задач – графовые нейронные сети.

Графовые нейронные сети

Графовая нейронная сеть (GNN) есть вариант нейронной сети, который работает напрямую со структурой графа, представляющей собой набор вершин и ребер. Граф G описывается множеством вершин (узлов) V и ребер E , которые он содержит: $G = (V, E)$. Идея графовой нейросетевой модели была впервые опубликована в 2009 году, в работе [1], расширившая существующие нейронные сети для обработки данных, представленных в графовых областях. Использование графовой нейронной сети дает возможность работать с графами, сохранять их пространственные отношения. Это достигается тем, что каждый узел (в случае химии это атом) представляется в виде вектора, значения которого обновляются с учетом соседних узлов, что позволяет учитывать топологию, и то, насколько значения этого вектора позволяют предсказывать свойства графа (в случае, когда граф описывает молекулу, примером такого свойства может быть липофильность) [2]. В связи с высокой производительностью в задачах анализа социальных графов, логистических карт и химических соединений, системы, основанные на варианте графовых нейронных сетей, получили большое распространение. Для нашей задачи наиболее интересным будет модификация GNN, называемая графовой сверточной сетью (GCN) [3]. Данная сеть работает с входной информацией, которой является граф связей между молекулой лигандом и белком [4]. Функцией такой сети является предсказание побочных эффектов взаимодействия лекарства и целевого белка, которые разбиты на классы.

Использование свертки позволяет получать сжатое, но информативное представление графа, которое уже может быть использовано как входная информация для полносвязной части нашей нейронной сети, функцию которой выполняет обычно многослойный персептрон. Далее

рассмотрим конкретное применение графовой сверточной сети в задаче предсказания активности андрогенного рецептора.

Описание задачи и архитектуры сети

Рецептором, активность которого требуется предсказать, является Андрогенный рецептор. Данный тип рецептора взаимодействует с тестостероном и дигидротестостероном, которые через этот сигнальный путь, влияют на целостность скелета, клетки сосудистого эндотелия и гладких мышц, а также на процесс ангиогенеза. В нашей задаче активность будет определяться просто – в виде ее отсутствия или наличия, т. е. размеченные активности будут принимать бинарные значения. Входными же данными будут являться строки SMILES, которые представляют собой популярный пример текстового представления химических соединений [5]. Для оценки качества предсказаний будем пользоваться ROC-AUC оценкой [6] на тестовом множестве, которое будет иметь размер в виде 20% от всех данных датасета Tox21 для NR-AR [7] т. е. 1453 соединения, а обучающее множество будет включать в себя 5812 соединений.

Архитектуру нейросетевой модели определим следующим образом:

- 1) Два графовых сверточных слоя по N выходных каналов;
- 2) Два полносвязных слоя по 256 и 2 нейронов.

Определенная выше неизвестная величина N отражает количество выходных каналов после каждого сверточного слоя и именно это значение нам нужно найти. Поиск будем осуществлять обычным перебором, начиная с 10-и каналов и заканчивая 130-ю каналами с шагом, равном 10, а результаты в виде ROC-AUC представим в виде таблицы.

Из таблицы 1 видно, что наилучшим качеством обладает нейросеть, имеющая 130 каналов в каждом сверточном слое. Однако оптимизация архитектуры сети, не ограничивается поиском количества выходных каналов в сверточных слоях. Более того, количество этих каналов может различаться. Кроме графовых сверточных слоев в нейросеть входит и полносвязный слой, настройка которого может сильно влиять на качество работы модели. В дальнейшем необходимо рассмотреть варианты поиска параметров архитектуры модели, основанные на неполном переборе всех вариантов, что значительно бы ускорило процесс получения качественной модели. Такими вариантами являются:

- 1) случайный поиск;
- 2) байесовские методы [8];
- 3) эволюционные вычисления, где наибольший интерес представляют генетические алгоритмы. Последние показали высокие результаты в задачах поиска нейросетевых архитектур, что продемонстрировано в работе NEAT [9].

Таблица 1 – Результаты обучения модели

Количество каналов	Обучающее ROC-AUC	Тестовое ROC-AUC
10	0,781	0,691
20	0,700	0,643
30	0,667	0,570
40	0,847	0,706
50	0,816	0,683
60	0,884	0,692
70	0,857	0,712
80	0,874	0,717
90	0,859	0,723
100	0,887	0,757
110	0,870	0,718
120	0,739	0,651
130	0,907	0,771

Заключение статьи

В результате поиска оптимальной архитектуры нейросетевой модели была найдена наилучшая конфигурация в смысле величины метрики ROC-AUC. Данная величина получена на тестовом множестве, что позволяет говорить об удовлетворительном качестве предсказания активности андрогенного рецептора, т. е. решении задачи экстраполяции. В дальнейшем работа будет посвящена анализу более эффективных методов поиска архитектуры нейросетевой модели.

Литература:

1. Scarselli F. The Graph Neural Network Model // IEEE Transactions on Neural Networks. – 2009. – Vol. 20, N 1. – P. 61–80.
2. Dwivedi V. P. Benchmarking Graph Neural Networks [Electronic resource] URL: <https://arxiv.org/pdf/2003.00982.pdf> (accessed: 22.11.2021).
3. Kipf T. N. Semi-Supervised Classification with Graph Convolutional Networks // [Electronic resource] URL: <https://arxiv.org/pdf/1609.02907.pdf> (accessed: 22.11.2021).
4. Zitnik M. Modeling polypharmacy side effects with graph convolutional networks // Bioinformatics. – 2018. – Vol. 34, no. 13. – P. 457–466.
5. Weininger D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules // Journal of Chemical Information and Computer Sciences. – 1988. – Vol. 28, N 1. – P. 31–36.
6. Murphy A. The Finley Affair: A Signal Event in the History of Forecast Verification // Weather and Forecasting. – 1996. – Vol. 11, N 1. – P. 3–20.

7. URL: https://github.com/filipsPL/tox21_dataset (accessed: 22.11.2021).

8. Brochu E. A tutorial on Bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning // [Electronic resource]. URL: <https://arxiv.org/pdf/1012.2599.pdf> (accessed: 22.11.2021).

9. Stanley K.O. Evolving Neural Networks Through Augmenting Topologies // Evolutionary Computation. – 2002. – Vol. 10, N 2. – P. 99–127

ОСНОВАННОЕ НА СЕГМЕНТАЦИИ РАСПОЗНАВАНИЕ РЕДКИХ ИЛИ УНИКАЛЬНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ СИМВОЛОВ МАШИНОПИСНОГО ТЕКСТА СО СКАН-КОПИЙ ДОКУМЕНТОВ

А.А. Ржевский

**Новосибирский государственный технический университет,
г. Новосибирск, arseniirzhevskii964@gmail.com**

Научный руководитель: Попов А.А., д.т.н, профессор

В данной работе предложен метод распознавания машинописного текста, основанный на посимвольной сегментации слов. Проведен сравнительный анализ различных архитектур нейронных сетей в задаче распознавания двух классов машинописного текста. Полученные результаты демонстрируют возможность распознавания машинописного текста с большей точностью.

This paper presents a typewritten text recognition method based on character level word segmentation. A comparative analysis of various architectures of neural networks for recognition of two classes of typewritten text is presented. The obtained results demonstrate the ability to recognize typewritten text with greater accuracy.

В настоящее время многие компании, следуя всеобщему тренду на оптимизацию и цифровизацию, стремятся автоматизировать и оцифровать свой документооборот и полностью исключить человеческий фактор из этого процесса. Это становится возможным при внедрении технологии оптического распознавания символов (OCR, англ. Optical Character Recognition) в бизнес-процессы компании. Успешность такого подхода во многом зависит от качества распознавания машинописного текста.

В большинстве случаев под текстом понимается набор осмысленных слов и предложений, и при распознавании такого текста со скан-копий документов наилучший результат показывают искусственные нейронные сети с LSTM ячейками [1-2]. Это ячейки, которые обладают памятью и улавливают контекстную информацию в последовательности символов. Однако, в некоторых случаях распознавание слов нейросетями с использованием LSTM ячеек ведет к ухудшению качества. Такими словами являются различные наименования (организаций, email и т.д.), аббревиатуры, числа и цифро-буквенные номера. Все эти исключения представляют довольно большой практический интерес при решении задачи автоматизации документооборота.

Настоящая работа посвящена выбору подходящей архитектуры искусственной нейронной сети для решения проблемы распознавания редко встречающихся или уникальных последовательностей символов в машинописном тексте на скан-копиях документов.

Большая часть современных искусственных нейронных сетей распознавания текста создавалась для решения задачи распознавания текста с фотографий (STR, англ. Scene Text Recognition) [1-4] и имеет следующий набор модулей: выравнивание текста, извлечение признаков, декодирование. Отличительной чертой таких сетей является реализация этих модулей, а общей чертой – использование LSTM ячеек на этапах извлечения признаков и декодирования.

Использование таких сетей при решении задачи OCR показывает низкое качество распознавания как из-за наличия трансформационного слоя (слоя выравнивания), так и из-за неверных для OCR обучающих наборов данных [5-6]. После исправления этих недостатков полученная в итоге модель показывает высокую, однако все еще недостаточную точность распознавания текста на скан-копиях документов.

При исследовании ошибок распознавания такой модели можно выделить класс слов, у которых точность распознавания заметно хуже, чем у остальных, а именно слов, состоящих из редко встречающихся или уникальных сочетаний символов и/или цифр. Это объясняется наличием в модели LSTM ячеек и, как следствие, попыткой учесть контекстную информацию в тех словах, где ее нет. Логичным шагом для решения этой проблемы является отказ от LSTM ячеек как на этапе извлечения признаков, так и на этапе декодирования.

Однако, на наш взгляд, такой подход не является полностью оптимальным, поскольку при отсутствии LSTM ячеек и, как следствие, отсутствии этапа выделения контекстной информации в последовательности появляется смысл использовать подход, применяемый к локализации объектов [7-10]. А именно использовать вторым модулем FPN сеть

для более тонкого извлечения признаков и выполнения предсказания для каждого пикселя изображения. Такой подход позволяет обращать больше внимания на тонкие, но значительные детали в очертаниях символов, а также получать в результате посимвольную сегментацию без необходимости выполнения по ней распознавания, которое требует больших ресурсов как на создание тренировочного набора данных, так и на само распознавание. Архитектура предложенной нейронной сети изображена на рисунке 1.

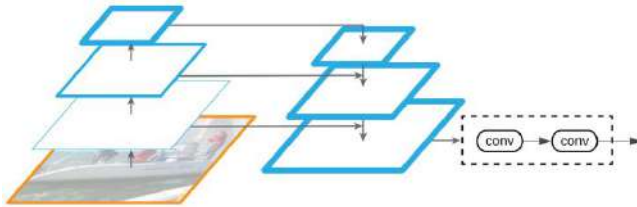


Рис 1. – Архитектура предложенной нейронной сети для распознавания текста

Таким образом, на выходе сети мы получаем изображение по размеру равное входному, каждый пиксель которого относится к одному из символов алфавита или фону. Для финального определения текста на изображении выделяются все связные компоненты (группы пикселей, относящиеся не к фону), и в них определяется, какому символу из алфавита соответствует больше всего пикселей. Confidence каждого символа рассчитывается по формуле

$$Conf = \max_{i \in A} C_i / N, \quad (1)$$

где C_i – количество пикселей связной компоненты, принадлежащих i -му символу алфавита, A – множество символов алфавита, N – общее количество пикселей, принадлежащих связной компоненте.

В таблице 1 представлена сравнительная точность распознавания обычного текста и текста, состоящего из различных наименований, аббревиатур, чисел и цифро-буквенных номеров, для 40 тестовых изображений.

Из полученных результатов видно, что качество распознавания машинописного текста со скан-копий документов искусственными нейронными сетями с использованием LSTM ячеек сильно ухудшается при распознавании текста без контекстной зависимости между символами, то есть текста, состоящего из редких и/или уникальных последо-

вательностей символов. В то же время, модели искусственных нейронных сетей без LSTM ячеек показывают стабильный результат на тексте различного вида. Однако, по точности на обычном тексте они уступают сетям, в которых используются LSTM ячейки. Таким образом, можно сделать вывод о целесообразности дальнейших исследований в этом направлении и применении искусственных нейронных сетей без LSTM ячеек в задаче распознавания текста без контекстной информации на уровне символов. Также, как видно из таблицы 1, предложенная в данной работе модель показала значительно более высокую точность распознавания такого текста.

Таблица 1 – Сравнительная точность алгоритмов распознавания обычного текста и текста, состоящего из редких и/или уникальных последовательностей символов

Способ распознавания	Точность распознавания	
	Обычный текст	Текст из редких и/или уникальных последовательностей символов
Движок tesseract с использованием нейронной сети с LSTM ячейками	0.927075	0.76658
Модель распознавания из работы [1] с модификациями из [5] и с LSTM ячейками	0.959234	0.84381
Модель распознавания из работы [1] с модификациями из [5], но без LSTM ячеек	0.88946	0.880725
Предложенная модель распознавания	0.91108	0.91528

Литература:

1. What is wrong with scene text recognition model comparisons? dataset and model analysis / B. Jeonghun, K. Geewook, L. Junyeop, P. Sungrae, H. Dongyoon, Y. Sangdoo, J. Oh Seong, L. Hwalsuk // In Proceedings of the IEEE International Conference on Computer Vision. – 2019. – P. 4715–4723.
2. Show, Attend and Read: A Simple and Strong Baseline for Irregular Text Recognition / H. Li, P. Wang, C. Shen, G. Zhang // Proceedings of the AAAI Conference on Artificial Intelligence. – July 2019. – Vol. 33. – P. 8610-8617
3. ASTER: An Attentional Scene Text Recognizer with Flexible Rectification / B. Shi, M. Yang, X. Wang, P. Lyu, C. Yao, X. Bai

// Proceedings of the IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2019. – Vol. 41. – P. 2035-2048.

4. *Canjie L., Lianwen J., Zenghuae S.* MORAN: A Multi-Object Rectified Attention Network for scene text recognition // Pattern Recognit. – 2019. – Vol. 90. – P. 109–118.

5. *Ржевский А.А.* Сравнение современных методов и подходов к поиску и распознаванию машинописного текста со скан-копий документов / А.А. Ржевский; науч. рук. А.А. Попов // Наука. Технологии. Инновации.: сборник научных трудов в 9 ч. – Новосибирск: Изд-во НГТУ, 2020. – Ч. 2. – С. 215–219.

6. *Ржевский А.А.* Распознавание машинописного текста со скан-копий документов на основе различных модификаций современных подходов / А.А. Ржевский, А.А. Попов // Труды 63-й Всероссийской научной конференции МФТИ. 23–29 ноября 2020 года. Прикладные математика и информатика. Москва : МФТИ, 2020. – С. 228-229.

7. Feature Pyramid Networks for Object Detection / T. Lin, P. Dollar, R. Girshick, K. He, B. Hariharan, S. Belongie // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – 2017.

8. *Sarshogh M.R., Hines K.E.* A Multi-Task Network for Localization and Recognition of Text in Images // International Conference on Document Analysis and Recognition (ICDAR). – 2019.

9. RSCA: Real-time Segmentation-based Context-Aware Scene Text Detection / J. Li, Y. Lin, R. Liu, C. Man Ho, H. Shi // IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). – 2021.

10. Mask R-CNN / K. He, G. Gkioxari, P. Dollar, R. Girshick // Proceedings of the IEEE International Conference on Computer Vision (ICCV). – 2017.

УПРАВЛЕНИЕ ПРИБОРОМ ИЗМЕРЕНИЯ ДЛЯ МИНИМИЗАЦИИ ОШИБКИ ИЗМЕРЕНИЯ ТЕМПЕРАТУРЫ ОБЪЕКТА

Б.С. Соловьев

Новосибирский государственный технический университет,
г. Новосибирск, b.solovev.2013@stud.nstu.ru
Научный руководитель: Карманов В.С., к.т.н., доцент

Реализован метод управления контактным датчиком в виде программного продукта. Используются методы синтеза оптимального управления по минимизации функционала отклонения.

A method for controlling a contact sensor has been implemented in the form of a software product. Methods for the synthesis of optimal control to minimize the deviation functional are used.

Введение

Фундаментальной проблемой при измерении температуры поверхности твердых тел контактным способом является проблема влияния датчика на объект измерений. Если речь идет о сравнении результатов измерений, полученных разными типами средств измерений, или в достаточно широком диапазоне условий окружающей среды, близость результата измерений к действительному значению температуры играет определяющую роль.

Естественным способом решения проблемы является использование активного датчика температуры с динамической компенсации тепловых потоков между ним и поверхностью объекта измерений. Данная задача рассматривается как задача управления, где динамической системой является контактный датчик, предназначенный для измерения температуры поверхности малых объектов. Необходимо минимизировать переход тепла от объекта к датчику (и наоборот), поскольку объект чаще всего имеет малые габариты. Соответственно необходимо минимизировать градиент температур между датчиком и объектом путём управления нагревателем и охладителем на контактном датчике. Данная работа основана на методике измерений «Изделия электронной техники миниатюрные. Температура поверхности. Методика измерений» [1].

Постановка задачи

Целью данной работы являются разработка методов управления нагревателем и охладителем на контактном датчике, созданном для измерения температуры поверхности особо малых (1 мм^2) объектов (рис.1). В связи с малым размером объекта необходимо при контакте

датчика с ним минимизировать как переход энергии (от объекта к датчику или наоборот), так и погрешность измерения.

Сам датчик состоит из штатива, нагревательного элемента, охлаждающего элемента, радиатора и двух датчиков термосопротивления (термопары). Также подключаются цифроаналоговый преобразователь (для подачи мощности на нагреватель и охладитель) и температурный преобразователь ТМ-4 для передачи данных с датчиков на компьютер через последовательный порт.



Рис. 1 – Контактный датчик (вверху) и объект измерения (внизу)

По мере уменьшения модуля разности температур на датчиках (термопарах) уменьшается расстояние между датчиком и объектом. По мере приближения термопары к объекту измерения увеличивается средний градиент температур между датчиками, который необходимо компенсировать, чтобы повысить точность и, что немаловажно, избежать повреждения изделия. Таким образом, разность температур стремятся свести к нулю посредством нагревателя, охладителя и управления ими.

Математическая модель

Был предложен подход использования динамической дискретной модели на основе предыдущих разработок [2], в которой состояние системы описывается вектором переменных состояния $t(k)=[t_1(k),t_2(k)]^T$, где $t_1(k)$ – значение температуры на первом датчике в момент времени k , $t_2(k)$ – значение температуры на втором датчике в момент времени k , ($k=0,\dots,N$).

Предложена математическая модель, которая имеет следующий вид:

$$\begin{bmatrix} t_1(k+1) \\ t_2(k+1) \end{bmatrix} = \begin{bmatrix} a_1(k)-a_2(k) & a_2(k) \\ 0 & b_1(k) \end{bmatrix} \cdot \begin{bmatrix} t_1(k) \\ t_2(k) \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ b_2(k) & b_3(k) \end{bmatrix} \cdot \begin{bmatrix} u_1(k) \\ u_2(k) \end{bmatrix} \quad (1)$$

где векторы $A(k)$ и $B(k)$ ($A=[a_1(k),a_2(k)]$, $B=[b_1(k),b_2(k),b_3(k)]$) – параметры, описывающие зависимость температуры от мощностей, вектор $U(k)$ ($U(k)=[u_1(k),u_2(k)]$) – подаваемая мощность в процентах. Каждую

итерацию происходит оценивание $A(k)$ и $B(k)$, чтобы наиболее точно воссоздать зависимость температур друг от друга и от подаваемой мощности.

Задача заключается в минимизации разницы температур, получаемых с датчиков на протяжении всей работы установки.

Методы решения

Поиск векторов $A(k)$ и $B(k)$ производится путём минимизации квадрата разности уже полученных данных и смоделированных при помощи оценённых векторов.

$$\sum_{j=1}^2 \sum_{i=0}^5 \left(t_j(k-i) - \hat{t}_j(k-i) \right)^2 \xrightarrow{A(k), B(k)} \min \quad (2)$$

В программной реализации для решения задачи оценивания векторов $A(k)$, $B(k)$ использовались такие методы как: поиск локального минимума в доверительной области; последовательное квадратичное программирование с ограничениями, ньютоновский метод решения системы Лагранжа; усеченный метод Ньютона с ограничениями; метод Бройдена-Флетчера-Гольдфарба-Шанно.

После получения оценок $A(k)$ и $B(k)$ вектор $U(k)$ оценивается путём минимизации функционала градиента температур (3).

$$\left| t_1(k+1) - t_2(k+1) \right| \xrightarrow{U(k)} \min \quad (3)$$

В результате работы программы удалось минимизировать разницу температур в течение всего времени эксперимента, см. рис. 2.

