

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

НАУКА ТЕХНОЛОГИИ ИННОВАЦИИ

**Сборник
научных трудов**

**г. Новосибирск, 03-07 декабря 2018 г.
в 9-и частях**

Часть 2

НОВОСИБИРСК
2018

УДК 62(063)
ББК 72.5я431
Н34

Н34 НАУКА. ТЕХНОЛОГИИ. ИННОВАЦИИ //Сборник научных трудов в 9 ч. / под ред. Гадюкиной А.В. – Новосибирск: Изд-во НГТУ, 2018. – Часть 2. – 282 с.

ISBN 978-5-7782-3733-9
Ч.2: ISBN 978-5-7782-3735-3

В сборнике публикуются материалы по научным направлениям: *«Информационные технологии математического моделирования и обработки данных»*

ОРГАНИЗАЦИОННЫЙ КОМИТЕТ

Председатель Оргкомитета:

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

Зам. председателя Оргкомитета:

Вострецов А.Г., д.т.н., профессор, проректор по научной работе НГТУ
Филатова С.Г., к.т.н., доцент, директор ИТЦ

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.социол.н., профессор, директор ИСТР
Рева И.Л., к.т.н., доцент, декан АВТФ
Ромм М.В., д.ф.н., профессор, декан ФГО
Саленко С.Д., д.т.н., профессор, декан ФЛА
Тимофеев В.С., д.т.н., доцент, декан ФПМИ
Хайруллина М.В., д.э.н., профессор, декан ФБ
Хрусталева В.А., д.т.н., профессор, декан РЭФ
Вильбергер М.Е., к.т.н., доцент, декан ФМА
Янпольский В.В., к.т.н. доцент, декан МТФ
Чернов С.С., к.э.н. доцент, декан ФЭН

УДК 62(063)
ББК 72.5я431

ISBN 978-5-7782-3735-3 (Часть 2)
ISBN 978-5-7782-3733-9

© Коллектив авторов, 2018
© Новосибирский государственный
технический университет, 2018

НАУЧНОЕ НАПРАВЛЕНИЕ – ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ И ОБРАБОТКИ ДАННЫХ

Секция МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ, АНАЛИЗ И ОБРАБОТКА ДАННЫХ

ПЛАТФОРМА ГИБРИДНОЙ РЕАЛЬНОСТИ С ОТКРЫТЫМ API

**А.А. Абдулов, З. А. Воловикова, М. Ф. Мустакимов,
А.Г. Балахчи, А.М. Веснин
ИГУ ФСИР Иркутск, Sanek_015@mail.ru**

Работа посвящена разработке тактильной матрицы (интерактивного мультимедийного стенда), которая выполняет распознавание размещения и перемещения объектов на ней, а также динамически проецирует реакцию системы в соответствии с определенными в ней законами. Обсуждаются элементы конструкции, алгоритмы работы программных модулей системы, а также приложений: игровых, обучающих, моделирующих

The work is devoted to the development of a tactile matrix (interactive multimedia stand), which performs recognition of the placement and movement of objects on it, and also dynamically projects the system's response in accordance with the laws defined in it. The article discusses the design elements, algorithms for the operation of the program modules of the system, as well as the applications: gaming, training, modeling

Текущее развитие IT-сферы направлено на соединение цифрового и реального мира. Технология дополненной реальности - это небольшой шаг в данном направлении, но для работы с ней обязательным является наличие дополнительных гаджетов, позволяющих увидеть виртуальный объект. Поэтому актуальна задача по разработке платформы с возможностью создания новых окружений и визуализаций, где физический и цифровой объекты существуют и взаимодействуют в реальном времени.

В данной работе Аппаратная часть представляет собой стол с прозрачной столешницей, сеткой-матрицей, предназначенной для размещения на ней маркеров, сами маркеры, видеокамеру, расположенную под столешницей, проектор или монитор. Программная - компьютер занимающийся распознаванием маркеров, управлением проектором (формированием изображения, которое будет наложено поверх стенда) и предоставляющий пользователю веб доступ к средствам HTTP API.

Такое решение позволяет совмещать слой реального мира - объекты выставленные на столе и слой виртуального мира - проекцию сформированную на стороне компьютера в зависимости от заложенной в приложении логики.

Система компьютерного распознавания реализована при помощи алгоритмов машинного обучения. Сначала находятся все клетки, в которых стоят маркеры, затем работает алгоритм (knn) k Nearest Neighbor, который классифицирует маркеры.

Для реализации наложения текстуры поверх объемных объектов выложенных на столе было написано небольшое приложение на языке программирования Processing. Оно позволяет калибровать выходное изображение под искажения формирующиеся при проецировании.

В качестве программных приложений были реализованы игры: «Drum машина» — при комбинировании на стенде разных маркеро-объектов проигрываются разные звуковые элементы (каждый звуковой элемент соответствует определенному маркеру). Сверху в зависимости от значения маркера на него наносится текстура. «ARBot» — система обучения программированию детей.

Такой интерактивный стенд имеет широкий спектр применения не только в обучающих системах и играх, но и для проведения и демонстрации различных экспериментов в области моделирования: например, размещение объектов в городских кварталах и их влияние на транспортную и пешеходную логистику.

Литература:

1. MACHINE LEARNING K-NEAREST NEIGHBORS (K-NN) ALGORITHM [Электронный ресурс] // OpenCV 3.0.0-dev documentation
2. Белогорцев Г. Игра «Жизнь» и моделирование естественного отбора [Электронный ресурс] // Хабрахабр, 2012. – Режим доступа:
3. Кушникренко А.Г., Рогожкина И.Б. Пиктомир: опыт обучения программированию старших дошкольников [Электронный ресурс] // <https://cyberleninka.ru/article/v/piktomir-opyt-obucheniya-programmirovaniyustarshih-doshkolnikov>

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ СМЕШАННОЙ КОНВЕКЦИИ В ПОРИСТОМ ГОРИЗОНТАЛЬНОМ КАНАЛЕ ПРИ НАЛИЧИИ ИСТОЧНИКА ЭНЕРГИИ

М.С. Астанина, М.А. Шерemet
Томский государственный университет,
г. Томск, astanina.marina@bk.ru

В работе проведено математическое моделирование смешанной конвекции в горизонтальном частично пористом канале с адиабатическими твёрдыми стенками. На нижней стенке расположен теплогенерирующий теплопроводный источник энергии. Над ним располагается пористый изотропный слой. В качестве рабочей среды рассматривалась ньютоновская теплопроводная жидкость, удовлетворяющая приближению Буссинеска. Математическая модель формулируется в безразмерных переменных «функция тока – завихренность скорости – температура». Численные исследования были проведены для различных значений определяющих параметров. Было показано влияние пористого слоя на эффективность теплоотвода энергии от тепловыделяющего элемента.

A mathematical modeling of mixed convection in an open horizontal partially porous channel with a heat-generating source was carried out. A heat-generating heat-conducting energy source is located on the bottom adiabatic wall. The working fluid was considered to be a Newtonian heat-conducting fluid which satisfies the Boussinesq approximation. The mathematical model is formulated in dimensionless variables "stream function - vorticity - temperature". Numerical studies were performed for various values of the governing parameters. It was shown that an addition of a porous insertion leads to an intensification of heat transfer from the surface of the energy source.

Эффективное охлаждение электронного оборудования позволяет продлить срок его службы и обеспечить выполнение задач предприятия. При повышении рабочей температуры электронного оборудования на 10°С относительно температуры окружающей среды срок его службы сокращается вдвое. Именно поэтому крайне важно позаботиться о правильном охлаждении устройств. В роли источников тепла могут выступать блоки питания, контроллеры, сервомеханизмы, трансформаторы, выпрямители тока, процессоры и серверы.

В настоящей работе проведено математическое моделирование активной системы охлаждения тепловыделяющего источника. Активная система охлаждения предполагает отвод тепла за счёт внешних факторов: ветер, вентиляторы и др.

Рассматривалась задача смешанной конвекции в открытом горизонтальном канале при наличии источника энергии постоянного тепловыделения. Схематично область решения показана на рисунке 1.

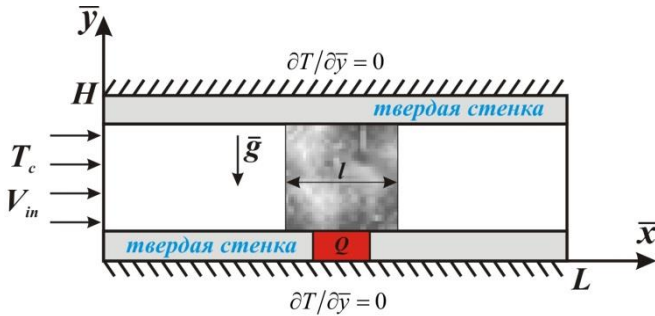


Рисунок 1 – физическая модель

Моделируемая область заполнена ньютоновской теплопроводной жидкостью, удовлетворяющей приближению Буссинеска. Вязкость рабочей среды зависит от температуры по экспоненциальному закону с параметром вязкости C [1]. Наружные поверхности горизонтальных стенок конечной толщины являются адiabатическими. Источник энергии расположен внутри нижней стенки. Течение жидкости и теплоперенос двумерны. На входе в канал индуцируется начальная скорость, образующая входящий поток холодной жидкости.

Математическая модель строится на основе безразмерных переменных «функция тока – завихрённость – температура». Использование такого подхода позволяет исключить из уравнений поле давления. Уравнения отдельно решаются для чистой среды, для пористой вставки, источника энергии и твёрдых стенок. Полученные уравнения решались методом конечных разностей на равномерной сетке [1, 2]. На границах раздела сред рассматривались граничные условия четвертого рода. Анализ результатов проводился по полученным распределениям изолиний функции тока и температуры, а также среднему числу Нуссельта и средней температуре, которые были рассчитаны на поверхности тепловыделяющего источника.

Численные эксперименты были проведены для различных значений определяющих параметров. Рассмотрены возможные комбинации параметров, позволяющие учесть многие особенности активной системы

охлаждения. На рисунке 2 показаны распределения изолиний функции тока и температуры для фиксированных теплофизических параметров: $Da = 10^{-5}$, $Re = 50$, $C = 1$ и разных размеров пористой вставки.

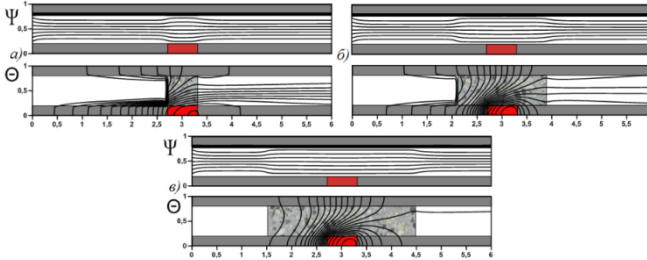


Рисунок 2 – Изолинии функции тока и температуры при $Da = 10^{-5}$, $Re = 50$, $C = 1$: а) $\delta = 0.6$, б) $\delta = 1.8$, в) $\delta = 3$

При любом размере пористой вставки течение сильно изменяет свою структуру в пористой зоне. При этом прогрев в этой области равномерный. Для минимального размера вставки (рисунок 8а) область канала после прохождения пористой зоны прогревается неравномерно. При $\delta = 1.8$ прогрев в области с чистой средой после пористой вставки становится более равномерным, изотермы симметрично распределены. Для максимального размера вставки (рисунок 8в) практически весь теплоём с источника происходит за счет теплопроводности внутри пористой среды и слабого течения, при этом твёрдые стенки прогреваются только в области вставки. При этом происходит более эффективный теплоём энергии с источника, средняя температура источника энергии уменьшается.

Таким образом, при моделировании активной системы охлаждения наиболее эффективным инструментом для улучшения теплоотвода от источника энергии является введение пористой вставки.

Представленные изыскания можно использовать не только для моделирования процессов охлаждения систем, но и для разработки более рациональных методов использования энергии.

Работа выполнена при финансовой поддержке Совета по грантам Президента РФ для молодых российских ученых (грант МД-2819.2017.8).

Литература:

1. *Astanina M.S., Sheremet M.A., Umavathi J.C.* Unsteady Natural Convection with Temperature-Dependent Viscosity in a Square Cavity Filled with a Porous Medium // *Transport in Porous Media*. –2015. – Vol. 110. –P. 113–126.

2. *Shenoy A., Sheremet M., Pop I.* Convective flow and heat transfer from wavy surfaces: viscous fluids, porous media and nanofluids. – Boca Raton: CRC Press, 2016. – 306 p.

РАСПРОСТРАНЕНИЕ ВОЛН В СПЛОШНОЙ НЕЛИНЕЙНОЙ СРЕДЕ

Д.С. Бардашов, А.Г. Быков, Д.В. Лосев

Национальный исследовательский Томский государственный университет, г. Томск, dark@gamelar.ru

Рассматривается задача распространения плоской волны с произвольной временной зависимостью в безграничной нелинейной среде. Решение этой задачи позволяет выявить общие закономерности преобразования формы волны и ее спектрального состава в зависимости от формы исходного сигнала и вида нелинейной характеристики. Предложенный подход, в отличие от классического, не приводит к неограниченному возрастанию амплитуды волны при больших значениях временной и пространственных переменных.

We consider the problem of propagation of a plane wave with an arbitrary time dependence in infinite nonlinear medium. The solution of this problem allows to reveal General regularities the transformation of waveform and its spectral composition depending on the original signal and nonlinear characteristics of medium. The proposed approach, unlike classical, does not lead to unlimited increase in the wave amplitude for large values of time and spatial variables

В последнее время большие усилия прикладываются для освоения терагерцового диапазона волн. Для преобразования частоты излучения предпринимаются попытки использования сосредоточенных нелинейных элементов [1-4], однако ввиду малого вклада одиночной нелинейности в энергетику сигнала это направление сталкивается с существенными трудностями. Одним из способов их преодоления является ис-

пользование протяженных нелинейных сред, в которых происходит накопление нелинейных эффектов, в результате чего возможно более гибкое управление формой и другими характеристиками сигнала.

Математически задачи анализа преобразования сигнала в сплошной нелинейной среде сводятся к решению нелинейных дифференциальных уравнений с частными производными, точные методы решения которых, за исключением простейших случаев, практически отсутствуют. Поэтому магистральным направлением исследования нелинейных сред является спектральный подход – своеобразный аналог метода комплексных амплитуд [5]. Он основан на представлении искомой зависимости в виде $\sum_n A_n e^{i\omega_n t}$, где «комплексные амплитуды» A_n зависят от

пространственных координат и находятся приближенными методами, основанными на малости изменения характеристик волны при распространении в среде, ω_n – исходная частота и ее кратные гармоники, выбираемые по виду нелинейной характеристики на основе физических представлений. Применение этого подхода возможно при 1) гармонической (узкополосной) временной зависимости сигнала; 2) полиномиальной аппроксимации нелинейной характеристики (обычно рассматриваются полиномы второго и третьего порядка).

Использование в ходе решения метода малых возмущений часто приводит к неограниченному возрастанию получаемой зависимости при увеличении временной и пространственных переменных. Подобный эффект хорошо известен в теории колебаний, где разработаны эффективные методы устранения соответствующих секулярных членов [6]. Кроме того, существующие подходы базируются на решении однородных уравнений, т.е. описанием возможных типов волн. Реализуемые же процессы адекватно описываются лишь неоднородными уравнениями с заданным источником. В данной работе делается попытка построения методов, свободных от указанных ограничений, на примере задачи о распространении излучения одномерного точечного источника с произвольной временной зависимостью в нелинейной среде.

Будем исходить из одномерного волнового уравнения в скалярном приближении

$$\frac{\partial^2 u}{\partial x^2} - \frac{1}{v^2} \frac{\partial^2 u}{\partial t^2} = \varphi(t)\delta(x) + \gamma_1 \frac{\partial^2 f(u)}{\partial t^2} + \gamma_2 \frac{\partial g(u)}{\partial t}, \quad (1)$$

где $\varphi(t)$ описывает временную форму сигнала, v – фазовая скорость в фоновой линейной среде, функции $f(u)$ и $g(u)$ описывают нелинейные свойства фазовой скорости и коэффициента поглощения среды, γ_1 и γ_2 – малые числовые параметры, характеризующие вклад нелинейных характеристик. Будем считать среду безграничной, а начальные условия однородными.

Поиск решения в виде прямого разложения по степеням малого параметра γ

$$u(x, t, \gamma) = u_0(x, t) + \gamma_1 u_1(x, t) + \gamma_2 u_2(x, t) + \dots$$

дает последовательность приближений, в частности

$$u_0(x, t) = -\frac{v}{2} \Phi(t - |x|/v) \chi(t - |x|/v), \quad (2)$$

где $\Phi(t) = \int_0^t \varphi(\tau) d\tau$, $\chi(t)$ – функция Хевисайда. Уже первые приближения

содержат слагаемые, пропорциональные $|x|$ (секулярные члены), что ограничивает применимость этих формул условиями $\gamma_1 |x| \ll 1$, $\gamma_2 |x| \ll 1$, т.е. областью вблизи источника.

Для устранения этой проблемы применим метод вариации параметров [7]. В качестве исходной формы выберем решение (2) невозмущенного уравнения (1) при $\gamma_1 = \gamma_2 = 0$. Будем варьировать амплитуду и время запаздывания сигнала (2) и введем в рассмотрение дополнительную аддитивную поправку. В результате решение уравнения (1) будем искать в виде

$$u(x, t) = -\frac{v}{2} [m(x, t) \Phi(\tau(x, t)) + s(x, t)] \chi(\tau(x, t)), \quad (3)$$

содержащем три подлежащих определению функции $m(x,t)$, $\tau(x,t)$ и $s(x,t)$, имеющие физический смысл коэффициента прохождения, запаздывания волны за счет изменения ее фазовой скорости нелинейной средой и рассеянной волны соответственно.

После подстановки представления (3) в уравнение (1) требуемые условия для нахождения неизвестных функций $\tau(x,t)$, $m(x,t)$ и $s(x,t)$ получим, приравнявая коэффициенты при $\delta'(\tau(x,t))$, $\delta(\tau(x,t))$ и $\chi(\tau(x,t))$. В первом приближении, пренебрегая взаимным влиянием искомых функций, были получены следующие результаты [7]:

$$\tau(x,t) \approx t - \frac{|x|}{v} \sqrt{1 + \gamma_1 v^2 F \left(-\frac{v}{2} \Phi \left(t - \frac{|x|}{v} \right) \right)}, \quad F(x) = f(x)/x,$$

$$m(x,t) = 1 + \frac{|x|}{\sqrt{1 + \gamma_1 v^2 F}} \left\{ \frac{\gamma_2 g}{\Phi} + \frac{\gamma_1 v \Phi}{2 \Phi Q} (F - f') \right\}, \quad Q = 1 - \frac{\gamma_1 |x| v^2 F' \cdot \Phi}{4 \sqrt{1 + \gamma_1 v^2 F}},$$

$$s(x,t) = w \Phi, \quad w = -\frac{v}{2} \int_0^t dt_1 \int_{x-v(t-t_1)}^{x+v(t-t_1)} dx_1 \left\{ 2\gamma_1 \left(\frac{\Phi}{\Phi} \right)' (f' - F) + \gamma_2 \frac{\Phi}{\Phi} (g' - G) \right\}.$$

В этих соотношениях для упрощения записи у функций Φ , Φ опущен аргумент $t - |x|/v$, «штрих» обозначает производную по всему аргументу, а нелинейные характеристики F, f, g зависят от аргумента $-v\Phi(t - |x|/v)/2$.

Предложенное решение задачи позволяет выявить общие закономерности преобразования формы волны в нелинейной среде в зависимости от формы исходного сигнала и вида нелинейной характеристики. Используемую для решения процедуру можно рассматривать как перенос методов нелинейной геометрической оптики [5] из спектральной во временную область.

Также можно отметить наличие в нелинейной среде ослабления/усиления волны даже при отсутствии у среды проводимости, что позволяет говорить об эквивалентной проводимости (переменного зна-

ка) нелинейной среды. Упоминание такого эффекта в литературе не обнаружено.

Литература:

1. Якубов В.П., Лосев Д.В., Мальцев А.И. Использование сверхширокополосного излучения для диагностики нелинейностей // Журнал радиоэлектроники (электронный журнал), 2000. № 3.
2. Якубов В.П., Лосев Д.В., Мальцев А.И. Диагностика нелинейностей по возмущениям рассеянного поля // Известия высших учебных заведений. Радиофизика. 2000. Т. 43. № 7. С. 645-651.
3. Якубов В.П., Шипилов С.Э., Сатаров Р.Н., Юрченко А.В. Дистанционная сверхширокополосная томография нелинейных радиоэлектронных элементов // Журнал технической физики. 2015. Т. 85. № 2. С. 122-125.
4. Быков А.Г., Лосев Д.В., Бардашов Д.С. Рассеяние волн нелинейным объектом // Известия высших учебных заведений. Физика. 2015. Т. 58. № 10-3. – С. 9-11.
5. Виноградова М.Б., Руденко О.В., Сухоруков А.П. Теория волн. – М.: Наука. Гл. ред. физ.-мат. лит., 1979. – 384 с.
6. Найфэ А. Введение в методы возмущений. – М.: Мир, 1984. – 526 с.
7. Лосев Д.В., Бардашов Д.С., Быков А.Г. Метод вариации параметров в задаче распространения волн в нелинейных средах // Журнал радиоэлектроники (электронный журнал), 2017, № 2.

СРАВНИТЕЛЬНЫЙ АНАЛИЗ НЕКОТОРЫХ АЛГОРИТМОВ НЕЛИНЕЙНОЙ ФИЛЬТРАЦИИ

Л.Ю. Витчикова, О.С. Черникова

**Новосибирский государственный технический университет,
г. Новосибирск, lbovvit@yandex.ru**

На примерах гауссовских нелинейных непрерывно-дискретных моделей проводится сравнительный анализ расширенного фильтра Калмана и сигма-точечного фильтра Калмана. Показано, что сигма-точечный фильтр предпо-

читательнее использовать в тех случаях, когда линеаризация модели состояний затруднительна.

The comparative analysis of the extended Kalman filter and the unscented Kalman filter is carried out on the examples of Gaussian nonlinear continuous-discrete models. It is confirmed, that the unscented Kalman filter is preferably used in cases where the linearization of the state model is difficult.

Введение

На практике часто возникает необходимость оценивания вектора состояния некоторой системы. Выбор алгоритма для решения поставленной задачи осуществляется в зависимости от математической модели объекта, а также характера случайных воздействий.

Если случайные воздействия подчинены гауссовскому закону распределения, то для данного случая применим расширенный фильтр Калмана (ЕКФ) и популярный в последнее время сигма-точечный фильтр Калмана (UKF).

В данной работе проводится сравнительный анализ двух алгоритмов нелинейной фильтрации для непрерывно-дискретных моделей: расширенного фильтра Калмана и сигма-точечного фильтра Калмана.

1. Постановка задачи

Модель гауссовской нелинейной непрерывно-дискретной системы в пространстве состояний имеет вид:

$$\frac{d}{dt}x(t) = f(x(t), u(t), t) + g(t)w(t), t \in [t_0, t_N]. \quad (1)$$

$$y(t_{k+1}) = h(x(t_{k+1}), v(t_{k+1}), t_{k+1}), k = 0, 1, \dots, N - 1. \quad (2)$$

Здесь:

$x(t)$ – n -вектор состояния, $u(t)$ – детерминированный r -вектор управления, $w(t)$ – p -вектор шума системы, $y(t_{k+1})$ – m -вектор измерения, $v(t_{k+1})$ – m -вектор шума измерения, $f(x(t), u(t), t)$, $g(t)$ и $h(x(t_{k+1}), v(t_{k+1}), t_{k+1})$ – нелинейные функции.

Предположим, что:

- случайный вектор $w(t)$ образует белую гауссовскую последовательность, для которой

$$E[w(t)] = 0, E[w(t)w^T(\tau)] = Q(t)\delta(t - \tau)$$

для любых $t, \tau \in [t_0, t_N]$ (здесь $\delta(t - \tau)$ – дельта-функция Дирака).

- случайный вектор $v(t_{k+1})$ образует белую гауссовскую последовательность (за исключением конечного числа точек $t_j \in \Omega_A$), для которой

$$\begin{aligned} E[v(t_{k+1})] &= 0, E[v(t_{k+1})v^T(t_{i+1})] = R(t_{k+1})\delta_{k,i}; \\ E[w(\tau)v^T(t_{k+1})] &= 0, k = 0, 1, \dots, N-1; \tau \in [t_0, t_N]. \end{aligned}$$

- начальное состояние $x(t_0)$ имеет нормальное распределение с параметрами

$$E[x(t_0)] = \bar{x}(t_0), E\{[x(t_0) - \bar{x}(t_0)][x(t_0) - \bar{x}(t_0)]^T\} = P(t_0)$$

и не коррелирует с $w(t)$ и $v(t_{k+1})$ при любых значениях k .

Для модели (1) - (2) с указанными априорными предположениями необходимо реализовать расширенный фильтр Калмана и сигма-точечный фильтр Калмана (UKF), а также исследовать целесообразность их применения.

2. Алгоритмы нелинейной фильтрации

Основная идея ЕКФ [1,2] состоит в аппроксимации нелинейной модели соответствующим линейным аналогом. Даже при вычислительной возможности нахождения матрицы частных производных для линеаризации исходной модели ЕКФ обеспечивает только первый порядок аппроксимации при оценивании параметров распределения вектора состояния. К сожалению, применение этого алгоритма к некоторым модельным структурам неприемлемо. Это связано с тем, что вектор-функция, стоящая в правой части уравнения (1), вычисляется достаточно сложно, в результате чего получение аналитических соотношений, необходимых для вычисления матриц Якоби, проблематично. Остроту вопроса в значительной мере снимает сигма-точечный фильтр Калмана [3-5]. UKF подвергает нелинейному преобразованию множество определенным образом вычисленных точек, на основе которых происходит оценивание параметров распределения вектора состояния. Указанный алгоритм не требует вычисления матриц Якоби, обеспечивает второй порядок аппроксимации для любой достаточно гладкой нелинейной функции системы.

3. Экспериментальные результаты

Проведем сравнительный анализ алгоритмов на следующих моделях:

Модель №1:

$$\frac{d}{dt} x(t) = -\sin(x(t)) + w(t), t \in [0,1],$$

$$y(t_{k+1}) = 0.5\sin(2x(t_{k+1})) + v(t_{k+1}), k = 0,1,\dots,99$$

Модель №2:

$$\frac{d}{dt} x(t) = \arctg(x(t)) + 0.1\exp(x(t)) - 1 + w(t), t \in [0,1],$$

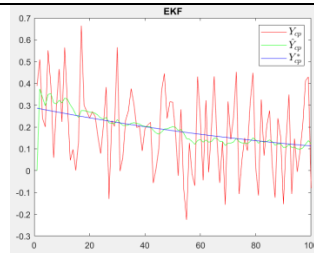
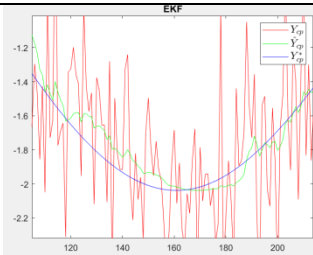
$$y(t_{k+1}) = \sin(x(t_{k+1})) + \arctg(x(t_{k+1})) + v(t_{k+1}), k = 0,1,\dots,99$$

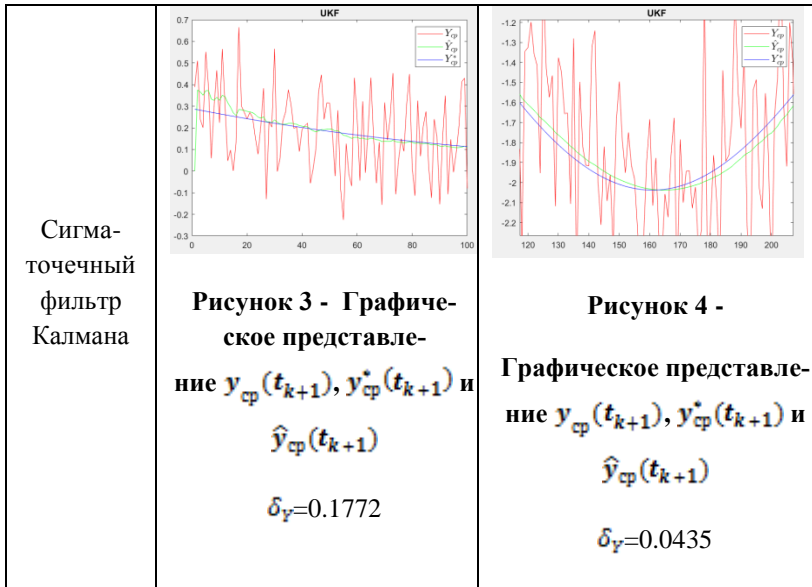
где $Q(t) = 0,01$; $R(t_{k+1}) = 0,02$; $\bar{x}(t_0) = 0$; $P(t_0) = 0,1$.

Результаты представим в Таблице 1, обозначив

$$\delta_y = \frac{\|Y_{cp}^* - \hat{Y}_{cp}\|}{\|Y_{cp}^*\|} = \sqrt{\frac{\sum_{k=0}^{N-1} (y_{cp}^*(t_{k+1}) - \hat{y}_{cp}(t_{k+1}))^2}{\sum_{k=0}^{N-1} (y_{cp}^*(t_{k+1}))^2}}$$

Таблица 1 – Численные результаты нелинейной фильтрации

| | Модель №1 | Модель №2 |
|----------------------------|---|---|
| Расширенный фильтр Калмана |  <p>Рисунок 1 - Графическое представление $y_{cp}(t_{k+1})$, $y_{cp}^*(t_{k+1})$ и $\hat{y}_{cp}(t_{k+1})$</p> <p>$\delta_y=0.2470$</p> |  <p>Рисунок 2 – Графическое представление $y_{cp}(t_{k+1})$, $y_{cp}^*(t_{k+1})$ и $\hat{y}_{cp}(t_{k+1})$</p> <p>$\delta_y=0.0971$</p> |



Заключение

На примерах некоторых модельных структур показано, что сигма-точечный фильтр предпочтительнее использовать в тех случаях, когда линеаризация модели состояний затруднительна, либо линейный аналог неточно описывает поведение системы. Следует отметить, что при возможности вычисления соответствующих матриц Якоби алгоритмы работают с сопоставимой точностью.

Литература

1. Огарков М.А. Методы статистического оценивания параметров случайных процессов / М.А. Огарков. – М.: Энергоатомиздат, – 1990. – 208 с.
2. Verhaegen M. Filtering and system identification: a least squares approach / M. Verhaegen, V. Verdult – Cambridge University Press, – 2007. – 422 p.
3. Julier S.J. A new extension of the Kalman filter to nonlinear systems / S.J. Julier, J. K. Uhlmann – in Proc. Of AeroSense: The 11-th Int. Symp. on Aerospace/Defence Sensing, Simulation and Control, – 1997. – 12p.

4. Särkkä S. On unscented Kalman filtering for state estimation of continuous-time nonlinear systems / S. Särkkä // IEEE Transactions on Automatic Control. –2007. – P. 11.
5. Ishihara S. Continuous-discrete robust UKF for nonlinear systems with parameter uncertainties / S. Ishihara, M. Yamakita // 48th ISCTE International Symposium on Stochastic Systems Theory and Its Applications. – 2016. – P. 6.

АВТОМАТИЗИРОВАННАЯ БАЗА ДАННЫХ «ПРОФИЛЬ»

И.О. Гелеверя, С.П. Куделин, Ю.В. Инатович
Уральский федеральный университет имени первого Президента
России Б.Н. Ельцина, г. Екатеринбург, igeleveria@urfu.ru

В УрФУ разработана автоматизированная база данных «Профиль», предназначенная для поиска калибровки валков - аналога при разработке калибровок с использованием экспертной системой «Технология сортовой прокатки». Созданная база данных позволяет быстро осуществить поиск калибровки профиля - аналога на основе специально разработанной системы функциональных формул.

In UrFU has developed an automated database "Profile" designed to search for pass design of rolls - analogue in the development of roll pass designs using the expert system "Technology of shape rolling". The created database allows you to quickly search for roll pass design profile-analogue on the basis of a specially developed system of functional formulas.

Для различных отраслей промышленности ежегодно выпускается около 2000 фасонных профилей отраслевого и специального назначения, причем сортамент этих профилей постоянно обновляется. Наличие наиболее близкой технологии – аналога помогает значительно сократить время разработки калибровки валков для прокатки новых профилей сортамента стана.

Для представления знаний о прокатке сложных фасонных профилей отраслевого и специального назначения формализовали известную технологическую классификацию [1]. Согласно этой классификации все освоенные профили разделены по общности принципов их калибровки на 16 классов и 43 группы.

При формализации использовали системную методологию классифицирования материальных объектов, согласно которой процесс клас-

сифицирования любых материальных объектов рассматривают как две взаимосвязанные процедуры: таксономию – разбиение объектов на группы (таксоны) по общности выполняемых ими функций, и мерономию – расчленение объектов на структурные части (мероны)[2-4]. Всё классифицируемое поле объектов в целом рассматривают как самый крупный таксон (таксономический универсум), который имеет иерархическую структуру и включает таксоны 1-го, 2-го и т. д. уровня.

С позиций такой методологии выделенные Н.Е.Скоруховым – Б.М.Илюковичем [1] 16 классов профилей являются таксонами первого уровня, а 43 группы, входящие в указанные классы, представляют таксоны второго уровня. С целью построения функциональных формул обозначили весь таксономический универсум буквой П, таксоны первого уровня – $i=01, 02, \dots, 16$, а таксоны второго уровня – $j=01, 02, \dots, 43$, тогда любой таксон в классифицируемом поле получает обозначение $П_{ij}$, например, обозначение $П0408$ кодирует полосовой профиль желобчатой формы.

Для того, чтобы описать строение каждого таксона (т.е. конструкцию профиля), выделили все составные элементы профилей (мероны): клиновидные и прямоугольные участки, выемки и выступы, зубья, желоба, ступеньки, гребни, фланцы и др. Причем каждый из этих меронов может иметь несколько разновидностей (модусов). Например, клиновидные участки профиля могут быть односторонними или двухсторонними, скошенными сверху или снизу и т.п. Каждому мерону присвоили цифровой номер, а каждый модус обозначили буквой. Сформированная таким образом совокупность всех возможных элементов фасонных профилей представлена в диссертации в виде таблицы и включает 21 мерон, каждый из которых может иметь до 20 модусов. Функциональная формула любого профиля включает обозначение таксона и приведенный в скобках кортеж буквенно-цифровых кодов меронов и модусов, разделенных точкой с запятой:

$П_{ij} (1Д; 2Ц; 3О, \dots)$.

Например, полосовой профиль с прямоугольными участками, с круглым выступом на верхней стороне и такой же выемкой снизу (Рисунок 1) относится к таксону $П0102$ и включает мероны 2Л (прямоугольный участок с левого края), 2П (прямоугольный участок с правого края), 6КВ (выступ полукруглый на верхней стороне) и 5КН (выемка полукруглая на нижней стороне). Соответственно этому функциональная формула имеет вид:

$П0102(2Л; 6КВ; 5КН; 2П)$.

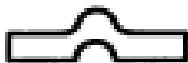


Рисунок 1 – Сложный профиль с функциональной формулой
П0102(2Л; 6КВ; 5КН; 2П)

Полученная формализованная классификация позволила поместить в память ЭВМ в виде функциональных формул все известные профили отраслевого и специального назначения.

Для графического представления фасонных профилей и калибров применили математическую модель, основанную на параметрическом описании кривых в форме Безье и реализованную Е.Ю.Слукиным и В.А.Шиловым в виде программы «Графок» [5]. По этой программе для автоматизированного построения профиля или калибра произвольной формы задают координаты X_i , Y_i характеристических вершин многогранника, описанного вокруг профиля и радиусы закругления R_i каждой характеристической вершины $i=1, 2, 3, \dots, n$, где n – число вершин. Для описания по такой методике калибровки валков составили файлы с координатами X_i , Y_i , R_i для каждого калибра, применяемого для прокатки каждого профиля описанного функциональной формулой. Эти координаты позволяют с помощью программы «Графок» отобразить на экране дисплея всю последовательность калибров для прокатки закодированного профиля.

На основе разработанной модели составлена база данных «Профиль», включающая следующие информационные и функциональные блоки (Рисунок 2):

1. Статическая база классов - таблица, в которой хранятся названия классов профилей по технологической классификации.
2. Статическая база групп - таблица, в которой хранятся названия групп профилей из той же классификации.

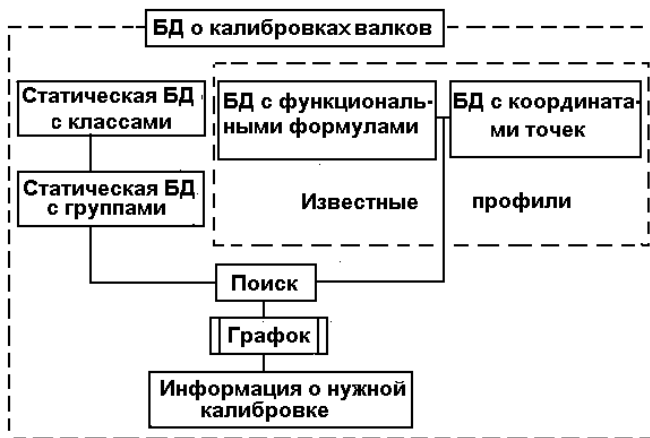


Рисунок 2 – Структурная схема базы данных "Профиль"

3. База с функциональными формулами - таблица, в которой хранятся функциональные формулы известных профилей.

4. База с координатами характеристических точек каждого калибра всех калибровок валков для прокатки профилей.

5. Функциональный блок «Поиск», в котором осуществляется поиск наиболее близких по форме чистовых профилей и соответствующих им калибровок.

6. Блок, содержащий программу "График" и позволяющий получить на экране монитора графическое отображение калибров при поиске нужной калибровки валков.

БД «Профиль» включена в экспертную систему «Технология сортовой прокатки [6] и позволяет при проектировании технологии прокатки нового фасонного профиля находить в автоматизированном режиме профиль - аналог и калибровку - аналог. При этом поиск профиля аналога по функциональной формуле производится с помощью метода решающих функций теории распознавания образов.

Литература:

1. Скороходов Н.Е., Илюкович Б.М., Шулаев И.П., Калибровка сложных профилей. М.: Metallurgy, 1979, 232 с.

2. Мартьянов Ю.А., Шилов В.А., Куделин С.П. База данных о калибровках валков для прокатки сложных профилей. Изв. Вузов. Черная металлургия. 1996. №6. С. 74–75.

3. Мартьянов Ю.А., Куделин С.П., Ваганов В.Н. Автоматизированная база данных по калибровкам прокатных валков. //Труды научно – техн. конф., посвященной образованию металлургического факультета УГТУ–УПИ. Екатеринбург: изд. УГТУ–УПИ. 1995. - С. 81.

4. Куделин С.П., Шилов В.А. Информационная система о калибровках валков для прокатки фасонных профилей. //Программные продукты и системы. 2000. №1. С. 47–48.

5. Шилов В.А., Слукин Е.Ю. Геометрическая модель калибров и профилей произвольной формы для диалоговых систем автоматизированного проектирования технологии сортовой прокатки // Известия вузов. Цветная металлургия, 1995, №2, с. 51 – 52.

6. Программный комплекс «Экспертная система технологии сортовой прокатки» / С.П.Куделин, В.К.Смирнов, В.А.Шилов, Ю.В.Инатович // Свидетельство №2001610602 от 25 мая 2001 г. М.: РАСПАТЕНТ, 2001.

МОДЕЛИРОВАНИЕ ГЕНЕРАТОРА СЕЙСМИЧЕСКИХ СИГНАЛОВ

Р.А. Гераськин, А.А. Игнатьева, А.А. Спектор
Новосибирский государственный технический университет,
г. Новосибирск, geraskin.ra@gmail.com a.a.ignateva@mail.ru

При разработке новых алгоритмов сейсмического обнаружения и классификации требуется их частая проверка. Из-за сложности проведения частых полевых испытаний была построена модель, формирующая множество сейсмических сигналов на основе реальных. В дальнейшем эти сигналы можно использовать для оценки новых алгоритмов.

When developing seismic detection and classification algorithms, their frequent verification is required. Due to the complexity of conducting frequent field tests, a model was built that forms a multitude of seismic signals based on real ones. In the future, these signals can be used to evaluate new algorithms.

Для обеспечения охраны территории предприятия, наряду с другими типами средств обнаружения, используются сейсмические системы об-

наружения (ССО). Принцип работы таких систем базируется на обнаружении сейсмических сигналов, возникающих в результате движения объектов в охраняемой зоне.

ССО включают в себя:

- сейсмические датчики, регистрирующие сигналы, исходящие от подвижных объектов;
- аппаратное обеспечение;
- программное обеспечение.

Наиболее важным из перечисленных составляющих является программное обеспечение, так как именно оно определяет, насколько успешно будет обнаружен подвижный объект на охраняемой территории.