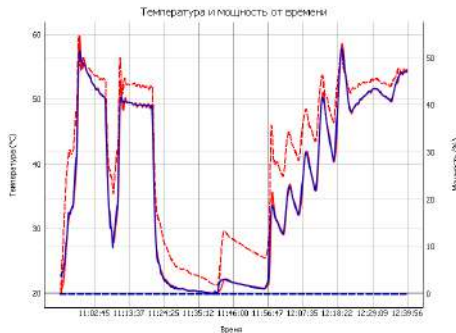


Рис.2 – Динамика изменения температур и используемой мощности в ходе эксперимента

На графике красной сплошной линией указаны значения температуры на нижнем датчике, синей сплошной – на верхнем. Пунктирными линиями показаны подаваемые мощности на нагреватель и охладитель (красным и синим цветом соответственно). На графике видно, что температуры датчиков почти одинаковы, за исключением резких перепадов, которые обусловлены изменением расстояния от объекта, а так же изменением температуры объекта в ходе эксперимента.

Эксперимент занял примерно 2 часа, в ходе которого температуру объекта меняли от 20 °С (комнатная температура) до 60 °С. Текущая установка, состоящая из контактного датчика и эмулятора объекта, ограничена температурами от 10 °С до 60 °С, что позволяет проводить дополнительные эксперименты по охлаждению.

Полученные результаты

Разработанные методы реализованы в виде программного комплекса, использующего клиент-серверную архитектуру на основе технологии сокетов TCP/IP, состоящего из несколько частей.

Поскольку мониторинг градиента температуры при измерении температуры поверхности это основной инструмент анализа состояния процесса, у оператора есть возможность наблюдать графики температур, мощностей и разности температур в реальном времени.

Заключение

В результате проведенных исследований:

1. Предложена математическая модель для синтеза оптимального управления стохастическими динамическими системами в таких предметных областях как термодинамика и точная калибровка измерительного оборудования.
2. Предложены различные пути развития для данной задачи, такие как: программные модификации для кроссплатформенности, смена моделей функционала управления, использование сторонних технологий и методов для моделирования.
3. Разработан программный комплекс, включающий в себя все необходимые программные компоненты управления аппаратурой для измерения температуры поверхности малых объектов.

Список литературы

1. 1029-RA.RU.311735-2020 Изделия электронной техники миниатюрные. Температура поверхности. Методика измерений (2020) // Западно-Сибирский филиал ФГУП «ВНИИФТРИ».
2. Соловьев Б. С. Динамическая модель управления воспроизведением реперных точек температурной шкалы / Б. С. Соловьев, В. С. Карманов, Д. А. Гривастов // Наука. Технологии. Инновации : сб. науч. тр. : в 9 ч., Новосибирск, 30 нояб.–4дек. 2020 г. – Новосибирск : Изд-во НГТУ, 2020. – Ч. 2. – С. 272–275.

ПРИМЕНЕНИЕ СРЕДЫ ANYLOGIC ДЛЯ МОДЕЛИРОВАНИЯ ТРАНСПОРТНО-ЛОГИСТИЧЕСКИХ ПРОЦЕССОВ

А.В. Солоха, И.А.Соболева

Новосибирский государственный университет экономики и управления «НИНХ», г. Новосибирск, solokha.nastya@yandex.ru

Научный руководитель: Соболева И.А., к.с.н., доцент

В статье рассмотрено понятие моделирования, которое сегодня является мощным аналитическим средством, включившим в себя весь спектр новейших информационных технологий. Разобран пример использования технологии в среде AnyLogic для транспортно-логистических процессов.

The article discusses the concept of modeling, which today is a powerful analytical tool that includes the entire spectrum of the latest information technologies. An example of using the technology in the AnyLogic environment for transport and logistics processes has been analyzed.

На современном этапе развития общества все больше возникает интерес к средствам имитационного моделирования. Актуальность обусловлена существенным технологическим развитием систем моделирования, которые сегодня являются мощным аналитическим средством, включившим в себя весь спектр новейших информационных технологий: развитые графические оболочки для конструирования моделей и интерпретации выходных результатов моделирования, мультимедийные средства и видео, поддерживающие анимацию в реальном масштабе времени и другие.

Имитационная модель представляет собой специальное программное обеспечение, позволяющее отобразить процессы, протекающие в моделируемой системе, с необходимым для решения поставленной задачи набором параметров и уровнем детализации. В результате прогона модели автоматически собирается статистическая информация о моделируемых процессах [1]. По-настоящему революционным шагом в развитии имитационного моделирования экономических процессов на современном этапе является внедрение в практику профессионального инструмента нового поколения построения имитационных моделей AnyLogic.

AnyLogic представляет собой гибкий инструмент имитационного моделирования, который допускает множество путей для решения поставленной задачи. При разработке модели на AnyLogic используются концепции и средства из нескольких классических областей имитаци-

онного моделирования: динамических систем, дискретно-событийного моделирования, системной динамики и агентного моделирования [4].

Основной технологией программирования в AnyLogic является визуальное программирование – построение с помощью графических объектов, пиктограмм иерархий структуры и поведения активных объектов. Для реализации специальных вычислений и описания логики поведения объектов AnyLogic позволяет использовать язык Java[2].

Демонстрацию возможностей применения AnyLogic при моделировании транспортно-логистических процессов рассмотрим на практической ситуации: со склада необходимо доставлять товар с помощью транспорта в розничные магазины по городу. Необходимо оценить оптимальное количество транспортных средств для данного процесса.

В данной имитационной модели представлены два основных агента. Это Склад и Магазин – розничный продавец товара. Дополнительным агентом является транспортное средство, осуществляющее доставку груза от склада к розничной точке. Таким образом, в случае потребности розничная точка размещает заказ складу, и заказ транспортируется в розничную точку. Создав агентов «Склад», «Транспорт» и популяцию агентов «Магазины» зададим им значения количества, поведения и остальных свойств.

Функционал программы позволяет подключить ГИС-карту и выбрать необходимую область, что дает возможность смоделировать процесс, приближенный к реальным условиям доставки. Розничная сеть магазинов рассмотрена на примере фактического расположения регионального склада и сети магазинов МТС в городе Новосибирске (рис. 1).

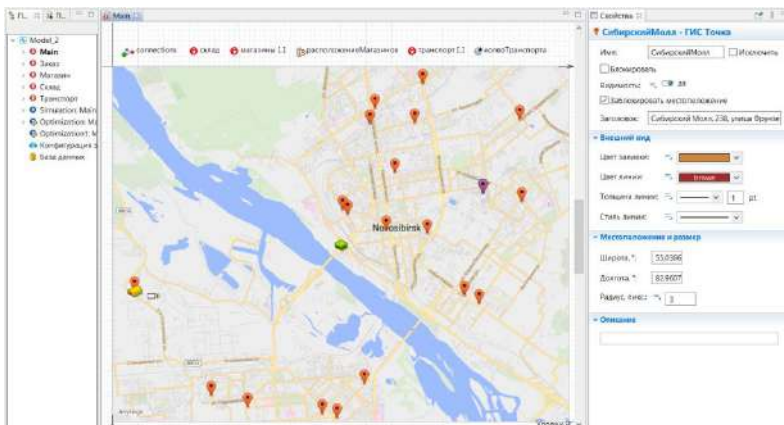


Рис. 1 – Размещение агентов на карте

В результате на карте появляется 20 точек, одна из которых склад. Далее, задается поведение агента «Магазин», а именно сам процесс заказа. Для этого используется диаграмма состояний. Формирование заявки (что отправляем и кому отправляем) оформляется фрагментом кода на языке Java (рис. 2).

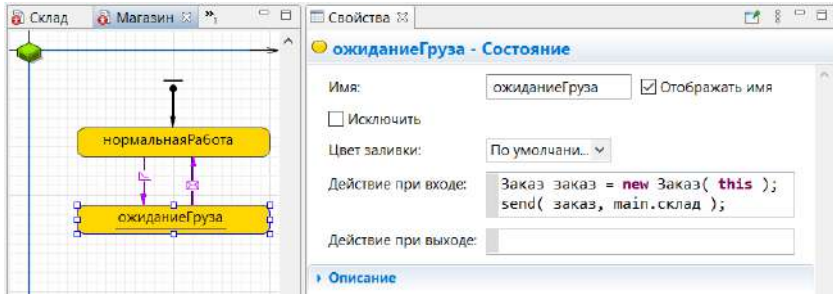


Рис. 2 – Диаграмма состояний

Для создания логики обработки заказов применяем библиотеку моделирования процессов. В результате получается следующая диаграмма процесса, которая представлена на рисунке 3.

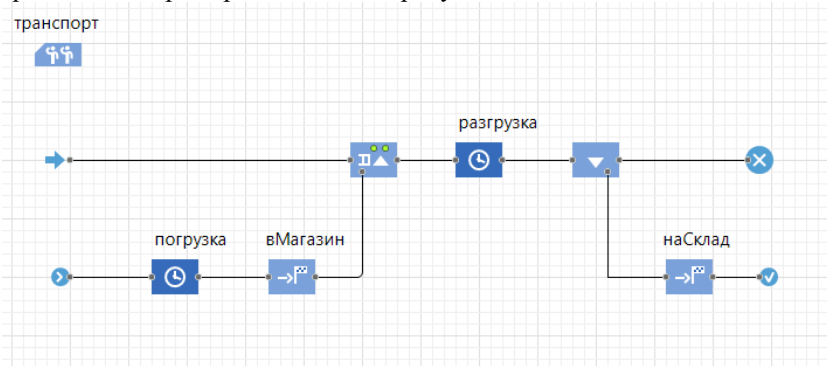


Рис.3 – Процесс обработки заказов

После создания модели – а иногда и в процессе разработки – AnyLogic позволяет исследовать структуру и понимать поведение системы, проверять, как она ведет себя при определенных условиях, сравнивать различные сценарии и оптимизировать ее. Когда оптимальное решение будет найдено, можно применять его в реальном мире [3].

Определим, какое количество транспорта необходимо для оптимального распределения ресурсов. Для этого нужно запустить модель и пе-

рейти внутрь агента «Склад», чтобы посмотреть текущую загрузку ресурсов. Результат моделирования показывает, что загрузка ресурсов колеблется в районе 32-35%. Это означает, что количество транспорта, которое используется (в данном случае 5), избыточное. Для того, чтобы оптимизировать количество транспорта, нужно создать эксперимент «Оптимизация» и задать новые свойства. Цель эксперимента – посмотреть, при каком количестве транспортных средств их загрузка будет максимальной. Результат представлен на рисунке 4.

Model_2 : Optimization

	Текущее	Лучшее
Итерация:	5	4
Функционал↑	0.44	0.825
Параметры		Copy best
колвоТранспорта	4	2

Рис.4 – Процесс оптимизации

Исходя из результатов этого эксперимента, видно, что лучшая итерация под номером 4, количество транспорта, достаточное для рационального использования в процессе доставки равно 2, при этом загрузка ресурсов около 82%. Для проверки функционирования модели, нужно подставить полученное значение в количество транспорта и запустить заново симуляцию процесса. В результате моделирования загрузка ресурсов колеблется в районе 80-82%, модель работает стабильно, т.е. нет очередей, задержек, сбоев. Следовательно, подставленное значение является оптимальным.

Таким образом, можно сделать вывод о том, что использование программного обеспечения AnyLogic в моделировании логистических процессов позволяет воспроизвести процесс с условиями, приближенными к реальности, а также добиться оптимального использования ресурсов, что способствует экономии финансовых средств.

Литература:

1. Григорьев, И. AnyLogic за три дня. Практическое пособие по имитационному моделированию/ Григорьев И. – Интернет-издание, 2016. – 202 с.
2. Киселева, М.В. Имитационное моделирование систем в среде AnyLogic: учебно-методическое пособие / М. В. Киселёва. Екатеринбург: УГТУ - УПИ, 2009. – 88 с.

3. Мезенцев, К.Н. Моделирование систем в среде AnyLogic6.4.1. Часть 2: учеб. пособие / А.Б.Николаев, К.Н.Мезенцев // МАДИ. – М.: 2011. 103 с.

4. Федотова, В.С. Технологии имитационного моделирования в системе AnyLogic // Царскосельские чтения, 2013. №XVII. – С.148-150.

СРАВНИТЕЛЬНЫЙ АНАЛИЗ МЕТОДОВ СИМУЛЯЦИИ РЕАЛИСТИЧНОГО ПОВЕДЕНИЯ ВИРТУАЛЬНЫХ АГЕНТОВ

В.Р. Тимшин

**Национальный исследовательский Томский политехнический
университет, г. Томск, vrt2@tpu.ru**

Научный руководитель: Видман В.В., ассистент

Наиболее распространенные методы симуляции реалистичного поведения при увеличении сложности разрабатываемого поведения трудно поддерживать и совершенствовать. В процессе поиска метода, лишённого этого недостатка, была выявлена связка методов, потенциально упрощающая разработку и не требующая машинного обучения.

The article is focused on the most common methods of realistic behavior simulation which are difficult to maintain and enhance when complexity of behavior increases. When searching for a method free of this disadvantage, a combination of methods which potentially simplifies the development and doesn't require machine learning has been found.

Введение

Симуляция поведения человеческого общества является трудоёмкой задачей, результаты решения которой актуальны для социального моделирования, включающего в себя вопросы организационного поведения, психологии, экономики, лингвистики, политики. Необходимость в симуляции событий с помощью компьютерных средств возникает в случаях, когда число реально происходящих событий исследуемого типа мало, их реконструкция затруднительна, но само исследование необходимо. Примером такого типа событий может быть чрезвычайная ситуация при пожаре, угрозе совершении преступления или иное социально-опасное происшествие, не поддающееся полноценной реконструкции в реальном мире. Учитывая сложность и неоднозначность человеческого поведения, а также его зависимость от сторонних факторов (болезнь, травма, влияние окружающего общества, общее самочувствие и т.д.), задача моделирования группы людей становится еще более сложной.

Существующие по этой теме работы [1-4] в основном направлены на исследование предполагаемых моделей, предметом текущей работы являются непосредственно способы реализовать перечисленные выше модели, сделать их более приспособленными к дальнейшему усовершенствованию. Целью работы является исследование и сравнение различных методов реализации моделей, симулирующих поведение виртуальных агентов (далее – методы).

Содержательная часть

Все указанные методы и их свойства подробно описаны в курсах, научных работах по искусственному интеллекту [5-7]. Для исключения менее удобных в реализации методов сначала будет проведено сравнение базовых методов и их модификаций, а затем – сравнение модификаций между собой.

1. Конечный автомат и GOAP

Конечный автомат легко реализовать на начальных этапах разработки, но, если число состояний увеличится до нескольких десятков, возможных условий перехода станет настолько много, что не только дальнейшая разработка, но и поддержка такой системы станет невозможной. Система GOAP, являющаяся модификацией конечного автомата, позволяет сгруппировать некоторые состояния для обеспечения лучшего контроля над ними и, хотя в начале разработки на реализацию отдельных состояний уйдет больше времени, в будущем это компенсируется более простой поддержкой архитектуры.

2. Дерево выбора и Utility AI

С деревом выбора и Utility AI ситуация аналогичная: модификация дерева выбора требует большего внимания на начальных этапах, но в будущем система будет более простой для поддержки и усовершенствования по сравнению с обычным деревом. Более того, Utility AI позволяет определить приоритет выполнения каждой задачи, что открывает возможность создания приоритетной очереди желаемых действий без повторного запуска системы выбора.

Как итог, использование модифицированных версий несколько усложняет разработку на начальных этапах, но существенно облегчает поддержание и доработку моделей в дальнейшем. По этой причине дальнейшего сравнения каких-либо методов с базовыми версиями проведено не будет.

3. GOAP и дерево поведения

Основным преимуществом GOAP перед деревом поведения является возможность перехода к произвольному состоянию без дополнительных проверок на возможность такого перехода. В случае дерева поведения необходимо или изначально создавать систему, которая будет проверять возможность перехода к узлу, или искать другой способ, позволяющий

получить желаемое поведение. Для дерева поведения писать многочисленные проверки затруднительно, в GOAP же система сама перейдет к требуемому состоянию при невозможности выполнения действия при заданных условиях.

Сравнение GOAP и Utility AI затруднительно по причине того, что первый метод уместно сравнивать скорее с деревом поведения, чем с деревом выбора, улучшением которого является Utility AI. Поскольку GOAP и Utility AI могут заменить дерево поведения и дерево выбора соответственно, появляется возможность сравнить данную связку с оставшимся методом.

4. GOAP + Utility AI и генетический алгоритм

Основным достоинством связки является возможность создания алгоритма, подходящего для сущности с любым набором параметров. Для генетического алгоритма необходимо создать такие ограничения, при которых развитие не остановится на сущности с сильнейшими характеристиками, и при этом будет стремиться к полному использованию любого набора предоставляемых сущностью возможностей. Более того, при добавлении нового свойства сущности или возможного действия, всю систему придется обучать заново, что является недостатком по сравнению с GOAP + Utility AI.

Выбрать лучший из двух методов для продвинутой разработки сложно из-за слабостей каждого из них по разным, не связанным непосредственно с разработкой критериям. Таким образом, среди описанных методов наилучшим для разработки должна быть либо связка Utility AI и GOAP, либо идея генетического алгоритма.

Заключительная часть

В данной работе были рассмотрены и сравнены методы симуляции реалистичного поведения виртуальных агентов, выявлена связка методов, модель из которых может стать самой удобной при детальной разработке сложного поведения сущностей. Среди всех методов связка из GOAP и Utility AI и генетический алгоритм определены как наиболее удобные и эффективные для крупной разработки подобных симуляций. Для определения лучшего метода для создания симуляции необходима реализация каждого из них на конкретном примере, что и планируется сделать в дальнейшем.

Литература:

1. Бекларян А.Л., Акопов А.С. Моделирование поведения толпы на основе интеллектуальной динамики взаимодействующих агентов // Бизнес-информатика. 2015. № 1 (31). С. 69–77.

2. "Melos" (motivesevolutionsimulation) Изучение поведенческих механизмов психики человека // Psyfactor URL: <https://psyfactor.org/lib/melos.htm> (дата обращения: 02.11.2021).

3. Моделирование поведения человека при помощи новой технологии Искусственного Интеллекта — MELOS // Psyfactor URL: <https://psyfactor.org/lib/melos2.htm> (дата обращения: 02.11.2021).

4. В Университете Аризоны разрабатывается компьютерная модель поведения человеческой толпы. [Электронный ресурс] // Центр гуманитарных технологий. — 02.11.2021. 14:50. URL: <https://gtmarket.ru/news/culture/2007/06/17/1021>

5. MITOPENCOURSEWARE URL: <https://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-034-artificial-intelligence-fall-2010/lecture-videos/> (дата обращения: 02.11.2021).

6. Анохин Александр Олегович, Катаев Александр Вадимович Конечно-автоматная модель управления поведением интеллектуальных агентов в обучающих играх // ИТНОУ: информационные технологии в науке, образовании и управлении. 2019. №4 (14). URL: <https://cyberleninka.ru/article/n/konechno-avtomatnaya-model-upravleniya-povedeniem-intellektualnyh-agentov-v-obuchayuschih-igrah> (дата обращения: 14.11.2021).

7. Донских А.К., Барабанов В.Ф., Гребенникова Н.И., Белых М.А. ОБЗОР АРХИТЕКТУРЫ СИСТЕМ УПРАВЛЕНИЯ ИНТЕЛЛЕКТОМ НА ОСНОВЕ ПОЛЕЗНОСТИ И ДЕРЕВА ПОВЕДЕНИЯ // Вестник ВГТУ. 2021. №2. URL: <https://cyberleninka.ru/article/n/obzor-arhitektury-sistem-upravleniya-intellektom-na-osnove-poleznosti-i-dereva-povedeniya> (дата обращения: 14.11.2021).

СРАВНЕНИЕ ТОЧНОСТИ МЕТОДОВ КЛАССИФИКАЦИИ ДОРОЖНЫХ ЗНАКОВ ОГРАНИЧЕНИЯ СКОРОСТИ ПРИ ЧАСТИЧНОМ ПЕРЕКРЫТИИ СИМВОЛОВ ЗНАКА

И.К. Харченко, И.Г. Боровской

**Томский государственный университет систем управления и радио-
электроники, г. Томск, igor.k.kharchenko@tusur.ru**

Проведено сравнение точности двух различных методов классификации в случае частичного перекрытия дорожного знака ограничения скорости: классификация знака по исходному изображению, а также

классификация отдельных символов знака. Выявлено, что для знаков ограничения скорости в 10, 30, 50, 70 и 100 км/ч при классификации отдельных символов с перекрытием области символов в 16% точность составляет от 69% до 91%; однако для знаков 20, 60 и 80 км/ч точность близка к случайной.

Comparison of the accuracy of two different methods in the case of partial overlap of a speed limit road sign is carried out: the classification of the sign by the source image, as well as the classification of individual sign symbols. It was found that for the signs of the speed limit area of 10, 30, 50, 70 and 100 km/h, when classifying individual symbols with overlapping by 16%, the accuracy ranges from 69% to 91%; however, the accuracy for signs 20, 60 and 80 km/h is close to random.

Введение

При распознавании дорожных знаков ограничения скорости нужно принимать во внимание условия внешней среды. Проблема заключается в том, что наличие таких естественных зашумлений, как плохие погодные условия (дождь, снег, пыль, дымка), солнечные блики, сильные потёртости и наклейки на знаке, способны значительно снизить качество распознавания знака.

Целью данного исследования является сравнение двух различных методов классификации для распознавания дорожных знаков при их частичном перекрытии: классификатора знака, а также классификатора отдельных символов знака. Объектом исследования является выборка изображений дорожных знаков ограничения скорости, предметом исследования – частичные перекрытия символов на изображениях дорожных знаков.

В качестве тестовой выборки была взята выборка "German Traffic Sign Recognition Benchmark" [1]. Она содержит 39210 тренировочных и 12631 тестовых изображений. Из данной выборки были взяты только изображения с дорожными знаками ограничения скорости от 20 до 120 км/ч, всего девять классов (в выборке нет знаков с ограничением скорости 40, 90 и 110 км/ч).

Описание сравниваемых методов классификации

Сравнение производилось с использованием двух классификаторов. Первый — это классификатор на базе свёрточной нейронной сети, который настроен на классификацию дорожного знака; его топология представлена на рисунке 1. Второй классификатор настроен на классификацию отдельных символов с дорожного знака от 0 до 8 (без цифры 4, поскольку её нет в выборке), его топология представлена на рисунке 2. Для обучения данного классификатора дополнительно создана ещё одна

производная выборка, состоящая из 8000 символов с дорожного знака (примерно по 1000 изображений для каждого символа).

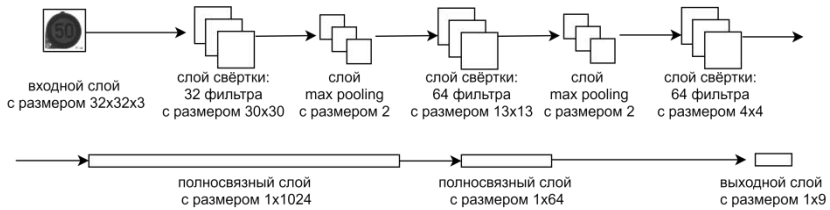


Рис.1 – Топология свёрточной нейронной сети для классификации дорожных знаков

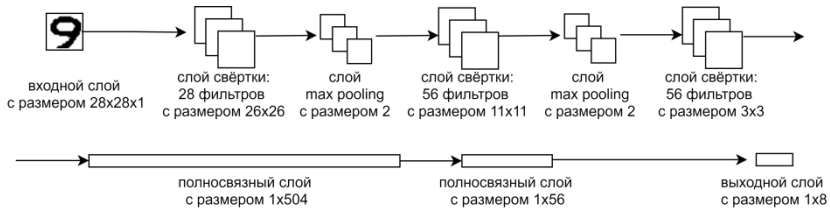


Рис.2 – Топология свёрточной нейронной сети для классификации символов дорожных знаков

Таким образом, в первом случае происходит классификация всего дорожного знака, а во втором случае на знаке определяются контуры символов, которые классифицируются по отдельности, и результатом классификации дорожного знака становится объединение результатов классификации символов.

При проверке обоих методов классификации на исходной выборке, у каждого изображения определяется граница области, в которой находятся все символы знака, после чего на данную область наносится перекрытие, которое реализовано в виде белого прямоугольника размером 2%, 4%, 8%, 10% и 16% от площади области символов знака соответственно (см. рис. 3).



Рис.3 – Примеры частичного перекрытия области символов дорожного знака

Также точность классификатора символов проверена на символах от 0 до 8 (кроме цифры 4, её нет в выборке): для каждого символа из выборки нанесено перекрытие с помощью белого прямоугольника с теми же размерами, что и в предыдущем случае.

Полученные результаты

Результаты сравнения точности обоих классификаторов при добавлении частичного перекрытия представлены в таблице 1.

В таблице 2 представлена точность классификатора символов при добавлении частичного перекрытия размером 16% от его площади в каждый из 8 возможных символов на знаке. Следует сделать замечание, что при перекрытии символа 8 площадью 10% точность составила 57,1364%.

Таблица 1 – Сравнение точности методов классификации

Размер области перекрытия, %	Классификатор символов дорожного знака		Классификатор дорожных знаков	
	Точность на тренировочной выборке, %	Точность на тестовой выборке, %	Точность на тренировочной выборке, %	Точность на тестовой выборке, %
2	96,124	95,2145	80,5783	82,4806
4	92,5581	89,7733	70,0967	70,8527
8	86,3566	81,6263	57,7649	55,1938
10	81,5504	80,1196	56,8785	48,5271
16	73,6434	71,5289	45,0617	34,7287

Таблица 2 – Точность классификатора символов при добавлении частичного перекрытия площадью 16%

Символ	Точность, %	Символ	Точность, %
0	69,5266	5	78,5379
1	91,1529	6	52,103
2	57,2296	7	71,1651
3	82,0054	8	19,3767

По результатам в таблице 1 видно, что классификатор символов способен корректно определять символ с частичным перекрытием знака на 16% в 71% случаев, тогда как классификатор дорожных знаков существенно теряет в точности уже после 4% перекрытия знака.

Из таблицы 2 следует вывод о том, что даже в ситуации, когда у знака ограничения скорости будет полностью перекрыта крайняя правая цифра (в рассматриваемом случае всегда цифра 0), при частичном пере-

крытии другого символа на знаках ограничения скорости в 10, 30, 50 и 70 км/ч вероятность правильного распознавания варьируется от 71% до 91%. Вышесказанное верно и для знака в 100 км/ч: точность распознавания при частичном перекрытии одного из трёх символов варьируется примерно на том же уровне, от 69% до 91%. Однако для знаков ограничения скорости в 20, 60 и 80 км/ч вероятность корректной классификации остаётся на уровне чуть выше 50%, т.е. близко к случайным результатам.

Таким образом, можно сделать вывод о том, что при классификации дорожного знака в условиях его частичного перекрытия до 16% большая точность достигается при распознавании отдельных символов знака. Сказанное верно для знаков ограничения скорости в 10, 30, 50, 70 и 100 км/ч. Данные результаты могут быть применены и к частичным перекрытиям других знаков: например, для распознавания конкретного символа внутри предупреждающих знаков.

Литература:

1. Stallkamp J., Schlipsing M., Salmen J., Igel C. The German Traffic Sign Recognition Benchmark: A multi-class classification competition // International Joint Conference on Neural Networks, IJCNN 2011, 03 oct. 2011: Proceedings. – San Jose, CA, 2011. – P. 1453–1460.

ИСПОЛЬЗОВАНИЕ ВОЗМОЖНОСТЕЙ МОБИЛЬНЫХ УСТРОЙСТВ ДЛЯ РАСПОЗНАВАНИЯ ЭМОЦИЙ ЧЕЛОВЕКА

М.В. Хнюнин

Новосибирский государственный технический университет,

г. Новосибирск, xnyunin.2016@stud.nstu.ru

Научный руководитель: Ганелина Н.Д., к.т.н.

Работа посвящена исследованию методов и разработке прототипа мобильного приложения для распознавания эмоций человека при помощи данных, получаемых от личных мобильных устройств, с целью анализа реакции человека на информацию, отображаемую на экране. В работе рассматриваются существующие подходы в области распознавания эмоций при помощи искусственных нейронных сетей и описываются текущие результаты работы.

The work is devoted to the study of techniques and the development of a prototype of a mobile application for recognizing human emotions using data

obtained from personal mobile devices in order to analyze a person's response to information displayed on the screen. The article considers techniques in the field of emotion recognition using artificial neural networks and reviews the current results of work.

С увеличением объёма информации, доступной человеку, возникла потребность лучше подбирать и адаптировать предоставляемый контент под конкретного человека. Социальные сети анализируют интересы пользователя для определения порядка и типа отображаемого контента; интернет-издания стараются предоставлять новости в порядке релевантности относительно возрастной категории читателя; таргетированная реклама вышла на первое место как наиболее эффективный инструмент маркетинга, значительно влияющий на лояльность потребителей к продукту [1]; всё больше набирает популярность интерактивное кино, где зрители выбирают действия героев сюжета [2]. Но на данный момент всё взаимодействие с человеком основано на его непосредственных действиях: оценки контента посредством кнопок «нравится», прохождение опросов, сбора информации о возрасте, интересах и т.п. Однако всё это можно автоматизировать и сделать менее заметным для потребителя при помощи оценки реакции человека за счёт распознавания его эмоций. Потребность определять реакцию человека на те или иные события существует давно, однако автоматизация распознавания эмоций не развита в связи со сложностью создания достаточно точных алгоритмов.

Человек выражает свои эмоции в основном при помощи мимики лица, речи и положения тела. Самым выразительным проявлением эмоций является изменение мимики лица. На данный момент существующие разработки по распознаванию эмоций используют в основном анализ фотографий лица, это наиболее очевидный и в то же время эффективный способ [3].

Для анализа фотографий с целью распознавания эмоций целесообразно использовать нейронную сеть. Существуют три подхода к обучению такой нейронной сети. Первый – это обучение на изображениях фрагментов лиц, таких как губы и глаза. Такой подход требует меньше данных для обучения, но в случае распознавания сложных эмоций он также не применим, т.к. при обучении практически не учитывается корреляция различных фрагментов. Второй - обучение на фотографиях лица целиком. Такой подход очень требователен к объёму данных для обучения и его сложно адаптировать под особенности выражений эмоций конкретным человеком. Третий вариант – использовать для обучения не сами фотографии лиц, а отклонения положения ключевых точек

миимики лица от нормали, где под нормалью понимается их положение при нейтральной эмоции. Такой подход требователен к яркости выражения эмоций, однако он может быть адаптирован к особенностям строения лица конкретного человека и потенциально может показать наибольшую точность. Но все перечисленные подходы обладают высокими требованиями к изображению, а именно к положению и удалённости лица на фотографии. Ведь при частичной видимости лица, искажению от поворота головы или различной удалённости от камеры сложно достичь высокой точности распознавания или вовсе невозможно.

В связи с тем, что в современном обществе с развитием личных мобильных устройств значительную часть информации люди стали получать посредством смартфонов и планшетов, элемент интерактивного взаимодействия наиболее востребован именно для этих устройств. Кроме того, человек смотрит на экран мобильного устройства под одним и тем же углом, на привычном для него расстоянии от экрана, а соответственно и положением головы, и расстоянием относительно фронтальной камеры. Таким образом, решается основная проблема применимости существующих алгоритмов распознавания эмоций на практике.

Во время разработки прототипа мобильного приложения для распознавания эмоций по фотографиям было выяснено, что достаточно сложно найти массив данных для обучения необходимого объёма. Существующие разработки в основном используют либо обучение на фотографиях лица целиком с различными конфигурациями нейронных сетей на одних и тех же данных для обучения [3]; либо подход с определением ключевых точек лица, но не обучением нейросети отклонениям этих точек, а простой алгоритмический подход, способный точно определить лишь яркую улыбку, удивление, грусть и злость [4]. Но для классификации эмоций учёные выделяют семь базовых эмоциональных состояний, сочетание которых позволяет охарактеризовать все возможные эмоции человека. Однако, на практике даже человеку сложно определить слабо выраженные мимикой эмоции. В связи с этим следует использовать большее количество факторов для определения состояний человека, близких к нейтральному. Так как в рамках моей работы используются камеры мобильных устройств, из доступных источников информации для уточнения результата распознавания эмоции есть микрофон того же устройства, при помощи которого можно получать запись голоса и на основе характеристик голоса повышать точность конечного алгоритма. Потенциально с развитием технологий носимых устройств и различных датчиков в них число факторов для повышения достоверности результата распознавания можно увеличивать. Например, в 2015 году российская компания ЕРАМ представила браслет, спо-

собный оценивать уровень стресса посредством измерения гальванического сопротивления кожи (GSR) [5]. В рамках данной работы было тестировано использование показателей пульса с фитнес-браслета как уточняющий фактор.

На текущем этапе работы удалось построить алгоритм распознавания эмоций, обученный по фотографиям лица на основе свободно доступного массива данных, размеченного для данной задачи. Данный алгоритм выделяет семь эмоциональных состояний человека. Кроме этого, был разработан алгоритм, способный распознавать эмоции по отклонениям ключевых точек лица, однако не удалось достичь приемлемой точности в связи с отсутствием достаточного объема данных, подходящих для обучения. Также, для повышения достоверности результата был разработан алгоритм для определения пяти эмоциональных состояний на основе анализа голоса в записи звука с микрофона устройства. На данный момент работа ведётся над объединением данных факторов с различной размерностью и созданием полноценного прототипа приложения, способного определять эмоциональную реакцию человека на отображаемый контент на экране мобильного устройства.

Литература:

1. *Мирончук А. М.*, Создание и продвижение нового продукта в индустрии интерактивных развлечений на виртуальном рынке : магистерская диссертация / А. М. Мирончук ; Уральский федеральный университет имени первого Президента России Б. Н. Ельцина, Институт "Высшая школа экономики и менеджмента", Кафедра маркетинга. — Екатеринбург, 2018. — 147 с. — Библиогр.: с. 142-147 (94 назв.).

2. *Сапарова А.*, Интерактивное кино: Игровой жанр или кино будущего? [Электронный ресурс] // Steppe - прогрессивный сайт о жизни, работе и увлечениях. — июль 2018. — URL: <https://the-steppe.com/gorod/interaktivnoe-kino-igrovoy-zhanr-ili-kino-budushchego> (дата обращения: 11.10.2021).

3. An End-to-End Visual-Audio Attention Network for Emotion Recognition in User-Generated Videos / Sicheng Zhao, Yunsheng Ma, Yang Gu, Jufeng Yang, Tengfei Xing, Pengfei Xu, Runbo Hu, Hua Chai, Kurt Keutzer. [Electronic resource] // Cornell University, arXiv. — 2020, february — URL: <https://arxiv.org/pdf/2003.00832v1.pdf> (accessed: 17.10.2021).

4. *Jayashree S. Katagihalli, SriKrishna Shastri C., Naveen S. Pagad.* Facial Landmark Detection and Localization Using Point Distribution Model // International Journal of Advanced Research in Computer Science and Electronics Engineering (IJARCSEE). — 2017, april. — Vol. 6, N 4.

5. EPAM разработал носимый стресс-трекер. [Электронный ресурс] // EPAM Systems, Inc. – 2015. – URL: <https://www.epam-group.ru/about/newsroom/in-the-news/2015/epam-developed-the-stress-tracker> (дата обращения: 17.10.2021).

МОДЕЛИРОВАНИЕ КОММУНАЛЬНОЙ ИНФРАСТРУКТУРЫ МУНИЦИПАЛЬНЫХ ОБРАЗОВАНИЙ С ИСПОЛЬЗОВАНИЕМ СТАТИСТИЧЕСКИХ МЕТОДОВ

А. О. Хуснутдинов, В.С. Карманов
Новосибирский государственный технический университет,
г. Новосибирск, evolext@gmail.com

Работа посвящена решению проблемы расчета характеристик коммунальных сетей при моделировании в условиях отсутствия или искажения исходных данных. В результате предложен и реализован метод, позволяющий замещать пропуски в исходных данных на основе аналогичных объектов.

The article focuses on the issue of calculating the parameters of utility networks during modeling in the condition of having gaps or distortions in the initial data. A method was suggested, that allows one to replace gaps in the initial data based on similar objects.

Введение

Одной из базовых потребностей современного человека является использование коммунальных систем: тепло-, водо-, газоснабжения, водоотведения и пр.

Для постройки оптимальной, с точки зрения обеспечения всех пользователей ресурсом и стоимости для заказчика, системы, необходимо предварительно создать её оптимальную модель с использованием специальных программных средств.

Специалисты сферы снабжения коммунальным ресурсом при выполнении расчетов сталкиваются с важной проблемой: исходные данные для расчетов могут быть искажены или иметь пропуски, это приводит к тому, что расчеты выполнить либо совсем невозможно, либо их результаты будут не соответствовать реальной системе.

Цель работы

Разработать методы расчета характеристик реальных коммунальных систем на основе их моделей, позволяющие решать задачу в условиях отсутствия или искажения исходных данных и реализовать в виде функций прототипа приложения для моделирования.

Математическая постановка задачи

Расчет потокораспределения в сети является основным для коммунальных сетей, в частности, для водопроводных сетей он заключается в определении величин расхода и потерь напора на всех участках сети.

Методы расчета базируются двух законах Кирхгофа [1, с. 44]:

- 1) сумма расходов, втекающих в каждый узел равна нулю (или утечке);
- 2) сумма падений напора на всех участках, образующих замкнутый контур, равна нулю (или сумме действующих напоров).

А также на эмпирической зависимости падения напора на участках сети от функции расхода вида

$$\Delta H = Sf(q), \quad (1)$$

где ΔH – значения падений напора на всех участках сети, S – гидравлические сопротивления участков сети, q – значения расхода на участках сети, функциональная зависимость от которых имеет общий вид

$$f(q) = (|q_1|^{\beta-1}q_1, \dots, |q_m|^{\beta-1}q_m), \quad (2)$$

где $f(q)$ – функция потерь от расхода, β – заданный показатель функции потерь, q_j – величина расхода на j -участке сети $j = \overline{1, m}$, m – общее число участков в сети. В частности, для водопроводных сетей функциональная зависимость принимает квадратичный вид, т.е. $\beta = 2$ [1, с. 46].

Классическая задача распределения потока гидравлической сети может быть записана как [2]:

$$\begin{cases} A^T H = Sf(q) \\ Aq = Q \end{cases}, \quad (3)$$

где A – матрица инцидентности графа сети без последней строки, H – значения напора на узлах сети, Q – расходы (утечки) в узлах.

Система (3) является нелинейной, её можно решить, например, методом Ньютона или, более эффективным для данной задачи – Методом Глобального Градиента (GGA) [3].

При решении системы (3) могут быть неизвестны отдельные компоненты вектора расхода в узлах Q . В реальной системе величина расхода имеет вероятностный характер – её значения зависят от разных факторов, воздействие которых заранее предсказать невозможно.

Пусть, неизвестна величина расхода Q_i на некотором -узле сети, а также имеются множество аналогичных рассматриваемому узлам (которые относятся к другим водопроводным сетям) и соответствующих им данных о величине расхода за некоторый период времени. Тогда, можно использовать в качестве неизвестного значения среднее значение расхода по аналогичным объектам.

Алгоритм замещения неизвестной величины Q_i для -узла:

- 1) подобрать данные о величине расхода на некоторых аналогичных рассматриваемому узлах;
- 2) с помощью критерия согласия определить по сформированной выборке теоретическое распределение $F(x)$ случайной величины расхода;
- 3) найти значение теоретического математического ожидания $E[Q_i]$ случайной величины расхода, формула для нахождения которого определяется видом выбранного распределения $F(x)$;
- 4) с помощью выбранного метода решения нелинейных систем уравнений выполнить гидравлический расчет.

Пример выполнения расчета

Пусть, имеется система из девяти узлов и девяти участков, схема которой изображена на рисунке 1.

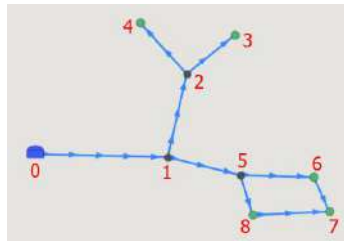


Рис.1 – Пример схемы водопроводной сети

Положим начальные гидравлические характеристики узлов сети значениями из *таблицы 1*, где Q_i – величина расхода (утечки) узла, H_i – значение напора воды на узле.

Таблица 1 – Гидравлические характеристики узлов

	Идентификаторы узлов								
	0	1	2	3	4	5	6	7	8
Q_i , л/с	–	0	0,001	0,7	0,7	0,0007	0,75	0,75	0,7
H_i , м	90	–	–	–	–	–	–	–	–

Результаты выполнения гидравлического расчета приведены в *таблице 2*, где ΔH_j – величина потерь напора на участке сети, q_j – расход участка.

Таблица 2 – Результаты гидравлического расчета

	Идентификаторы участков								
	0-1	1-2	2-3	2-4	1-5	5-6	6-7	5-8	8-7
ΔH_j , м	3,04 56	0,29 5	0,06 23	0,04 17	0,67 58	0,15 18	0,00 98	0,12 95	0,03 21
q_j , л/с	3.60 17	1.40 1	0.7	0.7	2.20 07	1.05 58	0.30 58	1.14 42	0.44 42

Исследовать точность предложенного подхода будем при разном количестве пропусков. В качестве замещающих значений будут использоваться математические ожидания случайной величины расхода, формулы для определения которых будут зависеть от выбранных законов распределения: $U(0; 0.002)$ – для промежуточных узлов и $\Gamma(0.37; 2)$ – для потребителей. Распределения выбраны таким образом, чтобы теоретическое математическое ожидание не совпадало точно с замещаемым значением расхода на узле, тем самым исследуемая задача была приближена к практической.

Полученные результаты сравним с найденными ранее точными результатами величин потерь напора и расхода по значениям максимальной абсолютной погрешности и максимальной относительной погрешности среди всех участков. Результаты оценки приведены в *таблице 3*, где Δ_{\max} – значение максимальной абсолютной погрешности, δ_{\max} – значение максимальной относительной погрешности.