Перед разработчиками программного обеспечения стоят несколько задач:

- разработка алгоритма обнаружения сейсмического сигнала, исходящего от объекта на фоне помех;
- определение местоположения подвижного объекта;
- классификация подвижного объекта.

При решении поставленных задач требуются частые проверки полученных алгоритмов. Проверки в реальных полевых условиях слишком затруднительны и дороги, а порой и вовсе невозможны. Альтернативой реальным испытаниям послужило использование математической модели для отработки различных алгоритмов, в частности алгоритмов обнаружения и классификации объектов.

Математическое моделирование предполагает построение и использование модели, способной на основе ограниченного количества реальных сейсмических сигналов различных классов объектов формировать множество псевдослучайных сигналов со схожими с реальными сигналами характеристиками. Такое моделирование позволяет более точно оценивать результат работы новых алгоритмов из-за значительного увеличения объёма тестовых сигналов.

Математическая модель представляет из себя генератор сейсмических сигналов четырёх классов подвижных объектов: человека, группы людей, крупного животного (лошадь), автомобиля. Входными параметрами генератора является отношение сигнал-шум, а также длина требуемого сигнала. Результатом работы модели является четыре сигнала, описанных выше классов, заданной длительностью и отношением сигнал-шум.

Генератор работает следующим образом. Изначально в модель загружаются участки реальных сейсмических сигналов, записанных на полигоне. Далее, по формулам (1), (2) и (3) рассчитываются характеристики каждого из заложенных сигналов, где B_x – отсчёты автокорреляционной функции сигнала [1].

$$B_x = \begin{pmatrix} B_x(0) & B_x(1) & \dots & B_x(J-1) \\ B_x(1) & B_x(0) & \dots & B_x(J-2) \\ \dots & \dots & \dots & \dots \\ B_x(J-1) & B_x(J-2) & \dots & B_x(0) \end{pmatrix} \quad (1)$$

$$B_n = \|B_x(1), B_x(2), \dots, B_x(J)\|^T \quad (2)$$

$$A = B_x^{-1} B_n \quad (3)$$

Характеристики сигнала представлены вектором параметров (4). Количество параметров в векторе задаётся перед началом работы генератора числом, называемым порядком модели. Далее, полученные параметры сейсмических сигналов всех четырёх классов включаются в модель предсказания (5), результатом работы которой и являются отсчёты сейсмических сигналов [1].

$$A = \|a_1, a_2, \dots, a_j\|^T \quad (4)$$

$$x_i = \sum_{j=1}^J a_j x_{i-j} + \xi_i \quad (5)$$

Для классов человека, группы людей и крупного животного требуется дополнительно сформировать специальную огибающую для моделирования шагов. В качестве огибающей было принято использовать колокольную функцию Гаусса, описанную формулой (6).

$$s_i = \sum_k a_k(r_k) \cdot \exp\left(-\frac{(i-kT + \varepsilon_k)^2}{2\tau^2}\right) \quad (6)$$

Помимо перечисленного, в генератор был добавлен пятый канал для генерации сейсмического фона для проверки помехоустойчивости алгоритмов сейсмической классификации и обнаружения подвижных объектов.

Литература:

1. Спектор А. А., Райфельд М.А. Статистические задачи и методы пассивной сейсмической локации Statistical problems and methods of passive seismic location : монография / Новосибирск : Изд-во НГТУ, 2018. - 200 с - (Монографии НГТУ). - 3000 экз. - ISBN 978-5-7782-3665-3

РАЗРАБОТКА СИСТЕМЫ ВИЗУАЛИЗАЦИИ ЭНЕРГЕТИЧЕСКИХ БИОРИТМОВ ЧЕЛОВЕКА

А. В. Демьяненко, Е. В. Рабинович

**Новосибирский государственный технический университет,
г. Новосибирск, d-aart@ya.ru**

В статье рассматривается разработанный вариант визуализации для динамической модели энергетических биоритмов человека, описывается структура, и полученный вариант визуализации.

The article discusses the developed version of visualization for a dynamic model of human energy biorhythms, describes the structure, and the resulting version of visualization.

В настоящее время предпринимаются попытки построения физических и математических моделей, адекватных основным понятиям и концепциям восточной медицины вследствие того, что европейской медициной все больше признается ее высокая эффективность.

Ряд моделей рассматривает человеческий организм с позиции теории открытых нелинейных динамических систем. Повышенное внимание уделяется моделям распространённых в биофизике, которые обладают хаотическим режимам динамики из-за их особой чувствительности к слабым внешним воздействиям. Одной из моделей, базирующихся на вышеупомянутых принципах, является модель энергетических биоритмов человеческого организма [1,2,3].

В докладе рассматриваются особенности разработанного варианта визуализации модели энергетических биоритмов человеческого организма на основе предложенной модели [4,5]. Описана структура модели, связи между ее компонентами, а также циклы обмена энергией между компонентами. Предложено графическое представление диссипативных структур (ДС) их уровней энергии в них, а также связей между всеми ДС и уровнями.

Для моделирования и визуализации таких сложных процессов возникла необходимость в создании визуального представления данной модели и ключевых процессов в ней.

Результат визуализации представлен в виде графа, состоящего из пяти диссипативных структур, произведена визуализация обмена энергией между структурами, показаны пути обмена и циклы, а также визуализация всех обменов в виде таблицы с обозначением соответствующих признаков (ДС была активна, ДС закрыта, ДС не активна). Пример результата визуализации представлен на рисунке 1.

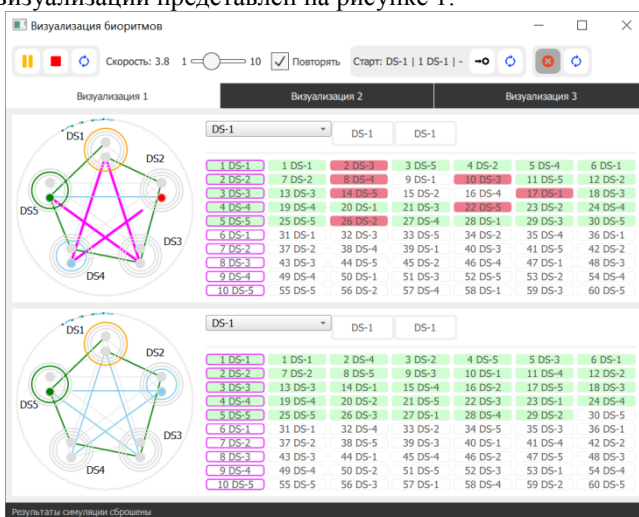


Рисунок 1 Главный экран системы визуализации

Помимо вышеописанного, был визуализирован вариант работы модели при наличии закрытых диссипативных структур и обратный обмен между ДС.

Литература:

1. Гласс Л., Меки М. От часов к хаосу. Ритмы жизни. – М.: Мир, 1991. – 248 с.
2. Goldberger A. L. Nonlinear dynamics, fractals, and chaos: applications to cardiac electrophysiology. // Ann. Biomed. Eng. 1990, v. 18, №2, p. 195 - 209.
3. Небрат В. В., Рабинович Е. В. Исследование механизмов лечебного воздействия традиционной восточной медицины с использованием информационной ЭМАТ технологии // Информационные системы и технологии: Труды междунар. Научн.-техн. Конф. – Новосибирск. НГТУ. – 2003. Т. 3. С. 48 - 53.
4. Небрат В. В., Рабинович Е. В. Управление диссипативными состояниями человека // Международная конференция: «Наука и будущее: идеи, которые изменят мир». – Москва. – ГГМ им. Верданского РАН 14-16.04. 2004. - С. 133-134
5. Рабинович Е. В., Небрат В. В. Динамическая модель акупунктуры. // Третий международный тихоокеанский конгресс по традиционной медицине: Материалы Конгресса. – Владивосток, Сеул. – 2006. – С. 17 – 18.

РАЗРАБОТКА СИСТЕМЫ ВИЗУАЛИЗАЦИИ ЭНЕРГЕТИЧЕСКИХ БИОРИТМОВ ЧЕЛОВЕКА

А.В. Демьяненко, Е.В. Рабинович

**Новосибирский государственный технический университет,
г. Новосибирск, d-aart@ya.ru**

В статье рассматривается разработанный вариант визуализации для динамической модели энергетических биоритмов человека, описывается структура, и полученный вариант визуализации.

The article discusses the developed version of visualization for a dynamic model of human energy biorhythms, describes the structure, and the resulting version of visualization.

В настоящее время предпринимаются попытки построения физических и математических моделей, адекватных основным понятиям и концепциям восточной медицины вследствие того, что европейской медициной все больше признается ее высокая эффективность.

Ряд моделей рассматривает человеческий организм с позиции теории открытых нелинейных динамических систем. Повышенное внимание уделяется моделям распространённых в биофизике, которые обладают хаотическим режимом динамики из-за их особой чувствительности к слабым внешним воздействиям. Одной из моделей, базирующихся на вышеупомянутых принципах, является модель энергетических биоритмов человеческого организма [1,2,3].

В докладе рассматриваются особенности разработанного варианта визуализации модели энергетических биоритмов человеческого организма на основе предложенной модели [4,5]. Описана структура модели, связи между ее компонентами, а также циклы обмена энергией между компонентами. Предложено графическое представление диссипативных структур (ДС) их уровней энергии в них, а также связей между всеми ДС и уровнями.

Для моделирования и визуализации таких сложных процессов возникла необходимость в создании визуального представления данной модели и ключевых процессов в ней.

Результат визуализации представлен в виде графа, состоящего из пяти диссипативных структур, произведена визуализация обмена энергией между структурами, показаны пути обмена и циклы, а также визуализация всех обменов в виде таблицы с обозначением соответствующих признаков (ДС была активна, ДС закрыта, ДС не активна). Пример результата визуализации представлен на рисунке 1.

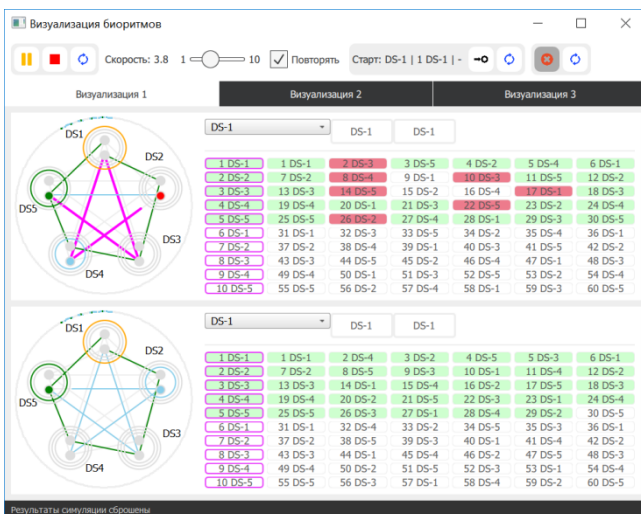


Рисунок 1 Главный экран системы визуализации

Помимо вышеописанного, был визуализирован вариант работы модели при наличии закрытых диссипативных структур и обратный обмен между ДС.

Литература:

1. Гласс Л., Меки М. От часов к хаосу. Ритмы жизни. – М.: Мир, 1991. – 248 с.
2. Goldberger A. L. Nonlinear dynamics, fractals, and chaos: applications to caediac electrophysiology. // Ann. Biomed. Eng. 1990, v. 18, №2, p. 195 - 209.
3. Небрат В. В., Рабинович Е. В. Исседование механизмов лечебного воздействия традиционной восточной медицины с использованием ионформационной ЭМАТ технологии // Информационные системы и технологии: Труды междун. Научн.-техн. Конф. – Новосибирск. НГТУ. – 2003. Т. 3. С. 48 - 53.
4. Небрат В. В., Рабинович Е. В. Управление диссипативными состояниями человека // Международная конференция: «Наука и будущее: идеи, которые изменяют мир». – Москва. – ГГМ им. Верданского РАН 14-16.04. 2004. - С. 133-134

5. Рабинович Е. В., Небрат В. В. Динамическая модель акупунктуры. // Третий международный тихоокеанский конгресс по традиционной медицине: Материалы Конгресса. – Владивосток, Сеул. – 2006. – С. 17 – 18.

МОДЕЛЬ ОБЪЕКТА УПРАВЛЕНИЯ ДЛЯ РОБОТОТЕХНИЧЕСКОЙ СИСТЕМЫ МЕДИЦИНСКОГО НАЗНАЧЕНИЯ

В.Г. Дружинин^{1,2}, В.А. Морозов¹

¹ Кафедра физической механики Санкт-Петербургского государственного университета, ² Центральный научно-исследовательский и опытно-конструкторский институт робототехники и технической г. Санкт-Петербург, vasily.dr.mob@gmail.com, v.morozov@spbu.ru

В данной работе представлена разработанная математической модель, описывающий отклонение медицинской стальной инъекционной иглы при ее движении в фантоме мягких тканей (имитация тканей человека). Данная модель необходима для обеспечения корректировки работы роботизированной системы при проведении операций брахитерапии или схожих операций, где необходимо высокоточное позиционирование кончика иглы в теле пациента. Поскольку кончик иглы является асимметричным, то при движении в тканях игла будет деформироваться, отклоняться от прямолинейного движения. Таким образом, внедряя и поворачивая иглу вокруг своей оси, можно провести кончик иглы по заданной траектории. Разрабатываемая модель необходима для корректировки движения иглы в тканях человека или для прогнозирования оптимальных мест прокола. В работе рассмотрена общая постановка задачи и решена одна из подзадач, а именно, расчет отклонения кончика иглы при поступательном движении в однородном материале. Приведены результаты моделирования для различной плотности материала и различного угла острия иглы. Проведено сравнение экспериментальных данных с результатами, полученными при расчете с помощью разработанной модели.

In this paper, a mathematical model describing the deviation of the medical steel needle when it moves in the phantom of soft tissues (human tissues). This model will expect for use in the operational mode, where high precision positioning of the tip of the needle is necessary. The needle has asymmetric tip and because when the needle moves, it to deviate from rectilinear motion. Thus, by inserting and rotating the needle around its axis, it is possible to move the tip of the needle along a predetermined path. Consequently, having developed the model with its help, it will be possible to track the

movement of the needle in the tissues or to predict optimal puncture points. The general formulation of the problems is considered and one of the sub-tasks is solved, as well as the available calculation of the needle tip deflection during the translational movement in a homogeneous material. The article shows the model used and the simulation results for different material densities and different needle tip angles. Comparison of the experimental data with the results obtained in the calculation using the proposed model is given.

В настоящее время в медицине для выполнения операций различной сложности широко применяются робототехнические комплексы. На сегодняшний день в ЦНИИ РТК разработан макет роботизированной системы «ОнкоРОБОТ» для брахитерапии рака предстательной железы (РПЖ) [1, 2]. Данная процедура проводится посредством внедрения микро-источников радиоизлучения в предстательную железу максимально близко к опухоли. Основная сложность заключается в подведении кончика иглы к целевой точке (опухоли) при проведении операции.

Преимущества использования роботов по сравнению с традиционными методами заключаются в том, что роботизированный манипулятор способен обеспечить высокую точность наведения инструмента и его контролируемого силового воздействия, что позволяет рассчитывать не только на повышение качества освоенных в настоящее время операций, но и создание базиса для разработки принципиально новых хирургических технологий. Другим важным преимуществом является отсутствие прямого контакта врача с радиоактивными источниками, что позволит обезопасить медицинский персонал от сопутствующего радиационного облучения.

Из-за своих геометрических особенностей и прилагаемых нагрузок в процессе выполнения операции игла деформируется, что приводит к отклонению иглы от прямолинейного движения.

Постановка задачи

Необходимо разработать модель для определения положения кончика ассиметричной медицинской инъекционной иглы (рис. 1) в системе координат манипулятора *Охуз* при поступательном и вращательном движении иглы.

Данная задача является объемной и междисциплинарной. Для ее решения нужно затронуть несколько разделов механики и учитывать большое количество начальных и граничных условий. В данной работе будет рассмотрен первый этап разработки модели, а именно: «Расчет движения иглы в плоскости *Оху*, деформация иглы в зависимости от поступательного движения».

Модель

На рисунке 1 представлены возможные силы, действующие на иглу:

- F_t – сила, действующая на кончик иглы;
 - F_f – сила трения, возникающая при движении иглы внутри ткани;
- ни;
- $w(x)$ – распределенная нагрузка (сила, которую оказывает ткань на поверхность иглы).

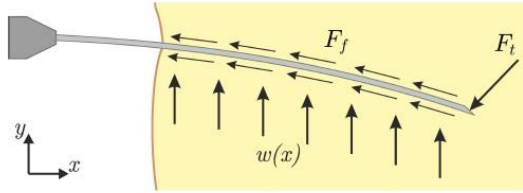


Рисунок 1 – Перемещаемая игла в тканях человека

Уравнение равновесия сил может быть представлено в следующем виде:

$$\vec{F}_{needle} = \vec{F}_t + \vec{F}_f + w(x)$$

где F_{needle} – сила, с которой внедряется игла.

В данном случае игла представляется в виде консольной балки с жесткой заделкой, с одной стороны. Местом закрепления будем считать место прокола. Тогда получается, что с ростом времени длина балки будет увеличиваться. Таким образом, при каждом шаге по времени будет рассчитываться новое отклонение и суммируется с предыдущим расчетом:

$$y_{all} = \sum_n^{n-1} y_n + y_n, \quad (2)$$

где n – текущая итерация моделирования; y_{all} – суммарное отклонение иглы при ее движении; y_n – отклонение иглы на текущем шаге времени.

Размер элементарных клеток (молекул) во много раз меньше, чем габаритные размеры иглы. Скорость перемещения иглы в среде достаточно низкая, а плотность среды достаточно высокая по сравнению с воз-

духом. Исходя из этого, для моделирования внешней силы F_{needle} при перемещении иглы в тканях человека можно использовать силу лобового сопротивления [3] $F_{needle} = C \cdot (\rho v^2 / 2) \cdot S$. Для расчета самого отклонения будем использовать формулу [4]:

$$y_n = \frac{Fl(t)^3}{3EJ_x} \quad (3)$$

где F – проекция силы лобового сопротивления на плоскость, ортогональную оси иглы; J_x – осевой момент инерции; $l(t)$ – длина иглы, находящейся в тканях человека; t – время; E – модуль Юнга.

На рисунке 2 представлены экспериментальный и расчетный графики зависимости отклонения иглы от скорости. Из рисунка видно, что эти зависимости несильно отличаются друг от друга. Можно сделать вывод, что данный подход применим для определения отклонения при движении иглы.

Неточность моделирования можно объяснить большим количеством допущений, которые на следующих итерациях разработки модели будут учтены.

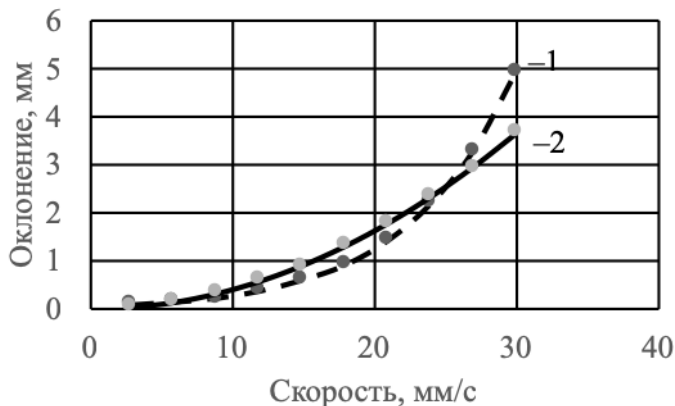


Рисунок 2 – Графики зависимости отклонения кончика иглы от скорости, 1 – результаты эксперимента, 2 – результаты моделирования

Заключение

В данной работе разработана модель, описывающая отклонение иглы в реальном времени. Для построения модели использовался новый подход. Сравнение результатов моделирования с экспериментальными данными показало работоспособность предложенного метода. Можно рассматривать данный этап работы, как задел для выполнения дальнейших этапов разработки модели.

Литература:

1. Управление роботом для брахитерапии на основе информации ультразвукового датчика/ Н.А. Грязнов, Г.С. Киреева, В.В. Харламов, К.Ю. Сенчик, Д.В. Новицкий, С.А. Никитин. // Робототехника и техническая кибернетика. 1(10). 2016. С 67-71.
2. Перспективы использования оригинальной роботизированной системы для брахитерапии рака предстательной железы / Н.А. Грязнов, Г.С. Киреева, В.В. Харламов, К.Ю. Сенчик, Д.В. Новицкий, С.А. Никитин. // Вестник хирургии им. И.И. Грекова. Том 176, выпуск 1. 2017. С 107-111.
3. Идельчик А. Е. Справочник по гидравлическим сопротивлениям/ Под ред. М.О. Штейнберга – 3-е изд., издательство Машиностроение 1992, С469 – 500.
4. Икрин В.А. Сопротивление материалов с элементами теории упругости и пластичности. Издательство АСВ 2004. С 159-165.

МОДЕЛИРОВАНИЕ КИНЕТИКИ РАДИАЦИОННОЙ СУШКИ ТОМАТНЫХ СЛАЙСОВ

Дяченко Э.П., Дяченко Н.П., Иванова М.И.

**Всероссийский научно-исследовательский институт овощеводства -
филиал Федерального государственного бюджетного научного
учреждения «Федеральный научный центр овощеводства»,
Московская обл., д. Верея, amed-nauka@yandex.ru**

На основе результатов экспериментально-аналитических исследований выполнено моделирование кинетики сушки томатных слайсов при объемном радиационном энергоподводе. Математически описана зависимость скорости радиационной сушки томатных слайсов от варьируемых факторов.

Based on the results of experimental and analytical studies, the kinetics of drying of tomato slices with a volumetric radiation energy supply was simulated. The dependence of the rate of radiation drying of tomato slices on variable factors is described mathematically.

Благодаря своим органолептическим и функциональным свойствам томат на сегодняшний день является одним из наиболее потребляемых овощных продуктов. Однако сезонность сырья и малые сроки хранения плодов томата при постоянной потребности в нем обуславливают необходимость развития традиционных и поиска новых способов консервации. Радиационная сушка в силу ряда своих преимуществ является сравнительно простым и одним из наиболее перспективных методов консервации томата [1, 2, 3, 5].

Между тем для анализа механизма переноса влаги внутри исследуемого материала, моделирования рациональных режимов обезвоживания, распределения температур в слое высушиваемого продукта и расчетов конструкции сушильного аппарата необходимо изучение кинетики сушки и её математическое описание.

Моделирование кинетики радиационной сушки томатных слайсов осуществлялось на основе эмпирических данных. В качестве исследуемых объектов использовались томаты сортов «Рио-Гранде», «Торпеда», «Каспар F1». В результате исследований установлены основные факторы, влияющие на интенсивность (удельную производительность процесса сушки): толщина нативного слайса h_n , = 1–4 мм; длина волны инфракрасных лучей $\lambda=1,41–1,67$ мкм и суммарная плотность падающего теплового потока с обеих сторон слайса $E_p = 1,84–2,98$ кВт/м². Оценка

качества сушеного томата проводилась согласно [4] по следующим критериям: внешний вид, цвет, вкус, запах, консистенция, массовая доля влаги.

На основе аппроксимации экспериментальных кривых обезвоживания расчетным методом получена функциональная зависимость скорости сушки исследуемого продукта от варьируемых факторов.

Выполнено математическое описание кривых сушки полиномиальной зависимостью вида [2]:

$$\tau(w) = a \cdot w^6 + b \cdot w^5 + c \cdot w^4 + e \cdot w^3 + f \cdot w^2 + g \cdot w + h, \quad (1)$$

где τ – время сушки, с; a, b, c, e, f, g, h – эмпирические коэффициенты.

Для получения зависимости эмпирических коэффициентов от толщины томатного слайса применена аппроксимация полиномом третьей степени, при этом увеличение показателя степени полинома не оказывало существенного влияния на величину достоверности аппроксимации R^2 :

$$\begin{aligned} \tau(w) = & (a_1 \cdot h^3 + b_1 \cdot h^2 + c_1 \cdot h + e_1) \cdot w^6 + (a_2 \cdot h^3 + b_2 \cdot h^2 + c_2 \cdot h + e_2) \cdot w^5 + \\ & + (a_3 \cdot h^3 + b_3 \cdot h^2 + c_3 \cdot h + e_3) \cdot w^4 + (a_4 \cdot h^3 + b_4 \cdot h^2 + c_4 \cdot h + e_4) \cdot w^3 + \quad (2) \\ & + (a_5 \cdot h^3 + b_5 \cdot h^2 + c_5 \cdot h + e_5) \cdot w^2 + (a_6 \cdot h^3 + b_6 \cdot h^2 + c_6 \cdot h + e_6) \cdot w + \\ & + (a_7 \cdot h^3 + b_7 \cdot h^2 + c_7 \cdot h + e_7). \end{aligned}$$

В результате дифференцирования выражения (2) по dw , получена зависимость вида:

$$\begin{aligned} \frac{d\tau}{dw} = & 6 \cdot (a_1 \cdot h^3 + b_1 \cdot h^2 + c_1 \cdot h + e_1) \cdot w^5 + 5 \cdot (a_2 \cdot h^3 + b_2 \cdot h^2 + c_2 \cdot h + e_2) \cdot w^4 + \\ & + 4 \cdot (a_3 \cdot h^3 + b_3 \cdot h^2 + c_3 \cdot h + e_3) \cdot w^3 + 3 \cdot (a_4 \cdot h^3 + b_4 \cdot h^2 + c_4 \cdot h + e_4) \cdot w^2 + \\ & + 2 \cdot (a_5 \cdot h^3 + b_5 \cdot h^2 + c_5 \cdot h + e_5) \cdot w + (a_6 \cdot h^3 + b_6 \cdot h^2 + c_6 \cdot h + e_6). \end{aligned} \quad (3)$$

Таким образом, функциональная зависимость для скорости сушки томатных слайсов (3) для определенных участков интенсивности теплового потока E_m , кВт/м² имеет следующий вид:

$$\frac{dw}{d\tau} = \left[\begin{aligned} & (a_1 \cdot h^3 + b_1 \cdot h^2 + c_1 \cdot h + e_1) \cdot w^5 + (a_2 \cdot h^3 + b_2 \cdot h^2 + c_2 \cdot h + e_2) \cdot w^4 + \dots \\ & \dots + (a_3 \cdot h^3 + b_3 \cdot h^2 + c_3 \cdot h + e_3) \cdot w^3 + (a_4 \cdot h^3 + b_4 \cdot h^2 + c_4 \cdot h + e_4) \cdot w^2 + \dots \\ & \dots + (a_5 \cdot h^3 + b_5 \cdot h^2 + c_5 \cdot h + e_5) \cdot w + (a_6 \cdot h^3 + b_6 \cdot h^2 + c_6 \cdot h + e_6) \end{aligned} \right]^{-1}. \quad (4)$$

Значения эмпирических коэффициентов ($a_1, b_1, c_1, e_1, a_2, b_2, c_2, e_2, a_3, b_3, c_3, e_3, a_4, b_4, c_4, e_4, a_5, b_5, c_5, e_5, a_6, b_6, c_6, e_6$) представлены в таблице 1.

Таблица 1 - Значения эмпирических коэффициентов (ЭК) скорости радиационной сушки томатных слайсов для различных диапазонов плотности теплового потока E_n

| Э К | Плотность теплового потока E_n , кВт/м ² | | |
|--------|---|---------------------------|---------------------------|
| | $E_n \leq 1,84$ | $1,84 \leq E_n \leq 2,70$ | $2,70 \leq E_n \leq 2,98$ |
| a_1 | -25539,818 | 134103,508 | -1107,693 |
| b_1 | 127150,38 | -910857,845 | 59753,302 |
| c_1 | - 421570,933 | 1608089,356 | -399306,280 |
| e_1 | - 346155,694 | -1205095,920 | 34665,918 |
| a_2 | 61635,011 | -319331,580 | 18927,927 |
| b_2 | - 317499,628 | 2144213,606 | -266729,191 |
| c_2 | 984846,64 | -3747613,454 | 1207776,037 |
| e_2 | 698633,38 | 2698386,049 | -330906,611 |
| a_3 | -61688,827 | 283472,197 | -31516,837 |
| b_3 | 336392,55 | -1887049,883 | 347849,645 |
| c_3 | - 935480,946 | 3285161,805 | -1296403,920 |
| e_3 | - 454121,845 | -2262889,173 | 521819,319 |
| a_4 | 31789,940 | -116538,752 | 18185,706 |
| b_4 | - 187269,308 | 773033,269 | -181210,647 |
| c_4 | 462996,41 | -1357638,068 | 601514,345 |

| | | | |
|-------|-----------------|-------------|-------------|
| e_4 | 69675,686 | 886902,341 | -302318,892 |
| a_5 | -8236,204 | 22096,242 | -3908,095 |
| b_5 | 53140,881 | -147076,447 | 37311,438 |
| c_5 | - 121791,385 | 265377,155 | -115220,671 |
| e_5 | 26914,202 | -163138,472 | 68463,583 |
| a_6 | 841,094 | -1522,393 | 225,347 |
| b_6 | -5977,920 | 10273,020 | -2185,399 |
| c_6 | 13408,403 | -19588,970 | 6425,860 |
| e_6 | -7563,081 | 11064,593 | -4785,865 |

На рисунке 1 представлены примеры полученных с использованием разработанной модели (4) кривых скорости радиационной сушки томатных слайсов в зависимости от концентрации сухих веществ в продукте. Сингулярные точки на кривых соответствуют переходу к удалению влаги с другим качественным и энергетическим состоянием связи, что подтверждается термодинамическим анализом [1].

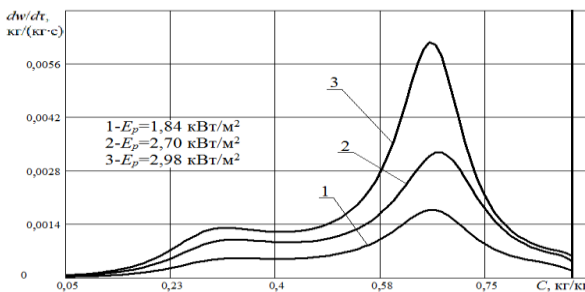


Рисунок 1 - Кривые скорости радиационной сушки томатных слайсов при $h_{нл} = 4 \text{ мм}$

Учитывая специфические свойства объекта сушки, минимальные значения толщины слайса и высокую интенсивность процесса, наиболее точное определение значений температуры в обезвоживаемом слое возможно расчетным путем [1]. Полученная модель кинетики радиационной сушки томатных слайсов (4) будет использована для анализа механизма переноса влаги внутри продукта, а также для расчета эволюции нестационарных температурных полей в слайсе в процессе сушки.

Литература:

1 Алексанян И. Ю., Буйнов А. А. Высокоинтенсивная сушка пищевых продуктов. Пеносушка. Теория. Практика. Моделирование. – Астрахань: Изд-во АГТУ, 2004. – 380 с.

2. Попов В. М., Афонькина В. А., Левинский В. Н. К вопросу об инфракрасной сушке томатов // Материалы LV международной научно-технической конференции «Достижения науки – агропромышленному производству» – Челябинск : ФГБОУ ВО Южно-Уральский ГАУ, 2016. – Ч. III. – 292 с.

3 Thin layer drying of tomato slices / Purkayastha M. Das, Mahanta C. L., Nath A., Deka B. C. // *Journal of Food Science and Technology* (c). – 2011. – Vol. 50, iss. 4. – P. 642–653. – doi: 10.1007/s13197-011-0397-x.

4 ОСТ 10 326-2003 «Томаты сушеные. Промышленное сырье. Технические условия».

5. Drying Characteristics of Tomato Slices and Mathematical Modeling / Sana Ben Mariem, Salah Ben Mabrouk // *International Journal of Energy Engineering* (c). – 2014. – Vol. 4(2A) – P. 17–24. – doi:10.5923/j.ijee.201401.03.

ПРОБЛЕМЫ ИСПОЛЬЗОВАНИЯ ЛОГИСТИЧЕСКОЙ РЕГРЕССИИ В ЗАДАЧЕ БИНАРНОЙ КЛАССИФИКАЦИИ И СПОСОБЫ ИХ РЕШЕНИЯ

А.А. Замащикова, В.Ю. Щеколдин
Новосибирский государственный технический университет,
г. Новосибирск, Zamashikova.N@mail.ru, raix@mail.ru

В данной работе предложен алгоритмический подход для предварительной подготовки данных для оценивания модели логистической регрессии. Предложены разнообразные спецификации моделей логистической регрессии с разными наборами действий по предварительной обработке данных. Произведен анализ построенных моделей, выявлены наиболее подходящие для описания изучаемого явления.

Algorithm for preliminary data preparation for estimating the logistic regression model has been proposed in the paper. Various specifications of logistic regression model have been proposed for different actions of data pre-processing. The analysis of the constructed models has been presented, and the most suitable models have been chosen for the description of data analysed.

Задача классификации является одной из самых распространённых задач, встречаемых на практике. Существуют параметрические и непараметрические подходы к её решению, при этом вторые являются на данный момент наиболее популярными [1]. Одной из причин такой распространённости непараметрических методов является то, что они предполагают наличие строгих требования к структуре и свойствам обрабатываемых данным. В то же время существует ряд задач, для которых параметрические методы оказываются предпочтительнее, например, в силу того, что разрабатываемые модели и результаты классификации, полученные на их основе, обладают простотой и удобством интерпретации, а также вследствие высокой вычислительной скорости таких методов. Поэтому исследование по подготовке данных для использования параметрических методов классификации является актуальным и будет рассмотрено в данной работе на примере решения задачи бинарной классификации с помощью модели логистической регрессии. Для её корректного использования входные данные должны удовлетворять следующим условиям: отсутствие мультиколлинеарности, отсутствие пропусков в данных и выбросов, независимость разных наблюдений, значимость входных признаков [2].

1. Предварительный отбор признаков

Первый этап заключается в использовании методов дисперсионного анализа для того, чтобы отсеять неинформативные признаки. Для этого значения каждого количественного показателя разбивались на две выборки в зависимости от класса объекта, к которому они принадлежат. Затем для каждой пары проверялась гипотеза о равенстве средних с помощью рангового критерия Манна-Уитни-Вилкоксона [3]. Для бинарных признаков использовался Z -критерий равенства долей [4].

2. Устранение мультиколлинеарности признаков

Для выявления и устранения мультиколлинеарности была использована мера, известная как фактор инфляции дисперсии [5], рассчитываемая как

$$VIF_j = 1/(1 - R_j^2),$$

где R_j^2 – коэффициент детерминации j -го признака относительно остальных, $j = 1, \dots, n$. Если хотя бы для одного из признаков этот коэффициент будет превышать 4, то делается вывод о наличии мультиколлинеарности. Вычислив факторы инфляции дисперсии для текущего набора признаков, удаляем тот, для которого этот коэффициент будет наибольшим, затем пересчитываем VIF для нового набора признаков. Продолжаем данную процедуру до тех пор, пока VIF для всех признаков не будет меньше 4.

3. Выявление и обработка аномальных значений признаков

Для поиска аномальных значений был применен подход, который часто используется при построении диаграмм размаха [6], т.е. к выбросам относились те значения признаков $x_{i,j}$, где $i = 1, 2, \dots, n$ – номер признака, $j = 1, 2, \dots, m$ – номер объекта, для которых не выполнялось одно из следующих условий:

$$x_{i,j} < Q_1 - k(Q_3 - Q_1), x_{i,j} > Q_3 + k(Q_3 - Q_1),$$

где Q_j - j -я квартиль соответствующей выборки. При этом параметр k выбирался таким образом, чтобы количество выбросов не превышало 20% от общего количества объектов в обучающей выборке. После этого значения признаков, которые были отнесены к выбросам, заменялись на минимальные и максимальные значения соответствующих им признаков без учета выбросов (проводилась так называемая винзоризация данных).

4. Преобразование признаков для обеспечения линейности модели регрессии

Одна из проблем построения логистической регрессии заключается в наличии признаков, которые находятся в нелинейной функциональной зависимости с откликом модели. Поэтому, чтобы уйти от нелинейности, было использовано преобразование Бокса-Кокса [8]:

$$x_i^\lambda = \begin{cases} (x_i + \alpha)^\lambda / \lambda, \lambda \neq 0, \\ \log(x_i + \alpha), \lambda = 0, \end{cases}$$

где параметр λ выбирается путем максимизации значений характеристики AUC (Area Under Curve) для модели классификации вида $y \sim x_i^\lambda$. Характеристика AUC представляет собой площадь под ROC-кривой [9], которая отображает зависимость числа верно классифицированных положительных примеров (S_e – чувствительности) от количества неверно классифицированных отрицательных примеров (S_p – специфичности).

5. Выбор наилучшей модели регрессии

Для построения корректной спецификации модели путем отбора наиболее информативных признаков были использованы два подхода: метод включения признаков (forward selection) и метод исключения признаков (backwards selection) [10]. Первый представляет собой подход, при котором сначала строится множество моделей от одного признака. Затем выбирается тот признак, который соответствует наибольшему значению статистики оптимизируемого критерия (в данной работе использовался AUC). После этого строится множество моделей от двух признаков путем добавления к наилучшей модели предыдущего этапа всех остальных возможных признаков, после чего также выбирается лучшая модель. Процедура продолжается до тех пор, пока добавление очередного признака не приведет к ухудшению AUC.

Второй подход аналогичен предыдущему, только отбор признаков начинается с полной модели, содержащей все исходные признаки, с последующим исключением признаков до тех пор, пока удаление очередного признака не приведет к ухудшению статистики оптимизируемого критерия.

Использование шагов обработки на реальных данных

Вышеописанные процедуры были использованы для решения задачи предсказания продления на будущий месяц контракта на рекламу фирмами, которые на данный момент являются рекламодателями. Обучающая выборка содержит количества фирм, которые продлили и не про-

длили контракт в соотношении 1.009, что свидетельствует о её сбалансированности. Это же соотношение для тестовой выборки составляет 1.20, что говорит о наличии умеренного дисбаланса. В таблице 1 представлены результаты классификации в зависимости от комбинаций реализованных процедур по подготовке данных, которые отмечены единицей. Номера шагов соответствуют номерам разделов с их описанием. Также представлены значения *AUC* для этих моделей для обучающей и тестовой выборки, а также по кросс-валидации с числом разбиений и повторов равным 5.

Предварительный отбор признаков, устранение мультиколлинеарности, выявление и обработка аномальных наблюдений, преобразование Бокса-Кокса позволяют значительно сократить число признаков, включаемых в модель, и повысить качество классификации. Метод включения обеспечивает построение модели с меньшим числом признаков и с наилучшим качеством классификации по сравнению с методом исключения, который, несмотря на это, также улучшает качество классификации.

Таблица 1. *AUC* logit-моделей с разными этапами подготовки данных

| № | 1 | 2 | 3 | 4 | 5 | <i>AUC</i> для обучающей выборки | <i>AUC</i> для тестовой выборки | <i>AUC</i> по кросс-валидации | Число признаков |
|----|---|---|---|---|----|----------------------------------|---------------------------------|-------------------------------|-----------------|
| 0 | 0 | 0 | 0 | 0 | 0 | 0.6979445 | 0.6772727 | 0.639951 | 61 |
| 1 | 1 | 0 | 0 | 0 | 0 | 0.6728262 | 0.6931818 | 0.6171985 | 47 |
| 2 | 0 | 1 | 0 | 0 | 0 | 0.6860451 | 0.6825758 | 0.6481239 | 31 |
| 3 | 1 | 1 | 0 | 0 | 0 | 0.6640288 | 0.6954545 | 0.6368242 | 22 |
| 4 | 0 | 1 | 1 | 0 | 0 | 0.6985338 | 0.6946970 | 0.6583107 | 31 |
| 5 | 1 | 1 | 1 | 0 | 0 | 0.6827566 | 0.7022727 | 0.6514242 | 22 |
| 6 | 0 | 1 | 1 | 1 | 0 | 0.6998794 | 0.6931818 | 0.6641383 | 31 |
| 7 | 1 | 1 | 1 | 1 | 0 | 0.6845869 | 0.7098485 | 0.6549344 | 22 |
| 10 | 0 | 1 | 1 | 1 | 1b | 0.6997168 | 0.6878788 | 0.6648463 | 27 |
| 11 | 1 | 1 | 1 | 1 | 1b | 0.6842285 | 0.7075758 | 0.6612101 | 20 |
| 12 | 0 | 1 | 1 | 1 | 1a | 0.6978194 | 0.6621210 | 0.6724543 | 20 |
| 13 | 1 | 1 | 1 | 1 | 1a | 0.6841011 | 0.6924242 | 0.6616623 | 17 |

В целом, в зависимости от целей исследования можно рекомендовать выбирать модели 12 и 13 как наиболее информативные.

Литература:

1. Yugal kumar, Sahoo G. Analysis of Parametric & Non Parametric Classifiers for Classification Technique using WEKA // I.J. Information Technology and Computer Science – 2012. – P. 43-49.