Таблица 3 – Оценки полученных результатов

Отсутствует данных	q		ΔH	
	Δ_{\max} , л/с	δ_{\max} , %	Δ_{\max} , м	δ_{\max} , %
25% (2 узла)	0,041	5,7143	0,0942	0,1088
50% (4 узла)	0,0513	5,7143	0,1118	0,129
75% (6 узлов)	0,0913	5,7143	0,1949	0,2262

Сравнивая результаты из *таблицы 3* для разного числа пропусков, следует вывод, что при достаточно точной оценке вида распределения неизвестной случайной величины можно получать низкую погрешность результатов даже при большом количестве пропусков в исходных данных.

Как повысить точность оценки распределения:

- определять распределения несколькими критериями согласия;
- использовать выборку как можно большей размерности;
- для каждого объекта подбирать вид распределения отдельно, а не универсальный для всех.

Заключение

Предложенный подход по замещению пропусков в исходных данных для расчета характеристик коммунальных сетей можно рекомендовать к использованию для решения практических задач, при условии более точного подбора вида распределения замещаемой случайной величины.

Литература:

1. Абрамов Н. Н. Водоснабжение : учебник для вузов / Н. Н. Абрамов ; – 2-е изд., доп. и перераб. – Москва :Стройиздат, 1974. – 480 с.
2. Меренков А. П. Теория гидравлических цепей : учеб. пособие / А. П. Меренков, В. Я. Хасилев ; Сибирский энерготехнический институт. – Москва : Наука, 1985. – 294 с.
3. Korelstein L. Modernization of Todini Global Gradient Algorithm for hydraulic analysis of networks with choked flow / Leonid Korelstein // E3S Web of Conferences. – 2020. – № 219 (01003).

ИМИТАЦИОННЫЕ МОДЕЛИ ПРИВОДА МЕХАНИЗМА ПЕРЕКЛЮЧЕНИЯ СТРЕЛОЧНОГО ПЕРЕВОДА

Е.А.Яковлева, Н.А. Маслов

**Сибирский государственный университет путей сообщения,
г. Новосибирск, namaslov@mail.ru**

Научный руководитель: Маслов Н.А., к.т.н., доцент

В работе рассмотрена принципиальная схема привода и управления стрелочным переводом с пусковым блоком и центральным приводом; разработана схема ее имитационной модели; разработаны схем имми-

тационных моделей привода механизма переключения стрелочного перевода с программируемым реле и логическим контроллером.

In this paper, we consider the principal circuit of the drive and control of the point relay with a starting unit and a central drive, developed the scheme of its simulation model, developed simulation models of the switch mechanism drive of the point relay with a programmable relay and a logic controller.

Цель работы - создание имитационных моделей привода механизма переключения стрелочного перевода в программе AutomationStudio.

Задачи работы: изучение принципиальной схемы привода и управления стрелочным переводом; разработка схемы имитационной модели привода и управления стрелочным переводом с пусковым блоком и центральным приводом; разработка схемы имитационной модели привода и управления стрелочным переводом с программируемым реле; разработка схемы имитационной модели привода и управления стрелочным переводом с программируемым логическим контроллером.

Новосибирск-Восточный - железнодорожная станция Новосибирского региона Западно-Сибирской железной дороги, расположенная в восточной части города Новосибирска. Станция занимается приемом и выдачей грузов повагонными и мелкими отправлениями, загружаемых целыми вагонами, только на подъездных путях и местах необщего пользования. На станции расположено более 50 стрелочных переводов, 80% из которых с электрической централизацией и 20% с ручным управлением. В электрической централизации станции для управления стрелочными электроприводами применяется наиболее распространённая двухпроводная схема с пусковым стрелочным блоком (Рисунок 1).[1] Особенность этой схемы заключается в том, что в ней используется привод с электродвигателем постоянного тока на рабочее напряжение 160 В.

Основная часть аппаратуры управления стрелочным переводом размещена на посту электрической централизации и сосредоточена в следующих блоках:

-«ПС»- пусковой стрелочный блок (содержит два одинаковых комплекта двухпроводной схемы управления стрелкой);

-«С» - контрольный стрелочный блок, который используется в схемах блочной маршрутной релейной централизации.

Для индикации положения стрелки на аппарате управления размещены следующие лампы:

-«зеленая»- показывает плюсовое положение стрелки;

-«желтая»- стрелка находится в минусовом положении;

-«красная»- стрелка в среднем положении.

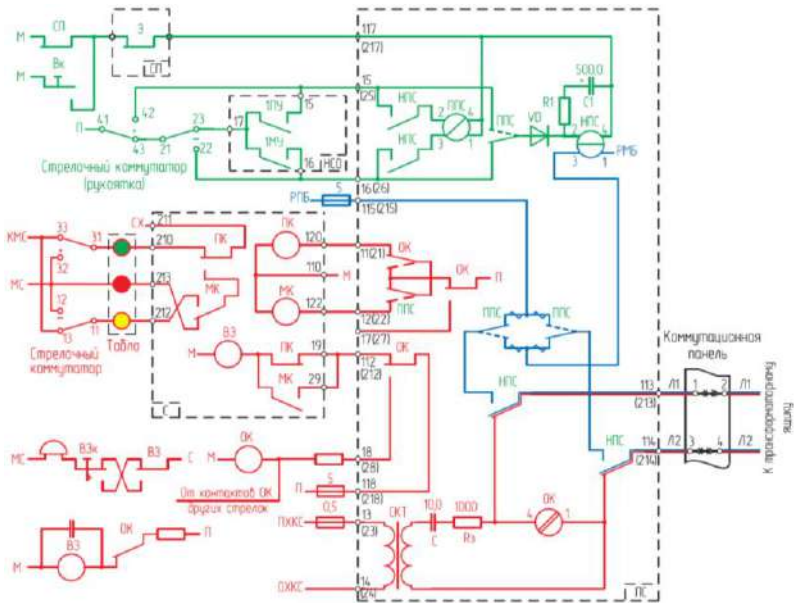


Рис.1 - Двухпроводная схема с пусковым стрелочным блоком

Управляющая цепь стрелки выполняет следующие задачи: включает пусковые приборы при нажатии кнопок управления (переводе коммутаторов или рукояток) или включении реле ПУ (МУ); исключает возможность перевода стрелки при занятой или замкнутой в маршруте стрелочной секции; фиксирует окончание перевода стрелки. Рабочая цепь подключает двигатель к источнику питания, позволяет произвести реверсирование электродвигателя при крайнем или промежуточном положении стрелки, контролирует окончания перевода. Контрольная цепь предназначена для непрерывного контроля трёх положений стрелочного перевода: плюсового, минусового и промежуточного.[1]

Для модернизации системы управления стрелочным переводом предлагается использовать программируемое интеллектуальное реле или программируемый логический контроллер.

Программируемые интеллектуальные реле являются одной из разновидностей программируемого логического контроллера. Принципиально программируемое реле работает аналогично логическому контроллеру и отличается от него малым объемом оперативной и программной памяти, что приводит к невозможности сложных математических вычислений. Кроме того, количество каналов ввода-вывода как цифровых, так и аналоговых у интеллектуальных реле также невелико, поэтому

область их применения достаточно ограничена.[2] Существует практика ограниченного применения программируемых реле в аппаратуре для местного перевода острых стрелок на подземном рельсовом транспорте шахт и рудников.

Программируемый логический контроллер - электронная вычислительная машина для автоматизации технологических процессов. Такие контроллеры могут в реальном времени собирать, преобразовывать, обрабатывать, хранить информацию и вырабатывать команды управления. Они имеют конечное количество входов и выходов, подключенных к ним датчиков, ключей, исполнительных механизмов к объекту управления. [2] Одним программируемым логическим контроллером можно заменить сотни механических или электрических реле, в т.ч. управляющих стрелочными переводами.

В программе AutomationStudio разработаны схемы имитационных моделей привода и управления стрелочным переводом: с пусковым блоком и центральным приводом, с программируемым реле и с программируемым логическим контроллером.

Разработанные имитационные модели могут быть использованы для:

- изучения вариантов систем управления механизмами переключения стрелочными переводами в ВУЗе;
- повышения квалификации кадров для железных дорог;
- моделирования неисправностей и их поиска при технической экспертизе;
- научных исследований, например, изучения особенностей функционирования вновь разрабатываемых систем управления стрелками в характерных и специфических условиях, в т.ч. аварийных.

Литература:

1. *Валиев Р.Ш., Валиев Ш.К.* Двухпроводная схема управления стрелкой с пусковым блоком ПС с центральным питанием. - Екатеринбург: Изд-во ООО «Вебстер», 2011 – 48 с.
2. ОВЕН. Оборудование для автоматизации. Каталог продукции. - М.: Изд-во «Овен», 2021 – 488 с.

**Секция ЭКОНОМИКО-МАТЕМАТИЧЕСКИЕ МЕТОДЫ,
СТАТИСТИКА И ЭКОНОМЕТРИКА**

**ОЦЕНКА ВРЕМЕНИ ДОЖИТИЯ ПАЦИЕНТОВ С БОКОВЫМ
АМИОТРОФИЧЕСКИМ СКЛЕРОЗОМ НА ОСНОВЕ
ХАРАКТЕРИСТИК ГРАФА ВОДОРОДНЫХ СВЯЗЕЙ**

Д.К. Воронкина

**Новосибирский государственный технический университет,
г. Новосибирск, voronik1801@yandex.ru**

Научный руководитель: Тимофеев В.С., д.т.н., доцент

В работе рассматривается задача построения модели оценки времени дожития пациентов с боковым амиотрофическим склерозом. Реализован алгоритм частичной регрессии наименьших квадратов на обучающей выборке, созданной из характеристик вершин графа, построенного для водородных связей пациентов. Проведена оценка точности модели.

The paper considers the problem of building a model for estimating survival time of patients with amyotrophic lateral sclerosis. The algorithm of partial least squares regression is implemented. For training the algorithm, a sample created on the basis of the characteristics of the patient's hydrogen bonds graph vertices was used. Accuracy of the model was evaluated.

Боковой амиотрофический склероз (БАС) – прогрессирующее, неизлечимое дегенеративное заболевание центральной нервной системы, приводящее к параличам и последующей атрофии мышц. В 20% случаев болезнь связана с мутациями гена супероксиддисмутазы–1 (SOD1). Изучением мутаций в данном гене занимаются ученые из института цитологии и генетики (ИЦиГ), предоставившие данные для построения модели прогнозирования времени дожития пациентов.

Данные содержат 182 признака водородных связей фермента SOD1 для 72 пациентов. Так как признаки водородных связей представлены донорно-акцепторным взаимодействием фермента с описывающим его весом, то для каждого пациента можно рассматривать фермент в виде графа.

Строятся три различных варианта графа, отличающиеся в представлении вершин:

- 1) Вершина представлена в виде полной водородной связи;
- 2) Вершина представлена в виде позиции аминокислотного остатка с субъединицей водородной связи;
- 3) Вершина представлена в виде позиции аминокислотного остатка без субъединицы водородной связи.

Предполагается, что течение болезни влияет на структуру водородных связей в ферменте SOD1. Идея заключается в выявлении наличия зависимости между временем дожития пациента и изменением структуры фермента, поэтому рассматриваются следующие характеристики вершин графа.

1) Степень вершины графа— это количество ребер, прилегающих к узлу [3, 4].

2) Коммуникабельность между парой вершин (u, v) в графе представляет собой сумму переходов разной длины, начинающихся в узле u и заканчивающихся в узле v [3, 4].

Из вычисленных характеристик строятся обучающие выборки для регрессионной модели. Рассматривается уравнение следующего вида [1]:

$$Y = XB + E ,$$

где B – неизвестные коэффициенты регрессии; E – случайная ошибка регрессии, X – матрица объясняющих переменных, составленная из вычисленных характеристик вершин, $Y = (y_1, y_2, \dots, y_n)^T$ – вектор наблюдаемых значений, составленный из времени дожития пациентов.

Число вычисленных характеристик вершин, выступающих объясняющими переменными, варьируется от 698 до 1068, что является проблемой, так как количество наблюдаемых значений равно 72. По этой же причине высока вероятность возникновения мультиколлинеарности между объясняющими переменными, что может привести к малоэффективному оцениванию параметров. Учитывая все особенности данных, задача решается методом частичных наименьших квадратов или partial least square (PLS). PLS регрессия выделяет небольшое количество латентных переменных, в пространстве которых связь между зависимой переменной и предикторами достигает максимального значения [2].

Алгоритм частичной регрессии наименьших квадратов [2]:

1. $X^{(0)} = X$ – инициализация X на нулевой итерации;
2. $w^{(0)} = X^T Y / \|X^T Y\|$ – вектор весов на нулевой итерации;
3. Для $k=0$ до N – цикл, где N – количество латентных переменных;
4. $t^{(k)} = X^{(k)} w^{(k)}$ – вектор проекции на k -й итерации;
5. $p^{(k)} = X^{(k)T} t^{(k)} / t^{(k)T} t^{(k)}$ – вектор нагрузки на k -й итерации;

6. $q = Y^T t^{(k)} / t^{(k)T} t^{(k)}$ – значение оценочного вектора коэффициентов регрессии;

7. Если $q = 0$, то завершить цикл;

8. Иначе $X^{(k+1)} = X^{(k)} - t^{(k)} p^{(k)T}$ – вычисление матрицы $X^{(k+1)}$ на k -й итерации;

9. $w^{(k+1)} = X^{(k+1)T} Y$ – вычисление вектора весов;

10. Составить матрицу W из векторов w , матрицу P из векторов p ;

11. $\hat{B} = W(P^T W)^{-1} q$ – оценка коэффициентов регрессии;

12. $E = q - P^T \hat{B}$ – вычисление случайных остатков регрессии.

Для оценки качества модели используются следующие показатели [1]:

1) Коэффициент детерминации:

$$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2},$$

где \hat{y}_i – значение оценки времени дожития пациента, \bar{y} – среднее значение времени дожития пациентов, n – количество пациентов;

2) Стандартная среднеквадратичная ошибка:

$$S = \frac{1}{n} \sqrt{\sum_i (y_i - \hat{y}_i)^2};$$

3) Кросс-валидация:

$$CV = \frac{1}{n} \sqrt{\sum_{i=1}^N (y_i - \hat{y}_{(-i)})^2},$$

где $\hat{y}_{(-i)}$ – полученное значение после оценки с учетом того, что отсутствует i -ая строка.

Вычисленные результаты оценки модели, с тремя различными вариантами построения графа, отличающиеся в представлении вершин представлены в следующих таблицах 1-3:

1) Вершина представлена в виде полной водородной связи;

Таблица 1 – Результаты оценки модели

Значение\Число компонент	5	7	10	12
Коэффициент детерминации	0.81	0.89	0.95	0.96
Среднеквадратичная ошибка	0.27	0.20	0.14	0.11
Кросс–валидация	0.75	0.80	0.92	0.96

2) Вершина представлена в виде позиции аминокислотного остатка с субъединицей водородной связи;

Таблица 2 – Результаты оценки модели

Значение\Число компонент	5	7	10	12
Коэффициент детерминации	0.88	0.95	0.98	0.99
Среднеквадратичная ошибка	0.22	0.13	0.06	0.03
Кросс–валидация	0.78	0.78	0.79	0.80

3) Вершина представлена в виде позиции аминокислотного остатка без субъединицы водородной связи.

Таблица 3 – Результаты оценки модели

Значение\Число компонент	5	7	10	12
Коэффициент детерминации	0.75	0.81	0.84	0.85
Среднеквадратичная ошибка	0.31	0.27	0.25	0.24
Кросс–валидация	0.66	0.68	0.71	0.73

Полученные результаты показывают, что построенная модель дает хорошее описание имеющихся данных. Увеличение количества латентных переменных (компонент), ведет к уменьшению ошибки оценки. Кросс–валидация с увеличением числа компонент растет. Можно сделать вывод о том, что изменение структуры водородных связей фермента SOD1 влияет на время дожитие пациентов.

Литература:

1. Тимофеев В.С., Фаддеев А.В., Щеголкин В.Ю. Эконометрика: учебник. – Новосибирск: Изд-во НГТУ, 2015. – 354 с.
2. Kee Siong Ng. A Simple Explanation of Partial Least Squares. – 2013.
3. Aric Hagberg, Dan Schult, Pieter Swart. NetworkX Reference [Electronic resource]. – 2021. – Release 2.6.2. URL: https://networkx.org/documentation/stable/_downloads/networkx_reference.pdf.
4. Омельченко А.В. Теория графов. – М.Ж МЦНМО, 2018. – 416 с.

ИМИТАЦИОННОЕ МОДЕЛИРОВАНИЕ ДВИЖЕНИЯ СУДОВ И РАБОТЫ ПОРТА

Д.А. Гринько

Новосибирский государственный технический университет,
г. Новосибирск, grinko.2020@stud.nstu.ru

Научный руководитель: Фаддеев А.В., к.т.н., доцент

Имитационное моделирование представляет собой процесс построения обобщенной компьютерной модели некоей системы с алгоритмическим описанием основных правил ее поведения и процессов. Имитационная модель, как правило, создается для исследования возможных сценариев развития системы при вариации определенных параметров. Целью данного исследования является создание имитационной модели для анализа работы речного порта. В качестве изменяющегося параметра выступают прибывающие суда с различными характеристиками (осадка судна, наличие груза, цель прибытия), которые меняют сценарий работы речного порта.

Simulation modeling is the process of constructing a generalized computer model of a certain system with an algorithmic description of the basic rules of its behavior and processes. A simulation model, as a rule, is created to study possible scenarios for the development of a system when certain parameters are varied. The purpose of this study is to create a simulation model for analyzing the operation of a river port. Arriving vessels with different characteristics (draft, cargo availability, purpose of arrival) act as a changing parameter, which change the scenario of the river port operation.

Водный транспорт является одним из ключевых элементов Мировой экономической системы. Объемы морских и речных перевозок постоянно возрастают, также повышаются требования к качеству перевозки грузов водным транспортом (своевременность, безопасность, надежность и т.д.). Современные водные транспортные системы включают транспортные средства (пассажирские суда, контейнеровозы, сухогрузы, наливные суда и др.) водные транспортные магистрали, а также морские (и речные) порты. Очевидно, что совершенствование инфраструктур водного транспорта является важным фактором ускоренного развития национальной и международной экономик [1-2]. Исходя из этого, имитационное моделирование движения водного транспорта является наиболее доступным и полезным способом развития судоходства.

Схема, согласно которой происходит движение судов по реке представлена на рис. 1.

Типы судов и их характеристики, которые используются в данной работе при генерации объекта «Судно», взяты из [3].

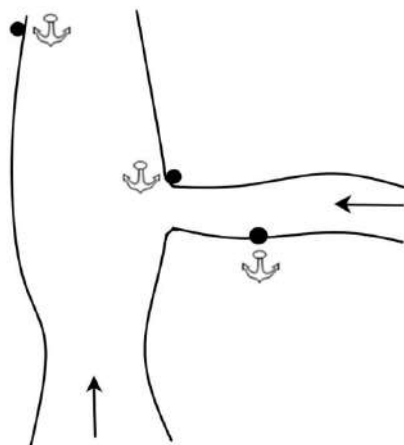


Рис. 1 – Схема речного пути

Для данного исследования были определены шесть сценариев работы речного порта. Начало любого сценария содержит условие загрузки судна (рис.2 а). Для судов, которые прибыли пустыми на терминал, предусмотрено два сценария (рис.2 б). Для загруженных судов реализуются остальные четыре сценария (рис.2 в).

Аббревиатуры для сценариев означают следующее:

- «EA» - Empty, Anchorage (Пустой, Якорная стоянка).
- «EB» - Empty, Berth (Пустой, Причал).
- «FAU» - Full, Anchorage, Unloading (Полный, Якорная стоянка, Разгрузка).
- «FAUL» - Full, Anchorage, Unloading, Loading (Полный, Якорная стоянка, Разгрузка, Загрузка).
- «FBU» - Full, Berth, Unloading (Полный, Причал, Разгрузка).
- «FBUL» - Full, Berth, Unloading, Loading (Полный, Причал, Разгрузка, Загрузка).

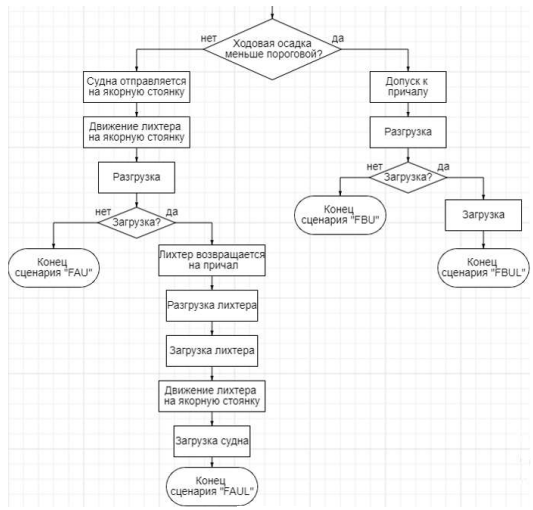
На первом этапе выбирается судно из данных [3] случайным образом, также случайно задается наличие груза у пришедшего судна и выбирается цель прибытия. Установив все необходимые параметры, происходит построение сценария согласно представленной выше блок – схеме. В окне вывода отображается порядок процессов, совершенных при обработке пришедшего судна, и время, затраченное на их выполнение.



a)



б)



в)

Рис. 2 – Блок – схема процессов

В конце выполнения программы строятся три гистограммы по различным параметрам (осадка судов, масса груза и продолжительность операций), а также круговые диаграммы времени для каждого прибывшего судна.

Таким образом, созданная модель позволяет случайным образом генерировать события из созданных сценариев при изменении входящих параметров. Результатом работы данной модели является списки сгенерированных событий и их графическое представление по временному параметру.

Литература:

1. Pipitsoulis, C. The EU eMaritime initiative – Single Window, with a view to the near future. In Logious Conference. Rotterdam. 2010.
2. Интеллектуальные технологии на водном транспорте // И.Г. Малыгин, В.И. Комашинский, М.А. Асаул // «Морские интеллектуальные технологии» - № 2 (36) – Т.2 - СПб.: НИЦ "МОРИНТЕХ", 2017. – С. 54-58.
3. Имитационное моделирование экономических процессов / А.А. Емельянов, Е.А. Власова // Материалы четвертой всероссийской научно-практической конференции по имитационному моделированию и его применению в науке и промышленности «Имитационное моделирование. Теория и практика» ИММОД. – СПб.: ОАО «ЦТСС», 2009. – С. 19-26.

НОВЫЕ ПОСТАНОВКИ ЗАДАЧИ МАРШРУТИЗАЦИИ ТРАНСПОРТА С УЧЕТОМ СОВРЕМЕННЫХ ТЕНДЕНЦИЙ

М. М. Заводченко

Новосибирский государственный технический университет,

г. Новосибирск, misha_zavodchenko@mail.ru

Научный руководитель: Карманов В.С., к.т.н., доцент

Предлагается и описывается новая постановка задачи маршрутизации транспорта учетом рейтинга исполнителей.

A new formulation of the transport routing problem is proposed and described, taking into account the rating of performers.

Введение

Задача маршрутизации транспортных средств (ЗМТ) это одна из хорошо известных задач комбинаторной оптимизации, в которой необходимо найти маршрут, а зачастую набор оптимальных маршрутов до заданного множества точек-потребителей, для определенного парка транспортных средств. Она относится к классу NP-трудных задач.

В настоящее время существуют классические, общеизвестные постановки задачи маршрутизации транспорта, к ним можно отнести следующие варианты: задача с учетом ограниченной грузоподъемности транспортных средств (ТС); оказание услуги потребителю в удобное для него время, так называемая задача с временными окнами; использование нескольких депо для обслуживания клиентов; возможность появления новых клиентов, которые должны быть добавлены в существующие или созданы новые маршруты (динамическая ЗМТ); задача с неоднородным парком транспортных средств[1-3]. Все выше перечисленные варианты задачи, интуитивно понятны и вытекают из самой сути ЗМТ. Однако с течением времени и развитием сферы услуг, частью которой является транспортная логистика, появляется необходимость в выделении новых постановок задачи, которые будут соответствовать требованиям конкретной компании или человека.

Задача маршрутизации транспорта с учетом рейтинга исполнителей

В современном мире и современной мировой экономике достаточно значимая роль отводится организациям и предприятиям, работающим в сфере услуг. Сегодня достаточно часто организации вводят различные системы и механизмы мотивации своих сотрудников для повышения критериев качества и количества выполненных работ. Как правило, мотивирующий фактор в таких схемах выражается в денежном эквиваленте. Очевидно, что размер дополнительных выплат зависит от некоторой оценки работодателем работы, которую выполнил сотрудник. Это могут быть как прямые выплаты, так и более сложные схемы мотивации персонала. Так, например, среди организаций по осуществлению монтажных работ по подключению абонентов к услугам связи популярен подход, в котором на основании некоторого рейтинга сотрудников происходит их поощрение, выражающееся в получении наиболее выгодного маршрута для выполнения заявок.

Выгодность маршрута определяется по его расстоянию, времени объезда, стоимости оказания услуг, включенных в маршрут, выражаемая в некоторой условной величине, например, в баллах. Очевидно, что чем меньше расстояние, время объезда маршрута и чем

выше стоимость услуг, оказанных при его объезде, тем он привлекательнее для исполнителя.

В связи с этим перед работодателем возникает вопрос о составлении «справедливых» маршрутов для каждого исполнителя на основании его рейтинга, при этом минимизируя свои затраты за счет снижения пройденного расстояния и времени объезда маршрутов. Именно такую постановку задачи и будем называть задачей маршрутизации транспорта с учетом рейтинга исполнителей. А так как задача будет применяться для сферы услуг, то в постановке задачи будем учитывать попадания моментов приезда и отъезда из точки в указанное и удобное потребителю временное окно.

Математической постановка задачи.

Для ее описания будем использовать следующие обозначения:

Взвешенный граф $G(V, E)$, где $V = \{v_0, v_1, \dots, v_n\}$ — множество вершин, v_0 — депо (склад);

$E = \{(v_i, v_j) \mid i \neq j\}$ — множество ребер;

C — матрица стоимости проезда, где c_{ij} стоимость проезда между вершинами v_i и v_j ;

Для каждой вершины v_i заданы следующие данные: временное окно $[tl_i; tr_i]$, в течение которого ее необходимо посетить, Ts_i — длительность нахождения в этой вершине (время обслуживания по нормативу), а также V_i — стоимость вершины (измеряется в условной величине, баллах); m — количество бригад, для которых необходимо сформировать маршруты (фактически этот параметр накладывает ограничение на количество маршрутов формируемых при решении задачи).

Моменты времени обозначим как t , а интервалы времени — T и введем следующие величины:

t_0 — момент времени начала движения ТС, момент выезда ТС из депо;

t_i^{arrival} — момент прибытия ТС в вершину i ;

$t_i^{\text{depatrure}}$ — момент отправления ТС из вершины i ;

Tv_i — интервал времени проезда ТС в i -ю вершину из $(i - 1)$ -ой вершины;

$Tc_i = Tv_i + Ts_i$ — общее время, потраченное на i -ю вершину;

$T_{\text{sum}}(i) = \sum_{j=0}^i Tc_j$ суммарное время, затраченное на посещение вершин с первой по текущую;

$M = \{M_1, M_2, \dots, M_m\}$ — набор маршрутов;

$$M_j = \begin{matrix} & v_0 & v_1 & v_2 & \dots & v_{n+1} \\ \begin{matrix} v_0 \\ v_1 \\ v_2 \\ \dots \\ v_{n+1} \end{matrix} & \begin{bmatrix} 0 & 0 & 1 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 1 \\ \dots & \dots & \dots & \dots & \dots \\ 1 & 0 & 0 & \dots & 0 \end{bmatrix} \end{matrix} - \text{маршрут};$$

$V_{lr}(M)$ – баллы в день $г$ для бригады l , где $l = 1..m$, $г = 1..7$

R_j недельный рейтинг для бригады j , где $j = 1..m$

$$x_{ijk} \begin{cases} 1, & \text{если ТС } k \text{ едет напрямую от } i \text{ к } j, \\ 0, & \text{иначе} \end{cases}$$

Используя обозначения выше, задача маршрутизации транспорта с учетом рейтинга исполнителей может быть записана следующим образом:

$$\min \sum_{i=0}^n \sum_{j=0}^n \sum_{k=0}^n c_{ij} x_{ijk} \quad (1)$$

Целевая функция (1) определяет суммарную стоимость всех маршрутов

$$\sum_{k=1}^m \sum_{i=0}^n x_{ijk} \geq 1, j = 0, \dots, n; \quad (2)$$

Ограничение (2) гарантирует, что каждого клиента посещает одно транспортное средство

$$\sum_{i=0}^n x_{ijk} = \sum_{i=0}^n x_{jik} = 1, j = 0, \dots, n; k = 1, \dots, m \quad (3)$$

Ограничение (3) показывает, что если транспортное средство k приезжает к клиенту j , то оно должно покинуть его [1].

$$\sum_{j=1}^n x_{0jk} \leq 1, j = 1, \dots, n; \quad (4)$$

Формула (4) гарантирует, что будет использовано не более k ТС [2].

$$tl_i \leq t_0 + T_{\text{sum}}(i - 1) + Tv_i \quad (5)$$

$$t_0 + T_{\text{sum}}(i) \leq tr_i \quad (6)$$

Условия (5)-(6) являются ограничениями на попадания моментов времени прибытия и отбытия ТС в заданное временное окно i -ой вершины.

$$x_{ijk} \in \{0,1\}, i, j, = 0, \dots, n; k = 1, \dots, m; \quad (7)$$

$$\sum_{i=1}^m \left| \frac{R_i}{\sum_{j=1}^m R_j} - \frac{\sum_{r=1}^7 B_{jr}(M)}{\sum_{r=1}^m \sum_{j=1}^7 B_{jr}(M)} \right| \rightarrow \min_m \quad (8)$$

Функция (8) определяет суммарное отклонение балльной стоимости маршрутов от текущего рейтинга исполнителей.

Методы решения

В связи с тем, что, задача маршрутизации транспорта с учетом рейтинга исполнителей является новым вариантом задачи, не существует и не были предложены методы для ее решения, которые дали бы хорошие результаты. Можно предложить несколько вариантов для ее решения.

Первый предлагаемый подход основан на идее минимизации величины отклонения стоимости маршрута от рейтинга бригады исполнителя (9)

$$\sum_i |\tilde{B}_i - \tilde{R}_i| \rightarrow \min \quad (9)$$

где B_i – стоимость построенного маршрута в баллах, R_i – рейтинг бригады. Величины \tilde{B}_i и \tilde{R}_i нормированы и приведены к 100-балльной шкале. В таком подходе к решению применяется любой из известных методов решения ЗМТ с временными окнами. Затем полученные маршруты распределяются по правилу (9).

В качестве альтернативного подхода распределения маршрутов можно предложить жадное распределение, в основе которого лежит идея " чем больше рейтинг, тем дороже заявка".

Заключение

Предложена новая постановка задачи маршрутизации транспорта, и подходы для ее решения. Требуется дальнейшие исследования по разработке методов решения такой постановки задачи, а также проверки их эффективности.

Литература

1. Пожидаев М.С. Алгоритмы решения задачи маршрутизации транспорта: дис. канд. техн. наук: 05.13.18 / Пожидаев М. С.; науч. рук. Костюк Ю. Л.; Томский. гос. ун-т. Томск, 2010. — 136 с.

2. Abdul Ghani, N., Shariff, S. and Zahari, S. (2016) An Alternative Algorithm for Vehicle Routing Problem with Time Windows for Daily Deliveries. *Advances in Pure Mathematics*, 6, 342-350. doi: 10.4236/apm.2016.65025.

3. Трофимов Д., Федуков А. Задача маршрутизации транспорта [Электронный ресурс] Режим доступа: https://www.lobanov-logist.ru/library/all_articles/55059/ — Загл. с экрана.

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ИНДИКАТОРА ТЕХНИЧЕСКОГО АНАЛИЗА

Е.Г. Корепанова¹, Ю.В. Филимоненко¹, Л.Е. Корепанов²

¹ Сибирский государственный университет науки и технологий
имени академика М.Ф. Решетнева, г. Красноярск,
korepanova.eg@gmail.com

²Балтийский государственный технический университет
«ВОЕНМЕХ» им. Д.Ф. Устинова, г.Санкт-Петербург,
korepanov0804@gmail.com

В статье рассматривается один из наиболее популярных биржевых индикаторов – инструмента для прогнозирования цены акции любой компании. Предстоит рассмотреть математическую модель, на которой основывается индикатор и его принцип работы.

The article discusses one of the most popular stock indicators - a tool for predicting the price of a stock of any company. It is necessary to consider the mathematical models on which the indicator and its principle of operation are based.

На сегодняшний день все больше и больше людей отказываются от размещений денежных масс на банковских вкладах в пользу инвестиций, из-за наиболее выгодного накопительного процента. В последнее 5 лет мы видим, что уровень инфляции составил около 20%. Тем самым население стремится, как минимум сохранить свои средства. Если посмотреть на уровень доходности разных продуктов, то мы видим следующую картину:

- депозиты – 5,5 % годовых,
- облигации государственного займа – 11,0 % годовых,
- акции США – 15 % годовых,
- российские акции в среднем – 17 % годовых.

Встает вопрос, - как копить на долгосрочные цели. Рассматривая список накоплений, в первых строчках можно назвать недвижимость, люди более зрелого возраста задумываются о безбедной старости, и конечно многие мечтают путешествиях. Чем раньше человек задумывается о инвестировании, предположим, когда дети очень маленькие, то по окончанию школы, можно рассчитывать на хорошую сумму инвестирования в обучения ребенка. Надо заметить инвестиции, это забота о будущем, грамотное и осознанное отношение к материальным ценностям и жизни в целом. В богатом арсенале инвестиционных инструментов, каждый выбирает более приемлемый вариант для себя. Можно ориентироваться на разные активы, составлять инвестиционный портфель под различные цели. У большинства нет никакого понимания, как работает инвестиционный портфель, какие инструменты, рычаги и как они устроены [1]. По аналитической статистике Московской биржи, за период 2020 года граждане нашей страны открыли счетов на 6,7 млн., что составило примерно 4.4 % населения страны. Конечно, до американских показателей нам еще далеко, там мы видим вовлеченность граждан в инвестиции составляет 55%.

Российские аналитики инвестиционных рынков констатируют факт увеличения финансовой грамотности россиян, тем самым граждане все чаще задумываются об эффективности использований своих капиталов. Важно начать инвестировать сейчас, как было ранее замечено, что процент по депозитным вкладам не дает гарантии даже на сохранение средств, не говоря уже о доходности. Решение граждан инвестировать так же поощряется государством, в налоговом вычете. Низкий порог входа позволяет начать с малого, а это от 500-100 рублей за одну ценную бумагу. Показатели ликвидности в инвестициях в разы выше, даже если взять к примеру, в недвижимости. На брокерских счетах можно так же вести валютные операции, освобождаясь от дополнительной комиссии коммерческого банка. Никто не знает, сколько пенсионных реформ ждет нас за ближайшие 30 лет и какая нам будет назначена пенсия от государства, и тем самым многие инвестируют как раз для того чтобы получать стабильный пассивных доход и пораньше выйти на пенсию. Развитие технологий привело к тому, что инвестировать стало легко и просто. Основная масса коммерческих банков и бирж, имеет свои консультационные подразделения и отделы в помощь начинающим инвесторам [2].

Инвестиции можно разделить на две категории: покупка акций компаний и приобретение облигаций займа. Рассмотрим наиболее прибыльную, а именно покупка акций – ценных бумаг, показывающих

долю лица на владения компанией. Каждая акция имеет свою цену, называемую котировкой, зависящую от так называемых рыночных сил, наибольшее влияние имеет закон спроса и предложения. Индикаторы производят определенные расчёты по заданным формулам, а затем автоматически выдают результаты в виде точек на графике, которые в последствии сами превращаются в кривую. Опытные инвесторы прислушиваются к результатам индикаторов как раз из чего и строиться их стратегия инвестирования.

Самым популярным и простым индикатором является «скользящая средняя» - кривая, которая в каждой точке является средним арифметическим определенного количества значений цены за промежутки времени. Её можно представить в виде формулы:

$$\text{Значение скользящей средней} = \frac{\text{Сумма цен за период времени}}{\text{Период времени}}$$

Цены двигаются по трендам, существует два типа трендов восходящий при котором цена становится выше и нисходящий при котором идет на снижение. Скользящая средняя является усредненным показателем самой цены, и ее направление на графике считается более объективным трендом, что помогает нам в выборе стратегии.

Рассмотрим принцип работы, скользящей средней на примере реального графика «График «Индекс Московской биржи»».



Рис.1 – График «Индекс Московской биржи»

На приведенном графике представлен скользящая средняя за период в 4 дня. Если график цены находится выше скользящей средней, то это может говорить о восходящем тренде, а если ниже, то нисходящий, что

представлено на «График «Индекс Московской биржи»». Именно по ней определяется изменение тренда, а исходя из этого определяем продавать или покупать акцию «Индекс Московской биржи».

Рассмотрим принцип работы «Скользящей средней» на примере таблицы с данными цен:

Дата	Цена (усл. ед.)
01.01.2021	100
02.01.2021	120
03.01.2021	110
04.01.2021	130
05.01.2021	140
Трен восходящий	
Если взять «Скользящую среднюю» за 3 дня	
С 01 по 03	$\frac{100+120+110}{3} = 110$
С 02 по 04	$\frac{120+110+130}{3} = 120$
С 03 по 05	$\frac{110+130+140}{3} = 127$
Каждое последующие число увеличивается, даже несмотря на колебание 03.01.2021, благодаря этому делаем вывод тренд восходящий	

Дата	Цена (усл. ед.)
01.01.2021	100
02.01.2021	80
03.01.2021	70
04.01.2021	80
05.01.2021	70
Нисходящий тренд	
Если взять «Скользящую среднюю» за 3 дня	
С 01 по 03	$\frac{100+80+70}{3} = 83$
С 02 по 04	$\frac{80+70+80}{3} = 77$
С 03 по 05	$\frac{70+80+70}{3} = 73$
Каждое последующие число уменьшается, даже несмотря на колебание 04.01.2021, благодаря этому делаем вывод тренд нисходящий	

Основным выводом по проведенному исследованию можно считать то, что в целом индикаторы технического анализа демонстрируют результаты на фондовой бирже при использовании дневных ценовых данных. Подводя итог, можно отметить, что скользящая средняя не предсказывает цену так как является отстающим показателем, а указывает только тенденцию направления цены. Она является наиболее простейшим индикатором, математическая модель проста, но результаты расчетов являются весьма полезными при выборе стратегии инвестирования [3].

Литература:

1. Бучко Ю.В., Применение технического анализа на современном российском фондовом рынке, – М. : Бизнес элайнмент, 2011. – С. 46-54.
2. Володин С.Н., Эффективность технического анализа на различных временных горизонтах инвестирования, – М. : Бизнес элайнмент, 2012. – С. 45-55.

3. Дьяконова Д.О., Применимость технического анализа для акций, различающихся по ликвидности, – М. : Бизнес элайнмент, 2012. – С. 77-89.

НАХОЖДЕНИЕ БИНАРНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ С МИНИМАЛЬНЫМ УРОВНЕМ БОКОВЫХ ЛЕПЕСТКОВ АВТОКОРРЕЛЯЦИОННОЙ ФУНКЦИИ

А.А. Кривецкий

**Новосибирский государственный технический университет,
г. Новосибирск, kriveczkij.2020@stud.nstu.ru
Научный руководитель: Карманов В. С., к.т.н.**

В современных радио- и оптических коммуникационных, навигационных и локационных системах широко используются методы расширения спектра. Эти методы позволяют существенно повысить помехоустойчивость указанных систем. Одним из основных методов является метод расширения спектра прямыми последовательностями.

Spread spectrum methods are widely used in modern radio and optical communication, navigation and location systems. These methods can significantly increase the noise resistance of signals (distinction). One of the main methods is the method of spreading the spectrum by direct sequences.

Введение

Суть метода расширения спектра заключается в повышении тактовой частоты модуляции [1], при этом каждому символу передаваемого сообщения ставится в соответствие некоторая бинарная последовательность, состоящая из -1 и 1.

Наиболее привлекательными с точки зрения помехоустойчивости являются последовательности с малым уровнем боковых лепестков (УБЛ) автокорреляционной функции (АКФ), задаваемой следующим выражением:

$$A_j = \left| \sum_{i=1}^{N-|j|} a_i a_{(i+|j|)} \right|, \quad a_j \in \{-1, 1\}, \quad j = -N, -N+1, \dots, 0, \dots, N, \quad (1)$$

где a_i – i -ый элемент бинарной последовательности; A_j – значение АКФ при j -ом сдвиге; N – длина сигнала.

Уровень боковых лепестков есть отношение модуля максимального значения АКФ при $j \neq 0$ к длине N сигнала, то есть:

$$L = \frac{\max |A_j|}{N}, \text{ где } j = 1, 2, \dots, N. \quad (2)$$

Малый уровень боковых лепестков достигается, в частности, в псевдослучайных M -последовательностях (УБЛ $= 1/\sqrt{N}$) и сигналах Баркера (УБЛ $= 1/N$). Последовательностей Баркера с длиной N больше 13 не существует [2]. В то же время прямой двоичный перебор показывает, что для N больше 13 существуют комбинации с УБЛ менее $1/\sqrt{N}$ [3]. Прямой перебор двоичных кодов при поиске последовательностей с минимальными УБЛ АКФ при $N > 20$ занимает значительное время. Поэтому задача поиска бинарных последовательностей с минимальным УБЛ является актуальной, о чем свидетельствуют исследования и публикации последних лет [4, 5].