2. King G., Zeng L. Logistic Regression in Rare Events Data// Political Analysis 9. – 2001. – P. 137–163.
3. Hollander M., Wolfe D.A. Nonparametric Statistical Methods. 2nd edition // New York: Wiley, – 1999. – P.106-125.
4. Eberhardt K.R., Fligner M.A. A Comparison of Two Tests for Equality of Two Proportions // The American Statistician – 1977. – Vol. 31, No. 4, P. 151-155.
5. Kutner M. H., Nachtsheim C. J., Neter J. Applied Linear Regression Models (4th ed.) // McGraw-Hill Irwin . – 2004. – P. 408.
6. McGill R., Tukey J.W., Larsen W.A. Variations of Box Plots // The American Statistician. – 1978. – P. 12–16.
7. Box G. E. P., Cox D. R. . An analysis of transformations // Journal of the Royal Statistical Society, Series B. – 1964. – 26 (2): P. 211–252.
8. Fawcett T. ROC Graphs: Notes and Practical Considerations for Researchers // Kluwer, Academic Publishers, Netherlands. – 2004. – P 1-38.
9. Hanley J.A., McNeil B.J. The meaning and use of the area under of receiver operating characteristic (roc) curve. In Radiology. – 1982. – 143 (1): P. 29–36.
10. Hocking, R. R. The Analysis and Selection of Variables in Linear Regression // Biometrics. – 1976. – 32.: P 1-49

СРАВНЕНИЕ НЕКОТОРЫХ РОБАСТНЫХ МОДИФИКАЦИЙ КРИТЕРИЯ МАКСИМАЛЬНОГО ПРАВДОПОДОБИЯ В ЗАДАЧЕ ПАРАМЕТРИЧЕСКОЙ ИДЕНТИФИКАЦИИ НЕПРЕРЫВНО-ДИСКРЕТНЫХ СИСТЕМ

С. О. Кулабухова, В.М. Чубич

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

Предложены робастные модификации критерия максимального правдоподобия для оценивания неизвестных параметров моделей непрерывно-дискретных систем при наличии аномальных наблюдений в измерительных данных. Произведено сравнение их эффективности на примере модели продольного движения самолета.

Robust modifications of the maximum likelihood criterion for unknown parameters estimation of models of continuous-discrete systems, taking into account outliers

in the measurement data, are proposed. The comparison of their efficiency on the example of aircraft longitudinal motion model is made.

В настоящее время математическое моделирование является важной составляющей современной науки. Методы определения моделей по результатам экспериментальных исследований являются предметом теории идентификации. Для решения задачи параметрической идентификации часто используют метод максимального правдоподобия, который позволяет находить оценки, обладающие хорошими асимптотическими свойствами. Однако на практике измерительные данные могут содержать аномальные наблюдения, которые могут возникнуть по причине сбоя во время съема измерительных данных, а также искажения при их передаче. При этом применение классического метода максимального правдоподобия ведет к смещению или даже расхождению оценок параметров и, как следствие, к некорректному результату идентификации. В таком случае, целесообразным является использование различных робастных модификаций. Данная работа посвящена сравнению трех таких модификаций критерия максимального правдоподобия применительно к непрерывно-дискретным системам.

Рассмотрим динамическую систему, описывающуюся следующими уравнениями в пространстве состояний:

$$\begin{aligned} \frac{dx(t)}{dt} &= f[x(t), u(t), t; \theta], \quad t \in [t_0, t_N], \quad x(t_0) = x_0; \\ y(t_{k+1}) &= h[x(t_{k+1}), t_{k+1}; \theta] + v(t_{k+1}), \quad k = \overline{0, N-1}, \end{aligned}$$

где $x(t)$ – n -вектор состояния; $u(t)$ – r -вектор управления; $y(t_{k+1})$ – m -вектор измерения; $v(t_{k+1})$ – m -вектор шума измерений; θ – s -вектор подлежащих оцениванию неизвестных параметров.

Будем считать шум измерений белой гауссовской последовательностью, причем

$$E[v(t_{k+1})] = 0, \quad E[v(t_{k+1})v^T(t_{j+1})] = \begin{cases} R, & k = j, \\ 0, & k \neq j. \end{cases}$$

Задача оценивания неизвестных параметров с помощью метода максимального правдоподобия предполагает решение следующей задачи нелинейного программирования

$$\hat{\theta} = \arg \min_{\theta \in \Omega_\theta} \chi(\theta; Y_1^N) = \arg \min_{\theta \in \Omega_\theta} -\ln L(\theta; y(t_1), \dots, y(t_N)), \quad (1)$$

где

$$\chi(\theta; Y_1^N) = \frac{N}{2} \ln 2\pi + \frac{1}{2} \sum_{k=0}^{N-1} \varepsilon^T(t_{k+1}) B^{-1}(t_{k+1}) \varepsilon(t_{k+1}) + \frac{1}{2} \sum_{k=0}^{N-1} \ln \det B(t_{k+1}) \quad (2)$$

и

$$\varepsilon(t_{k+1}) = y(t_{k+1}) - h(x(t_{k+1}), t_{k+1}; \theta).$$

Учитывая возможность наличия аномальных наблюдений в измерительных данных, модифицируем вычисление ковариационной матрицы ошибки прогнозирования $B(t_{k+1})$ в критерии идентификации (2), взяв за основу публикации [1-3]. По результатам проведенной работы составим таблицу 1. Значения присутствующих в ней параметров определены в таблице 2.

Для сравнения эффективности разработанных модификаций, проведем параметрическую идентификацию модели продольного движения самолета [4, с. 216].

Используя начальные условия, входной сигнал и значения коэффициентов из [4, с. 217], положим $N = 100$, $t_{k+1} = 2(k+1)c$, $k = 0, 99$, $R = \text{diag}(0.5, 0.1, 1, 0.01)$. Примем в качестве истинных значений параметров $\theta_{true} = (0.0164, 0.21, 0, 0.2, 4.4, -1)^T$.

Таблица 1 – Вычисление $B(t_{k+1})$ для различных модификаций

| Критерий идентификации | Соотношение для $B(t_{k+1})$ |
|-------------------------------------|---|
| Классический ММП | $B(t_{k+1}) = R$ |
| Робастная модификация на основе [1] | $c_i = \frac{\alpha_i}{(2\pi)^{m/2} (\det B_i)^{1/2}} \exp\left(-\frac{1}{2} \varepsilon^T(t_{k+1}) B_i^{-1} \varepsilon(t_{k+1})\right),$ $\omega_i = \frac{c_i}{c_1 + c_2}, \quad i = 1, 2; \quad B(t_{k+1}) = \sum_{i=1}^2 \omega_i (B_i + \varepsilon(t_{k+1}) \varepsilon^T(t_{k+1}))$ |

| | |
|-------------------------------------|--|
| Робастная модификация на основе [2] | $\alpha_i^{k+1} = \frac{1}{2} + \alpha_i^k, \beta_i^{k+1} = \beta_i^k + \frac{1}{2} (\varepsilon_i(t_{k+1}))^2, i = \overline{1, m},$ $B(t_{k+1}) = \text{diag} \left(\frac{\beta_1^{k+1}}{\alpha_1^{k+1}}, \dots, \frac{\beta_m^{k+1}}{\alpha_m^{k+1}} \right)$ |
| Робастная модификация на основе [3] | $\hat{\mathbf{B}}(t_{k+1}) = \frac{1}{k+1} \sum_{i=0}^k \varepsilon(t_{i+1}) \varepsilon^T(t_{i+1}), \lambda_R = \frac{\text{tr} \hat{\mathbf{B}}(t_{k+1})}{\text{tr} \hat{\mathbf{K}}(t_k)},$ $\hat{\mathbf{K}}(t_{k+1}) = \lambda_R \hat{\mathbf{K}}(t_k), B(t_{k+1}) = \hat{\mathbf{K}}(t_{k+1})$ |

Сравнение предложенных модификаций будем производить по значениям ошибки оценивания, которую будем вычислять по формуле

$$\delta_\theta = \frac{1}{s_1} \sum_{i=1}^{s_1} \frac{|\hat{\theta}_i^{cp} - \theta_{i,ucm}|}{|\theta_{i,ucm}|} + \frac{1}{s_2} \sum_{j=1}^{s_2} |\hat{\theta}_j^{cp}|.$$

Здесь $\hat{\theta}^{cp}$ – усреднённая оценка неизвестных параметров по всем запускам, s_1 – число ненулевых и s_2 – число нулевых истинных значений (у нас $s_1 = 5, s_2 = 1$).

С целью уменьшения влияния результатов от случайного характера наблюдений, осуществим 100 независимых запусков системы в программной среде Matlab, смоделируем данные наблюдений со случайным расположением аномальных наблюдений. Примем 10% в качестве процента загрязнения и дисперсию аномальных наблюдений $R_A = 10000R$. Определим в соответствии с таблицей 2 параметры рассматриваемых робастных модификаций критерия максимального правдоподобия и осуществим параметрическую идентификацию модели продольного движения самолета, представив численные результаты в таблице 3.

Таблица 2 – Параметры робастных модификаций

| | | |
|--|---|---------------------------|
| Модификация на основе [1] | Модификация на основе [2] | Модификация на основе [3] |
| $\alpha_1 = 0.9, \alpha_2 = 0.1,$ $B_1 = R, B_2 = 10000R$ | $\alpha^0 = \beta^0 = (1, 1, \dots, 1)$ | $\hat{K}(t_0) = R$ |

Таблица 3 – Результаты параметрической идентификации

| | Классический ММП | Робастная модификация на основе [1] | Робастная модификация на основе [2] | Робастная модификация на основе [3] |
|------------------|------------------|-------------------------------------|-------------------------------------|-------------------------------------|
| $\hat{\theta}_1$ | 0.0164 | 0.0159 | 0.0169 | 0.0174 |
| $\hat{\theta}_2$ | 0.2292 | 0.2173 | 0.2154 | 0.2155 |
| $\hat{\theta}_3$ | 0.0008 | -0.0009 | 0.0003 | 0.0000 |
| $\hat{\theta}_4$ | 0.2280 | 0.2205 | 0.2127 | 0.2185 |
| $\hat{\theta}_5$ | 4.3025 | 4.3090 | 4.3127 | 4.3765 |
| $\hat{\theta}_6$ | -1.0949 | -0.9938 | -1.0060 | -1.0049 |
| δ_θ | 0.071 | 0.040 | 0.030 | 0.038 |

Анализируя полученные результаты (см. последнюю строку таблицы 3) можно заметить, что использование предложенных робастных модификаций критерия максимального правдоподобия реально позволяет повысить качество параметрической идентификации. Для модели продольного движения самолета лучшей оказалась модификация, разработанная на основе работы [2]: при ее использовании ошибка оценивания уменьшилась в 2.37 раза, в то время как для модификации из [1] это величина равнялась 1.78, а для модификации из [3] – 1.87.

Литература:

1. Plataniotis K.N., Androutsos D., Venetsanopoulos A.N. Nonlinear filtering of non-Gaussian noise // Journal of Intelligent and Robotic Systems. – 1997. – V. 19. – P. 207–231.

2. Särkkä S., Nummenmaa A. Recursive noise adaptive Kalman filtering by variational Bayesian approximations // IEEE Transactions on Automatic control. – 2009. – Vol. 54. – P. 596 – 600.
3. Jwo D.-J., Chung F.-C., Weng T.-P. Adaptive Kalman filter for navigation sensor fusion // Sensor Fusion and its Applications. In TechOpen, – 2010. – P. 66 – 90.
4. Crassidis J.L. J.L. Junkins. Optimal Estimation of Dynamic. – Washington, DC: Chapman&Hall/CRC, 2004. – 586 p.

ПРИМЕНЕНИЕ МАТЕМАТИЧЕСКИХ МОДЕЛЕЙ КРОСС-ДОКИНГА В ЗАДАЧАХ СКЛАДСКОЙ ЛОГИСТИКИ

Е.А. Мальцева, В.С. Карманов
Новосибирский Государственный Технический Университет,
г. Новосибирск,

Рассматриваются вопросы применения математических моделей кросс-докинга в задачах складской логистики, проведены исследования их эффективности. Разработан программный модуль для расчета вариантов схем кросс-докинга.

There are questions of the application of mathematical models of cross-docking in warehouse logistics problems, which are considered, and studies of their effectiveness are carried out. A software module has been developed for calculating variants of cross-docking schemes.

Цель работы

Применение математических моделей кросс-докинга в задачах складской логистики, создание программного модуля для расчета вариантов схем кросс-докинга.

Введение

Кросс-докинг представляет собой процесс приёма и отгрузки товаров и грузов через склад (кросс-док) напрямую, без необходимости размещения товара для хранения на складе [1-3].

В определенной очередности множество грузовиков отгружаются в пункте приема (поток отгрузки). Все товары, которые были получены, помещают во временное хранилище. Одновременно с ними загружаются по очереди грузовики для приема (поток погрузки). Все входящие грузовики загружаются товарами из временного хранилища, если же в

нем нет нужных товаров, то грузовику придется ждать до того, как эти товары появятся. При этом грузовики могут отгружаться/загружаться только последовательно. Задача состоит в том, чтобы подобрать такой порядок подъезда и отъезда машин, (и, при необходимости порядок погрузки/разгрузки товаров), чтобы суммарное время операций было минимальным.

Математическая постановка задачи

Для построения модели введем обозначения:

M – общее время работы (промежуток времени); R – число входящих (выгружаемых) грузовиков; S – число выходящих (загружаемых) грузовиков; N – число типов товаров; D – время замены машины (промежуток времени); V – время перемещения товаров от выгрузки до погрузки (промежуток времени); r_{ik} – число единиц товара k , погруженного во входящем грузовике i , $i=1..R$, $k=1..N$; s_{jk} – число единиц товара k , погружаемое в выходящий грузовик j , $j=1..S$, $k=1..N$; t_{ij} – число единиц товара k , перемещаемое из машины i в машину j , $i=1..R$, $j=1..S$; v_{ij} – равен 1, если $t_{ij} > 0$; равен 0 в противном случае; C_i – время отъезда входящей машины i (момент времени); C_j – время отъезда выходящей машины j (момент времени); T_{ik} – норматив времени на погрузку единицы товара k (промежуток времени); T_{ok} – норматив времени на выгрузку товара k (промежуток времени).

Перестановку входящих машин будем обозначать PR, выходящих – PS. Также введем обозначения для перестановки товаров – PNR и PNS, для входящих и выходящих грузовиков соответственно. Таким образом, целевую функцию при решении оптимизационной задачи можно записать в виде

$$M(PR, PNR, PS, PNS) \rightarrow \min.$$

Математическая модель для расчёта суммарного времени на все операции погрузки/разгрузки будет иметь следующий вид (внесены уточняющие коррективы в формулы, приведённые в работе [2]):

$$M = C_{[S]},$$

где:

$$C_{[i]} = \begin{cases} \sum_{k=1}^N r_{[i]k} * T_{ik}, & \text{если } i=1 \\ C_{[i-1]} + D + \sum_{k=1}^N r_{[i]k} * T_{ik}, & \text{если } 2 \leq i \leq R \end{cases}$$

$$C_{[j]} = \max(h_1, h_2),$$

$$h_1 = \max_{1 \leq i \leq R} \{v_{[i][j]} (C_{[i]} - \sum_{k=1}^N (r_{[i]k} * Ti_k) + \sum_{k=1}^N (t_{[i][j]} * Ti_k) + V + D)\},$$

$$h_2 = C_{[j-1]} + D + \sum_{k=1}^N (s_{[j]k} * To_k).$$

Обозначения индексов в скобках [i] и [j] соответствуют величинам для машин i и j, имеющих данные индексы в перестановке.

Варианты постановок задачи и методы решения

В зависимости от исходных данных и условий задачи, можно выделить следующие подтипы:

1. погрузка/разгрузка со свободным выбором товара (например, для грузовика с открытым кузовом);
2. погрузка/разгрузка с фиксированным выбором товара (например, для грузовика с закрытым кузовом).

В качестве методов решения задачи рассмотрим методы случайного поиска и полного перебора. Метод полного перебора практически реализуем только на задачах малой размерности, в этом случае его можно применить как эталон для оценки эффективности найденного решения методом случайного поиска.

Пример расчёта на модельных данных

Рассмотрим применение разработанной модели на модельных данных небольшой размерности.

Таблица 1 – Исходные данные

| N | R | S | (r _{ik}) | (s _{jk}) | Ti | To |
|---|---|---|--------------------|--------------------|-------|-------|
| 5 | 5 | 3 | 2 5 6 | 8 5 6 | 2 2 9 | 4 3 6 |
| | | | 3 10 7 | 6 6 6 | | |
| | | | 13 5 2 | 5 7 3 | | |
| | | | 5 2 3 | 8 2 4 | | |
| | | | 8 3 5 | 4 5 4 | | |

Таблица 2 – Результаты расчёта

| Метод | Время счёта, сек | m inM | minPR | minPS | minPN R | minPNS |
|-------------------------|------------------|---------|--------------|--------------|---------|--------|
| Полный перебор | 2.6087e+004 | 4 22 | 1 2 4 5 3 | 1 2 3 4 5 | 2 3 1 | 1 2 3 |
| Случайный поиск (30% от | 307.51 | 4 94 | 1 3 4 5 2 | 1 2 3 4 5 | 1 2 3 | 1 2 3 |

| | | | | | | |
|--|--------|---------|--------------|--------------|-------|-------|
| общего числа перестановок) | | | | | | |
| Случайный поиск (70% от общего числа перестановок) | 6467.7 | 4 22 | 1 2 4 5 3 | 1 2 3 4 5 | 2 3 1 | 1 2 3 |

Для проведения исследований разработан прототип программного комплекса, который позволяет рассчитывать варианты схем кросс-докинга. Оценка трудоёмкости и эффективности рассмотренных методов проиллюстрирована на рис. 1-2, результаты расчёта – на рис. 3.

Заключение

Рассмотрены вопросы применения математических моделей кросс-докинга для решения задач складской логистики. Разработан программный модуль для решения поставленной задачи, при реализации которого использовано свободное программное обеспечение Octave.

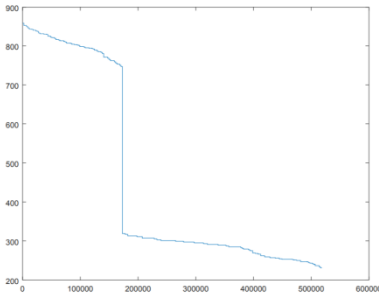


Рисунок 1 – Зависимость целевой функции от количества итераций при методе полного перебора

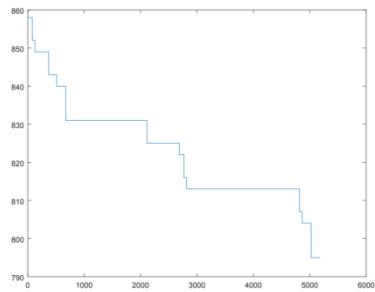


Рисунок 2 – Зависимость целевой функции от количества итераций при методе случайного поиска (30%)

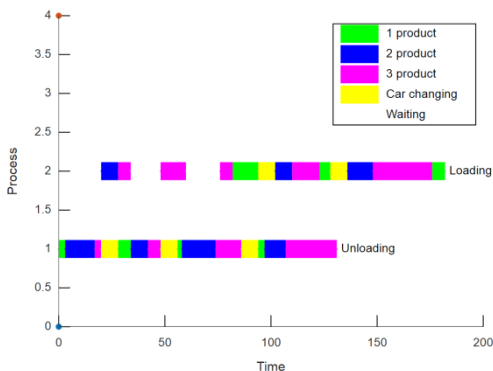


Рисунок 3 – Пример визуализации решения задачи кросс-докинга

Литература

1. Вельможин А.В., Гудков В.А., Миротин Л.Б., Куликов А.В. Грузовые автомобильные перевозки: учеб. для студентов вузов.– М.: Горячая линия-Телеком, 2006. – 559 с
2. Курейчик В.М., Рокотянский А.А. Решение задачи кросс-докинга на основе генетического алгоритма / Информатика, вычислительная техника и инженерное образование. – 2012. – № 2 (9)
3. Пензев В.Н. Кросс-Докинг [Электронный ресурс] / URL: <http://www.lscm.ru/index.php/ru/avtoram/item/1180/>

МАТЕМАТИЧЕСКОЕ ИССЛЕДОВАНИЕ ГЕОМЕТРИЧЕСКОЙ ТОЧНОСТИ ПОКОВКИ ПРИ ШТАМПОВКЕ ДЛЯ ОБЕСПЕЧЕНИЯ ПОСЛЕДУЮЩЕЙ БАЛАНСИРОВКИ КОЛЕНЧАТОГО ВАЛА*

А.В. Мартюгин
Научно-технический центр ПАО «КАМАЗ»,
г. Набережные Челны, avmartugin@gmail.com

В статье описан способ математического исследования геометрической точности поковки коленчатого вала и ограничения исходного дисбаланса, обеспечивающие успешную финишную балансировку вала.

The article describes a mathematical study of the geometric accuracy of the crankshaft forgings and the limitations of the original imbalance, ensuring successful final balancing of the shaft.

Кузнечный завод ПАО «КАМАЗ», который имеет в своем оснащении автоматическую линию «EUMUCO» (12000 т.с.) производства поковок коленчатых валов, не обладая дорогостоящим высокотехнологичным оборудованием для предварительной динамической балансировки, выступил как альтернативный поставщик заготовок коленчатых валов для производства двигателей отечественного автопроизводителя.

Требования к поковкам коленчатых валов нового поколения весьма жесткие, для них требуется обеспечить важную специальную характеристику – ограниченный дисбаланс поковки коленчатого вала. Задача усложняется тем, что балансировка коленчатого вала должна быть обеспечена без механической обработки противовесов. Ранее проблема дисбаланса решалась технологически, за счет механической обработки всего коленчатого вала, в том числе и противовесов.

В рамках данной задачи Кузнечный завод осуществил подготовку производства поковок коленчатых валов Р4 для нового современного производства средних рядных двигателей. Отличительной особенностью изготовления поковок коленчатых валов было введение в технологический поток первых операций механической обработки (торцовка и центровка), так как требовалось решить задачу локализации автокомпонента, обеспечивая полное соответствие и идентичность ранее поставляемого зарубежного аналога.

При осуществлении работ по подготовке производства выявилась проблема избыточного дисбаланса центрованных поковок и большого разброса дисбаланса в партии поковок, что приводило к значительным потерям и ставило производство таких валов на нижнюю границу рентабельности. Указанная проблема решалась при глубоком исследовании ключевых параметров штамповки и механической обработки, напрямую влияющих на качество балансировки коленчатого вала. Для проведения анализа использовался принцип «сквозного анализа технологии изготовления детали» [1].

Партии валов были измерены по различным геометрическим параметрам (величина изгиба и его направление, величинам недоштамповки и смещения, по диаметрам шатунных и коренных шеек и т.д.) и, в зависимости от влияния этих параметров на дисбаланс поковки коленчатого

*Работа выполнена под руководством профессора, д.т.н. И.М. Володина

вала Р4 (рядный четырехцилиндровый двигатель), отобраны специальные характеристики поковки и сформированы критические контрольные параметры штамповки для обеспечения балансируемости конечной детали [2, 3]. Однако, для разработки технологических ограничений отобранных параметров необходимо было набрать большой объём данных по дисбалансу для различных сочетаний параметров, что трудоёмко и затратно.

Было принято решение заменить затратный и недостаточно гибкий сбор статистики математическим моделированием влияния геометрических параметров поковки на балансируемость конечной детали. Для этого была создана 3D модель коленчатого вала в среде Siemens NX (лицензия), в которой при изменении любого, ключевого для балансируемости, параметра модель заново перестраивается. При перестроении промежуточно формируется 3D модель поковки с заданными параметрами, затем из нее виртуально «удаляется» часть материала в соответствии с чертежом детали (см. рис.1), при этом учитываются технологические особенности механической обработки.

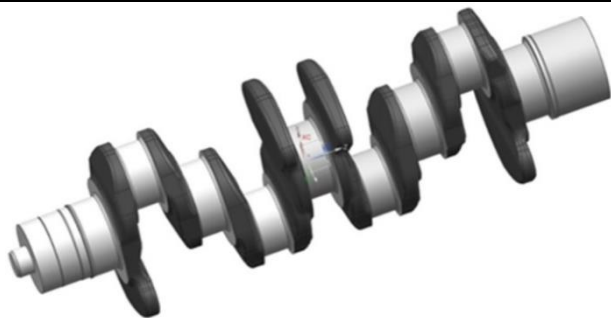


Рисунок 1 – Виртуально «обработанная» 3D модель поковки коленчатого вала Р4. Тёмно-серым цветом показаны необрабатываемые поверхности, светло-серым – обрабатываемые поверхности.

Ввод параметров поковки в среде Siemens NX был автоматизирован встроенным инструментом NXOpen, вывод параметров поковки и виртуально измеренных параметров полученной детали осуществлен в БД SQLite, анализ и вывод полученных данных осуществлен с помощью языка программирования R (см. рис.2).



Рисунок 2 – Общая схема взаимодействия компонентов программы технологического обеспечения балансировки деталей с необрабатываемыми поверхностями.

По результатам проделанной работы были приняты технологические решения: ужесточены, относительно чертежа поковки, допуски недоштамповки и изогнутости поковки; внесены изменения в штамповую оснастку для ограничения смещения по разъёму штампов, предупреждения значительного коробления полуфабриката при обрезке облоя, оптимизирована конфигурация формообразующих поверхностей штампа для горячей правки поковок.

Вывод: успешно решена сложная техническая задача обеспечения балансировки современного коленчатого вала грузового автомобиля с необрабатываемыми противовесами из поковки, получаемой горячей объёмной штамповкой. Достигнута возможность ограничения дисбаланса поковок с помощью контроля ключевых контрольных параметров штамповки и механической обработки, включая корректировку центровки по фактическим параметрам дисбаланса центрованных поковок. Указанное решение обеспечивает требования автопроизводителя к специальной характеристике автокомпонента – балансируемость коленчатого вала при механической обработке, а также использование риск-ориентированного подхода на стадиях проектирования и выпуска [4] поковок коленчатых валов средних рядных двигателей. Использование принципа «сквозного анализа технологии изготовления детали» при проектировании технологических процессов и штамповой оснастки позволяет получать поковки коленчатого вала повышенной геометрической точности, с оптимальными

параметрами для последующей балансировки при механической обработке [5].

Литература

1. Володин И. М. Система основных принципов проектирования процессов горячей объёмной штамповки и созданные на ее основе технологии / И. М. Володин, А. А. Ромашов // Кузнечно-штамповочное производство. Обработка материалов давлением. – 2008.– №9. – С. 19-25.
2. Мартюгин А. В. Технологическое обеспечение балансировки коленчатых валов большегрузных автомобилей / А. В. Мартюгин // Материалы Международной научно-практической конференции «Фундаментальные проблемы технических наук», 19 февраля 2014г. / научный центр «АЕ-ТЕРНА». – Уфа, 2014. – С. 97-100.
3. Программа для анализа балансируемости коленчатого вала // Патент России № 2014660945, 20.10.14. / А. В. Мартюгин
4. Касьянов С. В., Биктимирова Г. Ф. Технологический переход как ключевой процесс управления качеством продукции в соответствии с ИСО/ТС 16949:2009 // Автомобильная промышленность, №3, 2014. – С. 27-29.
5. I.V. Volodin, A.V. Martyugin. A new resource-saving method of forming crankshafts heavy vehicles / I.V. Volodin, A.V. Martyugin // Information materials are prepared for involvement of the Ministry of Education and Science of the Russian Federation in the Hannover Messe 2014. – Ганновер, ФРГ, 2014. – С. 16-18.

ОЦЕНКА ИЗМЕНЕНИЯ СОСТОЯНИЯ РАСТИТЕЛЬНОГО ПОКРОВА НЕФТЕДОБЫВАЮЩИХ ТЕРРИТОРИЙ С ИСПОЛЬЗОВАНИЕМ ДАННЫХ ДИСТАНЦИОННОГО ЗОНДИРОВАНИЯ ЗЕМЛИ

К.А. Маслов, О.С. Токарева
Томский политехнический университет,
г. Томск, kam20@tpu.ru

Целью данной работы является проведение оценки изменения состояния растительного покрова на территории Васюганского лесничества Каргасокского района Томской области за период с 2013 г. по 2018 г. С использованием данных со спутника Landsat 8 и спектральных индексов NDVI и dNDVI была

построена карта изменения состояния растительного покрова. Выполнен анализ факторов воздействия на состояние растительного покрова исследуемой территории.

The purpose of this research is to estimate the vegetation state changes on the territory of the Vasyugan forestry in the Kargasok district of the Tomsk region for the period from 2013 to 2018. Using Landsat 8 data and spectral indices NDVI and dNDVI, the map of change of vegetation cover was constructed. Then a factor analysis was performed.

Растительный покров территорий нефтегазодобычи Западной Сибири подвержен негативному влиянию механических повреждений вследствие строительства объектов инфраструктуры нефтегазовой отрасли, аварийных разливов нефти и загрязнения атмосферы в результате сжигания попутного газа [1].

В данной работе проведена оценка изменения состояния растительного покрова для участков территории Васюганского лесничества Каргасокского района Томской области (рис. 1). Исследуемая территория характеризуется слабой устойчивостью экосистем к внешним факторам, ландшафт территории определяется взаимопроникновением лесных и болотных выделов, преобладают темнохвойно-мелколиственные леса, основные леса и верховые болота [2].



Рисунок 1 – Расположение исследуемой территории: граница Васюганского лесничества (1), граница Каргасокского района (2), территория Томской области (3)

В настоящее время на территории Васюганского лесничества функционирует несколько нефтяных и газовых месторождений: Первомайское, Лонтыньяхское, Катильгинское, Западно-Катильгинское, Оленье, Столбовое, Южно-Черемшанское, Поселковое. Анализ состояния растительного покрова проведен в зонах воздействия месторождений и за

пределами этих зон. В качестве зон воздействия месторождений были выбраны буферные зоны радиусом 1.5 км вокруг объектов на их территории (кустовые площадки, продуктопроводы и т.д.).

Состояние растительного покрова оценивалось на основе значений спектрального индекса NDVI, рассчитываемого с использованием данных дистанционного зондирования Земли из космоса по формуле:

$$NDVI = \frac{NIR - RED}{NIR + RED},$$

где NIR – отражение в ближней инфракрасной области спектра, RED – отражение в красной области спектра.

Изменение состояния растительного покрова оценивалось с помощью значений индекса dNDVI:




$$dNDVI = NDVI_{post} - NDVI_{pre},$$

где $NDVI_{post}$ – значение NDVI для снимка более поздней даты, $NDVI_{pre}$ – значение NDVI для снимка более ранней даты.

Индексы NDVI и dNDVI рассчитывались по данным мультиспектральных космических снимков со спутника Landsat 8 уровня обработки L2, загруженных с сайта Геологической Службы США (USGS), за 2013 г. и 2018 г. Коллекция уровня обработки L2 предлагает готовые для проведения исследования данные, прошедшие атмосферную коррекцию, что позволяет сравнивать состояние растительного покрова по космическим снимкам с разными датами съемки. Анализ проводился с использованием инструментария геоинформационной системы QGIS.

Полученная карта dNDVI позволила провести оценку изменения состояния растительного покрова и выявить участки наиболее губительного воздействия антропогенных и природных факторов. Значения индекса dNDVI были классифицированы по степени воздействия на растительный покров территории (табл. 1). Фрагменты итоговой карты классификации, с обозначенными в соответствии с табл. 1 классами, приведена на рис. 2.

Таблица 1 – Классы значений dNDVI

| Класс | Значения dNDVI | Цвет на карте |
|------------------------|-----------------|---|
| Значительное ухудшение | $(-\infty; 4]$ |  |
| Серьезное ухудшение | $(-0.4; -0.25]$ |  |
| Умеренное ухудшение | $(-0.25; -0.1]$ |  |

| | | | |
|--------------------------|-------------|--|--|
| Незначительные изменения | (-0.1; 0.1] | | |
| Умеренное улучшение | (0.1; 0.25] | | |
| Серьезное улучшение | (0.25; 0.4] | | |
| Значительное улучшение | (0.4; +∞) | | |

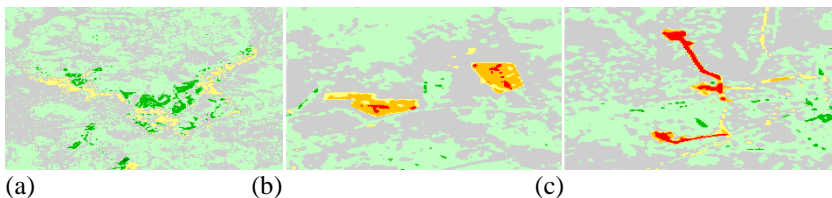


Рисунок 2 – Фрагменты карты dNDVI: для участка территории с гарью (а); для вырубок (b); для новых построенных объектов инфраструктуры (с)

В результате проведенного анализа было выявлено, что в зонах воздействия месторождений ухудшение состояния растительного покрова связано преимущественно с постройкой новых объектов инфраструктуры и рубками леса. Вне зон воздействия объектов нефтегазодобычи основным негативным фактором воздействия являются лесные пожары. Значительные ухудшения растительного покрова вызваны в большей степени постройкой новых объектов инфраструктуры.

Полученные в данной работе результаты могут быть использованы для прогнозов дальнейшего изменения состояния растительного покрова на исследуемой территории [3] и планировании мероприятий по восстановлению растительного покрова на нарушенных участках и использованию лесных ресурсов.

Литература:

1. Using MODIS NDVI products for vegetation state monitoring on the oil production territory in Western Siberia / Kovalev A., Tokareva O. // MATEC Web of Conferences. – 2016. – V. 48. – № 05003. – P. 1–4. URL: <https://doi.org/10.1051/matecconf/20164805003> (дата обращения: 07.11.2018)

2. Методика оценки воздействия техногенного химического загрязнения атмосферы на лесоболотные комплексы в нефтедобывающих

районах Западной Сибири / Полищук Ю.М., Токарева О.С. // Химия в интересах устойчивого развития. – 2002. – Т. 10. – № 5. – С. 659-668.

3. Аэрокосмический мониторинг лесного покрова / Седых В.Н. – Новосибирск: Наука.Сиб.отд-ние, 1991. – 239 с.

МОДУЛЬ АВТОКАЛИБРОВКИ ПРОГРАММНОГО КОМПЛЕКСА ДЛЯ БЕСКОНТАКТНОГО ИЗМЕРЕНИЯ 3D ГЕОМЕТРИИ СЛОЖНОПРОФИЛЬНЫХ ОБЪЕКТОВ

Н.Н. Овчинников, С.В. Двойнишников

**Институт теплофизики им. С. С. Кутателадзе СО РАН,
г. Новосибирск, iam@kolsha.ru**

В работе описан модуль автокалибровки программного комплекса для бесконтактного измерения 3D геометрии сложнопрофильных объектов методом оптической фазовой триангуляции. Модуль реализует алгоритм калибровки, основанный на построении гомографии между модельной плоскостью и ее изображением в приближении классической модели камеры обзора.

The paper describes the auto-calibration module of a software package for non-contact measurement of 3D geometry of complex objects by optical phase triangulation. The module implements a calibration algorithm based on the construction of a homography between the model plane and its image in the approximation of the classical pinhole camera model.

Трехмерные измерения имеют широкий спектр применения – от медицины, для проектирования протезов, до промышленности, для контроля качества производимого оборудования [1]. Со снижением стоимости камер расширяется сфера применения компьютерного зрения. В связи с этим, перспективно разрабатывать программное обеспечение для измерения трехмерной геометрии объектов, используя бытовые камеры.

В ИТ СО РАН разрабатывается программно-аппаратный комплекс для измерения трехмерной геометрии сложнопрофильных объектов с использованием двух цифровых камер, представляющих собой стерео-

пару, и проектора для формирования внешнего освещения [2]. На сегодняшний день реализован модуль поиска сопряженных точек на изображениях стереопары и модуль реконструкции 3D точек, который калибруется вручную. Поскольку калибровка проводится в ручном режиме, невозможно точно определить внутренние и внешние параметры камер и использовать их для реконструкции. Кроме того, ручная калибровка достаточно трудоемка по времени и требует от пользователя дополнительных знаний о работе системы в целом.

Для решения этой проблемы необходимо разработать модуль автоматической калибровки стереопары на основе набора снимков калибровочного шаблона. Разработанный модуль будет применен в программном комплексе для измерения трехмерной геометрии сложнопрофильных объектов.

Реализация калибровки выполнена методом, предложенным Z. Zhang [3]. Метод позволяет провести калибровку одной камеры без использования специальной аппаратуры, получить её внутренние и внешние параметры, а также радиальные и тангенциальные коэффициенты искажения изображения.

В реализованном методе калибровки используется классическая модель камеры обскура. Взаимосвязь между 3D точкой M и её проекцией на изображении m задается формулой:

$$sm = A[\mathbf{R} \quad \mathbf{t}]\mathbf{M} \quad (1)$$

где s произвольный фактор масштаба, (\mathbf{R}, \mathbf{t}) – это поворот и перенос, которые связывают пространственную систему координат с системой координат камеры и A матрица внутренних параметров камеры. Далее строится гомография между модельной плоскостью и её изображением, а после строятся два ограничения на внутренние параметры камеры. Затем эти ограничения переписываются в два однородных уравнения, решение которых уточняется с помощью метода максимального правдоподобия.

Калибровка стереопары реализована с помощью следующего алгоритма: сначала каждая камера калибруется обособленно по одному и

тому же шаблону, затем рассчитывается их смещение относительно друг друга.

В настоящее время разработан модуль калибровки на основе вышеописанного алгоритма и интегрирован с модулем реконструкции точек. Для проверки модуля проведен эксперимент. В качестве калибровочного шаблона использован паттерн шахматной доски, снятый обеими камерами с разных ракурсов. С полученными результатами проведена реконструкция трехмерной геометрии. Измерен трехмерный профиль плоского объекта с размерами 10 x 10 см. Установлено, что восстановленная трехмерная геометрия объекта не имеет пространственных искажений. Полученные результаты подтверждают корректность работы реализованной процедуры калибровки.

Таким образом, в работе проведен анализ существующих методов калибровки цифровых камер. Созданный модуль автокалибровки, реализует метод калибровки, основанный на построении гомографии между модельной плоскостью изображением в приближении классической модели камеры обскура. Модуль интегрирован в программный комплекс для измерения геометрии сложнопрофильных объектов.

Литература:

1. Меледин В.Г. Оптоэлектронные информационные системы для науки и промышленности // Интерэкспо Гео-Сибирь. 2014. Т. 5. № 1. С. 3-12.
2. Двойнишников С.В., Аникин Ю.А., Кабардин И.К., Куликов Д.В., Меледин В.Г. Оптоэлектронный метод бесконтактного измерения профиля поверхности крупногабаритных объектов сложной формы // Измерительная техника №1, 2016. – С. 17-22.
3. Z. Zhang. A Flexible New Technique for Camera Calibration. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(11):1330-1334, 2000. – 21 с.

ПРОВЕРКА СЛОЖНОЙ ГИПОТЕЗЫ О ВИДЕ ВИНЕРОВСКОЙ ДЕГРАДАЦИОННОЙ МОДЕЛИ

Е.А. Осинцева, Е.В. Чимитова

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

В данной работе рассматривается винеровская деградационная модель с линейной, степенной и экспоненциальной функцией тренда. Неизвестные параметры модели оцениваются методом максимального правдоподобия. Реализован алгоритм проверки сложной гипотезы о согласии опытного распределения с винеровской деградационной моделью. Проведено исследование сходимости распределения статистик Колмогорова, Крамера-Мизеса-Смирнова и Андерсона-Дарлинга к предельным распределениям соответствующих критериев. По результатам исследования сформулированы рекомендации о применении данных критериев.

In this paper, the Wiener Degradation Model with linear, power and exponential trend functions was considered. The unknown parameters are estimated by maximum likelihood method. The algorithm of goodness-of-fit testing was implemented. The investigation of the distributions of Kolmogorov, Cramer-von Mises-Smirnov and Anderson-Darling test statistics was carried out. As a result, the recommendations for application of goodness-of-fit tests were formulated.

Введение

Рост популярности деградационных моделей обусловлен повышенным вниманием к надежности изделий. Преимущество винеровской деградационной модели заключается в том, что ее можно использовать для деградационных данных с неположительными приращениями.

Проверка гипотезы о согласии опытного распределения с теоретическим законом на практике является важным этапом, поскольку неправильно подобранная модель может привести к неверным статистическим выводам.

Несмотря на то, что критериев для проверки статистических гипотез о согласии достаточно много, опыт ученых показывает, что данный процесс имеет ряд проблем. В первом случае, выбранный критерий может иметь небольшую мощность [1]. Во втором случае, сложность для использования критерия представляет отсутствие предельного распределения [2]. В третьем случае, на распределение статистики при проверке сложной гипотезы влияет ряд факторов: вид наблюдаемого закона, тип оцениваемого параметра и число оцениваемых параметров, используемый метод оценивания и другие [3].