Постановка задачи

Разработать метод перебора бинарных последовательностей, состоящих из -1 и 1, и найти такие последовательности, у которых уровень боковых лепестков автокорреляционной функции, задаваемой (1), минимален. То есть, необходимо найти минимальное значение величины L из (2), что эквивалентно задаче:

$$\max \left| \sum_{i=1}^{N-k} a_i a_{i+k} \right| \rightarrow \min, \quad k = 1, 2, \dots, N-1, \quad a_i \in \{-1, 1\}. \quad (3)$$

Методы решения

Последовательность с УБЛ менее $1/\sqrt{N}$ должна удовлетворять следующему критерию:

$$\left| \sum_{i=1}^{N-j} a_i a_{i+j} \right| \leq \sqrt{N}, \quad j = 1, 2, \dots, N-1, \quad a_i \in \{-1, 1\}. \quad (4)$$

Вычисления суммы и произведений в этом выражении требуют в разы больше вычислительных ресурсов, чем работа с булевым множеством и побитовыми операциями. Таким образом, для ускорения вычислений можно составить критерий, эквивалентный данному, но использующий операции булевой алгебры.

Пусть k – количество 0 или 1 в последовательности, образованной использованием булевой операции XOR к некоторой последовательности, и ее логическом сдвиге на j позиций. Тогда последовательность удовлетворяет условию (4), если выполняется критерий:

$$|N - j - 2k| \leq \sqrt{N}, \quad j = 1, 2, \dots, N-1. \quad (5)$$

При переборе необязательно рассматривать все комбинации. Для данных последовательностей справедливы вариации реверса и инверсии, таким образом, количество вариантов перебора можно сократить.

Результаты исследований

По описанному подходу было найдено 1688 последовательностей длины $N = 32$, удовлетворяющих условию (3). В данных последовательностях значения УБЛ автокорреляционной функции варьируются в пределах от -3 до 3. Из лучших результатов были выявлены две последовательности: $\{-1 -1 -1 -1 +1 +1 +1 +1 +1 -1 +1 +1 -1 +1 +1 -1 +1 -1 +1 +1 -1 -1 +1 +1 +1 -1 -1 +1\}$ и $\{-1 -1 +1 +1 -1 +1 +1 -1 -1 +1 -1 -1 -1 -1 -1 -1 -1 +1 +1 +1 -1 -1 -1 +1 -1 +1 -1 +1 +1 -1\}$. Графики АКФ для этих последовательностей приведены на рис. 1.

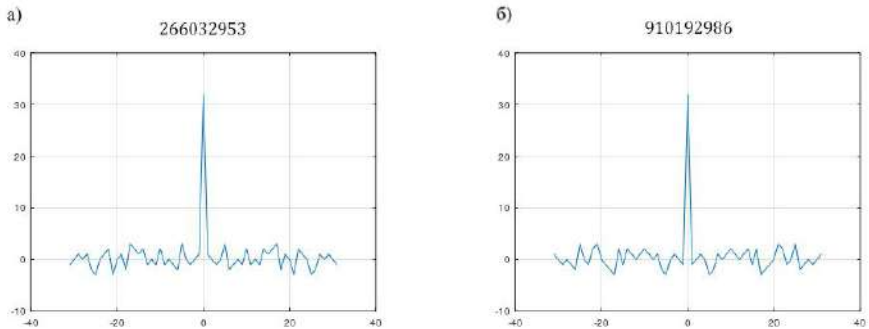


Рис. 1 – Автокорреляционная функция полученных последовательностей. а) для последовательности 1, б) для последовательности 2.

Сделав замену в данных последовательностях -1 на 0, можно интерпретировать их как двоичную запись порядкового номера

последовательности. Переводя в десятичный вид, получим соответственно номера последовательностей 266032953 и 910192986.

Время работы переборного алгоритма составило 10953.24 с, а время работы оптимизированного алгоритма - 222.19 с. Вычисления проводились на одном потоке. Таким образом, предлагаемый подход позволил повысить скорость вычислений в 50 раз.

Выводы и заключение

Разработанный подход поможет ускорить вычисления для нахождения последовательностей большей длины. Полученные последовательности могут найти широкое применение в различных коммуникационных и навигационных решениях.

Литература:

1. Слюсар В.И. Системы ММО: принципы построения и обработка сигналов. //Электроника: наука, технология, бизнес. – 2005. - № 8. – С. 53
2. Barker, R. H. "Group Synchronizing of Binary Digital Sequences." // In *Communication Theory*. London: Butterworth, pp. 273-287, 1953.
3. Чусов А.А., Ковылин А.А., Стаценко Л.Г. Миргородская Ю.В. Параллельный поиск сигналов с заданными взаимно и автокорреляционными свойствами на многопроцессорных платформах / Известия высших учебных заведений. Радиоэлектроника 2010. Т. 54, №8. С. 29-35
4. Radwa K. Hamad¹ , Ehab F. Badran¹ , Amira I. Zaki¹ , Said E. El-Khamy Comparison of Combined Multi-rate Barker Codes and Multi-rate Barker Codes For WOFDM Systems // SSRG International Journal of Electronics and Communication Engineering (SSRG-IJECE) 1.10 (2014): 7-14.
5. Ivan Tsmots, Oleg Riznyk, Vasyly Rabyk Synthesis Method and FPGA Realization of Quasi-barker Codes // International Scientific and Practical Conference "Electronics and Information Technologies" (ELIT-2018) : A-41 – A-44.

ИСПОЛЬЗОВАНИЕ БАЗ ДАННЫХ ДЛЯ УПРОЩЕНИЯ ДЕЯТЕЛЬНОСТИ ЭНЕРГОСБЫТОВОЙ КОМПАНИИ

А.А. Кулебякина¹, А.Е. Простакова², Е.В. Перезовова³
Российский государственный университет правосудия,
г. Нижний Новгород, ¹nastyu.kulebyackina5@gmail.com,
²katya.prostakova.05@inbox.ru, ³elena150886@ya.ru
Научный руководитель: Перезовова Е.В., ст.преподаватель

В данной статье используется система управления базами данных MS ACCESS для хранения, обработки и поиска данных энергосбытовой компании, также представлены отчёты по данным, используемые в работе юридического, расчётного и финансового отделов (бухгалтерские отчёты) энергосбытовой компании и проведёт статистический анализ данных.

This article uses the MS ACCESS database management system for storing, processing and searching data of an energy sales company, also presents data reports used in the work of the legal, accounting and financial departments (accounting reports) of an energy sales company and will conduct statistical data analysis.

База данных MS Access — это инструмент для структурирования, обработки и поиска большого объёма данных. В базах данных могут храниться сведения о товарах, заказах, различных типах документов и других данных. Это очень удобно если организация работает по нескольким предприятиям и ведёт учёт данных, а также расчёт по данным.

В данной статье используется система управления базами данных MS ACCESS для хранения, обработки и поиска данных энергосбытовой компании. А также представлены отчёты по данным, используемые в работе юридического, расчётного и финансового отделов (бухгалтерские отчёты) энергосбытовой компании.

Опишем деятельность энергосбытовой компании. Существуют ООО, АО, ЗАО (потребитель), которые использует электроэнергию для своего производства. Эту электроэнергию потребитель покупает у мелкой энергосбытовой компании. Сфера деятельности энергосбытовой компании-это договорная деятельность (составление договоров с потребителями), юридическая деятельность, расчётная деятельность и финансовая деятельность (бухгалтерия выставляет счёт по электроэнергии потребителю). Энергосбытовая компания рассчитывает сколько электроэнергии потратили потребители, соответственно реализовал энергосбыт и по какому напряжению ежемесячно. Электроэнергию учитывают приборы учета (ПУ). ПУ запитываются от ТП (точка поставки) эта точка поставки обозначается как правило

номером (ТП-323 «Щербинки»). Потребитель работает на различных мощностях и напряжениях. Существует несколько уровней напряжения ВН, СН1, СН2 и НН (высокое напряжение, среднее напряжение, низкое напряжение). Эти напряжения делятся по мощности, менее 670кВт, от 670 кВт до 10 МВт, не менее 10 МВт. По соответствующим напряжениям и ценовым категориям выставляется тариф и рассчитывается электроэнергия, потраченная потребителем за месяц. Все данные заносятся в базу данных.

В данной работе мы рассмотрели 5 договоров (на 5 потребителей).

Для обработки данных созданы 4 таблицы: Договор энергоснабжения, Технические характеристики, Данные для бухгалтерии и Контактные данные, которые связаны между собой.

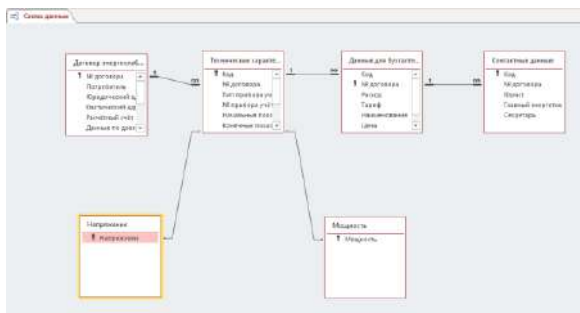


Рис. 1 – Все таблицы в «Схеме данных»

Также были созданы 3 запроса на выборку: Потребитель, Юридический отдел и Бухгалтерский отдел для получения данных.

Для выбора организации создали запрос с параметром. Для этого в запросной форме, в условиях отбора написали следующее: Like [Введите организацию] & "*". Во втором – данные о номере договора, потребителе, юридическом адресе, подписантах, сроках оплаты, дате начала и дате окончания договора. В третьем – о номере договора, наименовании, расходе, тарифе, цене и цене с НДС.

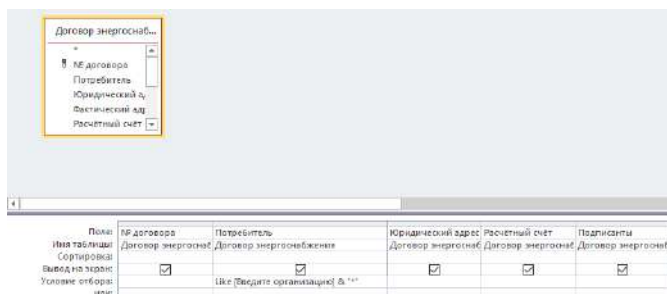


Рис. 2 – Запрос Потребитель в режиме «Конструктор»

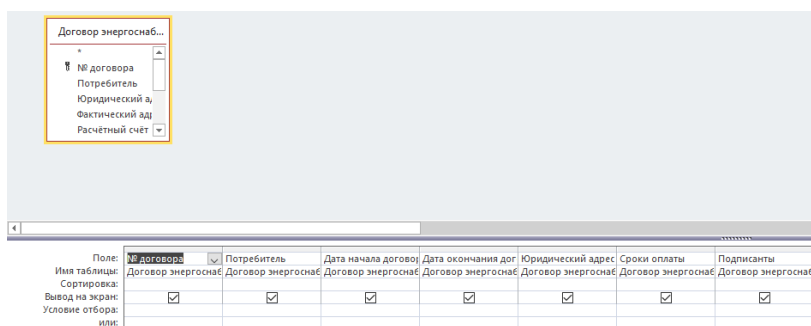


Рис. 3 – Запрос Юридический отдел в режиме «Конструктор»

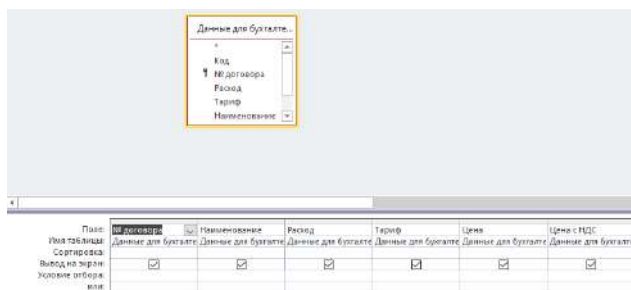


Рис. 4 – Запрос Бухгалтерский отдел в режиме «Конструктор»

Созданы отчёты по 2 запросам: Юридический отдел и Бухгалтерский отдел.

Юридический отдел	
№ договора	80482989
Потребитель	ООО "Мультисервис Плюс"
Дата начала договора	-
Дата окончания договора	31.12.2016
Юридический адрес	141495, обл. Московская, г. Углы

Рис. 5 – Отчёт Юридический отдел
(представлена верхняя часть отчёта)

Бухгалтерский отдел	
№ договора	90070005
Наименование	электроэнергия (НН)
Расход	1302
Тариф	1650 (3 за 1 кВт/ч)

Рис. 6 – Отчёт Бухгалтерский отдел
(представлена верхняя часть отчёта)

Для анализа работы энергосбытовой компании был создан отчёт расхода за год для определенного предприятия.

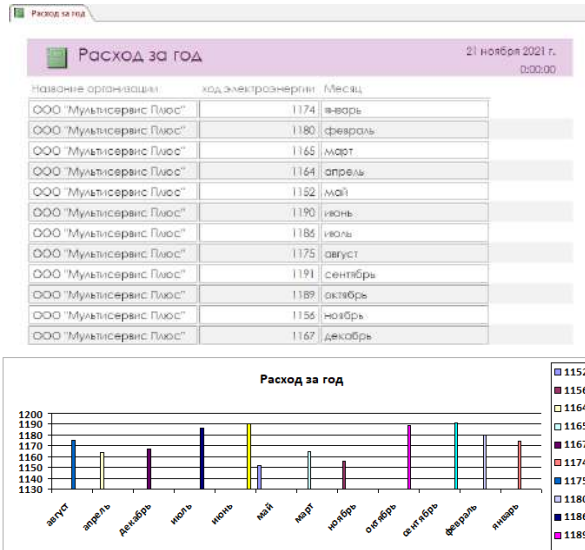


Рис. 7 – Таблица по отчёту Расход за год

Это позволит всем пользователям наглядно проследить за годовым расходом одной из организаций, а именно ООО "Мультисервис Плюс".

Таким образом, системы управления базами данных имеют широкое практическое применение и могут использоваться в организациях, занимающихся различной деятельностью.

Литература:

1. Тарарасов В.Л. РАБОТА С БАЗАМИ ДАННЫХ В ACCESS 2010. ЧАСТЬ 1: Учебно методическое пособие. – [электронный ресурс]. – Нижний Новгород: Нижегородский госуниверситет, 2014. – 126 с.
2. Договор электроснабжения с потребителем, приравненным к тарифной группе «население» № 90012932. Интернет-ссылка: <http://remiksuk.ru/upload/medialibrary/e5e/e5e8e3c275704f1e04ec6bddee6ce5ca.pdf>.

ВЫБОР ВЕСОВОЙ МАТРИЦЫ ПРИ ПОСТРОЕНИИ ПРОГНОЗИРУЮЩЕЙ МОДЕЛИ КООРДИНАТЫ ПОЛЮСА ЗЕМЛИ МЕТОДОМ НАИМЕНЬШИХ КВАДРАТОВ

П.С. Ларионов

Новосибирский государственный технический университет,
г. Новосибирск, pslarionov@mail.ru

Научный руководитель: Чубич В.М., д.т.н., профессор

В работе рассматриваются четыре варианта выбора весовой матрицы, использующиеся во взвешенном методе наименьших квадратов при построении прогнозирующей модели одного параметра вращения Земли. На основе сравнения абсолютных погрешностей предсказания по уровню вероятности 0.95 определяется наиболее эффективный вариант выбора весовой матрицы.

The paper considers four weight matrix options chooses that are used in the weighted least squares method when constructing a predictive model of one Earth rotation parameter. Based on absolute prediction errors at a probability level of 0.95 comparison, the most effective option for choosing a weight matrix is determined.

Повышение точности прогнозирования параметров вращения Земли (к ним относятся мгновенные координаты полюса Земли и расхождение шкал всемирного и всемирного координированного времени) способствует более качественному функционированию глобальных навигационных спутниковых систем.

В настоящее время задачей качественного прогнозирования параметров вращения Земли занимаются в различных странах (Австрия, Германия, Китай, Польша, Россия, Турция, Сербия, США, Франция и др.). При этом используются спектральный сингулярный анализ, вейвлеты, нейронные сети, анализ временных рядов, фильтр Калмана и другие подходы [1].

В данной работе для построения прогнозирующей модели координаты полюса Земли X_p используется взвешенный метод наименьших квадратов, даются рекомендации по возможному выбору весовой матрицы.

Данные о параметрах вращения Земли предоставляет официальный мировой поставщик – Международная служба вращения Земли. Эти данные в виде временных рядов $Y^T = \{y(t_1), y(t_2), \dots, y(t_N)\}$ по каждому параметру можно получить с сайта Международной службы вращения Земли <https://www.iers.org>.

В соответствии с [2-4] нам необходимо оценить параметры $\theta = [\theta_1, \dots, \theta_8]^T$ следующей статической модели

$$y(t_k) = \theta_1 + \theta_2 t_k + \theta_3 \cos \frac{2\pi t_k}{T_1} + \theta_4 \sin \frac{2\pi t_k}{T_1} + \theta_5 \cos \frac{2\pi t_k}{T_2} + \theta_6 \sin \frac{2\pi t_k}{T_2} + \theta_7 \cos \frac{2\pi t_k}{T_3} + \theta_8 \sin \frac{2\pi t_k}{T_3} + \xi(t_k), \quad k = 1, 2, \dots, N. \quad (1)$$

Следуя [5], примем $T_1=182.62$, $T_2=365.24$, $T_3=435.00$.

Задача оценивания неизвестных параметров математической модели (1) является задачей линейного регрессионного анализа, для которой взвешенный метод наименьших квадратов дает следующий результат (см., например, [2,3,6])

$$\hat{\theta} = (F^T W F)^{-1} F^T W Y. \quad (2)$$

В [2] предлагается **четыре варианта выбора весовой матрицы** при построении прогнозирующих моделей различных параметров вращения Земли. Приведем краткое описание этих вариантов.

В первом случае

$$W = \begin{pmatrix} W_1 & O & O \\ O & W_2 & O \\ O & O & W_3 \end{pmatrix}^\alpha,$$

$$W_1 = \text{diag} \left(\underbrace{\frac{1}{3}, \frac{1}{3}, \dots, \frac{1}{3}}_{\frac{N}{3}} \right), \quad W_2 = \text{diag} \left(\underbrace{\frac{2}{3}, \frac{2}{3}, \dots, \frac{2}{3}}_{\frac{N}{3}} \right), \quad W_3 = \text{diag} \left(\underbrace{1, 1, \dots, 1}_{\frac{N}{3}} \right).$$

Во втором случае

$$W = \begin{pmatrix} W_1 & O & \dots & O \\ O & W_2 & \dots & O \\ \dots & \dots & \dots & \dots \\ O & O & \dots & W_{10} \end{pmatrix}^\alpha,$$

$$W_1 = \text{diag} \left(\underbrace{\left(\frac{1}{10}, \frac{1}{10}, \dots, \frac{1}{10} \right)}_{\frac{N}{10}}, W_2 = \text{diag} \left(\underbrace{\left(\frac{1}{9}, \frac{1}{9}, \dots, \frac{1}{9} \right)}_{\frac{N}{10}}, \dots, W_{10} = \text{diag} \left(\underbrace{(1, 1, \dots, 1)}_{\frac{N}{10}} \right) \right).$$

В третьем случае

$$W = \text{diag} \left(\underbrace{\left(\left(\frac{1}{t_1} \right)^\alpha, \left(\frac{1}{t_2} \right)^\alpha, \dots, \left(\frac{1}{t_N} \right)^\alpha \right)}_N \right).$$

В последнем случае

$$W = \text{diag} \left(\underbrace{\left(\left(\frac{T-t_1}{T} \right)^\alpha, \left(\frac{T-t_2}{T} \right)^\alpha, \dots, \left(\frac{T-t_N}{T} \right)^\alpha \right)}_N \right), \quad T = t_N - t_1.$$

Отметим, что во всех приведенных вариантах выбора весовой матрицы присутствует параметр α , оптимальное значение которого находится численно.

Найденные при помощи соотношения (2) оценки параметров используются в дальнейшем для прогнозирования параметров вращения Земли по соответствующим формулам.

Проведем сравнительный анализ влияния варианта выбора весовой матрицы на качество прогнозирования параметра вращения Земли X_p .

В [3] отмечается, что оптимальное качество оценивания параметров вращения Земли достигается при использовании измерительных данных за 10 лет. В связи с этим при расчетах было выбрано $N = 3650$.

Как и в [2], будем прогнозировать параметр X_p в интервале с 1 января 2004 года по 15 января 2012 года на 10, 20, ..., 100 дней вперед. В отличие от указанной работы (в ней использовался критерий средней абсолютной ошибки), вычислим абсолютные погрешности (размерность которых – угловые миллисекунды) по уровню вероятности 0.95 (5% худших результатов отброшено). Результаты приведем в Таблице 1.

Из таблицы видно, что третий вариант выбора весовой матрицы с оптимальным параметром α дает более точные прогнозы.

В дальнейшем предполагается построение аналогичных таблиц для двух других параметров вращения Земли, сопряжение полученных статических моделей вида (1) с построенными на остатках авторегрессионными моделями и разработка новых методов на основе нейронных сетей и фильтра Калмана.