В данной работе будут рассмотрены три наиболее популярных критерия: Колмогорова, Крамера-Мизеса-Смирнова и Андерсона-Дарлинга. Это связано с тем, что в основе винеровской деградиционной модели лежит нормальное распределение, для которого получены предельные распределения статистик [3]. В данном контексте возникает вопрос: можно ли воспользоваться предельным распределением статистик вышеперечисленных критериев для проверки гипотезы о согласии данных с винеровской деградиционной моделью?

Таким образом, целью данной работой является исследование распределения статистик Колмогорова, Крамера-Мизеса-Смирнова и Андерсона-Дарлинга для винеровской деградиционной модели с линейной, степенной и экспоненциальной функцией тренда и формулирование рекомендаций для дальнейшего их использования.

1. Описание винеровской деградиционной модели

Случайный процесс $Z(t)$, характеризующий процесс деградации исследуемых изделий, называется деградиционным винеровским процессом, если данный процесс центрирован, с независимыми приращениями $\Delta Z(t) = Z(t + \Delta t) - Z(t)$ и данные приращения подчиняются нормальному распределению с функции плотности:

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

где параметр сдвига $\theta_1 = \mu(\rho(t + \Delta t) - \rho(t))$, а параметр масштаба $\theta_2 = \sigma \sqrt{\rho(t + \Delta t) - \rho(t)}$, $\sigma > 0$, $\rho(t)$ – некоторая возрастающая функция: $M(Z(t)) = \mu \cdot \rho(t)$.

Обозначим выборку приращений деградиционного показателя, полученную в результате наблюдения деградиционных процессов n объектов следующим образом:

$$\mathbf{X}_n = \{(\Delta z_{1j}, \Delta z_{2j}, \dots, \Delta z_{nj}), j = \overline{1, M}\},$$

где M – количество измерений деградиционного показателя для одного объекта, Δz_{ij} – приращение деградиционного показателя от t_{j-1} до t_j .

В данной работе рассмотрим $\rho(t)$ в качестве линейной, степенной и экспоненциальной функции тренда: $\rho(t) = t$, $\rho(t) = t^\gamma$, $\rho(t) = e^{\gamma t}$.

Неизвестные параметры модели оцениваются методом максимального правдоподобия [4].

2. Проверка сложной гипотезы о виде деградационной модели

Определим верную гипотезу H_0 в следующем виде:

$$H_0 : F_{X_n}(t) \in \{F_{Norm}(t; \sigma, \mu, \gamma), \sigma, \mu, \gamma \in \Omega\}.$$

К выборке приращений X_n классические критерии согласия неприемлемы, поскольку элементы данной выборки в общем случае не являются одинаково распределенными. Для этого необходимо ввести следующее преобразование приращений деградационного показателя:

$$U_j^i = F_{Norm}(\Delta Z_{ij}; \mathfrak{E}, \mathfrak{E}, \mathfrak{E}, \mathfrak{E}), \quad i = \overline{1, n}, \quad j = \overline{1, k}.$$

При справедливости гипотезы H_0 :

$$U_j^i \square \text{Uniform}(0, 1), \quad i = \overline{1, n}, \quad j = \overline{1, k_i}.$$

Таким образом, задача проверки гипотезы H_0 сводится к проверке гипотезы о равномерном распределении случайных величин U_j^i , $i = \overline{1, n}$, $j = \overline{1, k_i}$.

Обозначим через $U_{(1)}^* \leq U_{(2)}^* \leq \dots \leq U_{(M)}^*$, $M = \sum_{i=1}^n k_i$ элементы вариационного ряда, построенного по полной выборке:

$$U_M = \{U_j^i, \quad i = \overline{1, n}, \quad j = \overline{1, k_i}\}.$$

Статистика критерия Колмогорова с поправкой Большева:

$$S_K = \frac{6nD_M + 1}{6\sqrt{M}},$$

где

$$D_M = \max(D_M^+, D_M^-), \quad D_n^+ = \max_{1 \leq i \leq M} \left\{ \frac{i}{M} - U_{(i)}^* \right\},$$

$$D_n^- = \max_{1 \leq i \leq M} \left\{ U_{(i)}^* - \frac{i-1}{M} \right\}.$$

Статистика критерия Крамера-Мизеса-Смирнова:

$$S_{\omega} = M \omega_n^2 = \frac{1}{12M} + \sum_{i=1}^M \left\{ U_{(i)}^* - \frac{2i-1}{2M} \right\}^2.$$

Статистика критерия Андерсона-Дарлинга:

$$S_{\omega} = M \Omega_n^2 = -M - 2 \sum_{i=1}^M \left\{ \frac{2i-1}{2M} \ln(U_{(i)}^*) + \left(1 - \frac{2i-1}{2M} \right) \ln(1 - U_{(i)}^*) \right\}.$$

Аппроксимации предельных распределений рассматриваемых статистик при проверке сложной гипотезы о нормальном распределении и оценивании параметров методом максимального правдоподобия представлены в [3].

В данной работе методами компьютерного моделирования проведено исследование распределений статистик при проверке гипотезы о виде винеровской деградиационной модели при различных функциях тренда и сетках по времени. Показано, что на равномерной сетке в случае линейной функции тренда, распределения статистик вышеперечисленных критериев сходятся к соответствующим аппроксимациям предельных распределений. Однако на неравномерной сетке и нелинейных функциях тренда распределения статистик уже отличаются от соответствующих аппроксимаций, полученных в [3].

Заключение

В результате исследований было показано, что в общем случае при проверке сложной гипотезы о виде винеровской деградиационной модели возможно использование критериев согласия типа Колмогорова, Крамера-Мизеса-Смирнова и Андерсона-Дарлинга необходимо моделировать условные распределения статистик в интерактивном режиме. Вместе с тем, при использовании линейной функции тренда оправдано использование аппроксимаций предельных распределений статистик вышеперечисленных критериев.

Литература

1. Лемешко Б. Ю. Критерии проверки отклонения распределения от равномерного закона. Руководство по применению. / Б. Ю. Лемешко, П.Ю. Блинов – М.: НИЦ ИНФРА-М, 2015. – 182 с. DOI: 10.12737/11304
2. Лемешко Б.Ю. О применении и мощности непараметрических критериев согласия Купера, Ватсона и Жанга / Б.Ю. Лемешко, А.А. Горбунова // Измерительная техника. 2013, № 5. – С. 3-9. 4. Лемешко Непараметрические критерии согласия.

3. Лемешко Б. Ю. Непараметрические критерии согласия: Руководство по применению: Монография / Б. Ю. Лемешко. – М.: НИЦ ИНФРА-М, 2014. – 163 с. DOI: 10.12737/11873
4. Осинцева Е.А. Информационная матрица Фишера для параметров винеровской деградационной модели / Е.А. Осинцева, Чимитова Е.В. // Обработка информации и математическое моделирование: Материалы российской научно-технической конференции. – Новосибирск: СибГУТИ, 2017. – Т. 1 – С. 100.

ЭФФЕКТИВНОСТЬ ПРИМЕНЕНИЯ РЕПИТЕРА В СЕТИ LTE ДЛЯ УСТРОЙСТВ IoT

Переверзина Т.В.

**Сибирский государственный университет
телекоммуникаций и информатики
г. Новосибирск, pereverzina.tatna@mail.ru**

Репитеры могут являться одним из важных элементов сотовой сети связи, обеспечивая расширение зоны радиопокрытия в зданиях, тоннелях, подземных автостоянках и местах с низким уровнем сигнала. В данной работе рассмотрена эффективность использования репитеров в сети сотовой связи LTE для устройств IoT (Internet of things – Интернет вещей).

Repeaters can be one of the important elements of a cellular communication network, providing an extension of radio coverage in buildings, tunnels, underground parking lots and low-signal areas. This paper examines the effectiveness of using repeaters in an LTE cellular network for IoT devices (Internet of things).

По состоянию на май 2018 года в мире насчитывалось 48 коммерческих мобильных сетей Интернет вещей и большинство таких сетей развернуто на базе протоколов LoRaWAN, SigFox или NB-IoT в Западной Европе, США и в развитых странах Азиатско-Тихоокеанского региона. Применительно к России, в ближайшие годы количество межмашинных подключений (Machine-to-Machine, M2M) будет обеспечиваться главным образом за счет развернутых уже сотовых сетей 4G, в силу имеющихся площадей покрытия таких сетей, а также благодаря тому, что часть поставляемых на рынок M2M-устройств поддерживает стандарт LTE.

скоростью не более 300 кбит/с, может быть организована до 50 км, при использовании репитера радиусом действия около 1 км (рис.2).

На рисунке 2 показан результат работы приложения на C#, реализованного специально для расчета по спроектированной модели, на экране отображены все параметры оборудования и коэффициенты усиления, используемые при расчете, и по полученным данным спроектирован график. На графике изображены две кривые чувствительности, при пересечении которых получаем максимальную дальность связи, учитывая рельеф местности, постройки зданий, препятствий и ограждений.

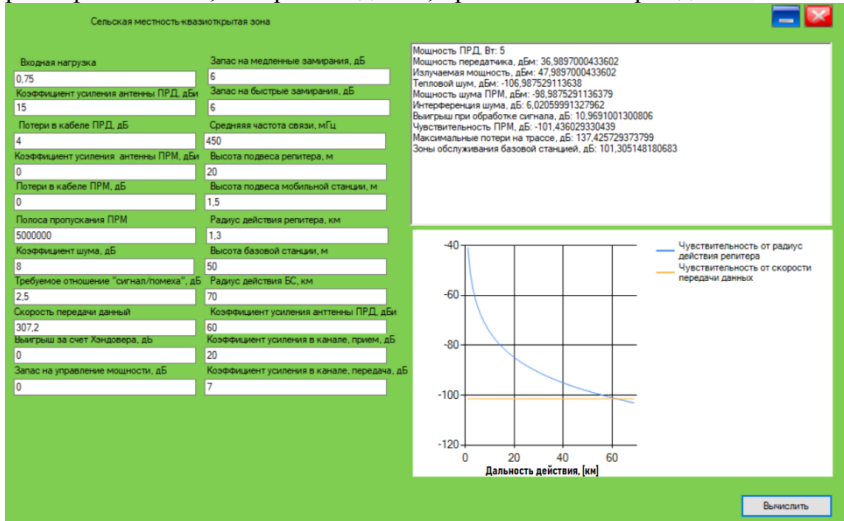


Рисунок 2. Дальность действия БС+репитер

Таким образом, применение репитеров в сетях сотовой связи LTE может рассматриваться как один из эффективных способов использования устройств M2M с целью обеспечения удаленного контроля над теми или иными технологическими процессами (производством) на территориях, где присутствие квалифицированного персонала экономически неэффективно.

Литература

1. *Переверзина Т.В.* использование репитеров для малонаселенных территорий в сети LTE//Современные проблемы телекоммуникации. СибГУТИ, Новосибирск, 2017 г. С. 131
2. *Переверзина Т.В.* оценка зоны радиопокрытия репитера в сети LTE на ранней стадии проектирования//Научно-техническое творчество молодежи. СибГУТИ, Новосибирск, 2018 г.
3. Платонов А. Российский институт потребительских испытаний: [Электронный ресурс] URL: <http://www.ripi-test.ru/3278-internet-nadache>
4. Бабков В.Ю., Цикин И.А. Сотовые системы мобильной радиосвязи. Учебное пособие. Санкт-Петербург: «БХВ-Петербург», 2013г. 432 с.

ИССЛЕДОВАНИЕ ВЕРОЯТНОСТИ ОБНАРУЖЕНИЯ НОВЫХ ГЕНОМНЫХ АССОЦИАЦИЙ ПРИ ПОЛНОГЕНОМНОМ МЕТА-АНАЛИЗЕ АССОЦИАЦИЙ

Д.В. Поверин, С.Н. Постовалов

Новосибирский государственный технический университет,
г. Новосибирск, foxlandg@gmail.com

В данной работе проводилось сравнение вероятности обнаружения новых ассоциаций с помощью гамма-комбинирования p -value и объединения выборок, полученных в независимых экспериментах. В качестве базового критерия ассоциаций использовался критерий хи-квадрат, а для комбинирования использовался гамма-тест, с параметром формы равным 5 и параметром масштаба равным 2. В ходе экспериментов брались различные объёмы выборок, количество групп, уровни значимости критерия, расстояния между гипотезами. Было установлено, что шанс найти новые локусы, ассоциированные с наблюдаемым заболеванием, при проведении мета-анализа с помощью комбинирования p -value может быть довольно высок при определённых условиях.

In this paper, the probability of detecting new associations was compared using gamma- p -value combining and combining samples obtained in independent experiments. The chi-square test was used as a base test association, and a gamma test was used for the combination, with a shape parameter of 5 and a scale parameter of 2. During the experiments, various sample sizes, number of groups, significance levels of the criterion, and the distances between hypotheses were taken. It was found that

the chance to find new loci associated with the observed disease, when conducting a meta-analysis using p-value combination, can be quite high under certain conditions.

Полногеномный анализ ассоциаций в наши дни является одним из важных и перспективных направлений исследований в биомедицине, которые направлены на выявление связи между геномом человека и наблюдаемыми заболеваниями. В основе данного метода лежит проверка гипотезы независимости между генотипом и наличием заболевания. В последнее время в связи проведением одинаковых исследований разными группами ученых, появилась необходимость объединять их результаты.

Мета-анализ – это статистический подход анализа данных на основе объединения результатов независимых экспериментов. Целью данного подхода является выявление и оценка степени согласованности или расхождения результатов проведённых экспериментов при наличии статистической неоднородности или гетерогенности. Например, в [1] при проведении полногеномного мета-анализа ассоциаций было найдено 17 новых локусов, ассоциированных с болезнью Паркинсона, которые в предыдущих исследованиях были отброшены.

В данной работе исследованы факторы, влияющие на вероятность обнаружения новых ассоциаций при комбинировании результатов полногеномного анализа ассоциаций, а также при объединении выборок независимых экспериментов. При исследовании использовалось компьютерное моделирование по методу Монте-Карло.

По методике проведения мета-анализ делится на следующие основные категории: комбинирование достигаемого уровня значимости *p-value*, анализ размера эффекта при большой гетерогенности, байесовский мета-анализ и прочие [2].

Существуют два основных подхода проведения мета-анализа. Первый подход – это анализ первичных данных экспериментов, что по сути своей является объединением всех анализируемых данных, что на практике далеко не всегда возможно. Вторым подходом является обобщением результатов (вторичных данных) независимых экспериментов, на основе которого делаются выводы.

Идея подхода комбинирования критериев с помощью *p-value* заключается в следующем. Пусть проводится k независимых экспериментов. В итоге мы получаем k независимых выборок X_1, X_2, \dots, X_k . Затем для каждой полученной выборки вычисляется достигаемый уровень значимости *p-value* с использованием статистики базового критерия (в данном случае это критерии ассоциаций), причём для разных выборок мо-

гут применяться статистики как одинаковых базовых критериев, так и разных, желательно в таком случае выбирать более мощные и робастные критерии. В результате получается выборка p_1, \dots, p_k и далее вычисляется статистика комбинированного критерия [3]:

$$T = \sum_{i=1}^k F_i^{-1}(1 - p_i), \quad (1)$$

где $F_i^{-1}(\square)$ обратная функция распределения, а p_i оценка достигаемого уровня значимости (*p-value*) i -го эксперимента. Также для получения потенциально более мощного комбинированного критерия можно рассмотреть в качестве статистики вычисление *взвешенной* суммы квантилей [3]

$$T_w = \frac{\sum_{i=1}^k w_i F_i^{-1}(1 - p_i)}{\sqrt{\sum_{j=1}^k w_j^2}}. \quad (2)$$

Данный подход проверки гипотез используется, когда отсутствует информация об элементах наблюдаемых выборок и как минимум известны лишь только объёмы этих выборок и результаты (*p-value*) проверки гипотезы на этих выборках.

Пусть проводится полногеномный анализ ассоциаций в k группах исследователей, в каждой из которых имеются рабочая и контрольная выборки индивидуумов по n человек каждая. Одновременно проводится проверка ассоциации по m маркерам, из которых d являются ассоциированными с заболеванием.

Для мета-анализа используется комбинирование *p-value* с помощью гамма-теста с параметром масштаба $b = 2$ и параметром формы равным $a_i = 5$. В качестве базового критерия выбран критерий Хи-квадрат, т.к. он наиболее часто встречается на практике.

Будем моделировать k выборок в соответствии с гипотезой H_a , параметры которой заданы в табл. 1. На первом этапе эта гипотеза проверяется базовым критерием с уровнем значимости α . На втором этапе гипотеза проверяется комбинированным критерием (гамма-тестом) по *p-value*, вычисленным на первом этапе. Всего возможно четыре случая

принятия решения о наличии ассоциации маркера с заболеванием, они приведены в табл. 2. Нас особенно интересует случай 2, когда комбинированный критерий находит ассоциацию, которая не была найдена ранее ни одной группой исследователей.

Чтобы оценить вероятность этого случая используется метод Монте-Карло [4]: процедура моделирования повторяется $N = 100\,000$ раз, и подсчитывается количество случаев 2. Оценка вероятности обнаружения новых ассоциаций равна отношению числа случаев 2 к числу повторений N .

Комбинируемые p -value вычисляются с помощью критерия ассоциаций хи-квадрат.

Таблица 1 – Гипотеза ассоциации при аддитивной модели наследования

| Альтернативная гипотеза | Модель наследования | K | p_a | ψ_1 | ψ_2 | ρ_{KL} |
|-------------------------|---------------------|------|-------|----------|----------|-------------|
| H_a | Аддитивная | 0.05 | 0.5 | 1,5 | 2,25 | 0,081 |

Таблица 2 – Исходы мета-анализа ПГАА

| Случай № | Описание | Базовый тест | Комбинированный тест |
|----------|---|--|--|
| 1 | Комбинированный критерий подтверждает ассоциацию, которая была найдена ранее хотя бы одной группой исследователей | основная гипотеза отвергнута (найдена ассоциация) | основная гипотеза отвергнута (найдена ассоциация) |
| 2 | Комбинированный критерий находит ассоциацию, которая не была найдена ранее ни одной группой исследователей | основная гипотеза не отвергнута (ассоциация не найдена) – ошибка второго рода | основная гипотеза отвергнута (найдена ассоциация) |
| 3 | Комбинированный критерий отбрасывает ассоциацию, которая была найдена ранее хотя бы одной группой исследователей | основная гипотеза отвергнута (найдена ассоциация) | основная гипотеза не отвергнута (ассоциация не найдена) – ошибка второго рода |
| 4 | Ассоциация не была найдена | основная гипотеза не отвергнута | основная гипотеза не отвергнута (ассоциация не найдена) |

| | | | |
|--|--|---|---------------------------------------|
| | | (ассоциация не найдена) – ошибка второго рода | ция не найдена) - ошибка второго рода |
|--|--|---|---------------------------------------|

Уровень значимости статистического критерия (вероятность ошибки первого рода) с коррекцией Бонферрони составил 10^{-8} .

По результатам проведённых исследований было установлено, что комбинированный гамма-тест позволяет найти с большой долей вероятности мутации со значимой ассоциацией, которые в отдельных экспериментах по ошибке были отброшены, что свидетельствует о том, что вероятность ошибки второго рода при гамма-комбинировании значительно меньше чем при использовании базового критерия хи-квадрат.

По полученным результатам наблюдается тенденция, что с ростом объёма выборки в каждой группе снижается риск получить ложноположительные результаты при достаточно низких уровнях значимости, то есть мощность гамма-теста растёт. Так же видно, что при значениях уровня значимости 10^{-5} и меньше частота случая 2 значительно выше частоты случая 1 при объёмах выборок 500 и 1000 и при этом частота случая 3, когда комбинирование отбрасывает уже найденные мутации, практически равна нулю.

Отсюда можно сделать вывод, что при рассмотренной альтернативе комбинирование с высокой долей вероятности позволит найти новые ассоциированные с наблюдаемым заболеванием локусы, которые ранее были отброшены другими экспериментаторами. Однако при $n = 100$ частоты случаев 1 и 2 довольно низкие что свидетельствует о том, что, если объём выборки в каждой группе недостаточно большой, то шанс найти новые мутации очень мал и смысла комбинировать результаты таких экспериментов нет смысла. Так же было установлено, что при использовании робастных базовых критериев, вероятность найти новые геномные ассоциации с увеличением объёмов выборки на каждом эксперименте растёт быстрее, чем при использовании неустойчивых базовых критериев. Это связано с тем, что в условиях неопределённости модели наследования неправильный выбор базового критерия приводит к увеличению числа ложных срабатываний и в целом понижается значимость p-value полученного в отдельном эксперименте, и поэтому результат мета-анализа в таком случае становится менее состоятельным и значимым.

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. 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.
3. 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.
4. Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход : [монография] : монография / Б. Ю. Лемешко, С. Б. Лемешко, С. Н. Постова-лов, Е. В. Чимитова. - : Новосибирск : Изд-во НГТУ, 2011. - 888 с.

АВТОМАТИЗИРОВАННЫЙ АНАЛИЗ ГРАФИЧЕСКИХ ПОЛЬЗОВАТЕЛЬСКИХ ИНТЕРФЕЙСОВ

К.С. Пыжлаков, М.А. Бакаев

**Новосибирский государственный технический университет,
г. Новосибирск, pyzhlakov@ Rambler.ru**

В эпоху информационных технологий ежедневно появляются различные программные продукты. Для реализации взаимодействия между пользователем и приложением преимущественно используется графический интерфейс. На то, как отображается интерфейс программы на экране устройства, влияет множество факторов, таких как: разрешение экрана, операционная система, веб-браузер и т.п. Цель статьи состоит в рассмотрении существующих подходов для анализа пользовательских интерфейсов.

In the era of information technology, various software products appear daily. The interaction between the user and the application is mainly performed via a graphical

interface. The way the program interface is displayed on the device screen is influenced by many factors, such as: the screen resolution, the operating system, the web browser, etc. The purpose of the article is to review existing approaches to analyzing user interfaces.

Для того, чтобы разработчик тратил как можно меньше времени на тестирование отображения интерфейса в различных условиях, существуют специализированные программы, которые анализируют пользовательские интерфейсы, определяют расположение объектов на них, предоставляют различные оценочные параметры и дают рекомендации по улучшению интерфейса.

Для анализа интерфейсов используются различные подходы. Можно выделить два основных, которые будут рассмотрены в данной статье. Первый подход основан на анализе графического изображения интерфейса, а второй основан на анализе программного кода интерфейса.

Первый метод называется сопоставлением шаблонов (pattern matching). Данный метод основан на попикельном сравнении фиксируемого объекта (шаблона) со всей областью изображения, на котором будет искаться данный объект. У этого метода есть ограничение, связанное с тем, что он ищет объекты, идентичные шаблону, то есть если объект будет повернут или иметь другой цвет, данный метод не сможет его опознать. Единственное решение этой проблемы состоит в том, чтобы составить большое количество шаблонов для каждого объекта, которые будут содержать в себе различные варианты его преобразования. Но это решение дает существенный минус, который связан с увеличением времени выполнения данного алгоритма за счет перебора всех шаблонов по каждому объекту поиска.

Вторым методом является анализ контуров. В контуре объекта имеется необходимая информация, которая позволяет распознать объект. Контур – это замкнутая монохромная линия, которая описывает границы объекта на изображении [2]. Имея правильный контур объекта, задача сопоставления шаблонов сводится к задаче сопоставления контуров. Одним из основных способов определения контуров объекта является способ, основанный на резком изменении цвета изображения.

Метод сопоставления с шаблоном, описанный ранее, ищет абсолютные совпадения точек шаблона с точками изображения, но если изображению изменить масштаб или развернуть относительно параметрам шаблона, то метод не сможет его определить. Чтобы решить эту проблему, можно использовать метод поиска особых точек. Особая точка (key point) – это небольшая область, которая определенным образом выделяется на изображении. Чтобы найти точку, можно воспользоваться некоторыми методами по их поиску, например углы или блобы, то есть небольшие области одинаковой яркости, достаточно четкой границей, выделяющиеся на общем фоне. Характеристикой для особой точки служит так называемый дескриптор, который находится по заданной окрестности точки, как направления градиентов яркости разных частей этой окрестности.

Чтобы найти объект на изображении используя этот метод, необходимо сделать следующие: вычисляем особые точки и их дескрипторы для объекта; на изображении определяем особые точки и их дескрипторы; сравниваем дескрипторы особых точек объекта и дескрипторы особых точек, найденных на изображении, если они равны, то на данной области изображения находится объект поиска.

Вторым подходом для анализа пользовательского интерфейса является анализ исходного кода. Код программы – это текстовая информация, поэтому чтобы ее проанализировать, можно обратиться к методам анализа текстовой информации, а именно к Text Mining, который подразумевает структурирование входных текстовых данных, определение шаблонов и интерпретацию результатов анализа [3].

Существует несколько методов извлечения информации из текста. Первый метод называется Feature (Entity) Extraction. Он извлекает слова или группу слов, которые имеют смысловую значимость для описания всего текста. Второй метод, под названием Extraction, является более сложным. Он прослеживает различного рода связи между извлеченными сущностями. Последний метод Relationship, Event and Fact Extraction является самым сложным вариантом извлечения информации. Он извлекает сущности, распознает факты и события, а также извлекает информацию из этих фактов.

Рассмотренные подходы анализа пользовательского интерфейса кардинально отличаются друг от друга, но их объединение даст разработчикам графического интерфейса возможность оперативно сравнить результаты того, что изображается на экране устройства, с тем, что написано в коде программы. Следовательно, будет сокращено время, затраченное на отладку работы интерфейса и на его тестирование.

Литература:

1. Красильников Н.Н. Цифровая обработка изображений. / Н.Н. Красильников – М.:Вузовская книга, 2001. – 320 с.
2. Фурса Н. Е. Метод поиска объектов на изображении с помощью контурного анализа по заданным характеристикам их контуров //Открытые информационные и компьютерные интегрированные технологии. – 2014. – №. 64. – С. 172-176.
3. Барсегян, А. А. Технологии анализа данных Data Mining, Visual Mining, Text Mining, OLAP/ А. А. Барсегян, М.С. Купрянов, В. В. Степаненко, И. И. Холод. – 2-е изд., перераб. и доп. – СПб.: БХВ-Петербург, 2007. – 384 с.

АНАЛИЗ МЕТОДА РАЗРАБОТКИ АЛГОРИТМОВ КОМБИНАТОРНОЙ ГЕНЕРАЦИИ НА ОСНОВЕ ДЕРЕВЬЕВ И/ИЛИ

Ю.В. Шапля, А.С. Репкин, Д.В. Кручинин
Томский государственный университет систем
управления и радиоэлектроники,
г. Томск, repkin572@gmail.com

В данной работе представлен анализ метода разработки алгоритмов комбинаторной генерации на основе деревьев И/ИЛИ, описаны проблемы, возникающие при применении данного метода. Рассмотрен пример определения биекции с помощью данного метода, приведены рекомендации по улучшению метода

This abstract presents an analysis of the method of developing algorithms for combinatorial generation based on AND / OR trees, describes problems arising from the application of this method. An example of the definition of a bijection using this method is considered, recommendations are given for improving the method

Введение

Комбинаторное множество – это конечное множество, элементы которого имеют некоторую структуру и имеется процедура построения элементов этого множества [1]. Элементы комбинаторных множеств (комбинаторные объекты), таких как сочетания, перестановки, размещения, разбиения, графы, деревья и так далее, играют важную роль в математике и информатике, а также имеют множество приложений [2,3].

Целью данного исследования является анализ методов разработки алгоритмов комбинаторной генерации и применение одного на практике.

Анализ методов разработки алгоритмов комбинаторной генерации

Процедура генерации комбинаторных объектов является ключевой при решении задачи обхода всех элементов некоторого комбинаторного множества. Существует несколько основных подходов к разработке алгоритмов комбинаторной генерации, каждый из которых обладает определенными недостатками: часть из них (поиск с возвратом [4], ЕСО-метод [5]) направлены только на последовательную генерацию комбинаторных объектов; большинство требуют, чтобы рассматриваемый объект был представлен в специальном виде (например, метод Б. Я. Рябко [6]), что трудно выполнимо ввиду отсутствия соответствующих методик. После сравнения методов можно выделить метод на основе деревьев И/ИЛИ, имеющий универсальный характер и предназначенный не только для построения алгоритмов последовательной генерации, но и для алгоритмов ранжирования и генерации в соответствии с рангами, что является существенным плюсом при работе с задачами глобальной нумерации и получения доступа к объекту по его номеру.

Основные проблемы актуальны непосредственно перед использованием метода: необходимо построить дерево И/ИЛИ, число вариантов которого должно совпадать со значением функции мощности комбинаторного множества. Также проблемой является определение биекции между комбинаторным множеством и деревом И/ИЛИ.

Метод разработки алгоритмов комбинаторной генерации на основе деревьев И/ИЛИ

Основой данного метода является представление комбинаторных множеств в виде структуры дерева И/ИЛИ, с помощью которого в дальнейшем можно построить алгоритмы последовательной генерации комбинаторных объектов, их ранжирования и генерации в соответствии с их рангами.

Деревом И/ИЛИ называется дерево, содержащее узлы двух типов: И-узел и ИЛИ-узел (рисунок 1) [7]. Треугольником обозначается рекурсия.

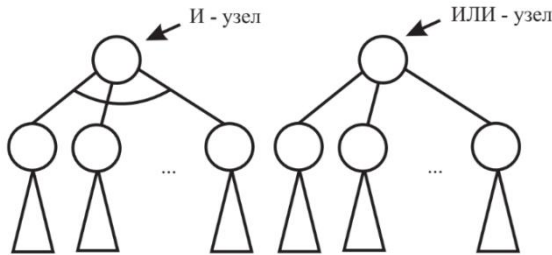


Рис. 1. Схематическое изображение узлов дерева И/ИЛИ

Вариантом дерева И/ИЛИ называется дерево, получаемое из данного путем отсечения всех дуг, кроме одной, у всех ИЛИ-узлов. Для ранжирования комбинаторного объекта, представленного в виде структуры дерева И/ИЛИ, необходимо предварительно провести сопоставление варианта v дерева D с самим деревом. Далее требуется рассмотреть все узлы варианта v и для каждого из них вычислить значение $l(z)$, соответствующее некоторому номеру узла z , для которого $0 \leq l(z) < w(z)$, где $w(s)$ – число вариантов в поддереве узла.

Эффективность данного метода показана на примере генерации различных комбинаторных множеств: перестановки, сочетания, разбиения, разложения, композиции, числа Фибоначчи, числа Каталана, деревья, выражения формальных языков [1]. В свою очередь сложность применения заключается в необходимости предварительной подготовки рассматриваемого комбинаторного множества и отсутствии формальных методик определения биекции.

Под процедурой определения биекции понимается перечень действий, направленных на поиск взаимно однозначного соответствия между множеством всех вариантов дерева И/ИЛИ и всеми элементами комбинаторного множества. В качестве рекомендации предлагается следующая последовательность действий: нужно рассмотреть изменения, происходящие в структуре комбинаторных объектов при переходе от одного узла дерева И/ИЛИ к другому, а затем отразить данные изменения в биекции. Например, это могут быть изменения параметров комбинаторного объекта при переходе с одного уровня дерева на другой или при переходе по разным ветвям дерева.

Пример определения биекции для перестановки из n элементов

Перестановка n элементов – это упорядоченный набор всех элементов множества из различных элементов.

Функция мощности множества всех перестановок n элементов определяется следующим выражением, принадлежащим алгебре $\{\mathbb{N}, +, \times, R\}$:

$$f(n) = P_n = nP_{n-1}.$$

где $P_n = n!$ и $P_1 = 1$.

На рисунке 2 представлено дерево И/ИЛИ для множества перестановок n элементов.

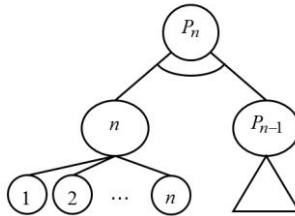


Рис. 2. Дерево И/ИЛИ для множества перестановок n элементов

Для определения биекции проанализируем изменения, происходящие в структуре перестановок при переходе от одного узла дерева И/ИЛИ к другому.

Выбор одного из n листов-сыновей узла по левой ветви дерева показывает набор возможных вариантов позиции размещения элемента под номером n в перестановке из n элементов. Соответственно, номер листа, выбранного в варианте дерева, указывает позицию элемента под номером n в перестановке n элементов. Далее данный элемент под номером n не учитывается, и аналогичные рассуждения применяются к перестановке $n-1$ элементов (правая ветвь дерева).

Заключение

В настоящее время существует несколько основных подходов к разработке алгоритмов комбинаторной генерации. Анализ метода разработки алгоритмов комбинаторной генерации на основе деревьев И/ИЛИ показал, что метод обладает следующими преимуществами:

– универсальность (отсутствует привязка к конкретному комбинаторному множеству);

– возможность построения алгоритмов ранжирования и генерации в соответствии с их рангами.

При этом сложность применения данного метода заключается в необходимости предварительной подготовки рассматриваемого комбинаторного множества и необходимости определения биекции между комбинаторным множеством и деревом И/ИЛИ.

Для определения биекции необходимо провести анализ изменений в структуре комбинаторных объектов при переходе от одного узла дерева И/ИЛИ к другому и отразить данные изменения в биекции.

Работа выполнена при финансовой поддержке РФФ в рамках научно-го проекта 18-71-00059

Литература

1. Кручинин В.В. Методы, алгоритмы и программное обеспечение комбинаторной генерации: дис. ... д-р техн. наук: 05.13.11 / Кручинин Владимир Викторович. – Томск, 2010. – 163 с.

2. Knuth D.E. The Art of Computer Programming, Volume 4A: Combinatorial Algorithms, Part 1. – Addison-Wesley, 2011. – 883 p.

3. Ruskey F. Combinatorial generation, Working version (1j-CSC 425/520) [Электронный ресурс]/ Department of Computer Science University of Victoria Victoria, B.C. V8W 3P6 CANADA. Дата публикации в Интернет: 10.10.2001. – 311 с. – URL: <http://www.1stworks.com/ref/ruskeycombgen.pdf> (дата обращения: 30.10.2018).

4. Рейнгольд Э. Комбинаторные алгоритмы: Теория и практика / Э. Рейнгольд, Ю. Нивергельт, Н. Део. – Москва: Мир, 1980. – 476 с.

5. Barucci E. ECO: A methodology for the enumeration of combinatorial objects / E. Barucci, A. Del Lungo, E. Pergola, R. Pinzani // Journal of Difference Equations and Applications. – 1999. – 5. – P. 435–490.

6. Рябко Б.Я. Быстрая нумерация комбинаторных объектов // Дискретная математика. – 1998. – 10, № 2. – С. 101–119.

7. Кручинин В.В. Методы построения алгоритмов генерации и нумерации комбинаторных объектов на основе деревьев И/ИЛИ. – Томск: В-Спектр, 2007. – 200 с.

ПРЕДПОСЫЛКИ К РАЗРАБОТКЕ БОРТОВОЙ АВТОМАТИЧЕСКОЙ СИСТЕМЫ КОНТРОЛЯ И ОБЕСПЕЧЕНИЯ КАЧЕСТВА ПРОТИВОГОЛОЛЁДНОЙ ОБРАБОТКИ ПОКРЫТИЙ

Я.С. Садовникова, К.П. Мандровский
Московский автомобильно-дорожный государственный
технический университет (МАДИ),
г. Москва, jana.sadovnikova@yandex.ru

В статье рассмотрены основные аспекты и перспективы создания бортовой автоматической системы контроля и обеспечения качества противогололёдной обработки покрытий жидким реагентом. Дана краткая характеристика факторов, определяющих качество и эффективность обработки покрытий реагентом. Указаны особенности предлагаемой системы, отличающие её от автоматических систем контроля, которыми оснащены эксплуатируемые дорожные и аэродромные распределители реагентов. Предполагается, что бортовая автоматическая система будет учитывать конструкцию распределительного оборудования, режимы его функционирования, внешние (погодные) условия, интенсивность движения, категорию дороги и тип покрытия. Задачи системы состоят в расчёте ширины полосы обработки, равномерности распределения реагента и параметров зоны перекрытия (если используются два диска) на протяжении всего процесса противогололёдной обработки. При фиксации наличия неравномерности, появления существенной деформации зоны (полосы) обработки или при поступлении данных о смене метеорологической обстановки система производит расчётным путём поиск способа обеспечения качественной обработки. В число таких способов входит: изменение режима работы оборудования, его расположения относительно покрытия, скорости движения машины или фактического расхода реагента. Таким образом осуществляется экономичное расходование реагента при сохранении рекомендуемых норм расхода, требуемой ширины полосы обработки и равномерности распределения реагента.

In the article the main aspects and perspectives of creating an on-board automatic system for monitoring and ensuring the quality of anti-ice treatment of coatings with a liquid reagent are considered. A brief description of the factors determining the quality and efficiency of coating treatment with a reagent is given. Specific features of the proposed system are distinguished, which distinguish it from automatic control systems, which are used for operating road and airfield reactant distributors. It is assumed that the on-board automatic system will take into account the design of the distribution equipment, its operation modes, external (weather) conditions, traffic intensity, road category and type of coverage. The tasks of the system are to calculate the processing bandwidth, the uniformity of the reagent distribution, and the parame-

ters of the overlap zone (if two discs are used) throughout the whole process of anti-ice treatment of coatings. When fixing the presence of unevenness, the appearance of a significant deformation of the processing zone (strip), or when data are received on changing the meteorological situation, the system calculates the search for a way to ensure quality processing. Such methods include: changing the mode of operation of the equipment, its location relative to the coating, the speed of the machine or the actual consumption of the reagent. Thus, the economical consumption of the reagent is carried out while maintaining the recommended flow rates, the required processing bandwidth and the uniformity of the reagent distribution.

Достижение качественной противогололёдной обработки (ПГО) дорожных и аэродромных покрытий невозможно без учёта ряда факторов, которые составляют три группы: метеорологические, конструктивные и эксплуатационные. Последняя группа включает в себя режимы работы распределительного оборудования (частоту вращения диска, давление подачи реагента), высоту диска над покрытием, скорость движения машины для распределения реагента (МРР) и отдельные транспортно-эксплуатационные показатели дороги такие, как класс, категория, тип покрытия, интенсивность движения. Из метеорологических факторов следует отметить влияние солнечной радиации (освещённости участков покрытия), температуры и влажности воздуха, скорости и направления ветра на характер и продолжительность реакции противогололёдного реагента (ПГР) с покрытием. Понятие «конструктивные факторы» включает в себя диаметр диска, длину лопаток диска, тип форсунок, угол наклона диска к его оси, схемы расположения форсунок (на штанге или в кожухе диска). Принятие во внимание всех вышеперечисленных факторов при подготовке и проведении ПГО является гарантией обоснованного назначения нормы расхода реагента и рационально подобранных методов и режимов внесения реагентов и, как следствие, экономичного и качественно осуществлённого процесса ПГО покрытий. В бортовых системах автоматического контроля, которыми оснащаются МРР, предусмотрено поддержание постоянным заданной в начале процесса ПГО нормы расхода, а также осуществление контроля ширины полосы обработки и асимметричности распыления ПГР. Одновременный учёт всех трёх групп факторов такими системами не производится, следствием чего является перерасход реагента или появление огрехов при ПГО. Кроме того, в таких системах нет возможности оценить оперативными методами равномерность распределения ПГР по длине полосы обработки (по обрабатываемой площади покрытия). Равномерность (расстояние между каплями жидкого реагента) при идеальных

условиях может существенно отличаться от фактической величины расстояния между соседними каплями ПГР и характером их распределения по зоне обработке [1]. Перспективна разработка бортовой системы, которая могла бы осуществлять на протяжении всего процесса ПГО расчёт ширины полосы обработки, зоны перекрытия, образываемой при использовании двух дисков, равномерности нанесения ПГР. Основы математического обеспечения работы предлагаемой системы изложены в работе [2]. Программный модуль как компонент программного обеспечения системы представлен в [3]. Механизм действия предлагаемой системы иллюстрирует рисунок 1.

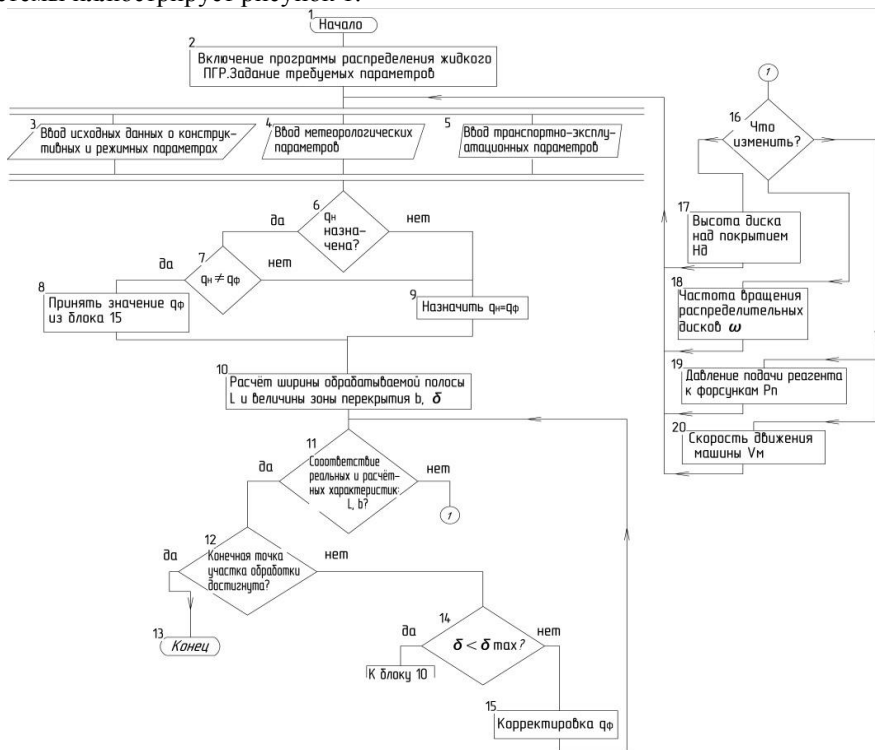


Рисунок 1 – Алгоритм работы бортовой системы контроля и обеспечения эффективности ПГО

Принцип функционирования системы состоит в следующем: в начале работы (блоки 1, 2) задаются требуемые параметры рабочего процесса: ширина полосы обработки l , величина зоны перекрытия b . Затем производится ввод информации о конструктивных, эксплуатационных, метеорологических факторах и дорожной обстановке (блоки 3-5). Исходя из этих данных выбирается рекомендованная норма расхода реагента q_n (блоки 6, 9). В процессе распыления ПГР производится расчёт ширины полосы обработки l , зоны перекрытия дисков b и численно определяется неравномерность распределения ПГР как отклонение δ фактического расстояния между каплями от расстояния, рассчитанного при идеальных условиях (блок 10). Если требуемые значения l и b не достигнуты, а также в случае, если отклонение расстояния между каплями при реальных условиях δ от расстояния между каплями, соответствующего идеальным условиям распыления, превышает допустимую максимальную величину δ_{\max} (то есть $\delta > \delta_{\max}$), система осуществляет поиск решений (блоки 11-20) для обеспечения заданных значений δ [4]. При этом системой предусмотрено варьирование режимных параметров рабочего оборудования (блоки 17-20) и фактической величины расхода реагента $q_{\text{ф}}$ (блоки 7, 8, 15).

Результат функционирования подобной системы: назначение оптимального расхода при сохранении равномерности распыления жидкого реагента. Возможная область применения системы – корректировка режимных и эксплуатационных параметров для обеспечения качественной ПГО покрытий, то есть осуществление оперативного мониторинга качества ПГО. Возможно применение системы при математическом и имитационном моделировании процессов ПГО для определения рациональных конструктивных параметров рабочего оборудования на этапе проектирования.

Литература:

1. Мандровский К.П., Садовникова Я.С. Влияние скорости машины на равномерность распределения противогололёдных реагентов // Механизация строительства. 2018. - Т. 79, № 4. - С. 60-64.
2. Mandrovskiy K.P., Sadovnikova Y.S. Characteristics of the droplet motion of a liquid antifreeze reagent. Magazine of Civil Engineering. - 2018. - No. 03. - Pp. 14–26.
3. Программа построения зоны распределения противогололёдного реагента по покрытию: ЭВМ 2018619402 / Кустарев Г.В., Мандров-

ский К.П., Садовникова Я.С. - № 2018616710; заявл. 27.06.18; опубл. 06.08.18.

4. Мандровский К.П., Садовникова Я.С. Предпосылки к разработке методики обеспечения эффективности противогололёдной обработки покрытий // Вестник Московского автомобильно-дорожного государственного технического университета (МАДИ). – 2018. - № 3(54). – С. 54-61.

ПОДХОД К ОПИСАНИЮ ВЗАИМНОГО ВОЗДЕЙСТВИЯ ОБЩЕСТВЕННЫХ СИСТЕМ

В.М. Саклаков

**Национальный исследовательский
Томский политехнический университет,
г. Томск, romanov_ky@mail.ru**

В работе рассматриваются существующие разрывы в описании общественных систем. Предложен новый подход к их анализу, позволяющий описывать сами общественные системы и формируемые ими процессы и явления.

The paper considers the existing gaps in public systems description. A new approach to their analysis has been proposed, which makes it possible to describe the public systems, processes and phenomena they form.

Общественные системы разного уровня и сложности через имеющиеся у них институты [1] формируют и корректируют процессы взаимного воздействия друг на друга, задавая соответствующие вектора движения данных систем. В конечном итоге, каждая серия воздействий на определенную часть системы может приводить к увеличению, снижению или полной утрате ее целостности (устойчивости). Каждая система, благодаря такому движению, стремится занять нишу, которая в большей степени позволит ей сохранить устойчивость. При этом необходимо понимать, что каждая отдельно взятая система является одновременно подсистемой и надсистемой ряда других систем [2], т. е. между ними могут возникать противоречия или наоборот - консолидация. На более низком уровне обобщения в процесс воздействия включено огромное количество агентов данных систем. Для множества сфер общественной жизни [3] актуальным является выявление и прогнозирование процессов

генерации и диффузии воздействующего контента и реакции на него агентов, а также систем, в рамках которых они ведут свою деятельность.

С широким распространением информационных систем и технологий специфика сбора, обработки и анализа данных изменилась в следствие сверхбольшого объема их генерации и накопления различными системами, средами и сообществами. Одним из ярких примеров подобных сообществ являются социальные интернет-сети – потенциально один из наиболее крупных и значимых источников эмпирических данных для проведения исследований различной направленности. Однако, помимо явных преимуществ, такие данные имеют и очевидное ограничение, вызванное спецификой формируемых социальными сетями информационных потоков. Цифровой след [4], оставляемый огромным количеством пользователей, в большей степени пригоден для анализа именно социальных систем, т. е. к нему в полной мере можно применить инструментарий социологической науки.

Существующая проблемная ситуация позволяет сформулировать **цель настоящей работы**: формирование нового подхода к анализу общественных систем, позволяющего выделять и описывать в структурированном виде непосредственно сами общественные системы и формируемые ими процессы и явления. Поставленная цель имеет системный характер и будет достигаться и уточняться поэтапно, в рамках настоящей и будущих работ.

Опишем предлагаемый автором подход к процессу взаимного воздействия общественных систем, в том числе социальных – см. рисунок 1. В первую очередь стоит выделить непосредственно сами общественные системы, имеющие определенную внутреннюю структуру. Каждая из таких систем имеет некоторый вектор движения в направлении ниши, в которой она будет наиболее эффективна [5]. В процессе движения возникает конкурентное и кооперационное взаимодействие с другими системами, как правило с помощью знакового представления данных [6]. Стремясь к достижению собственных точек эффективности, они будут оказывать влияние друг на друга, корректируя вектора движения.

Каждая система обладает как минимум следующим набором значимых условных параметров: (1) вес, (2) вектор движения, (3) время реакции на внешнюю среду (группы воздействий других систем) и (4) ее адекватность с точки зрения ее продвижения к точке эффективности. При этом необходимо понимать, что под воздействием внешней среды точки эффективности могут со временем меняться. При разработке данной концепции автор руководствовался пониманием происхождения

некоторых гуманитарных наук от базовой науки – физики [7]. В настоящее время рассматривается применимость каждого из данных параметров непосредственно в процессе проектирования и разработки программного обеспечения.

В рамках настоящей работы не ставится задача классификации общественных систем. Здесь они представлены лишь схематично для отражения принципов их взаимодействия. В дальнейшем планируется уточнение предлагаемой методики на основе интеллектуального анализа данных, получаемых из различных социальных сетей.

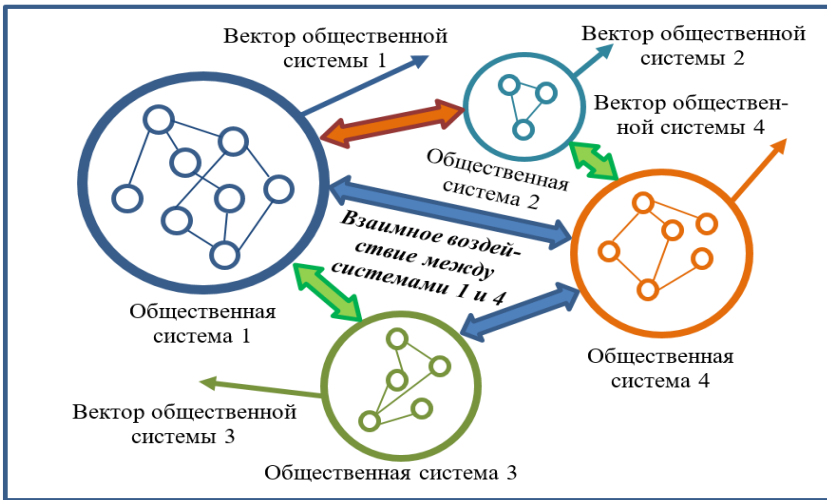


Рисунок 1 – Подход к описанию процесса взаимного воздействия общественных систем

Литература

1. Норт Дуглас. Институты, институциональные изменения и функционирование экономики \ Пер. с англ. А.Н. Нестеренко; предисл. и науч. ред. Б.З. Мильнера. - М.: Фонд экономической книги "Начала", 1997. - 180 с. - (Современная институционально-эволюционная теория)
2. Тарасенко Ф. П. Прикладной системный анализ: учебное пособие / Ф. П. Тарасенко. – Москва: КРОНУС, 2017. – 220 с.
3. Добрынин А.П., Черных К.Ю., Куприяновский В.П. Цифровая экономика - различные пути к эффективному применению технологий

(BIM, PLM, CAD, IoT, Smart City, Big Data и другие) // International journal of open information technologies. Т. 4, №1, 2016. с.4-11

4. Лушин Е. А. О термине «электронно-цифровые следы» // Расследование преступлений: проблемы и пути их решения №4, 2017. с. 161-163.

5. Р. Акофф О целеустремленных системах / Р. Акофф, Ф. Эмери – М.: Книга по Требованию, 2012. – 270 с.

6. Еремченко Е. Н. Концепция знака в контексте неогеографии // Информационные и математические технологии в науке и управлении №1, 2016. с. 49-54

7. Дорошенко М. Е. Анализ неравновесных состояний и процессов в макроэкономических моделях. М.: Экономический факультет МГУ, ТЕИС, 2000. – 206 с.

РАЗРАБОТКА МОДУЛЯ ПОСТОБРАБОТКИ ДАННЫХ ЛАЗЕРНОГО ДОПЛЕРОВСКОГО ИЗМЕРИТЕЛЯ СКОРОСТИ

Д.О. Семенов, С.В. Двойнишников

**Институт теплофизики им. С.С. Кутателадзе СО РАН,
г. Новосибирск, SeemDima@gmail.com**

Разработан программный модуль постобработки данных лазерного доплеровского измерителя скорости. Создан интерфейс приложения для взаимодействия с пользователем, и спроектировано ядро программного модуля, позволяющее загружать данные, полученные в экспериментах, выполнять фильтрацию и визуализацию.

A software module of post-processing of a laser Doppler velocimeter data has been developed. An application interface has been created for interacting with the user, and a program module core that allows loading, performing, filtering and visualizing data obtained in experiments has been designed.

Одними из самых точных устройств для измерения скорости и длины в производстве в настоящее время являются лазерные доплеровские измерители скорости (ЛДИС), работа которых основана на эффекте Доплера [1]. Данные измерители предназначены для бесконтактного измерения скорости потоков жидкости и газов, а также твердых и диффузно отражающих объектов. ЛДИС отличаются высоким разрешением и широким диапазоном измерения скорости. Современные ЛДИС пред-

ставляют собой сложные оптико-электронные измерительные комплексы и системы, сочетающие в себе передовые технические решения. В настоящее время ЛДИС широко используются в промышленных областях и научных исследованиях.

В ИТ СО РАН более 10 лет развиваются ЛДИС серии ЛАД-0ХХ [2-3]. Они состоят из оптоэлектронного модуля и компьютера, обеспечивающего управление измерительным комплексом и проведение автоматизированного эксперимента. В программном комплексе ЛАД данные сохраняются в базу данных в специализированном формате. Эти данные представляют собой запись, в которой содержится набор значений измеренных скоростей с дополнительными параметрами: пространственные координаты точки, в которой было выполнено измерение, время начала и конца измерения, количество зафиксированных в данной точке измерений, значения скорости, коэффициенты достоверности (мера качества зарегистрированного сигнала), направление и другие параметры. Для обработки и анализа полученных в экспериментах данных требуется разработать удобный инструментарий. Цель данной работы: разработка модуля постобработки данных лазерного доплеровского измерителя скорости. Модуль постобработки должен выполнять следующие функции:

- визуализация - построение необходимых графиков для визуального представления данных;
- фильтрация - наложение различных пороговых фильтров по времени, скорости, частоте, достоверности. Операции фильтрации должны применяться как на отдельные записи в некоторой точке, так и на всю базу данных текущего измерения;
- спектральный анализ – построение спектра сигнала для визуальной оценки частотного состава зарегистрированного сигнала.
- восстановление данных в случае неправильных действий пользователя.

Для создания программного модуля постобработки лазерного доплеровского измерителя скорости требуется разработать расширяемую архитектуру, реализовать взаимодействие с существующим модулем управления базой данных, создать интуитивно понятный графический пользовательский интерфейс.

Разработан основной интерфейс приложения для взаимодействия с пользователем, представляющий собой оконное приложение. Элементами этого окна является набор функций, позволяющих загрузить дан-

ные из файловой системы, визуализировать структуру данных в виде дерева. Кроме того, реализована поддержка различных функций постобработки, являющихся элементами этого дерева или динамически расширяемого контекстного меню. Реализовано независимое взаимодействие с существующим модулем базы данных путем использования абстрактного интерфейса этого модуля. Спроектирована архитектура, позволяющая динамически расширять функциональность приложения в построении графиков и фильтрации экспериментальных данных. Такая расширяемость достигается путем построения архитектуры, предусматривающей реализацию простого подключения дополнительных модулей визуализации и фильтрации независимо от основного алгоритма работы программы. Реализована возможность визуализации данных в виде двумерных графиков. Создан интерфейс для модулей фильтрации и визуализации, позволяющий подключать различные модули.

Разработано ядро модуля постобработки данных лазерного доплеровского измерителя скорости, позволяющее загружать данные, полученные в экспериментах, выполнять фильтрацию и визуализацию. В дальнейшем будут реализованы модули фильтрации, отображения и спектрального анализа с полной реализацией требуемого функционала.

Литература

1. Дубнищев Ю.Н. Ринкевичюс Б.С. Методы лазерной доплеровской анемометрии // М.: Наука, 1982.- С. 304.
2. Меледин В.Г. Лазерная доплеровская полупроводниковая анемометрия для науки и промышленности // Деловая слава России. Межотраслевой альманах для организаторов производства. - № 3(41), выпуск 2013. – С.23-26.
3. Меледин В.Г., Кротов С.В., Бакакин Г.В., Двойнишников С.В., Наумов И.В., Павлов В.А., Рахманов В.В., Садбаков О.Ю. Лазерная доплеровская анемометрия горячего проката // Современные проблемы науки и образования. – 2013. – № 2; URL: www.science-education.ru/108-8812.

ПОДХОД К ОПТИМИЗАЦИИ ПРОЦЕССА ПЛАНИРОВАНИЯ ПРОВЕДЕНИЯ ГЕОЛОГО-ТЕХНИЧЕСКИХ МЕРОПРИЯТИЙ

В.П. Соломатин, В.Ю. Гришаев, К.Р. Марупова, В.М. Саклаков

Томский политехнический университет

г. Томск, solomatin.vl.p@gmail.com

В статье рассматривается проблема анализа промысловых данных и планирования геолого-технических мероприятий на нефтяных месторождениях. Приводятся предложения по автоматизации процесса планирования геолого-технических мероприятий.

The article deals with the problem of analyzing field data and planning geological and technical measures at oil fields. Provides suggestions for automating the process of planning geological and technical measures.

Главным документом, регламентирующим процесс разработки нефтяных и газовых месторождений, является проект разработки месторождения. В данном документе содержится вся информация о месторождении, варианты разработки и мероприятий, проведение которых необходимо для качественной эксплуатации месторождения. На данный момент, в среднем, разработка одного проектного документа занимает 1 календарный год, в течение которого более 20 сотрудников проектного института и множество смежных департаментов работают над созданием проекта разработки [1].

Одной из важных составных частей проектно-технологического документа является планирование проведения геолого-технических мероприятий (ГТМ) - мероприятий, направленных на интенсификацию и увеличению добычи углеводородов.

Некоторые существующие месторождения содержат от двух тысяч скважин. При этом их разработку нужно планировать оптимальным образом на двадцать и более лет вперед. Однако на данный момент времени процесс обработки и анализа данных даже по таким крупным месторождениям осуществляется с достаточно низкой степенью автоматизации. Данный фактор обуславливает требование больших временных затрат специалистов на выполнение работ в ручную и не добавляет ценности проекту разработки месторождения.

В применяемых программных комплексах, используемых для проектирования разработки нефтегазовых месторождений, автоматизация до определенной степени коснулась этапов сбора и первичной обработки данных. Отсутствие полной автоматизации на всех этапах планирования

ГТМ является узким местом всего бизнес-процесса разработки проекта месторождения. Следствием данного фактора является низкая производительность в процессе анализа данных показателей разработки всего месторождения.

Целью настоящей работы является описание бизнес-процесса создания проекта разработки месторождения одного из томских проектных институтов. Описываемые в рамках настоящей работы предложения могут значительно снизить временные затраты отдельных специалистов на анализ и планирование проведения ГТМ. Высвободившиеся человеческие ресурсы могут быть привлечены к разработке других проектов месторождений или выполнению иной деятельности. Достижение данной цели также позволит создать дополнительную ценность для недропользователей в виде снижения риска ошибок, вызванных человеческим фактором. Таким образом, данная работа позволит снизить общие внутренние издержки в расчете на один проект.

Данная цель является комплексной и будет достигаться поэтапно. На данный момент для достижения цели необходимо решить следующие задачи:

1. Провести анализ существующего процесса планирования геолого-технических мероприятий, выявить его «узкие места».
2. Выработать предложения по их устранению с помощью современных информационных технологий.

Схема анализа и планирования ГТМ одного из томских проектных институтов представлена на рисунке 1.

Детализация процесса планирования ГТМ:

1. Для получения сведений о планировании проведения ГТМ и прогнозирования его результатов вводятся данные о показателях разработки в единую таблицу.

2. Происходит декомпозиция исходных данных на типы ГТМ. Таким образом мы получаем таблицу с необходимыми данными для каждого типа ГТМ.

3. На основании полученной таблицы производится анализ эффективности каждого типа ГТМ путем сравнения проектных и фактических показателей разработки для уже выполненных мероприятий, анализа районов проведения ГТМ.

4. На следующем этапе проводится синтез исследований эффективности ГТМ, результатом которого является получение представления об эффективности использования каждого вида ГТМ. Под эффективностью авторы понимают профицитное соотношение результатов и затрат.

5. Далее проводится прогнозирование дополнительной добычи от проведения ГТМ.

6. Следующим шагом является планирование проведение каждого вида ГТМ с обоснованием причин такого выбора и обоснованием отказа от других видов ГТМ.

7. На основании плана осуществляется прогноз возможных результатов проведения каждого вида ГТМ.

8. На заключительном этапе происходит составление аналитической записки о планировании проведения ГТМ и прогнозирования результатов их проведения с обоснованием причин принятых нами решений и причин отказа от каких-либо видов мероприятий. [2]

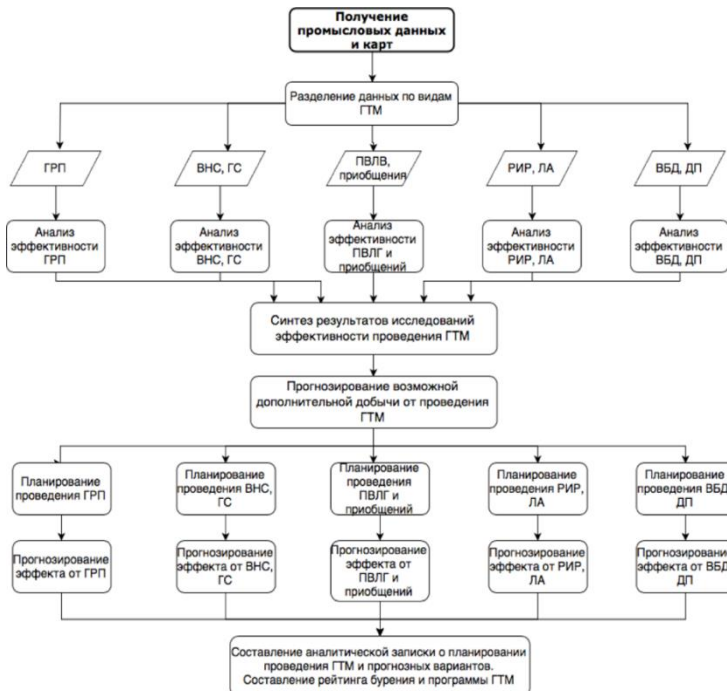


Рисунок 1 - Схема анализа и планирования ГТМ

| Легенда | |
|---------|--|
| ВНС | - ввод новых скважин, ГС - горизонтальные скважины, ПВЛГ - перевод на вышележащий горизонт, |
| РИР | - ремонтно-изоляционные работы, ЛА - ликвидация аварий, ВБД - вывод из бездействия, ДП - доп. перфорация |

Авторы предлагают оптимизировать существующий бизнес-процесс следующим образом:

На 2 этапе происходит потеря времени за счет ручного разделения данных по видам ГТМ и представления в необходимом и удобном для использования формате. Данную проблему можно решить за счет автоматического приведения всех данных к стандартизованному виду, пригодному для дальнейшего использования.

Этап 3 является также затратным по времени за счет ручного анализа и поиска зависимостей между показателями разработки, параметрами пласта и его свойствами.

Этап 5 и 7 требуют значительных временных затрат за счет ручной обработки данных и произведения расчетов, которые в будущем времени планируется осуществлять в автоматическом режиме.

Шаг 8 также вызывает большой интерес в плане автоматизации из-за необходимости внедрения алгоритма автоматического построения форм Госплана, рейтинга бурения и программы ГТМ. В данный момент на преобразование данных к установленным формам тратится большое количество дней.

Авторы предлагают ввести автоматизацию обработки информации на этапах 2, 3, 5, 7 и 8 используя технологии, позволяющие легко интегрироваться с входящим потоком данных, в частности язык расширяемой разметки xml, используемый во многих стандартах, в частности WITSML, PRODFML и других. В настоящее время процесс обработки данных проводится вручную. Авторы планируют выявить зависимости между показателями разработки и параметрами пласта, построив соответствующую математическую модель, разработать и внедрить программные модули, осуществляющие данную обработку на языке Python. Также необходимо разработать алгоритм автоматического преобразования конечных данных к виду установленного образца [3].

Литература

1. Нургалиева Р.А., Романова М.Ю. Анализ эффективности проведения геолого-технических мероприятий на объекте ЮВ1(1) Мегийонского месторождения // Международный студенческий научный вестник. – 2016.
2. Уметбаев В.Г. Геолого-технические мероприятия при эксплуатации скважин: Справочник рабочего. – М.: Недра, 1988
3. Goldberg D. E. Genetic algorithms in search, optimization, and machine learning. Reading, MA: Addison-Wesley. 1989.

АППРОКСИМАЦИЯ СЕЙСМИЧЕСКОГО ИМПУЛЬСА ПАРАМЕТРИЧЕСКОЙ МОДЕЛЮ

Н.Ю. Филипенко, Е.В. Рабинович
Новосибирский государственный технический университет,
г. Новосибирск, nikita1994@ngs.ru

Представлены результаты аппроксимации сейсмического импульса, возникающего при гидравлическом разрыве пласта, параметрической моделью. Приведены изображение сейсмограммы импульса и его аппроксимация параметрической моделью с найденными коэффициентами. Также представлены критерии оценки точности аппроксимации и их значения.

The results of approximation of a seismic pulse arising from hydraulic fracturing by a parametric model are presented. The image of the seismogram of the pulse and its approximation by a parametric model with the coefficients found are given. Also presented are the criteria for assessing the accuracy of the approximation and their values.

Гидравлический разрыв пласта (ГРП) является эффективным средством увеличения темпов отбора нефти из продуктивных пластов и более полной выработки месторождения. Данный метод добычи нефти основывается на закачивании жидкости в скважину для разрыва продуктивного пласта под действием давления воды. При проведении ГРП необходим мониторинг пространственного развития зоны трещиноватости. Мониторинг дает информацию для управления основными параметрами ГРП – объемом и темпами закачки жидкости в пласт.

В данной работе приведена аппроксимация сейсмического импульса параметрической моделью. [1]

Параметрическая модель (1).

$$s(X, t) = cA \frac{\cos(kX - \omega t + \theta)}{\operatorname{ch}\left(\frac{X - vt}{T}\right)} \quad (1)$$

В параметрической модели c – коэффициент пропорциональности, зависящий от вида эволюционного уравнения, A и T_0 – амплитуда и длительность импульса, $X = x - x_0$ – расстояние между сейсмоприемником и источником импульса, $\omega = kv$ – угловая частота, которая равна произведению волнового числа k и фазовой скорости волны v , θ – фазовый сдвиг волны. На рисунке 1 представлен участок сейсмограммы полученный при гидравлическом разрыве пласта и его аппроксимация параметрической моделью.

Для проведения аппроксимации был выбран участок с 1500 – 3000 мс, так как именно на этом участке содержится полезная информация для работы с сеймотрассой. На рисунке 1 представлено изображение сигнала и аппроксимируемого участка.

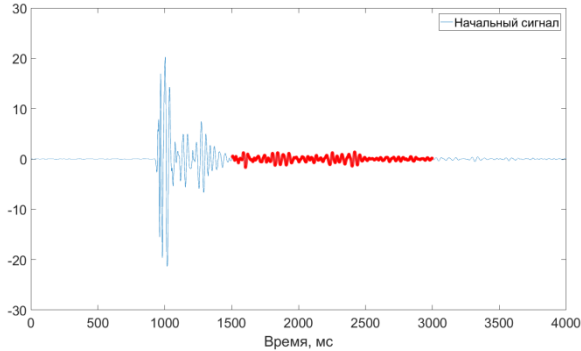


Рисунок 1 – Сеймотрасса реального сигнала длительностью 4000 мс и аппроксимируемая область (красным)

На рисунке 2 показана область сеймотрассы для которой проводилась аппроксимация. Участок длительностью 1500 мс. Область до 1500 мс является областью высоких амплитуд и на ней записано начало появления трещины, поэтому эта область не так важна для аппроксимации.

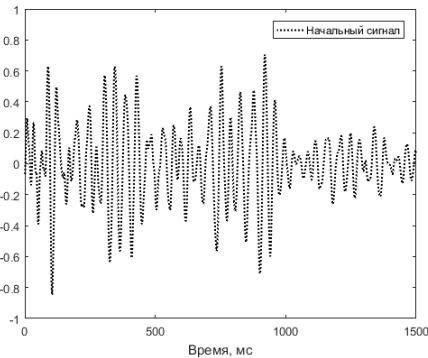


Рисунок 2 – Участок сеймотрассы 1500 – 3000 мс

На рисунке 3 показан начальный сигнал и его аппроксимация. По форме они практически идентичны. О высоком качестве аппроксимации говорят и значения критериев точности подбора, которые представлены в конце статьи.

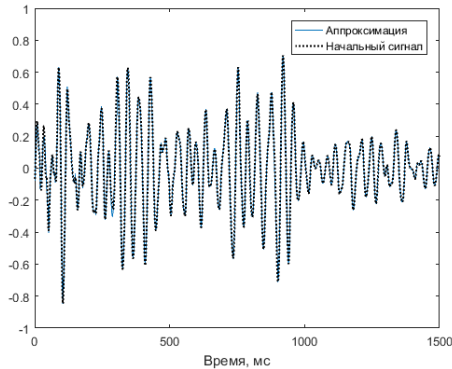


Рисунок 3 – Участок сейсмоотрассы и его аппроксимация параметрической моделью

Для того чтобы провести аппроксимацию участок протяженностью 1500 отсчетов был разбит на небольшие участки по 50 мс, каждый участок аппроксимировался отдельно и для него высчитывались коэффициенты параметрической модели и критерии точности подбора коэффициентов. После этого все участки объединялись в один.

Для того чтобы показать качество аппроксимации, приведены критерии точности подбора коэффициентов. Критерии точности подбора коэффициентов для реального сигнала (среднее значение):

- Сумма квадратов ошибок (SSE): 0.0123;
- Квадрат смешанной корреляции (R -square): 0.9994;
- Уточненный квадрат смешанной корреляции ($Adjusted R$ -square): 0.9994;
- Корень из среднего для квадрата ошибки ($RMSE$): 0.004599;
- Средняя абсолютная ошибка (MAE): 3.642e-04.

ЗАКЛЮЧЕНИЕ

Параметрическая модель позволяет аппроксимировать с высокой точностью реальный сейсмический импульс, возникающий при образовании трещин коллектора в результате ГРП.

Литература:

1. Модель сейсмического импульса, возникающего при гидравлическом разрыве пласта / Е.В. Рабинович, К.С. Ганчин, И.М. Пупышев, Г.С. Шефель // Математические структуры и моделирование. – 2014. №4. – С. 105 - 111

ИССЛЕДОВАНИЕ РАСПРЕДЕЛЕНИЙ СТАТИСТИК НЕПАРАМЕТРИЧЕСКИХ КРИТЕРИЕВ СОГЛАСИЯ ПО ВЫБОРКАМ БОЛЬШОГО ОБЪЁМА

Д.С. Халин, М.А. Семёнова

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

Объектом исследования являются критерии согласия, которые необходимо обновить для обработки выборок большого объема. Целью работы является разработка и внедрение алгоритма для применения критериев согласия к выборкам большого объема. В ходе работы были проведены исследования с целью выявления критических объемов выборок в различных экспериментальных условиях и выработки рекомендаций по применению критериев согласия Колмогорова, Крамера-Мизеса-Смирнова и Андерсона-Дарлинга на выборках большого объема. В результате работы были впервые определены особенности проверки гипотез о виде распределения в контексте выборок большого объема, разработан соответствующий алгоритм и сформулированы рекомендации по работе с ним.

The object of the study is the criteria of consent, which must be updated to process samples of large volume. The aim of the work is to develop and implement an algorithm for applying the acceptance criteria to large sample sizes. In the course of the work, studies were conducted with the aim of identifying critical sample sizes under various experimental conditions and developing recommendations on the application of the Kolmogorov, Kramer-Mises-Smirnov and Anderson-Darling consent criteria on large samples. As a result of the work, the features of testing hypotheses about the type of distribution in the context of large samples were first identified, a corresponding algorithm was developed, and recommendations for working with it were formulated.

Объемы данных растут с каждым днем, и иногда классические методы анализа данных начинают терпеть неудачу в таких условиях. Проблема исследуется применением критериев согласия к выборкам большого объема и заключается в том, что гипотеза, проверяемая на боль-

ших данных, отвергается, даже если она истинна. Основная причина этого – ограниченное количество значений наблюдаемой переменной, что часто бывает на практике. Чтобы решить эту проблему, мы предлагаем модифицировать существующие классические критерии согласия, такие как критерий Колмогорова, Крамера-Мизеса-Смирнова и Андерсона-Дарлинга, чтобы обеспечить возможность их применения к выборкам большого объема. Для этого мы должны провести несколько исследований.

Во-первых, необходимо исследовать распределения статистик непараметрических критериев согласия для определения критического объема выборки. Мы будем называть критическим объемом выборки такой объем, начиная с которого начинают получаться некорректные результаты проверки гипотезы. Мы нашли критические объемы выборок для всех трех критериев и для двух типов гипотез. Результаты показаны в таблице 1.

Таблица 1. Критические объёмы выборок в случае нормального распределения выборок и проверки простой и сложной гипотез.

| Распределение выборки | Тип гипотезы | Применяемый критерий | Крит. объём выборки |
|-----------------------|--------------|----------------------|---------------------|
| Нормальное | Простая | Колмогорова | 1000 |
| | | КМС | 7000 |
| | | Андерсона-Дарлинга | 9250 |
| | Сложная | Колмогорова | 500 |
| | | КМС | 2688 |
| | | Андерсона-Дарлинга | 2500 |

Далее, мы должны разработать алгоритм применения критериев согласия к выборкам большого объема. Он основан на разделении исходной выборки на подвыборки и проверке гипотезы на каждой подвыборке. Сам алгоритм заключается в следующем:

- 1) формирование выборки номеров наблюдений в исходной выборке, размера равного размеру подвыборки;
- 2) набор наблюдений из исходной выборки в подвыборку с номерами из сформированных на первом этапе выборок;
- 3) подсчет значений статистики по подвыборке;
- 4) повтор пунктов 1, 2 и 3 для всех подвыборок;

- 5) подсчет достигнутого уровня значимости;
- 6) принятие решения об отвержении или не отвержении гипотезы.

После этого мы должны выбрать метод вычисления достигнутого уровня значимости, который необходим, чтобы решить, следует ли отклонять гипотезу или нет. Мы решили рассчитать достигнутый уровень значимости как значение предельной функции распределения критерия в точке равной первому квартилю статистик по подвыборкам, так как в этом случае уровень значимости ведет себя адекватно и стабильно позволяет нам не отклонять верную нулевую гипотезу.

Определив выше критический объем выборок, мы, таким образом, задали максимально возможный объем подвыборок. Теперь же необходимо определить минимально возможный объем. Говоря о необходимости такого определения, будем руководствоваться тем, что объем подвыборки должен быть настолько большим, чтобы отвергать альтернативную гипотезу. Чтобы найти минимальные объемы, мы применим разработанный алгоритм для проверки гипотезы о нормальном распределении по распределенной по Лапласу образцу. Пример результатов показан в таблице 2.

Таблица 2. Результаты проверки простой гипотезы о нормальном распределении данных, имеющих распределение Лапласа, с помощью разработанного алгоритма на основе критерия Андерсона-Дарлинга.

| Объем подвыборки | Количество подвыборок | p-value |
|------------------|-----------------------|---------|
| 100 | 100 | 0,26 |
| | 200 | 0,21 |
| | 300 | 0,24 |
| ... | ... | ... |
| 300 | 100 | 0,03 |
| | 200 | 0,02 |
| | 300 | 0,03 |
| 350 | 100 | 0,00 |
| | 200 | 0,00 |
| | 300 | 0,00 |

После всех этих экспериментов мы можем дать рекомендации по применению разработанного алгоритма. Чтобы облегчить выбор разме-

ра подвыборки, объединим результаты нашего исследования в одной таблице.

Таблица 3. Интервалы объёмов подвыборок при проверке простой и сложной гипотез.

| Простая гипотеза | | | |
|--------------------|-------------|--------------------|-------------------------|
| Критерий | Колмогорова | Андерсона-Дарлинга | Крамера-Мизеса-Смирнова |
| Минимальный объём | - | 350 | 200 |
| Максимальный объём | 1000 | 9250 | 7000 |
| Сложная гипотеза | | | |
| Критерий | Колмогорова | Андерсона-Дарлинга | Крамера-Мизеса-Смирнова |
| Минимальный объём | 250 | 750 | 700 |
| Максимальный объём | 550 | 2500 | 2688 |

Стоит отметить, что не обязательно пользоваться минимальным или максимальным возможным объёмом подвыборок, однако точность полученных результатов будет тем выше, чем больше будет объём подвыборки. Тем не менее, в условиях ограниченности вычислительных и временных ресурсов может быть достаточно и минимальных объёмов подвыборок.

В ходе исследований было показано, что количество подвыборок не имеет большого влияния на принятие решения об отвержении или не отвержении гипотезы, а это значит, что у исследователя в этом вопросе появляется некоторая свобода. Поэтому мы рекомендуем взять как можно большее количество подвыборок, чтобы общее число наблюдений во всех подвыборках как минимум покрывало количество наблюдений в исходной выборке.

В заключение следует отметить, что проблемы применения критериев согласия к выборкам большого объёма были показаны впервые. Также были сделаны рекомендации по работе с такими выборками. Соответствующий алгоритм был разработан и реализован на языке програм-

мирования R. В будущем мы планируем оценить мощность разработанного алгоритма.

Литература:

1. Лемешко Б.Ю. Непараметрические критерии согласия : учеб. пособие / Б.Ю. Лемешко. – Новосибирск : Изд-во НГТУ, 2014. – 162 с.
2. Б.Ю. Лемешко. Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход: Монография / Б.Ю. Лемешко, С.Б. Лемешко, С.Н. Постовалов, Е.В. Чимитова. – Новосибирск : Изд-во НГТУ, 2011. – 888 с. (серия «Монографии НГТУ»)
3. Постовалов С.Н. Математическая статистика. Конспект лекций: учеб. пособие / С.Н. Постовалов, Е.В. Чимитова, В.С. Карманов. – Новосибирск: Изд-во НГТУ, 2014. – 140 с.
4. Лемешко Б.Ю., Постовалов С.Н. Прикладная статистика. Правила проверки согласия опытного распределения с теоретическим. Методические рекомендации. Часть II. Непараметрические критерии. – Новосибирск: Изд-во НГТУ, 1999. – 85 с.
5. Кобзарь А.И. Прикладная математическая статистика. Для инженеров и научных работников. / Кобзарь А.И. – М.: ФИЗМАТЛИТ, 2006. – 814 с.
6. Буховец А.Г. Статистический анализ данных в системе R. Учеб. пособие / А.Г. Буховец, П.В. Москалев, В.П. Богатова, Т.Я. Вирючинская; Под. ред. А.Г. Буховца – Воронеж: ВГАУ, 2010 – 124 с.

СРАВНЕНИЕ МЕТОДОВ ПАРАМЕТРИЧЕСКОЙ ИДЕНТИФИКАЦИИ МОДЕЛИ РАДИАЦИОННОГО ДАВЛЕНИЯ СОЛНЕЧНОГО ИЗЛУЧЕНИЯ ДЛЯ РЕШЕНИЯ ЗАДАЧИ ЭФЕМЕРИДНОГО ОБЕСПЕЧЕНИЯ

Ю.С. Четвертакова, О.С. Черникова

**Новосибирский государственный технический университет,
г. Новосибирск, julia_ch98@mail.ru**

На основе адаптивной модификации сигма-точечного фильтра Калмана проведено сравнительный анализ оценок параметров модели радиационного давления солнечного излучения на основе методов максимального правдоподобия и наименьших квадратов. Данный анализ методов оценивания позволил выявить наиболее точный метод оценки параметров.

Based on the adaptive modification of the Kalman sigma-point filter, a comparative analysis of the estimates of the parameters of the model of the radiation pressure of solar radiation is carried out on the basis of the maximum estimation and least squares methods. This analysis of assessment methods revealed the most accurate method for estimating parameters.

Качество эфемеридно-временного обеспечения для ГНСС технологий в значительной мере зависит от степени адекватности применяемых математических моделей, описывающих движение орбитальной группировки навигационных спутников (НС). При формировании таких моделей проблемным остается учет возмущений от радиационного давления на спутник солнечного излучения. Настоящая работа посвящена решению задачи идентификации модели радиационного давления на спутник солнечного излучения на основе применения современных математических методов.

Движение космического аппарата (КА) в инерциальной системе координат (ИСК) в гравитационном поле Земли можно описать следующей нелинейной непрерывно-дискретной моделью в пространстве состояний [1]:

$$\frac{d}{dt} \begin{pmatrix} r(t) \\ \dot{r}(t) \end{pmatrix} = f(r(t), \dot{r}(t), \theta) + w(t), \quad t \in [t_0, t_N],$$

(1)

$$s(t_{k+1}) = r(t_{k+1}) + v(t_{k+1}), \quad k = 0, 1, \dots, N-1.$$

(2)

Здесь $r(t)$ – вектор координат КА в ИСК; $f(\cdot)$ – нелинейная вектор-функция, учитывающая влияние на траекторию движения КА радиационного давления (РД) солнечного излучения, несферичности геопотенциала Земли, гравитационного воздействия Луны, Солнца и других планет; θ – вектор неизвестных параметров модели РД; $s(t_{k+1})$ – вектор измерения; $w(t)$ и $v(t_{k+1})$ – взаимно некоррелированные гауссовские векторы шумов системы и измерения соответственно с нулевыми математическими ожиданиями и ковариационными матрицами

$$E \left[w(t)w^T(\tau) \right] = Q(t)\delta(t - \tau), \quad E \left[v(t_{k+1})v^T(t_{i+1}) \right] = R(t_{k+1})\delta_{ki}.$$

Относительно функции $f(\cdot)$ в правой части уравнения состояния (1) заметим, что в настоящее время можно с незначительной погрешностью учесть несферичность геопотенциала и гравитационное влияние Луны, Солнца и других планет (см., например, [2]). Иначе обстоит дело с моделированием воздействия сил РД. Воспользуемся следующей моделью РД в объектоцентрической системе координат (ОСК) из [3]:

$$f_{RD} = \Lambda \cdot d \cdot r_{OS}^{-2} \cdot \left[x_{RP} \cdot (\theta_1 + \theta_2 \cos \kappa + \theta_3 \sin \kappa) + y_{RP} \cdot (\theta_4 + \theta_5 \cos \kappa + \theta_6 \sin \kappa) + z_{RP} \cdot (\theta_7 + \theta_8 \cos \kappa + \theta_9 \sin \kappa) \right]. \quad (3)$$

Данная модель применяется в центрах обработки международной службы International GNSS Service. Здесь Λ – фактор затмения; d – фактор, зависящий от формы НС, его массы, отражательной и поглощательной способности материалов его поверхности; r_{OS} – расстояние между спутником и Солнцем; x_{RP}, y_{RP}, z_{RP} – координаты НС в ОСК; κ – аргумент широты НС.