Таблица 1 – Абсолютные погрешности предсказания для различных вариантов выбора весовой матрицы

День предсказания	Вариант №1, ($\alpha = 1$)	Вариант №2, ($\alpha = 1$)	Вариант №3, ($\alpha = 1.5$)	Вариант №4, ($\alpha = 12$)
10	48,55	16,79	8,30	25,88
20	52,00	23,57	15,07	32,63
30	54,99	29,53	21,32	39,10
40	57,66	34,89	27,19	44,43
50	59,88	40,46	34,18	49,83
60	61,39	47,26	39,90	55,32
70	62,70	54,76	45,35	60,18
80	63,83	60,00	48,97	63,58
90	64,75	64,27	54,19	66,75
100	65,16	66,63	57,60	68,36

Литература:

1. Achievements of the Earth orientation parameters prediction comparison campaign / Kalarus M., Schuh H., Kosek W., Akyilmaz O., Bizouard Ch., Gambis D., Gross R., Jovanovic B., Kumakshev S., Kutterer H., Mendes Cerveira P. J., Pasynok S. L., Zotov L. // *Journal of Geodesy*. – 2010. – № 84. – P. 587–596.

2. Zhangzhen S., Tianhe X. Prediction of earth rotation parameters based on improved weighted least squares and autoregressive model // *Geodesy and Geodynamics*. – 2012. – N 3. – P. 57–64.

3. The application of a combination of weighted least-squares and autoregressive methods in predictions of polar motion parameters / Wu F., Deng K., Chang G., Wang Q. // *Acta Geodaetica et Geophysica*. – 2018. – Vol. 53, iss. 2. – P. 247-257.

4. Современные и перспективные информационные ГНСС-технологии в задачах высокоточной навигации / В. А. Бартнев, А.К. Гречкосеев, Д.А. Козорез, М.Н. Красильщиков, В.В. Пасынков, Г.Г. Себряков, К. И. Сыпало. – М.: Физматлит, 2014. – 192 с.

5. Xu X.Q. Zhou Y.H, Liao X.H. Short-term earth orientation parameters predictions by combination of the least-squares, AR model and Kalman filter // *Journal of Geodynamics*. – 2012. – Vol. 62. – P. 83–86.

6. Ермаков С.М., Жиглявский А.А. Математическая теория оптимального эксперимента. – М.: Наука, 1987. – 320 с.

КЛАССИФИКАЦИЯ ЗАШУМЛЕННЫХ ДАННЫХ ПРИ РАЗЛИЧНЫХ ОБЪЕМАХ ВЫБОРКИ

М.А. Сивак

**Новосибирский государственный технический университет,
г. Новосибирск, perelyaeva@ami.nstu.ru**

Научный руководитель: Тимофеев В.С., д.т.н., доцент

Рассмотрено решение задачи классификации зашумленных данных с использованием искусственных нейронных сетей. Построена модель классической нейронной сети (с квадратичной функцией потерь), а также пять робастных моделей: с использованием функций потерь Хьюбера, Рамсея, Уэлша, Эндрюса и "Fair". Все полученные модели включили в себя один скрытый слой. При проведении исследований рассматривались синтетические наборы данных различного объема. Проведен сравнительный анализ точности классификации построенных моделей, который показал, что использование робастного подхода дает значительный выигрыш в точности.

This study highlights solving the classification problem for noisy data. The ordinary neural network model (with the quadratic loss) has been obtained. Also five robust neural networks applying the Huber, Ramsey, Welsch, Andrews and "Fair" losses have been built. All the obtained models included one hidden layer. For the research, the simulated datasets of different size were considered. The comparative analysis of the results has shown that using the robust approach allows achieving much higher classification accuracy.

В настоящее время искусственные нейронные сети (ИНС) являются одним из наиболее широко известных инструментов машинного обучения. Они находят свое применение при решении множества задач, в том числе задачи распознавания образов или, по-другому, задачи классификации объектов. Однако на практике часто приходится сталкиваться с тем, что в анализируемых данных содержится некоторое количество нетипичных наблюдений (выбросов). Обычно это приводит к снижению точности построенной модели. В связи с этим авторами в [1] было предложено использовать робастный подход к построению искусственных нейронных сетей, позволяющий снизить негативное влияние выбросов.

ИНС представляет собой структуру, которая состоит из простых элементов, называемых нейронами. Нейроны связаны друг с другом посредством однонаправленных каналов (ребер) [2]. В данной работе рассматривались нейронные сети с одним скрытым слоем. Было построено пять робастных моделей с использованием различных функций потерь [3]: Хьюбера, Рамсея, Уэлша, Эндрюса и "Fair". Кроме

того, была построена классическая нейронная сеть (с квадратичной функцией потерь).

Целью данного исследования было изучить, как меняется точность работы построенных нейронных сетей при различном объеме анализируемых данных.

При проведении исследований использовались синтетические наборы данных, сгенерированные с использованием специальных инструментов языка Python 3.6 [4]. Все анализируемые наборы данных X имели одинаковую конфигурацию: объекты X_m , $m = 1, \dots, |X|$ описывались четырьмя признаками x_{mi} , $i = 1, \dots, 4$ и делились на три одинаковых по размеру класса y_k , $k = 1, \dots, 3$. Моделирование шума производилось аналогично тому, как показано в [1]. Оптимальные значения параметра β робастных нейронных сетей также выбирались в соответствии с рекомендациями, данными в [1]. Рассматривались наборы данных объемом 150, 450, 600 и 900 объектов при доле выбросов $\lambda = 0,25$.

Все наборы данных делились на обучающую и тестовую выборки, причем в обучающую выборку $L = \{X_1, \dots, X_{|L|}\}$ входило 80% объектов, а в тестовую выборку $D = \{X_{|L|+1}, \dots, X_{|X|}\}$ – 20% объектов. Поскольку в рамках данной работы классы объектов были сбалансированы, для оценки точности классификации использовалась следующая метрика:

$$\alpha = \frac{D_{corr}}{|D|} 100\% ,$$

где $|D|$ – объем тестовой выборки, D_{corr} – количество объектов, отнесенных к правильному классу.

Точность работы нейронных сетей фиксировалась после 100, 300 и 500 эпох обучения. Полученные значения представлены в Таблице 1. Здесь используются следующие обозначения столбцов: $|X|$ – объем набора данных, h – число эпох, $f1$ – функция потерь Эндрюса, $f2$ – функция потерь Рамсея, $f3$ – функция потерь “Fair”, $f4$ – функция потерь Уэлша, $f5$ – функция потерь Хьюбера, $f6$ – квадратичная функция потерь.

Полученные результаты показывают, что при увеличении числа объектов в наборе данных до 600 и до 900 точность классификации падает. При этом увеличение объема набора данных со 150 объектов до 450, как правило, не сказывается на точности классификации негативным образом. Кроме того, можно отметить слишком высокие значения метрики α для функций потерь Эндрюса и Уэлша при 150 объектах в наборе. Такие значения говорят о переобучении нейронных

сетей, в которых использовались указанные робастные функции потерь, то есть в данном случае для их обучения требуется менее 100 эпох. По сравнению с классической ИНС наибольший выигрыш в точности дает функция потерь Уэлша, наименьший – функция потерь Рамсея.

Таблица 1 – Точность классификации

X/	h	f1	f2	f3	f4	f5	f6
		$\beta = 4,7$	$\beta = 1,2$	$\beta = 0,4$	$\beta = 2,5$	$\beta = 0,2$	-
150	100	99,6	87,3	91,4	99,6	83,0	74,2
	300	98,0	89,5	91,4	99,9	90,9	85,9
	500	99,0	93,7	95,2	99,7	95,6	91,0
450	100	67,2	66,8	66,8	70,9	67,0	67,0
	300	96,2	90,8	92,2	97,1	92,3	87,6
	500	98,8	96,1	96,7	99,1	97,4	94,1
600	100	66,9	67,6	67,5	67,0	67,5	67,5
	300	95,3	91,8	93,0	95,9	93,9	87,3
	500	97,4	95,8	96,3	98,0	96,5	93,0
900	100	67,0	68,4	68,0	66,9	68,5	68,3
	300	93,5	89,1	90,2	95,4	91,3	88,0
	500	96,4	92,8	93,7	97,4	94,6	91,2

Проведенные исследования подтверждают целесообразность использования робастного подхода к построению нейронных сетей при работе с зашумленными данными, особенно в тех случаях, когда набор данных включает в себя достаточно большое число объектов.

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 20-37-90077.

Литература:

1. Сивак М.А., Тимофеев В.С. Построение робастных нейронных сетей с различными функциями потерь // Системы анализа и обработки данных. – 2021. – № 2 (82). – С. 67–83.
2. Bishop C. Neural Networks for Pattern Recognition. Oxford University Press, New York, US, 1995.
3. Сивак М.А. Исследование применимости робастных функций потерь в нейронных сетях // Сборник научных трудов НГТУ. – 2020. – № 4 (99). – С. 50–58. – DOI: 10.17212/2307-6879-2020-4-50-58.
4. How to Generate Test Datasets in Python with scikit-learn. URL: <https://machinelearningmastery.com/generate-test-datasets-python-scikit-learn/> (дата последнего обращения: 19.11.2021).

КОМПЛЕКСНАЯ ОЦЕНКА ЭФФЕКТИВНОСТИ ОНЛАЙН-МЕТОДОВ ПРИВЛЕЧЕНИЯ АБИТУРИЕНТОВ В ПЕРИОД ПАНДЕМИИ НА ПРИМЕРЕ СИБИРСКОГО ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА НАУКИ И ТЕХНОЛОГИИ ИМ. М.Ф. РЕШЕТНЕВА

Е.А. Бойцова, А.О. Соколова
Сибирский государственный университет науки и технологии
им. М.Ф. Решетнева, г. Красноярск
Научный руководитель: Федорова Н.В., доцент

Статья посвящена проблеме разработки эффективной стратегии привлечения абитуриентов в российские вузы в условиях пандемии и падении привлекательности офлайн мероприятий. Описана практика привлечения абитуриентов онлайн-методами в Сибирском государственном университете им. М.Ф. Решетнева. Сформулированы результаты применения новых подходов к организации работы, направленной на привлечение абитуриентов, с учетом особенностей целевых групп и использования имеющихся ресурсов.

The article is devoted to the problem of developing an effective strategy for attracting applicants to Russian universities in a pandemic and reducing the attractiveness of offline events. The practice of attracting applicants using online methods at the Siberian State University named after M.F. Reshetnev. The results of applying new approaches to organizing work to attract applicants, taking into account the characteristics of target groups and the use of available resources, have been formulated.

Последний год изменил привычный уклад во всех сферах общества. Пандемия коронавируса, локдаун и стремление минимизировать личные контакты привели к модернизации организации мероприятий. Не обошли изменения и университеты во всем мире. Несмотря на перевод занятий в дистанционный режим, процесс обучения продолжается. Также не прекратился набор абитуриентов на все специальности.

Методики личных встреч, такие как дни открытых дверей и ярмарки вакансий, не утратили своей актуальности, однако способы привлечения абитуриентов через Интернет и электронные каналы стремительно становятся самыми эффективными. Методы оффлайн-привлечения будущих студентов помогают не только эффективно справиться с этой задачей, но и помогают избежать массовых скоплений людей, что явно пойдет на пользу в борьбе с пандемией. Также, когда речь заходит о привлечении иностранных студентов, эти способы можно смело назвать необходимыми или даже незаменимыми.

Чтобы старшеклассник или будущий студент заинтересовались определенным местом обучения, необходимо предоставлять ему

актуальную информацию о событиях, происходящих в институте в целом, на кафедре и внутри каждой группы. Необходим свободный доступ ко всей информации помогающей абитуриенту, помимо доступности информация должна быть систематизирована и красиво оформлена. Благодаря глобальной сети Интернет выполнять эти условия становится в разы проще. Университеты могут использовать различные платформы и их возможности:

- веб-сайт университета;
- электронная почта;
- социальные сети («ВКонтакте» «Instagram»);
- медиа-сайты («Youtube»);
- блог-платформы («Twitter»);
- поисковые системы («Яндекс», «Google») [1].

Как показывает практика, интернет-способы привлечения абитуриентов охватывают больший круг людей, что является несомненным плюсом. В век развитых технологических возможностей, стремлении к экономии собственных ресурсов, в том числе времени, малое количество абитуриентов предпочтёт офлайн встречу онлайн-трансляции из университета. Доказательством является день открытых дверей в СибГУ в 2019 году. На собрание в инженерно-экономический институту пришло не более 10 человек, в то время как просмотры на онлайн-трансляциях с аналогичного мероприятия в 2020 году превышали 1500. Также стоит отметить, что затраты человеческих ресурсов в 15 раз больше при проведение традиционных мероприятий (согласно данным исследования результативности профориентационных мероприятий в СибГУ).

Новое поколение школьников переходит в тотальную цифровизацию, поэтому им намного проще воспринимать информацию через каналы глобальной сети Интернет [2]. Для это были созданы страницы в популярных сетях, как для университета, так и для его кафедр, там выкладывается вся актуальная информация по учебе, поступлению и важных событиях. Формат информации самый разный, но наиболее популярным является онлайн-трансляция, на которой абитуриенты, их родители и работодатели могут пообщаться с представителями университета. Так, в 2020 году онлайн-трансляция с директором Инженерно-экономического института набрала порядка 5000 просмотров, среди которых большинство было от школьников и абитуриентов. Помимо проведения онлайн-трансляций университет Решетнева активно ведёт страницы в социальных сетях, за которыми следят тысячи студентов, абитуриентов и школьников. Развитые и современно оформленные страницы привлекают целевую аудиторию

университета и делают её отношение к учебному заведению более лояльным. Например, информация по выездной приемной комиссии была просмотрена 395 тысяч раз, что несомненно превышает охват людей, получивших данную информацию с помощью оффлайн-методов [3].

Онлайн-мероприятия и ведение социальных сетей помогло в 2021 году первого сентября увидеть в стенах университета 3000 новых студентов, когда в 2019 году эта цифра была равно 1500. Эти данные показывают эффективность и важность онлайн-мероприятий по привлечению абитуриентов.

Литература:

1. Бердышев, С.Н. Секреты эффективной интернет-рекламы/ С.Н.Бердышев. – М.: Дашков и Ко, 2010 – 120 с.
2. Seopult – автоматизированная рекламная система. – Режим доступа: <http://seopult.ru/>
3. Отчет о результатах самообследования федерального государственного бюджетного образовательного учреждения высшего образования «Сибирский государственный университет науки и технологий имени академика М.Ф. Решетнева» за 2020 год г. Красноярск – Режим доступа: http://sibsau.ru/sveden/document/Otchet_po_samoobsledovaniyu_za_2020_pdf

РЕАЛИЗАЦИЯ МЕТОДОВ МОЗАИЧНОГО МОДЕЛИРОВАНИЯ СЛОЖНЫХ ГЕНЕТИЧЕСКИХ СЕТЕЙ

Б.С. Соловьев

Новосибирский государственный технический университет,

г. Новосибирск, b.solovev.2013@stud.nstu.ru

Научный руководитель: Карманов В.С., к.т.н., доцент

Разобран и реализован способ моделирования взаимодействий в геновых сетях (концентрация белка) на основе функции Хилла. Применены методы оценки параметров. Разработан прототип программного обеспечения.

A method for modeling interactions in gene networks (protein concentration) based on the Hill function has been analyzed and implemented. Methods for estimating parameters are applied. Software prototype developed.

Введение

Моделирование генных сетей является ключевым подходом в современных биомедицинских исследованиях, используется в системной биологии для изучения функционирования сложных биологических систем. Большинство существующих методов математического моделирования полезны для анализа хорошо изученных биологических процессов, для которых доступна информация о скоростях протекающих в них химических реакций. Однако сложные биологические процессы, такие как те, которые определяют фенотипические признаки организмов или процессы патологических заболеваний, включают в себя сложные перекрестные связи. Кроме того, внутренние детали взаимодействия часто отсутствуют. Разработан подход, который назван мозаичным сетевым моделированием, позволяющий комбинировать независимые математические модели генных регуляторных сетей и, тем самым, описывать сложные биологические системы. Преимущество этого подхода заключается в том, что он позволяет создавать интегрированную модель, несмотря на то, что информация о молекулярных взаимодействиях между частями модели (так называемые фрагменты мозаики) может отсутствовать. Моделирование генной сети мозаики может быть достигнуто путем непосредственной интеграции моделей, которые описывают динамику отдельных так называемых «кусочков мозаики».

Постановка задачи

Задача заключается в построении функционала управления первой генной сети второй сетью на основе анализа данных по возмущению элементов первой сети и отклику во второй сети.

Задаются две генные подсети, где первая генная подсеть влияет на состояние второй. Имеется некая зависимость между двумя подсетями, природа которой неизвестна. Необходимо смоделировать функционал зависимости второй подсети от первой для прогнозирования возможных последствий от воздействия на первую подсеть. Элементы внутри каждой из подсетей связаны друг с другом, образуя граф. Обе подсети включены в неизвестную глобальную сеть, являющуюся моделью генной сети какого-либо организма.

Исходная математическая модель

Математическая модель двух генных подсетей описана переменными и формулами взаимодействий между ними. За основу взяты материалы статей [1, 2].

Обобщённая интегрированная модель сети задается в виде системы уравнений

$$\left\{ \begin{array}{l} \frac{dY_i(t)}{dt} = P_i(Y_1, Y_2, \dots, Y_N) - D_i(Y_1, Y_2, \dots, Y_N), \\ \frac{dU_j(t)}{dt} = F_j(U_1, U_2, \dots, U_M) - \Phi_j(K_1, K_2, \dots, K_N, Y_1, Y_2, \dots, Y_N), \end{array} \right. \quad (1)$$

где первая сеть определяется набором функций, характеризующих концентрацию вещества в каждой вершине сети $[Y_1(t), Y_2(t), \dots, Y_N(t)]$, вторая сеть определяется набором $[U_1(t), U_2(t), \dots, U_M(t)]$. Функционал

$$\Phi(K_1, K_2, \dots, K_N, Y_1, Y_2, \dots, Y_N) = \sum_{j=1}^N K_j \cdot Y_j \quad \text{задает управление первой сети}$$

второй сетью.

Задача моделирования заключается в поиске значений K_1, K_2, \dots, K_N , при которых ошибка наблюдения будет минимальной. Поиск этих значений осуществляется на основе анализа данных по возмущению элементов первой модели и отклику элементов второй модели.

Приведены экспериментальные данные по ингибированию R элементов первой системы. Отклик второй системы представлен в виде набора измеренных концентраций ее элементов $\{\tilde{U}_j(t)\}^R, (j=0, 1, \dots, L)$, которые могут быть представлены в виде непрерывной функции или в виде дискретного набора концентраций. Для поиска значений параметров модели решалась задача минимизации отклонения расчётных данных от экспериментальных, представленного в виде суммы:

$$S = \sum_{j=1}^M \sum_{\tau=1}^T |U_j(\tau) - \tilde{U}_j(\tau)|, \quad (2)$$

где T – количество экспериментальных точек динамики \tilde{U}_j , $\tilde{U}_j(\tau) = \tilde{U}_j(\tau)_{t=\tau}$ значение рассчитываемой по модели динамики уровня экспрессии гена j в момент времени τ .

Задача поиска параметров функционала, при которых отклонение расчётных данных от экспериментальных было бы минимальным, решалась генетическим алгоритмом.

Генерация данных

Чтобы моделировать мозаичные генные регуляторные сети, правые части уравнений задавались зависимостью в виде функций Хилла

$$F_i(U_1, U_2, \dots, U_M) = \mu_p \frac{1 + \sum_{k \in U_i} \alpha_{i,k} Y^{h_{i,k}}(t)}{1 + \sum_{k \in D_i} \delta_{i,k} Y^{g_{i,k}}(t)} - \beta_i Y_i(t), \quad (3)$$

где U_i – множество элементов, положительно регулирующих активность синтеза элемента i ; D_i – множество элементов, отрицательно регулирующих активность синтеза элемента i ; $\alpha_{i,k}$ и $\delta_{i,k}$ – константы регуляции; $h_{i,k}$ и $\gamma_{i,k}$ – параметры нелинейности; β_i – константа деградации.

Чтобы смоделировать ингибирование элемента Y_p , вводится параметр ингибирования μ_p .

При моделировании параметры задавались случайным образом в диапазонах: $0 < \mu_p < 1$; $0 < \alpha_{i,k}, \delta_{i,k}, \beta_i < 2$; $h_{i,k}, \gamma_{i,k} \in \{0, 1, 2\}$. Начальные значения принимались равными 1 для всех элементов.

Полученные результаты

Моделирование производилась согласно предоставленным шаблонам графов подсетей и их комбинациям (рис. 1). Подобный подход значительно уменьшил количество используемых графов, которые ранее генерировались случайно.



Рис.1 – Примеры шаблонов генных подсетей

В ходе работы был разработан программный комплекс для работы с моделированием генных сетей в виде графов. В общий функционал входят:

- 1) генерация случайного графа;
- 2) загрузка готового, ранее подготовленного, графа;
- 3) оценка параметров линейной зависимости между подсетями.