Традиционно для оценивания параметров стохастических моделей в пространстве состояний применяют широко известный расширенный фильтр Калмана. К сожалению, в нашем случае этот подход неприемлем. Это связано с тем, что вектор-функция $f(\cdot)$ вычисляется достаточно сложно, в результате чего получение аналитических соотношений, необходимых для вычисления матриц Якоби проблематично. Остроту

вопроса в значительной мере снимает появившийся относительно недавно так называемый сигма-точечный фильтр Калмана [4], который применяется для нелинейных моделей и не предполагает проведение линеаризации. В данной работе использована непрерывно-дискретная модификация сигма-точечного фильтра из [5], адаптированная на наш случай с неизвестными ковариационными матрицами шумов системы и измерений.

Возьмем в качестве измерительных данных срочные эфемериды НС GPS от 14.07.2016г., полученные международной ГНСС-службой. В этом случае НС совершает более одного оборота вокруг Земли (проходит различные световые зоны). Расчет скорости НС в начальный момент времени произведем на основе срочных эфемерид с использованием интерполяции Эверетта. Оценивание параметров РД в модели (3), осуществим с помощью метода максимального правдоподобия и метода наименьших квадратов по данным траекторных наблюдений на участках полной освещенности и зон полутени. Сравним результаты работы методов:

Таблица 1 – Параметрическая идентификация модели радиационного давления

| | Метод максимального правдоподобия | Метод наименьших квадратов |
|------------------------------|--|--|
| Полученные оценки параметров | $\hat{\mathcal{F}} =$ (0.9490839520011050 | $\hat{\mathcal{F}} =$ (1.084603160 -0.035818168 |
| | - | |
| | 0.0473959646317777 | 0.124205012 |
| | - | |
| | 0.0173076224412937 | 0.035883076 |
| | - | |
| | 0.1361002751129270 | 0.032498382 |
| | - | |
| | 0.0061020037824392 | 0.012132207 |
| | - | |
| 0.0621088711448467 | 0.004283445 | |
| - | | |
| 0.0182516308874188 | 0.015582352 | |
| - | | |
| 0.1340714811990970 | 0.051604911) | |

| | | |
|-----------------------|------------------------------|------------------------------|
| | 0.1171232997331040) | |
| Точность оценивания * | $\mathcal{E}_S = 3.6685e-08$ | $\mathcal{E}_S = 3.7528e-08$ |

* Для сравнения качества прогнозирования траекторий орбитального движения спутника PG01 с финальной эфемеридой от 15.07.2016 г. вычислим \mathcal{E}_S для обоих методов по формуле:

$$\mathcal{E}_S = \frac{\sqrt{\sum_{k=0}^{N-1} \|s(t_{k+1}) - \mathcal{E}_P(t_{k+1})\|^2}}{\sqrt{N}}$$

где $\{s(t_{k+1}), k = 0, 1, \dots, N-1\}$ – финальная эфемерида НС;
 $\{\mathcal{E}_P(t_{k+1}), k = 0, 1, \dots, N-1\}$ – спрогнозированная траектория при $\theta = \theta$.

Проведенные исследования показали целесообразность применения методов параметрической идентификации (метода максимального правдоподобия) при прогнозировании траектории движения космического аппарата.

Работа выполнена при финансовой поддержке Министерства образования и науки РФ по государственному заданию (проект № 2.7996.2017/8.9).

Литература:

1. Дубошин Г.Н. Небесная механика. Методы теории движения искусственных небесных тел. М.: Наука, 1983. 352 с.
2. Бордовицына Т.В., Авдюшев В.А. Теория движения искусственных спутников Земли. Аналитические и численные методы : учеб. пособие. Томск : Изд-во Том. ун-та, 2007. 178 с.
3. Kouba J. A Guide to using international GNSS Service (IGS) Products. Ottawa, 2009. 34 p.
4. Julier S.J., Uhlmann J.K. A new extension of the Kalman filter to nonlinear systems // Proc. Of AeroSense: 11-th Int. Symp. on Aerospace/Defence Sensing, Simulation and Control, 1997. 12p.
5. Särkkä S. On unscented Kalman filtering for state estimation of continuous-time nonlinear systems // IEEE Transactions on Automatic Control, 2007. 11p.

АНАЛИЗ СИГНАЛОВ ЭЛЕКТРОМАГНИТНОГО ИЗЛУЧЕНИЯ СОПУТСТВУЮЩИХ РАЗРУШЕНИЮ ГОРНЫХ ПОРОД

Шапорева Г.О., Бизяев А.А., Яковицкая Г.Е.¹

¹Институт горного дела им. Н.А. Чинакала СО РАН
Новосибирский государственный технический университет,
г. Новосибирск, shganna@bk.ru, bizyaev@corp.nstu.ru

В работе представлены результаты экспериментального исследования параметров сигнала электромагнитного излучения сопутствующих разрушению горных пород. В ходе эксперимента исследовались образцы горной породы, которые при одноосном сжатии разрушались до полной потери несущей способности. Синхронно регистрировались сигналы электромагнитного излучения, нагрузки и перемещения. В структуре сигнала электромагнитного излучения выявлены параметры сигнала, соответствующие различным стадиям разрушения горной породы.

The paper presents the results of an experimental study of the electromagnetic emission signal parameters connected with the destruction of rocks. In the experiment, we investigated the rock samples, which were destroyed to the complete loss of bearing capacity in uniaxial compression. Signals of electromagnetic emission (EME), load and displacement were recorded synchronously. The parameters of the signal corresponding to different stages of rock failure are revealed in the structure of the electromagnetic emission signal.

Одним из перспективных направлений прогноза динамических проявлений горных пород является бесконтактный метод, основанный на интерпретации сигналов электромагнитного излучения (ЭМИ) [1], сопутствующего процессу разрушения горной породы. Для исследования сигналов ЭМИ в лабораторных условиях разрушались образцы горных пород, различных по своему геологическому составу, путем одноосного сжатия до полного разрушения. В процессе разрушения синхронно регистрировалась, нагрузка и перемещение, по которым определялись стадии нагружения и деформация, синхронно механическими параметрами регистрировался сигнал ЭМИ. В ходе экспериментальных исследований выявлены параметры сигнала ЭМИ характерные различным стадиям нагружения. Интерпретация стадий разрушения образцов горных пород по параметрам ЭМИ соответствует концентрационной модели [2]. Показано, что на различных стадиях разрушения образцов горных пород, меняются спектр сигнала, интенсивность одиночных импульсов, энергия сигнала в единицу времени, что соответствует стадиям схлопывания микротрещин, их концентрации в магистральную трещину

и её росту. Выявлены параметры сигнала ЭМИ, соответствующие лавинообразному разрушению горной породы характеризующую стадию предразрушения. Показано, что для различных, по механическим и геологическим свойствам образцов горных пород параметры сигналов ЭМИ отличаются, но характерность изменения стадий остается постоянной, что можно использовать как общий критерий при построении оборудования прогнозирования динамических проявлений горного давления по сигналам ЭМИ.

Таким образом, показано, что применение метода ЭМИ для прогноза изменения напряженно-деформированного состояния массива горных пород требует предварительного анализа геологического состава рудного тела.

Литература:

1. Курленя М.В., Вострецов А.Г., Кулаков Г.И. и др. Регистрация и обработка сигналов электромагнитного излучения горных пород. Новосибирск: СО РАН, 2000. 230 с.

2. Куксенко В.С., Махмудов Х.Ф., Манжиков Б.Ц.. Концентрационная модель разрушения твердых тел и прогнозирование катастрофических ситуаций крупномасштабных объектов // Физико-технические проблемы разработки полезных ископаемых, №. 4, 2010. С. 29-40.

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ КОНЦЕНТРАЦИОННЫХ ПРЕДЕЛОВ

Д.А. Шевченко, А.Л. Осипов
Новосибирский государственный университет
экономики и управления,
г. Новосибирск, alosip@mail.ru

Исследованы математические модели для предсказания концентрационных пределов воспламенения химических веществ различных классов, основанные на использовании информационных и зарядовых дескрипторов. Данные типы дескрипторов вычисляются автоматически по структурной формуле вещества. Установлено, что полученные в статье результаты для прогнозирования кон-

центрационных пределов воспламенения дают незначительную погрешность по сравнению с экспериментом.

Mathematical models for predicting the concentration limits of the ignition of chemicals of various classes, based on the use of information and charge descriptors, are investigated. These types of descriptors are calculated automatically by the structural formula of the substance. It is established that the results obtained in the article for predicting the concentration limits of ignition give an insignificant error in comparison with the experiment.

В статье исследуется новый подход к предсказанию концентрационных пределов воспламенения химических веществ, который основан на использовании зарядовых и информационных дескрипторов.

В работах [1- 2] были представлены исследования по предсказанию пожароопасных свойств химических веществ на ограниченных объемах выборки для определенных классов с помощью регрессионных и нейросетевых моделей [3-4].

Целью работы является создание эффективных моделей предсказания концентрационных пределов для различных классов химических веществ. Для выполнения этой цели решаются следующие задачи: автоматическое вычисление зарядовых и информационных дескрипторов, связанных с пожарными свойствами; разработка регрессионных моделей установления связи между этими дескрипторами и концентрационными пределами; создание компьютерных программ, для реализации моделей; создание баз данных, которые содержат экспериментальные данные для проверки эффективности разработанных моделей и выбранного способа описания.

С целью выявления связи пожароопасных свойств химических соединений с их структурой предлагается метод, использующий в качестве факторного признака молекулы – среднее число электронов на

внешней оболочке атомов в молекуле: $Z = \sum_{i=1}^N n_i Z_i / N$, где n_i – число

атомов i -ого сорта с числом валентных электронов Z_i (т.е. число электронов на внешней оболочке i -ого атома).

Суммирование выполняется по всем атомам в молекуле; N – общее число атомов. Дальнейшее исследование связи структура – пожароопасные свойства связано с информационной функцией Шеннона:

$H = -\sum_{i=1}^N p_i \log_2 p_i$, где $p_i = n_i / N$, причем для p_i выполняются со-

отношения: $0 \leq p_i \leq 1$, n_i – число атомов сорта i в молекуле, N – общее число атомов в молекуле. Отношение n_i / N дает доленое участие i -ого сорта атома в молекуле.

Для выборки из 763 химических элементов различного класса [3] было вычислено количество атомов, доленое участие каждого атома в структуре вещества, а затем вычислено среднее число электронов на внешней оболочке атомов в молекуле Z .

Для поиска зависимостей было отобрано 649 веществ из указанной выборки. Оставшаяся часть выборки (114 элементов) была использована в качестве тестовой для прогноза нижнего концентрационного предела воспламенения (НКПВ). Для прогнозирования зависимости НКПВ от информационного фактора H , на этих же экзаменационной и тестовой выборках были автоматически вычислены доленое участие каждого атома в молекуле и значение $\log_2 p_i$, а затем вычислено значение информационной функции Шеннона.

Для прогнозирования зависимости НКПВ от Z и от H было разработано по 9 уравнений регрессии: линейная ($y = a + bx$); квадратичная ($y = a + bx + cx^2$); кубическая ($y = a + bx + cx^2 + dx^3$); логарифмическая ($y = a + b \ln x$); гиперболическая ($y = a + \frac{b}{x}$); степенная ($y = ax^b$); показательная ($y = ab^x$); S-кривая ($y = \exp(a + \frac{b}{x})$); экспоненциальная ($y = ae^{bx}$).

По каждому из этих уравнений были вычислены средние квадратические ошибки, выбраны модели с наименьшей ошибкой. В обучении была использована выборка объемом 649 элементов.

Была вычислена средняя квадратическая ошибка по каждой модели. Получены следующие наилучшие модели: для зависимости НКПВ от Z : линейная функция: $y = 0,972Z - 0,918$ и квадратичная:

$y = 0,2Z^2 - 0,098Z + 0,467$; для зависимости НКПВ от H : линейная

функция: $y = 1,786H - 0,774$ и квадратичная:

$$y = 0,905H^2 - 0,529H + 0,66.$$

Полученные модели были протестированы на экзаменационной выборке, состоящей из 114 элементов, и квадратичные модели показали наилучшую ошибку – 0,529 и 0,481 для зависимости по Z и по H соответственно.

Предложены и исследованы регрессионные модели для прогнозирования концентрационных пределов воспламенения химических веществ. Показано, что разработанные регрессионные модели, в которых используются информационные и зарядовые дескрипторы, дают качественный прогноз концентрационных пределов воспламенения.

Литература:

1. Осипов А.Л., Криветченко О.В. Компьютерная оценка нижнего концентрационного предела воспламенения химических веществ // В мире научных открытий. - Красноярск: Научно-инновационный центр. - 2013. - №10.1(46) (Математика. Механика. Информатика). - С. 34-45.
2. Осипов А.Л., Криветченко О.В. Математическое моделирование концентрационных пределов воспламенения // International Journal of Advanced Studies. - Volume 8. - № 1-2 - 2018. - P. 135-141.
3. Осипов А.Л., Трушина В.П. Информационная система с Web-интерфейсом для расчета показателей пожарной опасности// International Journal of Advanced Studies. - Volume 7. - № 4-2. - 2017. - P. 85-92.
4. Осипов А.Л., Трушина В.П., Павлик И.О. Моделирование концентрационных пределов на основе нейронных сетей // International Journal of Advanced Studies. - 2016. - Т. 6. - № 2. - С. 67-78.

МЕТОД ПОСЛЕДОВАТЕЛЬНОГО ОБНАРУЖЕНИЯ РАЗЛАДКИ ВО ВРЕМЕННЫХ РЯДАХ

А.О. Шерстобитова, Т.В. Емельянова
Национальный исследовательский Томский
государственный университет,
г. Томск, annasherstobitova06@gmail.com

Рассматривается проблема разделения временных рядов произвольной природы X_t (стохастических, детерминированных или смешанных) на сегменты, порожденные одним механизмом, а также обнаружения момента смены τ одного механизма генерации другим.

We consider the problem of separating time series of arbitrary nature X_t (stochastic, deterministic or mixed) into segments generated by one mechanism, as well as detecting the moment of changing τ one generation mechanism by another.

В задачах обработки временных рядов часто встречается ситуация, когда временной ряд порожден различными механизмами генерации, однако у исследователя нет никакой априорной информации об этих механизмах. Для извлечения адекватной информации из массива данных необходима предварительная сегментация ряда на однородные подмассивы данных, без чего нельзя строить математические модели, оценивать параметры и т.д.

Пусть $X = \{x(t)\}_{t=1}^N$ - временной ряд с неизвестными моментами изменения механизма генерации $t_i, i = 2, \dots, k$. Механизмы генерации ряда неизвестны и могут быть стохастическими, детерминированными или смешанными. Отрезки ряда $[t_i, t_{i+1}], t_1 = 1, t_{k+1} = N$, которые генерируются одним и тем же механизмом, называем однородными. Б. С. Дарховским была сформулирована и доказана следующая теорема [1], согласно которой связь ε -сложности функции гельдерова класса, заданной конечным набором своих отсчетов на равномерной решетке, с погрешностью аппроксимации ε характеризуется парой действительных чисел (A, B) , называемых коэффициентами сложности [1].

Ключевая идея предлагаемой методологии заключается в предположении о том, что на i -м сегменте однородности $[t_i, t_{i+1}]$ временного ряда X для $t_i \leq t, t + n < t_{i+1}$ коэффициенты сложности удовлетворяют соотношению

$$R(j) = R_i + \xi^i(j),$$

где R_i – математическое ожидание последовательности $R(j)$ на отрезке $[t_i, t_{i+1}]$, а $\xi^i(j)$ – последовательность случайных величин с нулевым математическим ожиданием. Иными словами, средние значения коэффициентов ε -сложности временного ряда остаются постоянными на интервалах его однородности, а изменение механизма генерации ряда в моменты $t_i, i = 2, \dots, k$, приводит к изменению этих средних значений.

Таким образом, если предположение справедливо, задача сегментации временного ряда сводится к задаче обнаружения «разладки» по среднему значению в диагностической векторной последовательности $R(j)$.

Для решения последней задачи предлагается использовать следующее семейство статистик

$$Y(s, \delta) = \left(\frac{(N_1 - s)s}{N_1^2} \right)^\delta \left(\frac{1}{s} \sum_{k=1}^s z(k) - \frac{1}{N_1 - s} \sum_{k=s+1}^{N_1} z(k) \right), \quad (1)$$

где $0 \leq \delta \leq 1, 1 \leq s \leq N_1 - 1, N_1 = \left[\frac{N}{n} \right], Z = \{z(k)\}_{k=1}^{N_1}$ – реализация компонент диагностической последовательности R [2].

На практике часто необходимо знать номер наблюдения, в котором происходит смена механизма генерации либо фиксируется аномальное поведение данных, не характерное для наблюдаемого временного ряда. В связи с этим предлагается модификация метода сегментации Б. С. Дарховского, которая позволяет получить конкретный номер наблюдения, с зафиксированной аномалией. В качестве семейства F методов аппроксимации используются кусочно-полиномиальные функции до 10 порядка включительно. В каждом эксперименте генерируется временной ряд, затем выполняется разделение полученного временного ряда на «окна» размера n . Каждое «окно» обрабатывается методом наименьших квадратов и определяются коэффициенты зависимости (A, B) по значениям $\log \varepsilon, \log S$. Далее выполняется обнаружение момента изменения механизма генерации («разладки») ряда. При помощи алгоритма, основанного на статистике (1), определяются моменты «разладки».

Рассматривается модель ARCH(2) вида

$$X_t = \varepsilon_t \sqrt{\alpha_0 + \sum_{i=1}^q \alpha_i X_{t-i}^2}, \quad (2)$$

где $\varepsilon \sim N(0,1)$, $\alpha = (0.5, 0.3, 0.6)$. Объем выборки для эксперимента выберем равным $N = 1\,000$, размер окна $n = 100$. Полученный ряд имеет неравномерное поведение, поэтому в качестве аномалий выбираются скачки, намного превышающие среднее значение ряда.

Будем применять к временному ряду метод сегментации Б. С. Дарховского, в результате чего, получаем окна 2, 3, 5, 6 с возможными аномалиями. Далее будем рассматривать каждое идентифицированное окно в пересечении с соседними следующим образом

$$X_t = \{x_i \in X: i = \overline{nt - 50 : n(t+1) + 50}\}, \quad (3)$$

где t – идентифицированный номер окна. Таким образом, будем рассматривать четыре полученных временных ряда для обнаружения аномального наблюдения. Для этого необходимо рассматривать каждое наблюдение в новом сгенерированном ряде отдельно и проверять следующие 3 условия

$$X_{1i} > \overline{X_{1i-1}} + l; X_{1i} > X_{1i-1} + l; X_{1i} > X_{1i+1} + l, \quad (4)$$

где l – контролирующий коэффициент.

Таким образом, если наблюдение в окне удовлетворяет 3 условиям из (4), то наблюдение принимается в качестве аномального.

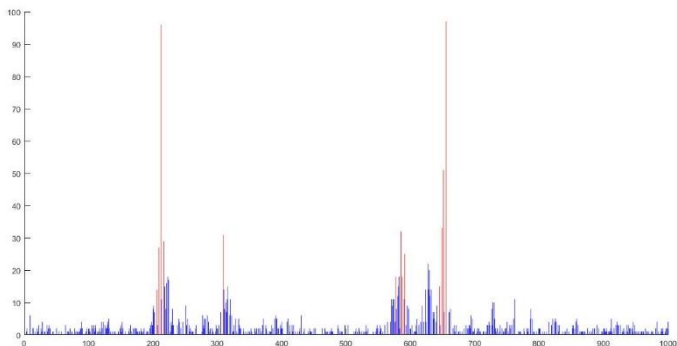


Рис. 1. Обнаружение аномальных наблюдений.

На рисунке красным обозначены обнаруженные аномальные наблюдения.

Таким образом, алгоритм сегментации Б. С. Дарховского позволяет не только выполнить сегментацию временного ряда на подмассивы, порожденные различными механизмами генерации, но и выполнить обнаружение аномалий во временном ряде.

Литература:

1. Дарховский Б.С., Пирятинская А. Новый подход к проблеме сегментации временных рядов произвольной природы, Тр. МИАН, 2014, т. 287, с. 61-74
2. Brodsky В.Е., Darkhovsky В. S. Non-parametric statistical diagnosis: Problems and methods. Dordrecht: Kluwer, 2000.

АЛГОРИТМ ОПРЕДЕЛЕНИЯ ДЕФЕКТОВ СВЧ МИС С ИСПОЛЬЗОВАНИЕМ ГИСТОГРАММ УЧАСТКОВ МИКРОФОТОГРАФИИ

Б.В. Ширяев, А.Ю. Ющенко, А.В. Безрук
Научно-исследовательский институт
полупроводниковых приборов,
г. Томск, shiryayev_bv@niipp.ru

В статье представлен алгоритм автоматизированного визуального контроля МИС. Микрофотография разбивается на участки в соответствии с чертежом фотошаблона, формируются их гистограммы. Отклонения среднеквадратического значения от среднего значения и количество существенных максимумов гистограммы свидетельствуют о наличии дефектов в соответствующих областях.

In the article an algorithm present for automated visual control MIC. Micrograph is divided into sections according to the drawing of the photomask, their histogram is created. Deviations of the RMS value from the mean and number of histogram peaks indicate defects in the areas.

Визуальный контроль при производстве СВЧ МИС заключается в контроле качества изготавливаемых устройств по их внешнему виду с использованием специальных оптических приборов. Изготовленное устройство по визуальному признаку оценивается в нескольких категориях, наиболее важные из которых точность воспроизведения критических размеров элементов (точность повторения фотошаблона устройства), качественный анализ поверхностей элементов СВЧ МИС и отсутствие повреждений и загрязнений.

Автоматизация визуального контроля заключается в создании программного продукта, который повышает скорость и качество детектирования дефектов СВЧ МИС. В качестве основы алгоритма выступила гистограммная обработка участков микрофотографий СВЧ МИС [1-2].

Алгоритм определения дефектов СВЧ МИС построен на сравнении фотографии устройства, полученной в автоматическом режиме, с чертежом фотошаблона одного устройства в электронном формате (далее чертёж), который применяется для его изготовления. В идеальном случае чертёж и фотография устройства должны полностью совпадать, в реальности они имеют незначительные отличия из-за особенности технологии изготовления и контроля: искажённая форма элементов, рассовмещение слоёв устройства, механические повреждения (царапины,

касания зондов), недостаточное разрешение камеры визуального контроля и другие (рис.1).

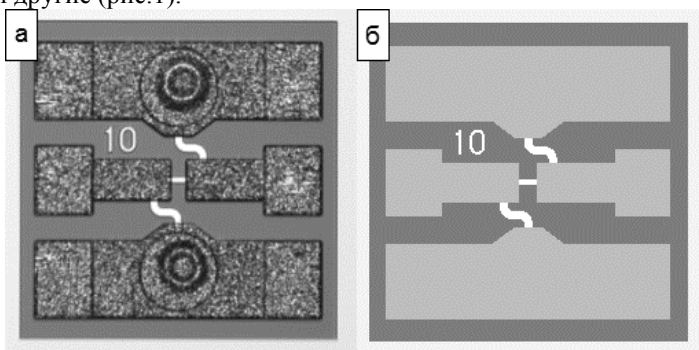


Рис. 1. Изображение готового устройства (а) и маски, полученная из чертежа устройства (б)

Общая последовательность алгоритма определения дефектов сведена к выполнению следующих этапов:

1) поиск двух точек совмещения, заранее указанных в алгоритме визуального контроля (противоположные углы СВЧ МИС);

2) расчёт масштаба, положения и состояния чертежа с использованием найденных координат точек совмещения на микрофотографии устройства, и их зафиксированное положение на чертеже;

3) растеризация векторного чертежа, объединение групп пикселей, характеризующих определённые области на чертеже (область положки, область металлизации и другие);

4) выделение группы пикселей подобных областей из изображения устройства в соответствии с областями фотошаблона. Области выделения подготавливаются на этапе настройки алгоритма;

5) построение гистограмм для каждой группы пикселей;

6) анализ полученных гистограмм: при удовлетворении условия годности устройство помечается как годное, при несоответствии – брак. Условие годности формируется на подготовительном этапе настройки алгоритма;

7) выделение пикселей, не удовлетворяющих условию годности позволяет определить положение дефекта.

В основе алгоритма лежит оценка гистограмм областей микрофотографии. Предполагается, что распределение количества пикселей от

интенсивности каждой компоненты цветного пространства RGB принимает нормальную форму. Для этих распределений рассчитывается среднее значение и среднеквадратическое отклонение (рис. 2).

Вычисление среднего значения позволяет определить смещение цвета всей области анализа, что повышает качество определение дефектов. Это относится к прозрачным плёнкам, из-за зависимости их цвета от толщины. Фиксация допустимого отклонения среднего значения цвета области задаёт максимальное отклонение толщины прозрачной плёнки.

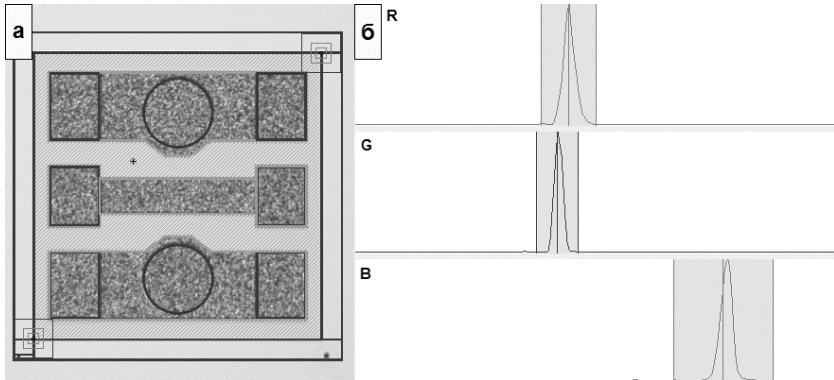


Рис. 2. Схематическое изображение выделенной области (а) и цветовая гистограмма её пикселей (б)

Вычисление среднеквадратического отклонения позволяет оценить область на неоднородность покрытия. Форма гистограммы и количество существенных максимумов, уровень которых заранее задан на этапе подготовки, свидетельствуют о наличии дефектов в области анализа.

Тестирование разработанного алгоритма показало, что средняя скорость процесса автоматизированного визуального контроля по сравнению с ручным увеличивается в 8-10 раз. Совпадение результатов предложенного алгоритма и ручного визуального контроля составляет 93%.

Литература:

1. Форсайт Д., Понс Ж. Компьютерное зрение. Современный подход. – М.: Издательский дом «Вильямс», 2004. – 928 с.
2. Лукьяница А.А., Шишкин А.Г. Цифровая обработка видеоизображений. – М.: «Ай-Эс-Эс Пресс», 2009. – 518 с.

**ПОСТРОЕНИЕ ПОВЕРХНОСТЕЙ СЛОЕВ ГЕОЛОГИЧЕСКОЙ
МОДЕЛИ НЕФТЯНОГО МЕСТОРОЖДЕНИЯ
ДЛЯ МОДЕЛИРОВАНИЯ ПРОЦЕССА
МНОГОФАЗНОЙ ФИЛЬТРАЦИИ**

А.М. Гриф, М.Г. Персова
Новосибирский государственный технический университет,
г. Новосибирск, alexgrif@inbox.ru

Рассматривается подход к построению поверхностей слоев геологической модели с использованием сглаживающих функций с адаптивной регуляризацией. Приводятся результаты моделирования, подтверждающие эффективность данного подхода.

This paper presents an approach to building surfaces of layers of a geological model using smoothing functions with adaptive regularization. The results of modeling confirming the effectiveness of this approach are given.

Введение. Качество управления технологическим процессом разработки нефтяного месторождения напрямую зависит от адекватности строящихся или адаптируемых геолого-гидродинамических моделей [1, 2]. Зачастую, подобные модели на макро-уровне задаются в виде набора слоев, имеющих различные петрофизические характеристики. Опорными точками таких слоев являются точки траектории инклинометрии скважины. Граничные точки задаются на опорной траектории пользователем-экспертом с учетом результатов геофизического исследования скважин. Построение поверхностей слоев с использованием триангуляции Делоне или формального применения сглаживающих сплайнов без использования сильной регуляризации может приводить к пересечениям границ слоев по вертикали и, как следствие, к некорректной геометрии сеток для моделирования. В работе рассмотрен подход к построению набора слоев геологической модели с помощью сглаживающих функций и адаптивной процедуры удаления возможных «перехлестов».

Математический аппарат. Пусть для первичного слоя геологической модели задан набор координат $\{\tilde{x}_j, \tilde{y}_j, \tilde{z}_j\}$, где \tilde{z}_j – значения функ-

ции глубины $z(x, y)$ в точке траектории скважины с координатой $(\tilde{x}_j, \tilde{y}_j)$, $j = \overline{1, k}$.

Построим линейную сглаживающую функцию $P(x, y)$, аппроксимирующую функцию z . Представим P в виде кусочно-линейной функции

$$P(x, y) = \sum_{i=1}^n q_i \psi_i$$

где $\psi_i(x, y)$ – глобальные линейные базисные функции [3]. Данная функция определена на регулярной сетке $X \times Y$.

Коэффициенты $\{q_i\}$ будем находить из условия минимизации функционала

$$F(\mathbf{q}) = \sum_{j=1}^k \omega_j (P(\tilde{x}_j, \tilde{y}_j) - \tilde{z}_j)^2 + \sum_{i=1}^n \alpha_i \int_{\Omega_i} \left(\frac{\partial P(x, y)}{\partial x} \right)^2 + \left(\frac{\partial P(x, y)}{\partial y} \right)^2 d\Omega.$$

В результате дифференцирования функционала $F(\mathbf{q})$ по каждому искомому параметру q_i и приравнивания частных производных к нулю, получим СЛАУ

$$\mathbf{A}\mathbf{q} = \mathbf{b},$$

где матрица \mathbf{A} и коэффициенты правой части \mathbf{b} определяются соотношениями

$$\begin{aligned} \mathbf{A} = & \sum_{j=1}^k \omega_j \psi(\tilde{x}_j, \tilde{y}_j) \psi^T(\tilde{x}_j, \tilde{y}_j) + \\ & + \sum_{i=1}^n \alpha_i \int_{\Omega_i} \left(\frac{\partial \psi(x, y)}{\partial x} \frac{\partial \psi^T(x, y)}{\partial x} + \frac{\partial \psi(x, y)}{\partial y} \frac{\partial \psi^T(x, y)}{\partial y} \right) d\Omega, \\ \mathbf{b} = & \sum_{j=1}^k \omega_j \psi(\tilde{x}_j, \tilde{y}_j) \tilde{z}_j. \end{aligned}$$

Для границ первичных слоев построим функции $P_m(x, y)$, $m = \overline{1, M}$. Причем коэффициенты регуляризации α_i выберем достаточно малыыми.

Далее для пар соседних границ слоев (1-2, 2-3, и т.д.) будем производить следующие действия.

1. Для всех точек сетки $(x_p, y_p) \in X \times Y$ вычислим значения $p_1 = P_1(x_p, y_p)$ и $p_2 = P_{1+1}(x_p, y_p)$.

2. Для тех точек (x_p, y_p) , в которых $p_1 \ll p_2$, увеличим значения регуляризации α_i для содержащих эти точки конечных элементов $\{\Omega_i\}$ (для функций P_1 и P_{1+1}).

Будем повторять данное действие для всех слоев до тех пор, пока не останется пересечений.

Результаты моделирования. Тестирование предложенного метода проводилось на данных реального нефтяного месторождения. На рис. 1 представлена развертка по скважинам для небольшой области месторождения, включающей 20 скважин. По данной области были заданы 11 геологических слоев (включая нефтенасыщенный, водонасыщенный, непроницаемый пропласток).

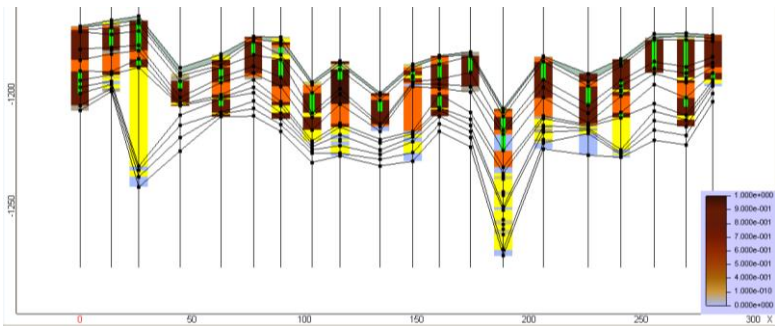


Рисунок 1 – Развертка скважин геологической модели с заданными граничными точками

По выбранным граничным точкам были построены поверхности слоев. На рис. 2 представлен разрез геологической модели с сеткой для гидродинамического моделирования, построенной предложенным способом.

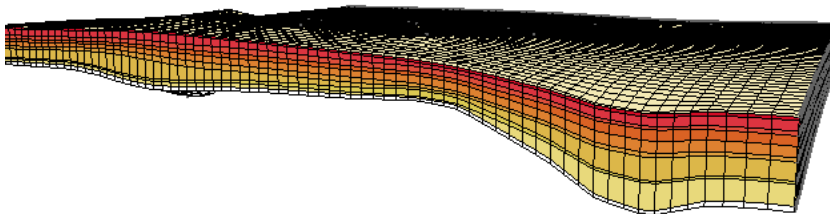


Рисунок 2 – Разрез геологической модели

На рис. 3 представлен разрез геологической модели в окрестности скважины №5 для трех способов построения: использование малой регуляризации ($\alpha = 0.01$), использование средней регуляризации ($\alpha = 0.1$) и с применением предложенного подхода.

Из рисунка видно, что использование малой регуляризации при построении геологической модели не гарантирует отсутствие пересечения границ: в модели пересекаются предпоследний и последний слой. Использование средней регуляризации избавляет геологическую модель от пересечений границ, но, вместе с тем, усредняет толщины слоев, что влечет за собой неверное построение гидродинамической модели (например, вместе с утончением нефтенасыщенных и водоносных слоев происходит утолщение пропластков). Применение адаптивной регуляризации устраняет пересечение границ, а так же не влечет к значимому изменению толщин целевого горизонта.

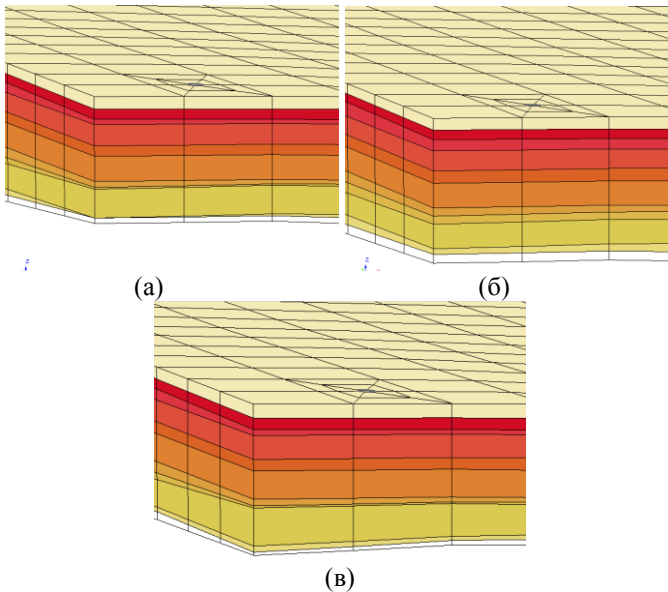


Рисунок 3 – Разрез модели с использованием малой (а), средней (б) и предложенной адаптивной регуляризации (в)

Работа выполнена при финансовой поддержке гранта Президента Российской Федерации для государственной поддержки молодых российских ученых – докторов наук (№ гранта МД-3516.2017.5).

Литература:

1. Clemens T. Optimizing Water-Injection Design in a Shallow Off-shore Reservoir / Clemens T., Kienberger G., Persaud M., Suri A., Sharma M.M., Boschi M., Overland A.M. // SPE production & operations – 2017 – Vol. 32 – №4 – pp. 551-563
2. Flow balancing in FEM modelling of multi-phase flow in porous media/ Marina G. Persova, Yuri G. Soloveichik, Alexander M. Grif, Ilya I. Patrushev // Актуальные проблемы электронного приборостроения (АПЭП–2018) = Actual problems of electronic instrument engineering (AREIE–2018): междунар. науч.-техн. конф., Новосибирск 2-6 окт. 2018 г. : в 12 т. – Новосибирск : Изд-во НГТУ, 2018. – С. 205-211.
3. Соловейчик Ю.Г., Рояк М.Э., Персова М.Г. Метод конечных элементов для решения скалярных и векторных задач : учеб. пособие / Сер. «Учебники НГТУ», 2007. - 896 с.

ИССЛЕДОВАНИЕ ВЛИЯНИЯ ТЕХНОГЕННЫХ ОБЪЕКТОВ НА РАСЧЁТ НАЧАЛЬНОГО ПОЛЯ ДЛЯ РЕШЕНИЯ ЗАДАЧ ГЕОЭЛЕКТРИКИ

Р.В. Гумалевский, Ю.Г. Соловейчик
Новосибирский государственный технический университет,
г. Новосибирск, gumalevskij@mail.ru

Работа посвящена разработке программного обеспечения, реализующее расчёт начального поля для решения задач геоэлектрики в областях с техногенными объектами.

The work is devoted to the development of software that implements the calculation of the initial field for solving the problems of geoelectrics in areas with man-made objects.

С развитием технологий добычи трудноизвлекаемых запасов нефти появилась возможность извлекать нефть, оставшуюся возле освоенных месторождений или в освоенных месторождениях. Перед самой добычей необходимо провести геофизические исследования, чтобы убедить-

ся в целесообразности развития этого месторождения. На этапе геофизического исследования возникает проблема точного расчёта характеристик электромагнитного поля в расчётной области, вследствие присутствия в данной области различных техногенных объектов, в частности, магистральных транспортных трубопроводов.

Таким образом, при решении задач электроразведки появляется необходимость учитывать наличие таких объектов в расчётной области, их местоположение, размеры и параметры.

К настоящему времени разработано множество технологий электроразведки, различающихся способом возбуждения и приёма поля. Рассмотрим технологии с так называемым контролируемым источником. Для возбуждения постоянного поля обычно используют незаземлённую токовую петлю.

В данной работе представлены алгоритмы и вычислительные схемы для моделирования расчёта начального магнитного поля с учётом техногенных объектов в расчётной области.

В данном случае в трёхмерной задаче магнитостатики напряженность магнитного поля \vec{H} ищется в виде:

$$\vec{H} = \vec{H}^c - \text{grad } U, \quad (1)$$

где \vec{H}^c – напряженность магнитного поля от токов \vec{j} в вакууме, определяемая соотношением по закону Био-Савара [1]

$$\vec{H}^c(r) = \frac{1}{\mu_0} \vec{B}^c(r) = \frac{1}{4\pi} \int_{\Omega} \frac{\vec{j}^c(r') \times (r-r')}{\|r-r'\|^3} dr', \quad (2)$$

а U – неполный скалярный магнитный потенциал. При этом $\text{rot} \vec{H}^c = \vec{j}$ по определению \vec{H}^c , и тогда, учитывая, что $\text{rot grad } U = 0$, убеждаемся, что \vec{H} автоматически удовлетворяет уравнению $\text{rot} \vec{H} = \vec{j}$:

$$\text{rot} \vec{H} = \text{rot} \vec{H}^c - \text{rot grad } U = \vec{j}. \quad (3)$$

Поэтому для нахождения поля \vec{H} достаточно решить одно скалярное уравнения с одной неизвестной скалярной функцией U :

$$\text{div}(\mu \vec{H}^c) - \text{div}(\mu \text{ grad } U) = 0, \quad (4)$$

в котором μ – относительная магнитная проницаемость среды. При представлении \vec{H} в виде (1) это уравнение полностью эквивалентно $\text{rot} \vec{B} = 0$.

Уравнение (4) удобнее переписать в виде:

$$-\operatorname{div}(\mu \operatorname{grad} U) = -\operatorname{div}(\mu \vec{H}^c). \quad (5)$$

Уравнение (5) будем решать с помощью метода конечных элементов [2]. Для этого разобьём расчётную область Ω на ячейки Ω_s [3]. В каждой ячейке будем считать магнитную проницаемость среды постоянной. В результате конечноэлементной аппроксимации получаем в расчётной области распределение скалярного поля U . Можно посчитать вектор магнитной индукции $\vec{B} = \mu_0 \mu \vec{H}$ с использованием соотношения (1).