Граф, задающий топологию сети, получается из множества изолированных вершин путём последовательного случайного добавления соединяющих вершины рёбер.

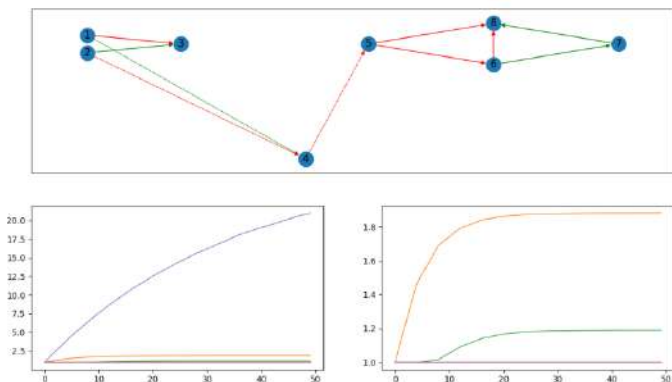


Рис.2 – Пример генерации и результатов моделирования

После генерации возможна визуализация графа сети. Далее происходит моделирование взаимодействий элементов подсетей согласно (1) и (3), после чего происходит оценка параметров модели и вывод результатов (рис. 2).

Заключение

В рамках проделанной работы была разобрана задача, которая непосредственно связана с синтезом оптимального управления стохастическими динамическими системами. Имеется ряд продвижений:

1. Реализован метод мозаичного моделирования, в дальнейшем планируется его модификация для синтеза оптимального управления.
2. Реализован метод для оценки качества работы методов структурной идентификации в задаче мозаичного моделирования.
3. Реализована возможность задания шаблонов подсетей.

Литература:

1. Popik OV, Petrovskiy ED, Mishchenko EL, Lavrik IN, Ivanisenko VA. Mosaic gene network modelling identified new regulatory mechanisms in HCV infection. *Virus Res.* 2016;218:71-78. doi:10.1016/j.virusres.2015.10.004
2. Petrovskaya OV, Petrovskiy ED, Lavrik IN, Ivanisenko VA. A study of structural properties of gene network graphs for mathematical modeling of integrated mosaic gene networks. *J Bioinform Comput Biol.* 2017;15(2):1650045. doi:10.1142/S0219720016500451

ПОДХОДЫ К ИЗМЕРЕНИЮ ЭФФЕКТИВНОСТИ ИНВЕСТИЦИОННОГО ПОРТФЕЛЯ

Тазутдинова Ю.Р., Федорова Н.В
Сибирский государственный университет науки и технологии
им. М.Ф. Решетнева,
г. Красноярск

В данной статье рассматриваются три подхода к оценке инвестиционного портфеля, которые учитывают помимо показателя прибыли, риск, являющийся неотъемлемой частью процесса инвестирования, и могут быть использованы для измерения эффективности как начинающими, так и опытными инвесторами.

This article discusses three approaches to evaluating an investment portfolio. These approaches take into account the risk that is an integral part of the investment process, in addition to the profit indicator. These techniques can be used to measure the effectiveness of novice and experienced investors.

Популярность инвестиций с течением времени только возрастает. Данная тенденция прослеживается особенно явно на фоне текущей экономической ситуации, когда способов эффективно распорядиться своими накоплениями становится все меньше. Но при этом достаточно большое количество начинающих инвесторов оценивают эффективность своего инвестиционного портфеля только по размеру полученной прибыли. Такой подход недостаточно полно оценивает результативность, ведь он не учитывает риск, которым сопровождается подобный способ получения дохода.

Существует множество более эффективных методик оценки результатов конкретного инвестиционного портфеля, например:

1. Коэффициент Йенсена – это формула расчёта, которая использует модели CAPM (Capital Asset Pricing Model) или оценки долгосрочных активов. Она основывается на том, что объем необходимой прибыльности на количество вложенных средств определяется уровнем риска по рынку в целом. Таким образом коэффициент показывает избыток прибыли, что приносит инвестиционный портфель, по сравнению с ожидаемой. Результаты интерпретируются следующим образом: чем больше значение коэффициента, тем выше прибыльность инвестиции, при учете риска. Расчетная формула выглядит следующим образом:

Jenshon's alpha = $PR - CAPM$, где PR – доходность портфеля, а $CAPM$ – сумма безрисковой процентной ставки и прибыльности рыночной безрисковой процентной ставки.

2. Коэффициент Трейнора – это составная формула, которая позволяет определить величину доходности по отношению к систематическому рыночному риску. Данный коэффициент представляет собой развитие портфельной теории Гарри Марковица. Расчет основывается на том, что предполагаемая будущая прибыль актива соотносима с его математическим ожиданием, а риск с дисперсией, оба показателя при этом равны, то есть имеет место волатильность цен, что считаются, основываясь на данных прошлых корректировок. Расчетная формула коэффициента имеет следующий вид:

$T = \frac{r_p - r_f}{\beta_p}$, где: r_p – доходность портфеля, r_f – безрисковая процентная ставка, β_p – коэффициент риска. Главное преимущество такого метода расчета в том, что он показывает, каким образом ценные бумаги изменяются внутри инвестиционного портфеля. С помощью данного подхода, инвестор способен определить долю значимости вклада каждой конкретной инвестиции в составе портфеля. Значение, полученное в результате использования коэффициента интерпретируется так: чем больше оно для конкретного актива, тем выгоднее он для процесса инвестирования собственных средств.

3. Коэффициент Шарпа – формула достаточно похожая на коэффициент Трейнора, основное их отличие заключается в том что в этой формуле риск – это стандартное отклонение инвестиционного портфеля, а не систематический рыночный коэффициент риска. Также при интерпретации результатов, с помощью коэффициента Шарпа оценивание происходит с учетом диверсификации. Эта формула подходит для оценки эффективности более диверсифицированных инвестиционных портфелей. Расчетная формула выглядит следующим образом:

Sharpe ratio = $\frac{PR - RFR}{SD}$, где PR – доходность портфеля, RFR – безрисковая процентная ставка, а SD – стандартное отклонение. При интерпретации результатов необходимо обратить внимание на то, что лучшим инвестиционным портфелем является не тот, что приносит наибольшее количество прибыли, а тот, что сочетает в себе прибыльность с приемлемым уровнем риска.

Таким образом, мы можем сделать выводы о том, что система оценивания инвестиционного портфеля только по показателю прибыли недостаточно полно раскрывает и показывает эффективность

инвестиционного портфеля, так как не учитывает один из важнейших показателей – риск.

Литература:

1. Инвестиции: учебник для вузов / под ред. Л.И. Юзвович, С.А. Дегтярева, Е.Г. Князевой. – Екатеринбург: Изд-во Урал. ун-та, 2016.– 359 с.

2. Habr – русскоязычный веб-сайт в формате системы тематических коллективных блогов с элементами новостного сайта. – Режим доступа: [<https://habr.com/ru/all/>]

3. Оценка доходности и риска финансовых инвестиций: учебник / под ред. Лисица М.И. СПб.: Издательство Университета при МПА ЕврАзЭС, 2018. – 23 с.

СИСТЕМА СБОРА ПОКАЗАТЕЛЕЙ ОБРАЗОВАТЕЛЬНЫХ ОРГАНИЗАЦИЙ

Л.А.Фадейкин

Новосибирский государственный технический университет,

г. Новосибирск, arjen.robben2012@yandex.ru

Научный руководитель: Аврунев О.Е.

Разработана система сбора данных характеризующих динамику развития образовательных организаций России. Накапливаются данные вузов и колледжей. Разработана база данных для построения аналитических запросов.

An information system has been developed that characterizes the dynamics of development of educational organizations in Russia. Data from universities and colleges are being accumulated. A database for constructing analytical queries has been developed.

Поставлена задача автоматического сбора данных интегральных ежегодных показателей деятельности образовательных организаций высшего образования (ВУЗов) и среднего профессионального образования (колледжей). Собранные и приведенные к общей согласованной структуре данные должны сохраняться в базе данных.

Программная система на основе собранных данных, должна отвечать на следующие вопросы:

1. Как менялось соотношение учреждений ВО и СПО по годам – кол-во / контингент?

2. Как менялось соотношение учреждений ВО государственной и негосударственной формы - кол-во / контингент;
3. Как менялось соотношение контингента очного, очно-заочного и заочного ВО по годам;
4. Как менялось соотношение обучающихся по программам бакалавриата и магистратуры;
5. Как соотносится демографическая ситуация с выделением бюджетных мест –Соотношение кол-ва мест и кол-ва 18-летних в этом году в регионе;
6. Как соотносится демографическая ситуация с выделением бюджетных мест по отношению к контрактным местам;
7. Как соотносится проходной балл ЕГЭ (по рейтингу) - востребованности направления подготовки у абитуриентов к увеличению бюджетных мест;
8. Как менялись приоритеты государства в подготовке студентов за бюджетные средства по разным направлениям;
9. Как менялись предпочтения студентов, выбирающих высшее образование на контрактной основе;
10. Как менялось соотношение количество студентов иностранцев, обучающихся в российских вузов;
11. Как менялась география студентов иностранцев, обучающихся в российских вузов.

Исходные данные загружались из следующих источников:

- Информационно-аналитические материалы по результатам проведения мониторинга эффективности деятельности образовательных организаций высшего образования - <https://monitoring.miccedu.ru/>
- Сайт Министерства просвещения РФ - <https://edu.gov.ru>
- Сайт Министерства науки и высшего образования РФ - <https://minobrnauki.gov.ru>

Обработка данных включает следующие операции:

Парсинг xls-файлов в json формат по данным СПО/ВПО.

Формат xls-файлов представляет собой таблицы, содержащие данные о количестве студентов по разным показателям, связанные с приемом, выпуском и общей информацией о студентах среднего профессионального и высшего профессионального образования за 2013-2020 года.

Структуры файлов делятся на старый – с 2013 по 2015 года и новый – с 2016 по 2020. В общем случае формат данных старых и новых годов немного отличаются. Для каждого формата данных был разработан отдельный парсер, который формировал json-файл с результирующими

Структура базы данных представляет собой таблицы, имеющие по смыслу такой же формат как и таблицы из xls-файлов.

Схема базы данных представлена на рисунке 1.

В настоящее время ведется работа по реализации веб-интерфейса системы, позволяющего выполнять онлайн визуализацию накопленных данных.

ОРГАНИЗАЦИЯ СМЕШАННОГО ОБУЧЕНИЯ ПРОТОТИПИРОВАНИЮ В ПРОФИЛЬНОМ РЕСУРСНОМ ЦЕНТРЕ

Е.А. Мамаева

**ФГБОУ ВО «Вятский государственный университет»,
г. Киров, maevaekathy@gmail.com**

В статье рассматривается организация обучения аддитивным технологиям в условиях смешанного обучения. Затрагиваются кадровые проблемы развития аддитивных технологий. Организация обучения школьников в рамках сетевого профильного класса позволит повысить мотивацию обучающихся к изучению аддитивных технологий и повысит качество образования по данному направлению.

The article discusses the organization of training in additive technologies in a blended learning environment. The author touches upon the personnel problems of the development of additive technologies. The schoolchildren's teaching within the framework of a network profile class will increase the motivation to study additive technologies and improve the quality of education in this area.

Согласно Стратегии развития аддитивных технологий в Российской Федерации, отрасль аддитивных технологий играет важную роль в достижении национальных целей и стратегических задач развития страны. Кадровой проблемой отрасли является неполное соответствие профессиональных компетенций работников предъявляемым квалификационным требованиям по всем существующим направлениям аддитивных технологий [1].

Вопрос подготовки кадров является особенно актуальным, потому что в данный момент в России существует очень маленькое количество учебных заведений, специализирующихся на подготовке специалистов в сфере аддитивных технологий [2].

Для привлечения внимания абитуриентов к новым профессиям необходимо осуществлять профориентационную работу со школьниками. Предварительное ознакомление с аддитивными

технологиями возможно в рамках урочной и внеурочной деятельности по дисциплине информатика. Также федеральный государственный стандарт среднего общего образования по предмету «Технология» содержит в ряде других следующие направления: компьютерное черчение, промышленный дизайн; 3D-моделирование, прототипирование, технологии цифрового производства в области обработки материалов, аддитивные технологии. Обучение данным направлениям требует наличия не только оборудования, но и квалифицированных педагогических кадров.

С сентября 2019 года в Кирове были созданы 7 профильных ресурсных центров, основной целью деятельности которых является содействие созданию условий для получения качественного образования независимо от места жительства обучающихся; развитие сети образовательных организаций, реализующих подготовку обучающихся по профильным программам среднего общего образования.

В профильных ресурсных центрах были созданы сетевые профильные классы для обучающихся всей Кировской области. Любой школьник может обучаться в сетевом профильном классе навыкам той профессии, которую он хочет получить.

В профильном ресурсном центре КОГОАУ ВТЛ обучающиеся могут начать изучение аддитивных технологий. Учебный план рассчитан на 2 года обучения и содержит 11 элективных курсов. Обучение проходит в смешанном формате: изучение элективных курсов по профильной направленности в дистанционной форме во время учебы и в очной форме в каникулярное время.

По дисциплине «Прототипирование» предусмотрено 6 часов учебных занятий в каждом полугодии. Программа состоит из 2 разделов. При изучении первого раздела «Основы 3D-моделирования» обучающиеся знакомятся с основами программ для создания трехмерных моделей и создают свои трехмерные объекты. Второй раздел программы «3D-печать» знакомит школьников с особенностями работы на новейшем оборудовании, способами подготовки объектов для качественной печати на 3D-принтере.

Работа в дистанционном формате организована с применением различных цифровых сервисов и платформ. Для каждого обучающегося имеется доступ к СДО Moodle, где размещены различные курсы по выбранному профилю. Курс по аддитивным технологиям содержит презентации с теоретических и практических занятий, видеозаписи отдельных фрагментов по работе с технологиями, интерактивные задания для обучающихся. Кроме того, контрольно-измерительные

материалы, расположенные на образовательной платформе, позволяют быстро проводить срез имеющихся знаний или организовывать самоконтроль знаний по дисциплине. Два раза в неделю обучающиеся сетевого класса посещают онлайн-занятия с преподавателем на платформе Pruffme, где они могут лично задавать вопросы. Проведение занятий на данной платформе позволяет использовать следующие преимущества: возможность демонстрации презентаций и рабочего стола преподавателя, проведение различных опросов. Позже записи занятий можно найти в СДО Moodle в курсе по дисциплине и повторно посмотреть.

На очном обучении в каникулярное время обучающиеся осваивают работу с оборудованием для 3D-печати.

Таким образом правильная организация смешанного обучения, применение авторского подхода и использование ИОС в образовательной организации способствуют повышению мотивации обучающихся и формированию профориентации, снижают цифровое неравенство и повышают качество образования по аддитивным технологиям.

Литература:

1. Распоряжение Правительства РФ от 14 июля 2021 г. N 1913-р «Об утверждении Стратегии развития аддитивных технологий в РФ на период до 2030 года» [Электронный ресурс]. – Режим доступа: <https://www.garant.ru/products/ipo/prime/doc/401404208/#1000> (дата обращения: 01.11.2021).

2. Петров П.А., Сапрыкин Б.Ю., Петров М.А., Гневашев Д.А. Аддитивные технологии в системе среднего профессионального и высшего образования // РИТМ машиностроения. – 2021. – № 4. – С. 84–88.

СОДЕРЖАНИЕ

НАУЧНОЕ НАПРАВЛЕНИЕ – ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ И ОБРАБОТКИ ДАННЫХ

Секция МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ, АНАЛИЗ И ОБРАБОТКА ДАННЫХ

Боева В.А.	Новосибирск	3
Вохмин А.А.	Новосибирск	8
Грициенко И.В.	Новосибирск	12
Дворецкая В.К.	Новосибирск	17
Добшинский Н.И.	Москва, Северск	21
Дорошенко Е.А., Титаренко Е.Ю.	Томск	25
Дунаева К.Е.	Новосибирск	30
Тырышкин А.В., Жуань С.	Томск	36
Кайда А.Ю.	Томск	39
Клементьев Д.С., Терехов А.В.	Новосибирск	45
Кондратьев И.А.	Новосибирск	49
Крохина А.Е.	Новосибирск	53
Логачев Е.С., Калпакова Ю.А.	Новосибирск	60
Лузянин А.В.	Новосибирск	64
Новиков С.П.	Таганрог	66
Олейник Н.С.	Новосибирск	71
Осинцева Е.А., Чимитова Е.В.	Новосибирск	77
Перевалов К.Е., Богданов А.А., Черникова О.С.	Новосибирск	82
Поверин Д.В., Кушмуратов С.И., Постовалов С.Н.	Новосибирск	87
Чирихина А.С., Данченко И.Г., Черникова О.С., Марарескул Т.А.	Новосибирск, Железногорск	91
Чурквейдзе Д.И.	Москва	96

**Секция ЧИСЛЕННОЕ МОДЕЛИРОВАНИЕ ФИЗИЧЕСКИХ
ПРОЦЕССОВ В ТЕХНОЛОГИЯХ И ПРИРОДНЫХ ЯВЛЕНИЯХ**

Алтынбекова Г.Ж., Овчинникова А.С., Персова М.Г.	Новосибирск	101
Булатова А.З., Солнышкина О.А., Фаткуллина Н.Б.	Уфа	105
Гриф А.М.	Новосибирск	108
Дьячкова М.С., Гриф А.М.	Новосибирск	111
Кантарбаева А.И.	Томск	115
Ларкин Д.О.	Томск	119
Леонович Д.А., Соловейчик Ю.Г.	Новосибирск	123
Любимов Д.Н.	Новосибирск	126
Овчинникова А.С.	Новосибирск	130
Пастернак А.Д.	Новосибирск	135
Патрушев И.И.	Новосибирск	139
Патрушева Е.Е.	Новосибирск	143
Попова К.Е.	Северск	148
Рожкова Е.И.	Томск	152
Фархутдинов А.Р.	Казань	156
Фаткуллина Н.Б., Солнышкина О.А., Булатова А.З.	Уфа	158
Шанин Ю.В., Хоршев А.А., Бондарь А.С.	Санкт-Петербург	161

**Секция ТЕХНОЛОГИИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА,
МАШИННОГО ОБУЧЕНИЯ И ПРЕДСКАЗАТЕЛЬНОГО
МОДЕЛИРОВАНИЯ**

Адонин Н.Р.	Северск	166
Аксенова Д.В.	Новосибирск	171
Бережнев И.Е., Сокольникова Ю.А.	Якутск	174
Голованова Д.А.	Новосибирск	179
Добров Э.Р.	Новосибирск	183
Ершов П.К., Авдеев Т.В.	Новосибирск	189

Павлова А.И., Корж А.А.	Новосибирск	192
Кривецкий А.А.	Новосибирск	194
Кузьмина И.М.	Якутск	198
Макаров В.В.	Москва	201
Нестерова Е.Р.	Новосибирск	207
Пякилля Б.И.	Томск	211
Ржевский А.А.	Новосибирск	216
Соловьев Б.С.	Новосибирск	221
Солоха А.В., Соболева И.А.	Новосибирск	225
Тимшин В.Р.	Томск	229
Харченко И.К., Боровской И.Г.	Томск	232
Хнюнин М.В.	Новосибирск	236
Хуснутдинов А. О., Карманов В.С.	Новосибирск	240
Яковлева Е.А., Маслов Н.А.	Новосибирск	244

**Секция ЭКОНОМИКО-МАТЕМАТИЧЕСКИЕ МЕТОДЫ,
СТАТИСТИКА И ЭКОНОМЕТРИКА**

Воронкина Д.К.	Новосибирск	248
Гринько Д.А.	Новосибирск	252
Заводченко М.М.	Новосибирск	255
Корепанова Е.Г., Филимоненко Ю.В., Корепанов Л.Е.	Красноярск, Санкт-Петербург	260
Кривецкий А.А.	Новосибирск	264
Кулебякина А.А., Простакова А.Е., Перезовова Е.В.	Нижний Новгород	268
Ларионов П.С.	Новосибирск	273
Сивак М.А.	Новосибирск	277
Бойцова Е.А., Соколова А.О.	Красноярск	280
Соловьев Б.С.	Новосибирск	282
Тазутдинова Ю.Р., Федорова Н.В	Красноярск	287
Фадейкин Л.А.	Новосибирск	289
Мамаева Е.А.	Киров	292

НАУКА. ТЕХНОЛОГИИ. ИННОВАЦИИ

**XV Всероссийская научная конференция молодых ученых,
посвященная Году науки и технологий в России
г. Новосибирск, 06-10 декабря 2021 г.**

Сборник научных трудов в 10 частях

Часть 2

Под редакцией Соколовой Д.О.

Подписано в печать 15.12.2021. Формат 60×84/16. Бумага офсетная.

Уч.-изд. л. 17,43. Печ. л. 18,75. Тираж 100 экз. Заказ № Р-00575

Отпечатано в типографии

Новосибирского государственного технического университета
630073, г. Новосибирск, пр. К. Маркса, 20