При построении конечноэлементных аппроксимаций были рассмотрены различные возможности учёта источника электромагнитных полей (незаземлённой токовой петли): по закону Био-Савара, описывающему магнитное поле токов в проводниках, и пересчёт осесимметричного поля от петлевого источника на трёхмерную область.

Была проведена верификация программы для задачи магнитостатики с неполным скалярным потенциалом путём сравнения с двумерной осесимметричной задачей, перед этим для этих двух способов расчёта начального поля было проведено тестирование на полиномиальных функциях по пространству.

В трёхмерной задаче основное поле – это $\vec{B}^c = \mu_0 \vec{H}^c$; начальное поле – $\vec{B} = \mu_0 \mu \vec{H}$, где напряженность магнитной индукции \vec{H} вычисляется из формулы (1); аномальное поле – $\vec{B}^a = \vec{B} - \vec{B}^c$.

На примере модельной задачи проиллюстрируем возможность расчёта вектора магнитной индукции с помощью решения трёхмерной задачи магнитостатики с использованием неполного скалярного потенциала. Пусть через всю расчётную область пролегает техногенный объект с сечением $1 \times 1 \text{ м}^2$. Источником возбуждения электромагнитного поля является петля, имеющая радиус 250 м и силу тока 1 А, и расположенная на высоте 4 м над объектом так, что объект параллелен касательной к окружности петли и ближайшая точка объекта до петли находится в 30 метрах. Будем измерять z -компоненту магнитной индукции \vec{B} с учётом объекта и без него, чтобы оценить аномальное поле \vec{B}^a .

Обычно датчиками во время геоэлектроразведки измеряют поле на поверхности земли, поэтому измерять значения поле будем при $z = 0$. Также измерим значения магнитной индукции при $z = -1$, $z = -2$. Получим значения B_z^a вдоль линии на дневной поверхности от $x = 250$ до $x = 310$ в сечении при $y = 0$.

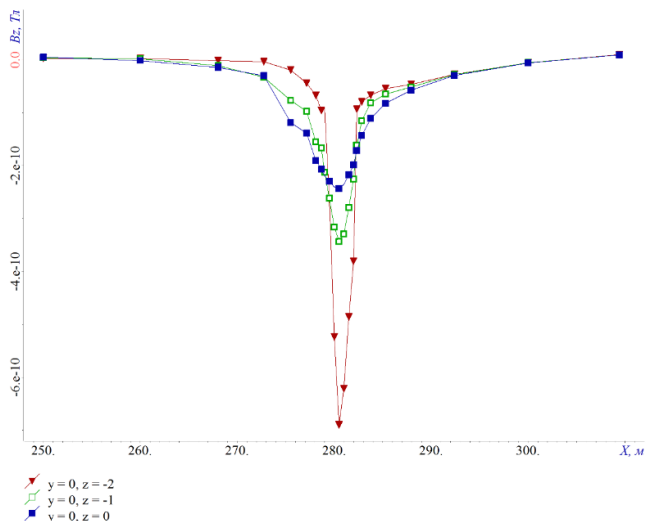


Рисунок 1 – График зависимости B_z^a от x в сечении $y = 0$

На рисунке 1 изображён график зависимости z -компоненты аномального поля \vec{B}^a при $y = 0$ для различных высот: $z = 0$, $z = -1$, $z = -2$.

Наблюдается возрастание абсолютных значений аномального поля объекта при $x = 280$, именно в том месте, через которое проходит техногенный объект. С удалением от техногенного объекта абсолютное значение характеристик аномального поля объекта уменьшается.

Литература:

1. Матвеев А.Н. Электричество и магнетизм: Учеб. пособие./ М.: Высш. школа, 1983 – 463 с.
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. Персова М. Г. Методы конечноэлементного анализа : конспект лекций / М. Г. Персова, Ю. Г. Соловейчик ; Новосиб. гос. техн. ун-т. - Новосибирск, 2015. – 203 с.

ПЛАВЛЕНИЕ ПРЕГРАДЫ КОНЕЧНОЙ ТЕПЛОПРОВОДНОСТИ ПРИ НАТЕКАНИИ СНИЗУ СТРУИ НАГРЕТОЙ ЖИДКОСТИ

О.О. Гусельникова, В.С. Бердников
Новосибирский государственный технический университет,
г. Новосибирск, cersimbod@yandex.ru

Представлены результаты численного моделирования процесса развития термогравитационной струи над линейным источником тепла и сопряженного конвективного теплообмена головной части струи с преградой конечной теплопроводности, в том числе в режиме плавления. Численные эксперименты проведены для жидкостей гептадекан и вода, для которой реализован учет зависимости плотности от температуры.

The paper presents results of numerical simulations of heat-gravitational jet originated over linear heat source and coupled heat exchange from the head of the rising stream with upper wall of finite heat conductivity, including consideration of melting in the wall. Numerical experiments have been conducted for heptadecane and water, for the latter temperature-density relationship is taken into account.

Исследования на основе современных методов геодинамики, в частности сейсмической томографии, позволяют выдвинуть гипотезу о конвективном характере динамики мантии Земли. Центральной теорией глобальных катастроф и образований залежей полезных ископаемых является теория мантийных плюмов – восходящих к дневной поверхности Земли локализованных потоков нагретого вещества. Комплексные исследования глубинных процессов также могут быть полезными для объяснения других природных явлений, характеризующихся вертикальным обменом жидкости со дна водоемов на поверхность. Таким образом, являются актуальными исследования процессов развития конвективной струи и ее взаимодействия с преградой конечной толщины и теплопроводности, в том числе в режиме плавления преграды.

Задача сопряженного конвективного теплообмена плавучей струи с верхней стенкой решается в двумерной осесимметричной постановке. Расчетная область представляет собой прямоугольную полость, заполненную жидкостью и ограниченную сверху твердой стенкой заданной толщины конечной теплопроводности. Источник тепла расположен на нижней стенке полости и имеет форму плоской полосы заданной ширины. Стенки полости жесткие адиабатические. На границе раздела сред в режиме теплопроводности поставлено условие идеального теплового

контакта, а в режиме плавления – условие Стефана [1]. На верхней границе верхней стенки поддерживается постоянная температура.

Математическая модель конвективного движения в жидкости описывается системой уравнений Навье-Стокса в безразмерных переменных температура T , вихрь скорости w и функция тока ψ :

$$\left\{ \begin{array}{l} \left(1 + \frac{\delta(T - T_{pc})}{Ste} \right) \frac{\partial T}{\partial t} + \frac{\partial \psi}{\partial y} \frac{\partial T}{\partial x} - \frac{\partial \psi}{\partial x} \frac{\partial T}{\partial y} = \frac{1}{Pr} \left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} \right) \\ \left(\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} \right) = -w \\ \frac{\partial w}{\partial t} + \left(\frac{\partial \psi}{\partial y} \frac{\partial w}{\partial x} - \frac{\partial \psi}{\partial x} \frac{\partial w}{\partial y} \right) - \left(\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} \right) = Gr \frac{\partial F(T)}{\partial x} \end{array} \right. ,$$

где $F(T)$ отражает зависимость плотности жидкости от температуры.

Скрытая теплота фазового перехода учитывается как дополнительный источник тепла, сосредоточенный на границе раздела сред. Параметры задачи задаются с помощью трех безразмерных комплексов – числа Прандтля Pr , числа Грасгофа Gr и числа Стефана Ste . Численные эксперименты проведены для жидкостей гептадекан с числом Прандтля $Pr = 36.73$ и вода с числом Прандтля $Pr = 12.99$. Высота слоя $H = 35$ мм. Для дискретизации системы использован метод конечных элементов на треугольниках.

Расчеты для жидкости гептадекан показывают, что наибольшая вертикальная компонента скорости наблюдается на оси струи в центральных слоях жидкости (Рисунок 1). Отрицательные значения скорости соответствуют нисходящим потокам жидкости, натекающим на источник тепла с боковых сторон.

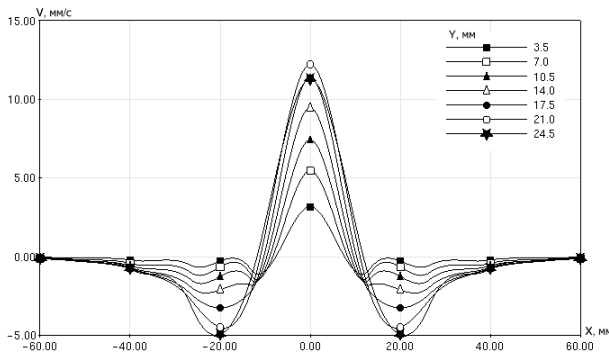


Рисунок 1 – Распределения вертикальной компоненты скорости по ширине слоя в момент касания струи верхней стенки, $Pr = 36.73$

В момент касания струи верхней стенки тепловой поток на границе раздела сред является локализованным на оси струи. По мере натекания струи на преграду наблюдается увеличение локального теплового потока в боковых областях верхней стенки (Рисунок 2).

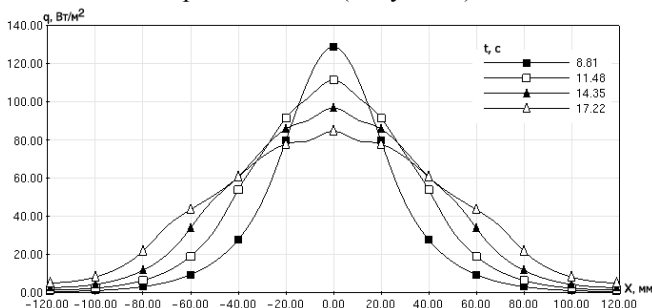


Рисунок 2 – Распределения локальных тепловых потоков на границе раздела сред

Для воды характерно свойство инверсии плотности: слой воды с температурой $4.029\text{ }^{\circ}\text{C}$ обладает наибольшим значением плотности [2]. В процессе развития течения верхний слой головной части струи, в процессе поднятия охлажденный до $4.029\text{ }^{\circ}\text{C}$ за счет теплоотдачи в окружающую среду, стекает с боковых сторон струи вниз (Рисунок 3 – а, 3 – б). Устойчивый восходящий поток формируется после достаточного прогрева жидкости на оси струи (Рисунок 3 – в).

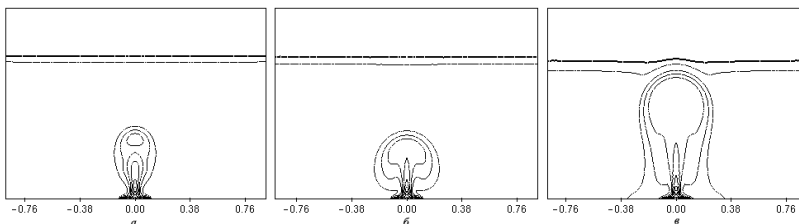


Рисунок 3 – Поле температуры для $Pr = 12.99$ в моменты времени: а – 25.86 с, б – 37 с, в – 97.86 с

Плавление верхней стенки происходит локально на оси струи (Рисунок 4), что следует из особенностей распределения локальных тепловых потоков на границе раздела сред, рассмотренных ранее.

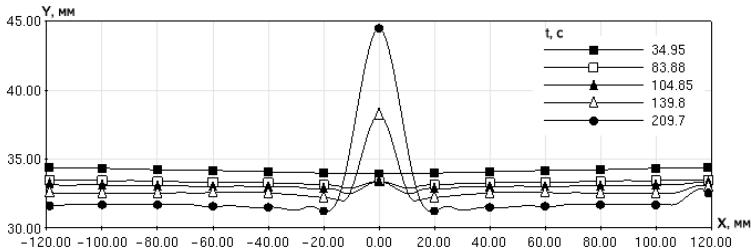


Рисунок 4 – Форма фронта плавления и кристаллизации

В рамках данной работы численно исследованы процессы развития термогравитационных струй в жидкостях гептадекан и вода. Для воды расчеты проведены с учетом зависимости плотности от температуры. Исследован процесс взаимодействия свободноконвективной струи с преградой конечной теплопроводности при натекании снизу. Полученные численно результаты предполагается использовать в качестве основы для планирования лабораторного эксперимента.

Литература:

1. Бердников В.С., Кислицын С.А., Митин К.А. Численное моделирование процессов роста кристаллов методом горизонтальной направленной кристаллизации из расплавов с различными числами Прандтля // Известия Российской академии наук. Серия Физическая. – 2017. Т. 81, № 10. – С. 1389–1394.
2. James N. Hewett, Mathieu Sellier. The pear-shaped fate of an ice melting front // 12th International Conference on CFD in Oil & Gas, Metallurgical and Process Industries, SINTEF 2017, Trondheim, Norway. May 30th – June 1st, 2017: Proceedings. – Trondheim, 2017. – P. 423–429.

РАЗРАБОТКА ПРОГРАММНОГО КОМПЛЕКСА ДЛЯ ЧИСЛЕННОГО МОДЕЛИРОВАНИЯ ДВУХМЕРНЫХ ЗАДАЧ СОПРЯЖЕННОГО КОНВЕКТИВНОГО ТЕПЛООБМЕНА

Н.И. Данилов, В.С. Бердников

Новосибирский государственный технический университет,
г. Новосибирск, nikita.daniloff@gmail.com

В данной работе была поставлена двухмерная задача сопряженного конвективного теплообмена для дальнейшего численного моделирования трехмерных задач, дано описание возникающих физических процессов, программных средств, использующихся для реализации метода конечных элементов и программного комплекса в целом, а также прототипирование и описание основных модулей комплекса.

In this paper we pose a two-dimensional problem of conjugate convective heat exchange for further numerical modeling of three-dimensional problems. Emerging physical processes, software tools used to implement algorithms for constructing adaptive grids and the software complex as a whole are described. In addition, prototyping and description of the main modules of the complex are also given.

Дана двухмерная задача конвективного теплообмена в ограниченных фрагментах горизонтального слоя жидкости (этиловый спирт, гептадекан), равномерно подогреваемого снизу. Горизонтальные границы изотермические нагретые до разных температур. Задана толщина боковых жестких стенок конечной теплопроводности (оргстекло). Относительные размеры расчетной области $2 \leq L/H \leq 10$. H – высота слоя жидкости, L горизонтальный размер. Расчеты выполнены при значениях чисел Прандтля $Pr = 16$ и $45,6$ в диапазонах чисел Рэлея $1000 \leq Ra \leq 20000$. Учтен сопряженный теплообмен с боковыми стенками.

Для численного моделирования подобных задач необходима реализация программного комплекса, который включает в себя:

- Построитель расчетной области (препроцессор);
- Расчетный модуль (с использованием различных подходов метода конечных элементов);
- Визуализатор полученного решения (с применением средств компьютерной графики).

Ключевыми требованиями к комплексу являются:

- Адаптивность конечноэлементных сеток (то есть их сгущение на переходном слое);
- Высокая точность базисных функций (в работе будут использоваться иерархические);

- Высокая производительность отрисовщика (что обеспечивают шейдеры OpenGL).

Литература:

1. В.И. Полежаев. Математическое моделирование конвективного теплообмена на основе уравнений Навье-Стокса. – М.:Наука, 1987.
2. А.В. Гетлинг. Конвекция Рэлея-Бенара. Структуры и динамика. – М.:Эдиториал УРСС, 1999, 248 с.
3. Г.З. Гершуни, Е.М. Жуховицкий, А.А. Непомнящий. Устойчивость конвективных течений. – М.:Наука, 1989, 320 с.
4. О. Зенкевич, К. Морган. Конечные элементы и аппроксимация. – М.:Мир, 1986, 318 с.
5. Д. Вольф. OpenGL 4 Язык шейдеров, настольная книга. – Пакт Паблицинг, 2013, 394 с.

МОДЕЛИРОВАНИЕ НЕСТАЦИОНАРНОГО ЭЛЕКТРОМАГНИТНОГО ПОЛЯ ДЛЯ РЕШЕНИЯ ЗАДАЧ ГЕОЭЛЕКТРИКИ В ОБЛАСТЯХ С ТЕХНОГЕННЫМИ ОБЪЕКТАМИ

И.И. Карасенко, Ю.Г. Соловейчик
Новосибирский государственный технический университет,
г. Новосибирск, ikar-95@yandex.ru

Работа посвящена разработке и программной реализации вычислительных схем для моделирования нестационарного электромагнитного поля для решения задач геоэлектрики в областях с техногенными объектами.

The work is devoted to the development and software implementation of computational schemes for modeling the non-stationary electromagnetic field in solving geoelectrics problems in environments with technogenic objects.

Введение. С развитием технологий добычи полезных ископаемых одной из проблем, возникающих в ходе геофизических исследований на частично освоенных месторождениях, становится наличие в исследуемых недрах различных техногенных объектов, таких как нефтепромысловые или транспортные трубопроводы, которые в определённой степени оказывают влияние на расчёт характеристик электромагнитного поля в заданной области. В связи с этим, при решении задач электроразведки

появляется необходимость учитывать наличие таких объектов в расчётной области, их местоположение и параметры. Полученная информация о характеристиках электромагнитного поля в среде, как результат проведения прямого моделирования, служит входными данными для дальнейших задач определения параметров вмещающей среды.

В настоящее время существует множество технологий электроразведки, различающихся способом возбуждения и приема поля. Рассмотрим метод электромагнитного зондирования становлением поля, основанный на изучении поля переходных процессов, которое возбуждается в земле при импульсном переключении тока в источнике. В качестве источника используется вертикальный магнитный диполь (незаземленная токовая петля). В качестве приемника также используется незаземленная петля. Измерения производятся на этапе становления поля, т.е. после выключения тока в источнике.

Математический аппарат. В данной работе представлены вычислительные схемы для моделирования нестационарного электромагнитного поля в областях с техногенными объектами. Для расчёта начального поля в работе применяется модель с использованием технологии выделения части поля от петлевого источника, которое рассчитывается при решении двумерной осесимметричной задачи. Для расчёта нестационарного поля на этапе становления, т.е. после выключения тока в источнике, в работе применяется модель на основе векторного уравнения, являющаяся основанием для использования векторного МКЭ.

Рассмотрим математическую модель, описывающую нестационарное электромагнитное поле:

$$\operatorname{rot} \left(\frac{1}{\mu} \operatorname{rot} \vec{\bar{A}} \right) + \sigma \frac{\partial \vec{\bar{A}}}{\partial t} = \vec{\bar{J}}^{\text{ст}}, \quad (1)$$

где $\vec{\bar{A}}$ – вектор-потенциал, связанный с индукцией магнитного поля $\vec{\bar{B}}$ соотношением

$$\vec{\bar{B}} = \operatorname{rot} \vec{\bar{A}}, \quad (2)$$

μ – коэффициент магнитной проницаемости, σ – удельная электрическая проводимость среды, $\vec{\bar{J}}^{\text{ст}}$ – вектор плотностей сторонних токов.

Начальным полем будем называть поле, установившееся в результате включения тока в источнике. Рассмотрим уравнение (1). Поле, описываемое вектор-потенциалом $\vec{\bar{A}}$, может быть представлено в виде

суммы двух полей, каждое из которых описывается «своим» вектор-потенциалом $\vec{\mathbf{A}}^0$ (основное поле) и $\vec{\mathbf{A}}^+$ (добавочное поле):

$$\vec{\mathbf{A}} = \vec{\mathbf{A}}^0 + \vec{\mathbf{A}}^+. \quad (3)$$

Вектор-потенциал $\vec{\mathbf{A}}^0$ описывает поле петлевого источника в одно-родной среде и удовлетворяет уравнению

$$\text{rot} \left(\frac{1}{\mu_0} \text{rot} \vec{\mathbf{A}}^0 \right) = \vec{\mathbf{J}}^{\text{ст}}. \quad (4)$$

При этом, поле $\vec{\mathbf{A}}^0$ вычисляется в результате решения двумерной осесимметричной задачи.

Вычитая из уравнения (1) уравнение (4), с учётом (3) получим:

$$\text{rot} \left(\frac{1}{\mu} \text{rot} \vec{\mathbf{A}}^+ \right) + \sigma \frac{\partial \vec{\mathbf{A}}^+}{\partial t} = \text{rot} \left(\left(\frac{1}{\mu_0} - \frac{1}{\mu} \right) \text{rot} \vec{\mathbf{A}}^0 \right) - \sigma \frac{\partial \vec{\mathbf{A}}^0}{\partial t}. \quad (5)$$

Таким образом, вектор-потенциал $\vec{\mathbf{A}}$ (начальное поле), который описывает поле вмещающей среды и трёхмерных объектов, находится в результате решения векторного уравнения (5).

Для расчёта нестационарного поля на этапе становления, т.е. после выключения тока в источнике, решается трёхмерная начально-краевая задача для уравнения (1), в котором $\vec{\mathbf{J}}^{\text{ст}} = 0$ во всей области:

$$\text{rot} \left(\frac{1}{\mu} \text{rot} \vec{\mathbf{A}} \right) + \sigma \frac{\partial \vec{\mathbf{A}}}{\partial t} = 0, \quad \vec{\mathbf{A}} \Big|_{t=t_0} = \vec{\mathbf{A}}^{t_0}, \quad (\vec{\mathbf{A}} \times \vec{\mathbf{n}}) \Big|_S = 0, \quad (6)$$

где $\vec{\mathbf{A}}^{t_0}$ – найденное начальное поле.

Модели на основе векторных уравнений (5),(6) корректно описывают электромагнитные поля в случаях, когда среда содержит неоднородности с изменёнными электрическими и магнитными свойствами, однако требуют разрывности нормальных составляющих вектор-потенциала на границах сред с различными значениями удельной проводимости. Это обуславливает применение векторного МКЭ [1], где в качестве базисных функций используются векторные базисные функции с разрывными нормальными составляющими.

Вычислительный эксперимент. Пусть исследуемому объёму соответствует расчётная область $\Omega = [-1000, 1000] \times [-1000, 1000] \times [-1000, 1000]$. Плоскость $z = 0$ соответствует земной поверхности. Поверхностный слой земли толщиной 20 метров, расположенный в координатах $-20 < z < 0$, имеет удельную

электрическую проводимость $\sigma=0,05$. Глубинный слой $z < -20$ имеет удельную электрическую проводимость $\sigma=0,005$. На глубине 3,5 метров в координате $x=300$ параллельно оси y расположен металлический техногенный объект. На поверхности земли расположен петлевой источник с радиусом 250 метров и центром, совпадающим с центром расчётной области. Ток в источнике равен 1 А. Периоды включения и выключения тока в источнике равны 100 мс.

На поверхности земли вдоль оси x расположено 6 приёмников, начиная с координаты $x=100$ с шагом в 100 метров.

Зависимость B_z от времени в приёмниках после выключения тока в источнике изображена на Рисунке 1.

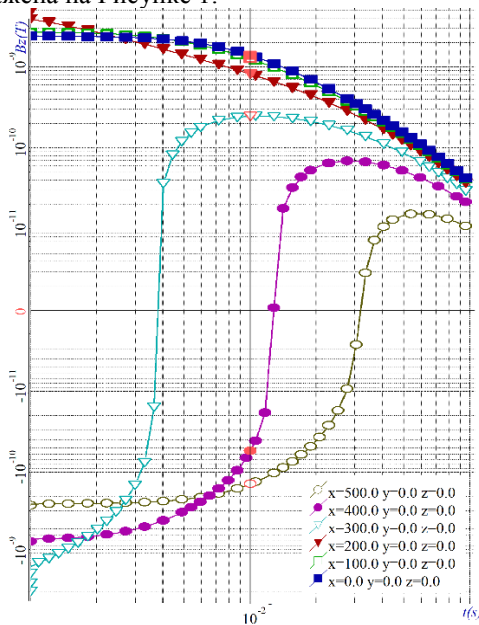


Рисунок 3. Зависимость B_z от времени в приёмниках.

Наличие техногенного объекта в расчётной области оказывает тем большее влияние на характеристики поля, чем ближе к объекту происходят замеры. На примере описанной модельной задачи наибольшее влияние наблюдается в приёмнике, расположенном над объектом в ко-

ординате $x = 300$ (разница 5,6% по сравнению со значениями, полученными в аналогичной задаче без объекта). Влияние на значения, полученные с других приёмников, уменьшается по мере их удаления от объекта.

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. Современные проблемы прикладной математики и информатики: конспект лекций / авт. М.Г. Персова, Ю.Г. Соловейчик. – Новосибирск : Изд-во НГТУ, 2011. – 58 с.
3. Персова М. Г. Методы конечноэлементного анализа: конспект лекций / М. Г. Персова, Ю. Г. Соловейчик ; Новосиб. гос. техн. ун-т. – Новосибирск, 2015. – 203 с.

ЧИСЛЕННОЕ МОДЕЛИРОВАНИЕ КОЛЕБАНИЙ УЛЬТРАЗВУКОВЫХ ВОЛНОВОДОВ НА ОСНОВЕ ДИНАМИКИ МАССИВА ЧАСТИЦ

А.Е. Кузובה, Д.Я. Суханов
Томский государственный университет,
г. Томск, ang_kuz93@mail.ru

Предлагается численная модель акустических процессов в твёрдых телах на основе интегрирования уравнения движения частиц в кубической объёмно – центрированной решётке. Взаимодействие частиц описывается зависимостью силы притяжения от расстояния между частицами с точкой равновесия на расстоянии равном периоду кристаллической решётки. Проведено моделирование колебаний ультразвукового волновода из алюминия и найдены его резонансные частоты.

A numerical model of acoustic processes in solids is proposed on the basis of integrating the equation of particle motion in a cubic volume-centered lattice. The interaction of particles is described by the dependence of the attractive force on the distance between particles with an equilibrium point at a distance equal to the period of the crystal lattice. Modeling of oscillations of an ultrasonic waveguide from aluminum is carried out and its resonant frequencies are found.

Численное моделирование акустических процессов позволяет

предсказать прохождение ультразвуковых сигналов через различные структуры и определить их резонансные частоты [1–4]. Мы предлагаем моделировать акустические процессы, представляя твердые тела в виде массива частиц, расположенных в кубической объемно-центрированной кристаллической решетке [5–6]. Частицы в рассматриваемой решетке размещаются периодически, но между частицами существуют два разных расстояния - d и $d\sqrt{3}/2$.

Каждая частица имеет собственную массу и силу взаимодействия со своими соседями. Для каждой частицы мы вычисляем скорость и координату путем численного интегрирования ускорения по времени. Ускорение частицы рассчитывается по второму закону Ньютона как отношение силы, действующей на частицу к ее массе. Масса частицы и сила взаимодействия задают тип материала.

Зависимость силы взаимодействия частиц от расстояния будем характеризовать формулой (1):

$$F(x) = \begin{cases} k \left(1 - \frac{d}{x}\right) \left(\sqrt{2} - \frac{x}{d}\right)^2 & \text{if } (x < d\sqrt{2}) \\ 0 & \text{if } (x \geq d\sqrt{2}) \end{cases}, \quad (1)$$

где d – расстояние между взаимодействующими частицами, коэффициент k подбирается исходя из типа материала и направления в кристалле для обеспечения изотропности или анизотропности взаимодействия.

Сила, действующая на определенную частицу, вычисляется как векторная сумма сил взаимодействия с соседними частицами. Масса и сила взаимодействия частиц определяет плотность и скорость звука конкретного материала, что позволяет связать моделирование с процессами в реальных материалах.

Мы провели численное моделирование цилиндрической структуры, состоящей из твердых сред. Цилиндр (волновод) изготовлен из двух цилиндров из алюминия диаметром 20 мм (скорость звука 5050 м / с, плотность 2712 кг / м³), длина первого цилиндра 172 мм, длина второго цилиндра 72 мм. Между цилиндрами размещается пьезоэлемент из пьезокерамики (PZT-8) диаметром 4 мм (скорость звука 3200 м / с, плотность 7600 кг/м³) (Рисунок 1).

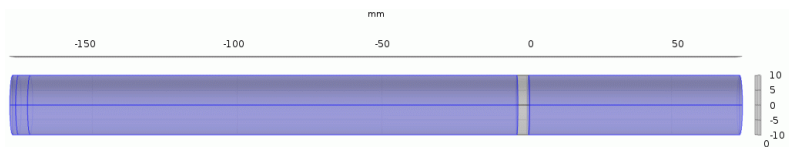


Рисунок 1 – Алюминиевый ультразвуковой волновод с пьезокерамическим элементом.

Для задания материалов с различными свойствами регулировались значения массы частиц и угол наклона кривой зависимости силы от расстояния вблизи точки равновесия. В качестве возбуждающего сигнала применялся короткий импульс. На конце волновода размещалась мониторинговая частица, которая регистрировала отклонения координаты и скорости частицы. В результате проведенного численного моделирования был получен резонансный частотный спектр моделируемой структуры. Во время моделирования контролируется координата $X(t)$ краевых точек. В результате численного моделирования мы получили функцию $X(t)$ и вычислили ее спектр. Спектр краевых пространственных колебаний представлен на (Рисунке 2).

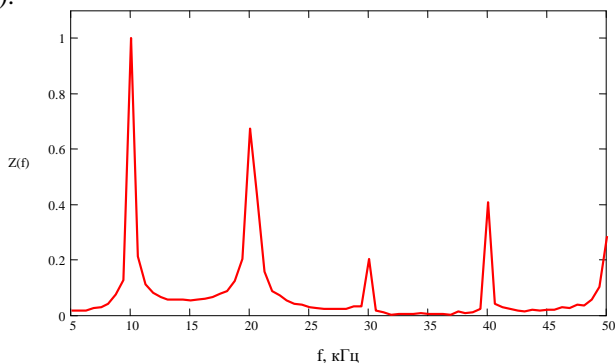


Рисунок 2 – Спектр импульсного сигнала на конце волновода

Для проверки достоверности полученных результатов численного моделирования, мы провели моделирование данной структуры в программе Comsol Multiphysics. На (Рисунке 3) представлен резонансный частотный спектр, который был получен в результате численного моделирования. Из полученных результатов численного моделирования видно совпадение резонансных частот, что

подтверждает достоверность и применимость предложенного метода численного моделирования.

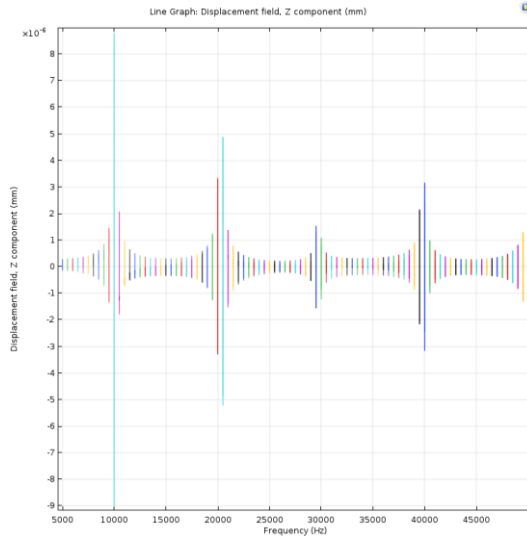


Рисунок 3 – Резонансный спектр частот на конце волновода в Comsol Multiphysics

Работа выполнена при поддержке Минобрнауки России в рамках проекта «Создание высокотехнологичного комплекса ультразвуковой хирургии» (Уникальный идентификатор проекта RFMEF157517X0163).

Литература:

1. Коробов А.И., Карабутов А.А., Сапожников О.А., Ультразвуковые волны в твердых телах: методическая разработка специального физического практикума кафедры акустики. – ФФ. МГУ, 2011. – 34 с.
2. Кривцов А.М., Кривцова Н.В. Метод частиц и его использование в механике деформируемого твердого тела // Дальневосточный математический журнал. – 2002. – №2. – С. 254 – 276.
3. Бархатов В.А. Решение волновых уравнений методом КРВО. Основные соотношения. // Дефектоскопия. – 2007. – №9. – С. 54 – 70.
4. Кузорова А.Е., Суханов Д.Я. Моделирование акустических процессов в твёрдых телах на основе взаимодействия частиц // Информационно-измерительная техника и технологии: материалы докладов VIII междунард. науч.-практ. конф. – Томск: Изд-во ТГУ,

2016. – С. 92–93.

5. Ашкрофт. Н., Мермин. Н. Физика твёрдого тела: учебник для вузов. В 2 ч. Ч. 1. – М.: Мир, 1979. – 458 с.

6. Греков. Ф. Ф., Рябенко. Г. Б., Смирнов. Ю. П. Структурная кристаллография: учебное пособие. – Л.:ЛГПИ, 1988. – 77с.

ПОСТРОЕНИЕ И ВИЗУАЛИЗАЦИЯ ЗАМКНУТЫХ СИЛОВЫХ ЛИНИЙ ВЕКТОРНЫХ ПОЛЕЙ

А.В. Кунц, И.М. Ступаков

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

Силовые линии — это кривые, обозначающие траекторию движения частицы с нулевой массой в данном поле. Закрытые силовые линии встречаются в соленоидальных магнитных полях. В данной работе производится рассмотрение различных методов построения этого типа силовых линий.

Field lines are curves that depict the path of a massless particle under the force of a given field. Closed-loop field lines are present in solenoidal vector fields, e.g. magnetic fields. In this work, different techniques of constructing a closed loop line are considered.

Построение силовых линий в трёхмерном пространстве является актуальной проблемой для многих областей, в частности, визуализации полей, являющихся результатом численного моделирования задач электромагнетизма.

Для корректного решения задачи построения силовых линий необходимо решить несколько проблем:

- 1) Эффективный и точный метод построения линии из заданной точки
- 2) Способ нахождения замкнутых линий и остановки алгоритма 1)
- 3) Способ нахождения распределения плотности силовых линий

Данная работа затрагивает варианты решения первых двух проблем.

Для основной задачи – построения линий, были выбраны несколько алгоритмов численного дифференцирования, а также специальные способы построения кривых, например, метод окружностей.

Выбранные алгоритмы:

- a) Метод касательных, или модификация метода Эйлера

Данный метод является простейшим способом построения кривой: в точке, от которой рассчитывается шаг, вычисляется значение поля, и делается шаг заданной длины вдоль вектора значения поля. Несмотря на скорость вычисления и простоту использования, этот метод на практике обладает существенным недостатком: в соленоидальных полях силовые линии имеют вторую производную своей позиции (или, что то же самое с точностью до константы, производную значения поля вдоль его направления вектора его значения), указывающую в одну и ту же область, являющуюся своеобразным центром этой кривой, на всём своём протяжении, что приводит к тому, что методы, неспособные принять во внимание подобную кривизну, а к ним относятся, в частности, метод касательных, вносят в последующие точки построимой кривой систематическое отклонение, с каждой точкой всё дальше удаляющее их от «центра».

Очевидным решением этой проблемы является следующий метод.

б) Метод последовательных дуг

В точке, являющейся исходной для текущего шага построения кривой, вычисляется направление поля и вторая производная позиции, описанная ранее. Используя эту производную, можно вычислить точку, которая бы служила центром окружности при условии постоянства модуля второй производной. Далее, вокруг этой точки строится дуга заданной длины, начало которой является исходной точкой, и эта дуга включается в построимую кривую. Данный метод использует численное дифференцирование для нахождения вектора второй производной, что усложняет вычисления, но позволяет совершать шаги большего размера по сравнению с методом касательных

с) Модификация метода Рунге-Кутты 4 порядка

Так как, в некотором смысле, описанная задача является решением дифференциального уравнения, описывающего позицию частицы с нулевой массой в заданном поле, следует предположить, что классические методы решения дифференциальных уравнений также могут быть использованы для решения данной задачи, с некоторыми модификациями. В частности, все вектора, вычисленные, как значения поля в некоторой точке, должны быть заменены на свои нормированные значения, для корректного учёта размера шага.

Для решения задачи определения замкнутости кривых были выбраны критерии пространственной близости, а также комбинации пространственной близости с близостью направления с последующей корректировкой построенной кривой.

Суть критерия заключается в вычислении расстояния от луча с направлением, совпадающем с направлением вектора значения поля в текущей точке кривой, до точки, из являющейся начальной для процедуры построения. Если и само расстояние от луча до начальной точки, и расстояние от текущей точки до наименее удалённой от начальной точки луча достаточно малы, то следует скорректировать текущий шаг и ослабить для следующей точки критерий остановки по расстоянию. Данная проверка позволяет максимально точно найти момент, когда кривая замыкается, и предотвратить лишние вычисления и ошибки визуализации (накрадывающиеся кривые).

РАЗРАБОТКА И РЕАЛИЗАЦИЯ МАТЕМАТИЧЕСКОЙ МОДЕЛИ РАСЧЕТА ТЕМПЕРАТУРНОГО ПОЛЯ В ДЫМОВЫХ ТРУБАХ

А.А. Набатова*, В.А. Гольцев*, Н.А. Метелкин**

***Уральский федеральный университет первого президента России
Б.Н. Ельцина, **ООО «Центр Экспертных Решений в Строитель-
стве», г. Екатеринбург, anabatova@urfu.ru**

Промышленная дымовая труба – сложное инженерное сооружение, ее эксплуатация связана с потенциальной повышенной опасностью. Для обеспечения долговечности, надежности и бесперебойной работы дымовых труб следует контролировать температурное поле этих объектов. Целью настоящей работы является разработка и реализация математической модели и создание расчетной программы по определению температурного поля в дымовых трубах любых конструкций. Разработанный программный продукт позволяет также автоматизировать ведение технической документации с возможностью экспорта отчетов в форматы MS Word и Excel.

Industrial chimney is a complex engineering structure, its operation is associated with potential increased danger. To ensure durability, reliability and smooth operation of chimneys, the temperature field of these objects should be monitored. The purpose of this work is the development and implementation of a mathematical model and the creation of design programs for resolving the temperature field in chimneys of any structures. The developed software product also allows you to automate the maintenance of technical documentation with the ability to export reports in MS Word and Excel formats.

Для контроля технического состояния дымовых труб весьма актуальна задача по расчетному определению температурного поля по сечению и высоте дымовой трубы. Обладая такими сведениями, можно прогнозировать, будет ли возможно появление конденсата в трубе, определять минимально возможные температуры выходящих из теплового агрегата газов с оценкой технического состояния подземных газохранилищ.

С теплотехнической точки зрения решение задачи по определению температурного поля в дымовой трубе сводится к решению уравнения баланса теплоты дымовых газов и уравнения стационарной теплопередачи от дымовых газов к окружающему трубу воздуху через многослойную цилиндрическую стенку. Исходными данными являются расход V и температура T_r газов при входе в трубу, температура T_0 и скорость w окружающего воздуха, а также геометрические размеры слоев стенки и теплофизические характеристики материалы. Расчетная схема к задаче приведена на рисунке 1.

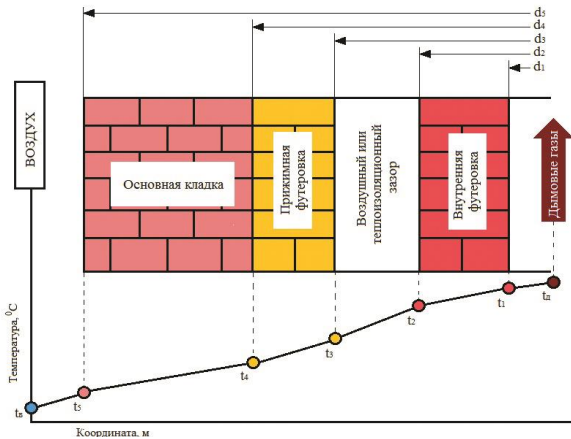


Рисунок 1 – Схема к расчету температур по радиусу дымовой трубы

Первоначально условно разбивают трубу на расчетные участки по высоте. Для каждого участка решают уравнение теплопередачи и определяют линейную плотность теплового потока Q_L , Вт/м:

$$Q_L = k_L \pi (T_r - T_0),$$

где k_L - линейный коэффициент теплопередачи для цилиндрической стенки, Вт/(м·К):

В свою очередь, линейный коэффициент теплопередачи, Вт/(м·К):

$$k_L = \frac{1}{\frac{1}{\alpha_1 d_1} + \left[\frac{1}{2\lambda_1} \ln\left(\frac{d_2}{d_1}\right) \right] + \left[\frac{1}{2\lambda_b} \ln\left(\frac{d_3}{d_2}\right) \right] + \left[\frac{1}{2\lambda_3} \ln\left(\frac{d_4}{d_3}\right) \right] + \left[\frac{1}{2\lambda_4} \ln\left(\frac{d_5}{d_4}\right) \right] + \frac{1}{\alpha_2 d_5}}$$

зависит от геометрических размеров и теплофизических свойств материалов дымоотводящей трубы. Методикой расчета предусмотрено уточнения теплофизических свойств продуктов сгорания движущихся в трубе, и материалов самой трубы в зависимости от температуры. Коэффициент теплоотдачи от дымовых газов к футеровке α_1 рассчитывают с учетом конвективной α_k и лучистой α_l составляющих. Коэффициент α_2 теплоотдачи конвекцией от внешней стенки трубы к воздуху определяют в зависимости от скорости ветра заданного района и высоты трубы по СНиП 2.01.01-82.

Для расчета теплообмена в воздушном зазоре для определения среднего коэффициента теплоотдачи применили расчетную формулу $0,105 \cdot (\text{Gr} \cdot \text{Pr})^{0,3}$ при $(\text{Gr} \cdot \text{Pr}) < 10^7$ либо $0,4 \cdot (\text{Gr} \cdot \text{Pr})^{0,2}$ при $(\text{Gr} \cdot \text{Pr}) > 10^7$.

Поскольку потери теплоты на расчетную длину участка приводят к уменьшению теплосодержания дымовых газов, движущихся в трубе, можно записать

$$\Delta T = Q_L / (V c_p),$$

где ΔT – падение температуры на метр длины газохода, °С; c_p – теплоемкость дымовых газов при постоянном давлении Дж/(м³·К).

Зная длину расчетного участка, уточняют температуру газов и вновь повторяют расчет до достижения требуемой точности. Один из фрагментов математической модели расчетного алгоритма представлен на рисунке 2.

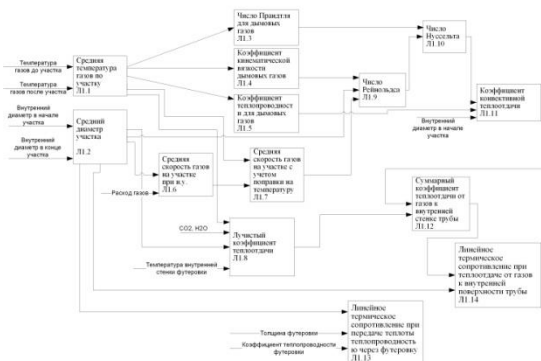


Рисунок 2 – Фрагмент начального расчета математической модели расчетного алгоритма

На кафедре теплофизики и информатики в металлургии разработан программный продукт, предназначенный для инженерно-технических работников организаций, занимающихся осмотром, обследованием и экспертизой промышленной безопасности промышленных дымовых и вентиляционных труб различной конструкции. При использовании разработки предполагается уменьшить затраты рабочего времени на теплотехнические расчеты и автоматизировать ведение технической документации (регистрацию договоров, натурных обследований и пр.) при техническом обслуживании промышленных труб.

Данный программный продукт по расчету температурного поля в дымовой трубе выполняет следующие функции:

- ввод, загрузка исходных данных и хранение их в базе данных;
- выполнение расчета;
- ведение справочника материалов.
- просмотр результатов расчета и построение диаграмм (графиков зависимостей);
- формирование отчета по расчету и экспорт его в форматы MS Word и MS Excel.

Из-за большой высоты дымовой трубы и изменения геометрических параметров и теплофизических свойств материалов по высоте трубы в программе использован зонный принцип расчета, который предполагает усреднение параметров по высоте трубы в пределах каждой зоны. Теп-

лотехнический расчет проводят в последовательности, начиная с нижней зоны (участка).

Пользователь производит загрузку данных из базы, корректировку и добавление данных в базу, выбор отображения результатов, формирование отчета, формирование запросов в документах.

Архив документов включает модуль просмотра, модуль добавления и редактирования каждого из документов, модуль формирования и экспорта отчетов. Все документы представлены в табличном виде и хранятся в базе данных.

При помощи разработанного программного продукта удалось подтвердить, что конденсат на внутренней поверхности кладки кирпичной дымовой трубы после котла ДКВР 20/13 в городе Сургуте образуется уже на первом расчетном участке при температуре окружающего воздуха 0°C . В конечном итоге было доказано, что разрушение данной дымовой трубы произошло из-за многократного оттаивания и замерзания образующегося конденсата в кладке.

Литература:

1. Дужих Ф.П., Осоловский В.П., Ладыгичев М.Г. Промышленные дымовые и вентиляционные трубы: Справочное издание/ Под ред. Ф.П. Дужих. – М.: Теплотехник, 2004 – 464 с.
2. Телегин А.С., Швыдкий В.С., Ярошенко Ю.Г. Теплообмен: Учебник для вузов/ Под ред. Ю.Г. Ярошенко. – М.: ИКЦ «Академкнига», 2002. – 455 с.

ИССЛЕДОВАНИЕ СПОСОБОВ ВЫБОРА ШАГА ПО ВРЕМЕНИ В ЧИСЛЕННОМ МЕТОДЕ МОДЕЛИРОВАНИЯ ПРОЦЕССОВ МНОГОФАЗНОЙ ФИЛЬТРАЦИИ

И.И. Патрушев, М.Г. Персова
Новосибирский государственный технический университет,
г. Новосибирск, patrushev.ii@mail.ru

В работе представлены различные подходы к решению проблемы выбора временного шага в численном методе моделирования задачи фильтрации.

The research highlights different approaches to choosing the time step value for the numerical method of multiphase flow modeling.

В настоящее время на многих российских месторождениях завершается разработка прямыми технологиями добычи, в ходе которых в коллекторе может оставаться значительное количество трудноизвлекаемой нефти. Ввиду этого, становятся особо актуальными задачи компьютерного моделирования месторождений и разработка новых вычислительных схем, позволяющих проводить высокоточные расчёты для сложных физических моделей.

Пусть дана некоторая расчётная область Ω , представляющая собой пористую среду, и её пространственная дискретизация $\{\Omega_e\}$. Будем считать, что в каждый момент времени все ячейки Ω_e полностью заполнены многофазной жидкостью. Решение задачи фильтрации заключается в аппроксимации движения многофазного потока в среде Ω . Описание численного метода решения этой задачи с использованием метода конечных элементов [1] представлено в работах [2-5]. Основной идеей предлагаемого подхода является перерасчёт новых объёмов фаз в ячейках Ω_e по известным втекающим и вытекающим объёмам фаз $V_{\Gamma_i}^m$ (m – номер фазы), проходящих через границы ячеек $\Gamma_{\Omega_e}^i$. Согласно закону Дарси [6], рассчитываются потоки перетекающих фаз $Q_{\Gamma_i}^m$, и объёмы $V_{\Gamma_i}^m$ рассчитываются по соотношению

$$V_{\Gamma_i}^m = Q_{\Gamma_i}^m \cdot \Delta t, \quad (1)$$

при этом считается, что за время Δt поток $Q_{\Gamma_i}^m$ вместе с порождающим его давлением значительно не изменяются.

Проблема выбора шага Δt является отдельной задачей моделирования, поскольку от величины шага зависит как качество аппроксимации задачи фильтрации, так и общее время, затрачиваемое на решение задачи.

Первым подходом является выбор шага Δt – общего для всей ячеек Ω_e . Естественным условием выбора шага по времени Δt является то, что суммарный вытекающий объём фазы из ячейки не должен превышать имеющийся объём фазы в ней. Записать это условие можно следующим соотношением

$$V_{out, \Omega_e}^m \leq mes(\Omega_e) \Phi S^m, \quad (2)$$

где Φ – пористость ячейки Ω_e , а S^m – насыщенность m -ой фазы в ячейке Ω_e . Из условия (2) видно, что выбор шага Δt зависит от малости величины насыщенности фазы S^m . То есть, если S^m – достаточно малая величина, то, руководствуясь этим условием, необходимо выбрать соответствующий малый шаг по времени, что может повлечь за собой резкое замедление расчёта процесса фильтрации. Поэтому введём специальные величины для контроля величины шага Δt : $S_{\max, \text{crit}}^m$ – максимальная критическая насыщенность m -ой фазы в ячейке Ω_e , которая определяет нижнюю границу насыщенности S^m , по которой выбирается шаг по времени Δt ; и $S_{\min, \text{crit}}^m$ – минимальная критическая насыщенность, которая определяет величину насыщенности S^m , при которой насыщенность фазы в ячейке настолько мала, что данная фаза будет вытесняться из ячейки вместе с остальными вытекающими фазами. Для тех фаз, которые не удовлетворяют условию (2), выполняется специальная процедура выталкивания фазы из ячейки. Суть процедуры выталкивания заключается в том, что выталкиваемая фаза полностью покинет ячейку пропорционально вытекающему потоку смеси. Таким образом, введён контроль за малостью выбора шага по времени Δt .

Условие (2) является в некотором смысле достаточным для корректности рассматриваемой схемы моделирования. Вместе с введением критериев $S_{\min, \text{crit}}^m$ и $S_{\max, \text{crit}}^m$ и процедуры выталкивания, шаг по времени может быть увеличен, но его величина не может превышать максимальный шаг по времени Δt^{\max} , который может быть найден из условия

$$V_{\text{out}, \Omega_e}^{\text{mix}} \leq \text{mes}(\Omega_e)\Phi, \quad (3)$$

где $V_{\text{out}, \Omega_e}^{\text{mix}}$ – объём перетекающей смеси. Очевидно, что выбор максимального шага повлечёт за собой увеличение количества ячеек с выталкиваемыми фазами, что, вообще говоря, может значительно нарушать качество аппроксимации решения.

Второй подход к выбору шага заключается в разбиении множества ячеек Ω_e на группы, в каждой из которых выбирается свой шаг по времени Δt_k . При этом каждый шаг из множества $\{\Delta t_k\}$ кратен

наименьшему шагу из этого множества. Для реализации такого подхода необходимы специальные алгоритмы согласования перетоков между ячейками, которые попали в разные группы.

Сравнение разных способ выбора шага по времени проводилось на модели газонефтяного месторождения (рис. 1) с одной добывающей скважиной.

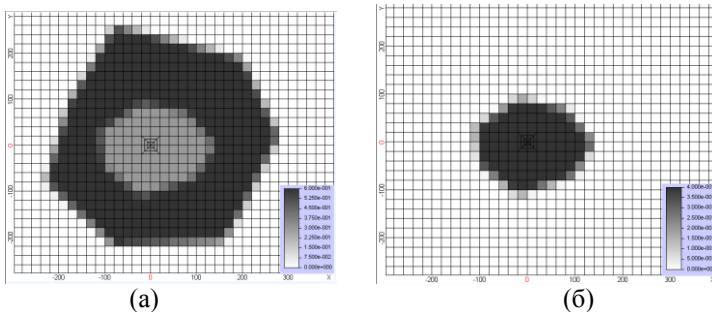


Рис. 1 – Распределение поля насыщенности газа (а) и нефти (б) в начальный момент времени.

Количество итераций по пересчёту состояния ячеек для разных способов подбора шага по времени приведены в таблице 1. Отметим, что результаты для данных вариантов расчётов отличаются не более на 0.3%. Из таблицы видно, что контроль за выбором шага по времени значительно сокращает вычислительные затраты. В то время как технология группирование на порядок понижает количество пересчётов состояний ячеек. В данном тесте максимальное за всё время расчёта количество групп равно 9, а минимальное – 3.

Таблица 1 – Сравнение вычислительных затрат для разных способ выбора шага по времени

| Способ | $S_{min,crit}^{gas}$ | $S_{max,crit}^{gas}$ | Количество итераций | Количество пересчётов состояний ячеек |
|-------------------|----------------------|----------------------|---------------------|---------------------------------------|
| Без группирования | 0 | 0 | 10134 | 34293456 |
| | 0.00 | 0.01 | 6835 | 23129640 |
| | 0.01 | 0.1 | 3286 | 11119824 |

| | | | | |
|---------------|-----------|------|-----|---------|
| Группирование | 0.00 1 | 0.01 | 500 | 1757738 |
|---------------|-----------|------|-----|---------|

Работа выполнена при финансовой поддержке гранта Президента Российской Федерации для государственной поддержки молодых российских ученых – докторов наук (№ гранта МД-3516.2017.5).

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. Nick H.M. and Matthäi S.K. A Hybrid Finite-Element Finite-Volume Method with Embedded Discontinuities for Solute Transport in Heterogeneous Media // *Vadose Zone Journal*. – 2011. – Vol. 10. – №. 1. – pp. 299-312. DOI: 10.2136/vzj2010.0015.

3. Numerical modeling of multi-phase flow for various junctions of water and oil saturated layers in 3-D porous media/ Persova M.G., Soloveichik YU.G., Patrushev I.I., Grif A.M. // *Actual problems of electronic instrument engineering (AREIE–2018): междунар. науч.-техн. конф., Новосибирск 2-6 окт. 2018 г.*– Новосибирск : Изд-во НГТУ, 2018. – С. 212-215.

4. Flow balancing in FEM modelling of multi-phase flow in porous media/ Marina G. Persova, Yuri G. Soloveichik, Alexander M. Grif, Ilya I. Patrushev // *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.

5. Исследование численного метода трёхмерного моделирования процесса многофазной фильтрации // И.И. Патрушев, М.Г. Персова, Ю.Г. Соловейчик // *Обработка информации и математическое моделирование : материалы Рос. науч.-техн. конф.* – Новосибирск : СибГУТИ, 2018. – С. 85-92. - ISBN 978-5-91434-042-8.

6. Дж. Коннор, К. Бреббиа Метод конечных элементов в механике жидкости. Пер. с англ. – Л.: Судостроение, 1979. – 264 с.

ВЛИЯНИЕ ОСОБЕННОСТЕЙ КОНСТРУКЦИИ МЭМС НА ЭЛЕКТРОФИЗИЧЕСКИЕ ХАРАКТЕРИСТИКИ

К.Г. Пельменев, Д.И. Остертак, В.П. Драгунов
Новосибирский государственный технический университет,
г. Новосибирск, pelkonst@mail.ru

В работе проводится теоретическое исследование влияния особенностей конструкции микроэлектромеханических систем (МЭМС) на электрическую ёмкость и электростатические силы с учётом конечной толщины электродов. С помощью метода конечных элементов получены зависимости ёмкости и силы от межэлектродного зазора для двух вариантов конструкции двухэлектродной МЭМС. Оценивается влияние геометрических размеров электродов: длины, ширины и толщины на исследуемые электрофизические параметры МЭМС. Приведены оценки погрешности расчётов исследуемых параметров при учете основания подвижного электрода.

Theoretical study of an effect of microelectromechanical systems (MEMS) design features on electric capacitance and electrostatic forces is carried out taking into account the fringing field effects. Dependences of the capacitance and electrostatic force on the interelectrode gap for two designs of two-electrode MEMS are obtained by means of finite element method. An influence of geometrical parameters: length, width and thickness on given MEMS parameters is evaluated. Estimations of calculation errors with taking into account the moveable electrode base are given.

При разработке электростатических МЭМС возникает необходимость в оценке электрической емкости и электростатических сил, действующих между различными частями конструкции. Как правило, для оценки емкости используется классическая формула для емкости идеального плоского конденсатора (ИПК), однако, такая оценка не позволяет учесть влияние краевых эффектов. В литературе встречаются аналитические выражения для расчёта ёмкостей с учётом краевых эффектов [1, 2], однако данные выражения имеют достаточно ограниченный интервал применимости. Произвести расчет электрофизических параметров в МЭМС с учётом краевых эффектов для конструкций любой сложности позволяет использование численных методов, таких как метод конечных (МКЭ). Исследования [3] показывают, что толщина электродов, не учитываемая при расчетах по ИПК, может вносить существенный вклад в электрическую ёмкость за счёт краевых эффектов.

В данной работе исследуются зависимости электрической емкости и электростатических сил от геометрических параметров системы с электростатическими силами.

тродами конечной толщины. В качестве основы для расчетов рассматривается двухэлектродная конструкция МЭМ-генератора энергии.

Конструкция двухэлектродной МЭМС представлена на рисунке 1(а). Подвижный электрод закреплён на основании с тем же электрическим потенциалом, что и он сам. Для того, чтобы показать влияние основания подвижного электрода на характеристики системы анализ зависимостей ёмкости C и электростатических сил F от межэлектродного зазора d проводился для двух моделей МЭМС, изображённых на рис. 1(б, в). В качестве параметров варьировались различные соотношения ширины b и толщины h электродов к максимальному межэлектродному зазору d_0 , а также рассматривались случаи $a \gg b$ и $a = b$, где a – длина электродов. Для расчётов МКЭ использовался программный пакет ANSYS. В качестве окружающей среды при моделировании был использован воздух. Рассматривался случай постоянного напряжения на электродах, в качестве граничных условий полагалось, что потенциал одного из электродов был 1 В, а другого 0 В.

На рисунках 2(а, б) представлены зависимости нормированной ёмкости C/C_0 и электростатической силы F/F_0 от нормированного зазора d/d_0 при различных отношениях b/d_0 и для бесконечно длинных прямоугольных электродов с $a \gg b$. Линии с пустыми маркерами соответствуют численным расчётам для модели без учёта основания подвижного электрода (см. рис. 1б), линии с заполненными маркерами – с учётом основания подвижного электрода (см. рис. 1в). Чтобы оценить степень влияния краевых эффектов, полученные зависимости $C(d)$ и $F(d)$ нормировались на соответствующие значения, получаемые в рамках модели ИПК: $C_0 = \varepsilon_0 \varepsilon ab/d$ для ёмкости и $F_0 = -\varepsilon_0 \varepsilon abV^2/2d^2$ для нормальной составляющей электростатической силы.

Анализ показывает, что с ростом соотношения d/b и увеличением толщины h электродов влияние краевых эффектов усиливается, так, например, в случае с $h/d_0 = 0,2$ численно рассчитанные значения ёмкости могут превосходить ёмкость ИПК почти в 1,9 раза, а значения силы – в 1,22 раза при $d/d_0 = 0,5$ и $b/d_0 = 1$. Отсюда следует, что в таких случаях, использование расчетов по модели ИПК дает серьезную погрешность и ее применение нежелательно.

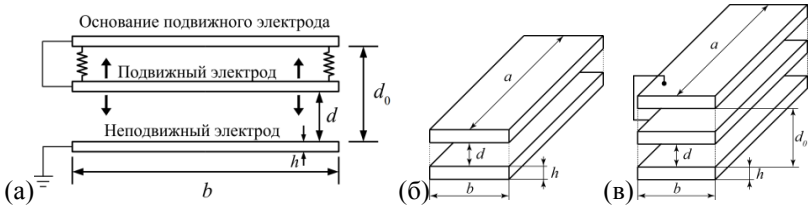
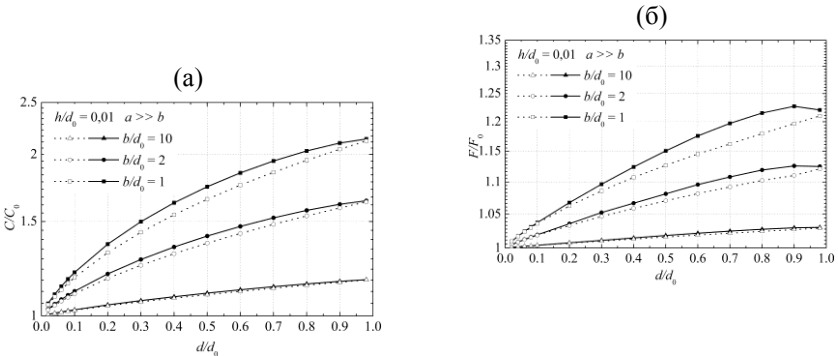


Рисунок 1 – Схематическое изображение двухэлектродной МЭМС

На рисунке 2 также можно видеть, что разница между зависимостями, рассчитанными с учётом и без учёта основания подвижного электрода, растёт с уменьшением отношения b/d_0 и наиболее сильно себя проявляет примерно при $d/d_0 = 0,5$ для емкости и при $d/d_0 = 0,75$, для нормальной составляющей электростатической силы.

На рисунках 2(в, г) представлены зависимости нормированной ёмкости C/C_0 и электростатической силы F/F_0 для квадратных электродов ($a = b$). Видно, что при переходе от случая $a \gg b$ характер влияния геометрических параметров системы на значения емкости и силы сохраняется, однако вклад краевых эффектов увеличивается. Так, например, в случае с $h/d_0 = 0,2$ численно рассчитанные значения ёмкости уже могут превосходить ёмкость ИПК почти в 3,1 раза, а значения силы – в 1,36 раза при $d/d_0 = 0,5$ и $b/d_0 = 1$. Видно, что если принимать во внимание длину электродов, то отклонение значений, получаемых по модели ИПК, от значений МКЭ оказывается еще более существенным.



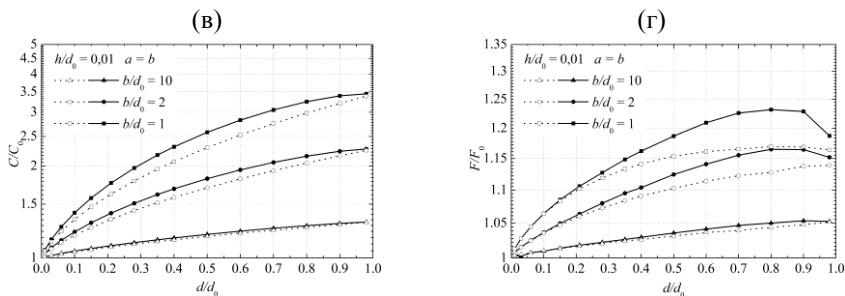


Рисунок 2 – Зависимости нормированной ёмкости C/C_0 и электростатической силы F/F_0 от нормированного зазора d/d_0 при различных параметрах:

для C/C_0 : (а) – $a \gg b$, (в) – $a = b$; для F/F_0 : (б) – $a \gg b$, (г) – $a = b$

Анализ показывает, что разница между ёмкостями и силами, рассчитанными с учётом и без учёта основания подвижного электрода при $a = b$, может достигать 11% для ёмкости при $h/d_0 = 0,01$, $d/d_0 = 0,5$ и $b/d_0 = 1$ и 5% для силы при $h/d_0 = 0,01$, $d/d_0 = 0,75$ и $b/d_0 = 1$. Это говорит о необходимости учитывать влияние основания подвижного электрода МЭМС в таких случаях при проведении расчетов.

В работе представлены результаты теоретических исследований влияния особенностей конструкции МЭМС на электрическую ёмкость и электростатические силы с учётом конечной толщины электродов. При помощи МКЭ рассчитаны зависимости $C(d)$ и $F(d)$ для МЭМС состоящей из двух электродов с учётом и без учёта влияния основания подвижного электрода.

Показано, что разница между параметрами C и F , рассчитанными с учётом и без учёта основания подвижного электрода, растёт с уменьшением ширины, длины и толщины электродов. Приведены оценки погрешности расчётов исследуемых параметров при учете основания подвижного электрода.

Литература:

1. Доржиев В.Ю., Драгунов В.П., Остертак Д.И. Расчёт ёмкости МЭМС в 2D-приближении // Сборник научных трудов НГТУ. 2010. № 4 (62). С. 73–80.

2. Драгунов В.П., Остертак Д.И. Электростатические взаимодействия в МЭМС с плоскопараллельными электродами. Часть I. Расчёт ёмкостей // Нано- и микросистемная техника. 2010. № 7. С. 37–41.

3. Остертак Д.И. Анализ электростатических взаимодействий в плоскопараллельных МЭМС с учетом краевых эффектов в 3D-приближении // Доклады академии наук высшей школы Российской Федерации. 2017. № 1 (34). С. 116–132.

ЧИСЛЕННОЕ МОДЕЛИРОВАНИЕ ЛЕВИТАЦИИ ЧАСТИЦ В УЛЬТРАЗВУКОВОМ ПОЛЕ

С. Росляков, Д.Я. Суханов
Томский государственный университет
г. Томск, sergea_1994_0@mail.ru

Разработана численная модель левитации частиц в ультразвуковом поле на основе метода конечных разностей во временной области. Силовое воздействие ультразвукового поля на частицы вычислено на основе потенциала Горькова, учитывается сила трения в газах для сферических частиц. С помощью технологии OpenCL реализовано параллельное вычисление траекторий множества частиц.

A numerical model for the levitation of particles in an ultrasonic field is developed based on the finite difference time domain method. The force effect of the ultrasonic field on the particles is calculated on the basis of the Gor'kov potential, and the friction force in gases for spherical particles is taken into account. With the help of OpenCL technology, a parallel calculation of the trajectories of a set of particles is implemented.

Вопросам управление частицами в ультразвуковом поле в последнее время уделяться большое внимание. Технология акустической левитации частиц в воздухе имеет перспективы применения в качестве бесконтактного пинцета для манипуляции малоразмерными объектами. Такие методы найдут применение в 3D печати, при производстве лекарств, для изготовления сложных химических примесей. В работе Горькова дано выражение для радиального давления на малой сферической частице[1]. В работе [2] произведен расчет акустического давления для частицах сложной формы. Акустическое поле может вызывать, как эффекты притяжения, так и отталкивания [3]. Исследования в области акустического воздействия на частицы

привлекают внимание исследователей из разных стран [4-7].

Для предсказания поведения частиц в ультразвуковом поле необходимо проведение численного моделирования воздействия ультразвукового поля на частицы.

Вычисление силы действующей на частицу в ультразвуковом поле можно осуществить с помощью формулы Горькова:

$$\mathbf{F} = -\nabla U, \quad (1)$$

$$\text{где } U = 2\pi R^3 \rho \left(\frac{\overline{p^2}}{3\rho^2 c^2} f_1 - \frac{\overline{v^2}}{2} f_2 \right) \text{ – потенциал Горькова,}$$

$$f_1 = 1 - \frac{c^2 \rho}{c_0^2 \rho_0}, \quad f_2 = 2 \frac{\rho_0 - \rho}{2\rho_0 + \rho}, \quad \rho \text{ – плотность среды (воздуха), } \rho_0 \text{ –}$$

плотность вещества частиц.

Траектория частиц моделируется на основе уравнений движения: $\mathbf{r}(t + \Delta t) = \mathbf{r}(t) + \mathbf{v}(t)\Delta t + \mathbf{a}\Delta t^2 / 2$ – координаты частицы в следующий момент времени, $\mathbf{v}(t + \Delta t) = \mathbf{v}(t) + \mathbf{a}\Delta t$ – скорость частицы, $\mathbf{a} = \mathbf{F} / m$ – ускорение частицы, m – масса частицы.

Было разработано программное обеспечение на основе технологии параллельного программирования OpenCL для одновременного вычисления траектории множества частиц в ультразвуковом поле. Было проведено моделирование левитации частиц пластика диаметром от 100 до 1000 мкм в ультразвуковом поле скрещенных стоячих волн на частоте 40 кГц в воздухе со скоростью звука 340 м/с. По вертикальной оси поле имеет огибающую в виде гауссоиды, что должно обеспечить левитацию в поле силы тяжести. Разработанная программа позволила рассчитать положения 2048 частиц со случайным начальным распределением координат (рисунок 1а) и получить положение частиц через 500 мс (рисунок 1б).

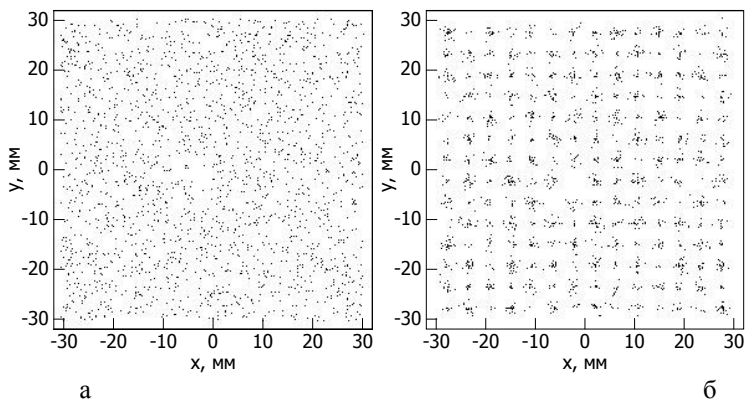


Рисунок 1 – Начальное распределение частиц (а) и через 500 мс (б) в ультразвуковом поле частотой 40 кГц и амплитудой 6 кПа.

Работа выполнена при поддержке гранта Российского научного фонда № 17-79-20051.

Литература:

1. Горьков Л.П. О силах, действующих на малую частицу в акустическом поле в идеальной жидкости// Доклады академии наук СССР. 1961, том 140 №1. С. 88-91.
2. Б. П. Шарфарец. Радиационное давление при рассеянии произвольного поля на включении сложной формы // Акустический журнал, 2010, том 56, № 6, с. 767-772
3. F.G. Mitri. Acoustic attraction, repulsion and radiation force cancellation on a pair of rigid particles with arbitrary cross-sections in 2D: Circular cylinders example // Annals of Physics. Vol. 386. – 2017. 1–14 DOI: 10.1016/j.aop.2017.08.023
4. M. Settnes, H. Bruus. Forces acting on a small particle in an acoustical field in a viscous fluid // Physical review E. Vol. 85, 016327. – 2012. DOI: 10.1103/PhysRevE.85.016327
5. Diego Baresch, Jean-Louis Thomas, Regis Marchiano. Three-dimensional acoustic radiation force on an arbitrarily located elastic sphere // The Journal of the Acoustical Society of America. Vol. 133, 25 (2013). DOI:10.1121/1.4770256.
6. Shilei Liu, Yanye Yang, Zhengyang Ni, Xiasheng Guo, Linjiao Luo, Juan Tu, Dong Zhang, Jie Zhang. Investigation into the effect of acoustic

radiation force and acoustic streaming on particle patterning in acoustic standing wave fields // Sensors 2017, Vol. 17, 1664; DOI: 10.3390/s17071664

7. P. B. Muller, R. Barnkob, M. Jakob Herring Jensen, H. Bruus. A numerical study of microparticle acoustophoresis driven by acoustic radiation forces and streaming-induced drag forces // The Royal Society of Chemistry. – 2012. DOI: 10.1039/c2lc40612h

МУЛЬТИПОЛЬНЫЙ МЕТОД ГРАНИЧНЫХ ЭЛЕМЕНТОВ С ИСПОЛЬЗОВАНИЕМ БАЗИСНЫХ ФУНКЦИЙ ВЫСОКОГО ПОРЯДКА

И.А. Толстобров, И.М. Ступаков
Новосибирский государственный технический университет,
г.Новосибирск, zortnaga@gmail.com

Метод граничных элементов является одним из методов численного моделирования. Использование быстрого мультипольного метода позволяет реализовывать метод граничных элементов, затрачивая меньшее, по сравнению с оригинальным методом, количество памяти и времени. Для повышения точности решения можно использовать базисные функции высокого порядка. В докладе рассказывается идея реализации мультипольного МГЭ с базисными функциями высокого порядка.

The boundary element method is one of the methods of numerical simulation. Usage of the fast multipole method allows to implement the boundary element method, while spending a small amount of memory and time. It is possible to use high order basis functions to improve the accuracy of the solution. The report describes ideas of the implementation of the multipole BEM with high order basic functions.

Методы численного моделирования используются в широком спектре областей для численного решения задач упругости, электромагнетизма, текучести и т.д.

Существуют различные методы численного моделирования: метод конечных разностей, метод конечных элементов (МКЭ), метод граничных элементов (МГЭ) и другие.

Метод граничных элементов позволяет находить решение дифференциальных уравнений в области, проводя дискретизацию лишь на границе области. Поэтому, при увеличении подробности сетки в k

раз, размерность получаемой СЛАУ растёт в k^2 раз, в то время как в МКЭ - в k^3 .

Однако, если учитывать что матрица в МГЭ является плотной, а в МКЭ – разреженной, сложность операции умножения матрицы на вектор после дробления сетки в k раз в каждом случае соответственно растёт в k^4 и k^3 раз. Этот факт затрудняет применение МГЭ на практике.

Для решения этой проблемы возможно использование методов, позволяющих не хранить напрямую СЛАУ в МГЭ, что улучшает показатели затрат как памяти, так и времени применения матричных операций. В частности, возможно использование быстрого мультипольного метода.

При заполнении граничноэлементной матрицы необходимо вычислять значения интегралов вида

$$\iint_{\Omega} \varphi(\vec{x}) \cdot \frac{1}{|\vec{x} - \vec{y}|} d\vec{x}$$

и

$$\iint_{\Omega} \psi(\vec{x}) \cdot \frac{\partial}{\partial n_{\Omega}} \frac{1}{|\vec{x} - \vec{y}|} d\vec{x},$$

полученных из граничноэлементных дифференциальных уравнений.

Здесь \vec{x} и \vec{y} – векторы в трехмерном пространстве, Ω – двумерная область в трехмерном пространстве, φ и ψ – полиномиальные базисные функции, а \vec{n}_{Ω} – внешняя нормаль к Ω .

Из-за неограниченности подынтегральной функции в окрестности $\vec{x} = \vec{y}$ численное взятие этих интегралов с достаточной точностью может стать весьма ресурсозатратной задачей.

Однако они могут быть вычислены аналитически, в случае, когда базисные функции являются константами. В работе [1] описывается алгоритм, позволяющий по средствам рекуррентных соотношений сводить к этой задаче процедуру вычисления таких интегралов с произвольными базисными функциями.

Если же $|\vec{x} - \vec{y}| \gg 0$, можно использовать стандартные процедуры численного интегрирования, получая достаточную точность без излишнего использования ресурсов.

Кроме того, в случае когда $|\vec{x} - \vec{y}| \gg 0$, не нужно напрямую хранить значения описанных интегралов, вместо этого вычисляя их посредством использования представления фундаментального решения в виде ряда из произведений шаровых функций. Используя октодереву, листьями которого являются граничные элементы, с помощью специальных формул пересчета возможно выполнять операции с матрицей, используя меньше времени и памяти, чем при хранении ее напрямую.

В докладе представлены сравнения точности и затраченного времени при решении задачи магнитостатики с использованием МКЭ, классического МГЭ и мультипольного МГЭ.

Полученные результаты иллюстрируют целесообразность использования вариаций МГЭ для решения задач численного моделирования. Несмотря на существование программных комплексов, решающих задачи магнитостатики, проблема уменьшения затрат памяти и времени и увеличения качества решения остается актуальной.

Литература:

1. Толстобров И.А., Ступаков И.М. Вычисление сингулярных интегралов для базисных функций высокого порядка в методе граничных элементов с применением рекуррентных соотношений «Обработка информации и математическое моделирование». Новосибирск, 2018. С 139-143
2. Greengard L., Rokhlin V. A new version of the fast multipole method for the Laplace equation in three dimensions //Acta numerica. – 1997. – Т. 6. – С. 229-269.
3. Steinbach O. Numerical approximation methods for elliptic boundary value problems. – New York : Springer, 2008.

МОДЕЛЬ ФОРМИРОВАНИЯ ГРАНИЦЫ РАЗДЕЛА «ПОКРЫТИЕ – ПОДЛОЖКА» ПРИ ЭЛЕКТРОВЗРЫВНОМ НАПЫЛЕНИИ БИОИНЕРТНЫХ ПОКРЫТИЙ

А.Д. Филяков, Д.А. Романов, С.А. Невский, К.В. Соснин
Сибирский государственный индустриальный университет,
г. Новокузнецк, filyakov.1999@mail.ru

Конструкционные материалы, используемые в биомедицинских применениях, должны обладать оптимальными механическими свойствами и биосовместимостью и должны оставаться химически стабильными в условиях жесткой враждебности. Применение объемного легирования биоинертных материалов для достижения необходимых требований часто является неэкономичным, а в некоторых случаях технически невозможным. Традиционные способы химико-термической обработки – это энергоемкие и длительные процессы, поэтому в последние годы находят развитие и применение новые способы упрочнения металлов и сплавов, основанные на использовании концентрированных потоков энергии. Практика показала, что эффективным инструментом упрочнения и защиты поверхности металлов и сплавов различного назначения могут служить многофазные плазменные струи продуктов электрического взрыва проводников. Электровзрывное напыление импульсными плазменными струями является одним из таких перспективных методов. Контактная граница, отделяющая расплавленный материал подложки и жидкий металл электрически взорванной фольги, при определенных условиях становится неустойчивой, что влечет за собой перемешивание потоков. В данной работе рассматривается получение и анализ дисперсионного уравнения, описывающего начальную стадию развития неустойчивости Рэлея – Тейлора в рамках уравнений Навье-Стокса с соответствующими краевыми условиями на примере процесса нанесения электровзрывного биоинертных покрытий системы Ti-Nb.

Engineering materials used in biomedical applications must have optimal mechanical properties and biocompatibility, and should remain chemically stable under severely hostile conditions. Application of volume-doped materials is often uneconomical and sometimes technically impossible. Traditional methods of thermochemical treatment are energy-intensive and long-lasting processes, therefore, in recent years, new methods of hardening metals and alloys, based on the use of concentrated energy flows, have been developed and applied. Experience has shown that multiphase plasma jets created by electric explosion are usable as effective instrument for hardening and protection of metals and alloys for various purposes. Electroexplosive deposition by pulsed plasma jets is one such promising method. The contact boundary separating the molten substrate material and the liquid metal of the electrically exploded film. In certain circumstances, boundary becomes unstable resulting in mixing streams. This article reviews obtaining and analyzing the dispersion equation describing the initial

stage of development of the Rayleigh-Taylor instability within the boundary condition of the Navier-Stokes equations with the corresponding boundary conditions using the example of applying electro-explosive bioinert binary Ti-Nb coatings.

Конструкционные материалы, используемые в биомедицинских применениях, должны обладать оптимальными механическими свойствами и биосовместимостью и должны оставаться химически стабильными в условиях жесткой враждебности. По сравнению с биомедицинскими нержавеющими сталями и сплавами кобальт-хром сплавы на основе титана (Ti) обладают отличной стойкостью к коррозии [1].

Применение объемного легирования биоинертных материалов для достижения необходимых требований часто является неэкономичным, а в некоторых случаях технически невозможным.

Традиционные способы химико-термической обработки – это энергоемкие и длительные процессы, поэтому в последние годы находят развитие и применение новые способы упрочнения металлов и сплавов, основанные на использовании концентрированных потоков энергии.

Электровзрывное напыление импульсными плазменными струями является одним из таких перспективных методов.

Развитие неустойчивости контактной границы, движущейся с постоянным ускорением – неустойчивость Рэлея-Тейлора, также может использоваться для изучения начальной стадии перемешивания при электровзрывном напылении металла на подложку.

Для формирования многофазных плазменных струй и исследования процессов ЭВЛ в настоящей работе использовали лабораторную электровзрывную установку ЭВУ 60/10М.

Первоначально была определена скорость плазмы на срезе сопла:

$$v = 6,75 \cdot 10^3 \left(\frac{CU^2}{2\tau_u} \right)^m, \quad (1)$$

C – емкость конденсатора;

U – зарядное напряжение;

m – показатель степени, может принимать значения от 0,05 до 0,25.

Считая соотношение (1) справедливым и для электровзрывных источников, примем $m = 0,05$.

Ускорение плазмы находим по формуле (2):

$$g = \frac{v^2}{2S} \quad (2)$$

v – скорость плазмы, а S – расстояние от среза сопла до поверхности обрабатываемой пластины.

При численном решении уравнений (3) было впервые для неустойчивости Кельвина-Гельмгольца получена дисперсионная зависимость, в которой наблюдаются два максимума [2] – один из них в нанодиапазоне. В [2] показывается что, зависимость декремента от длины для вязко-потенциальной жидкости совпадает с такой же зависимостью, полученной из дисперсионного уравнения (3):

$$G = \frac{-4z(z^2+1)}{zch(zd)sh(d) - c(zd)sh(d)} + \frac{z(z^4+2z^2+5) - (z^4-6z^2+1)th(d)th(zd)}{zth(d) - th(zd)}, \quad (3)$$

в широком диапазоне параметров для неустойчивости Кельвина-Гельмгольца. Поэтому проведем анализ неустойчивости Релея-Тейлора, основываясь на дисперсионном уравнении для вязко-потенциальной жидкости, где вязкость учитывается только на границе. В этом случае получается квадратное уравнение с комплексными коэффициентами:

$$a\omega_R^2 + 2b\omega + c = 0, \quad a = cth(kh)\rho_1 + cht(kH)\rho_2 \quad (5)$$

$$b = ik^2(\rho_1 v_1 cth(kh)), \quad c = g(\rho_2 - \rho_1)k - \sigma_0 k^3.$$

Подставим $\omega = \omega_R - i\alpha$ и получим систему для нахождения корней:

$$\begin{cases} a\omega_R^2 - a\alpha^2 - b\alpha + c = 0, \\ a\omega_R\alpha + b\omega_R = 0 \end{cases} \quad (6)$$

эта система имеет одно решение, что соответствует появлению одного максимума в зависимости инкремента неустойчивости α от λ , при решении (6) представлены на (Рисунок 1) для различных значений плотности второго слоя.

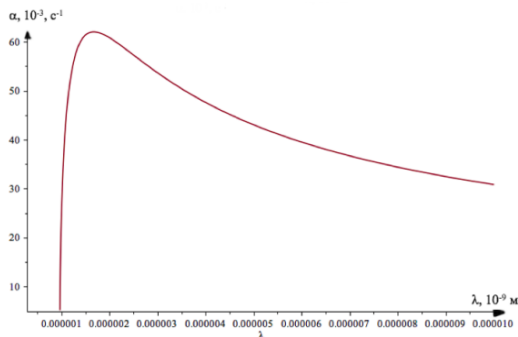


Рисунок 1 – Зависимости инкремента α от длины волны λ для покрытия системы Ti-Nb при значении зарядного напряжения равного 1,8 кВ

Необходимо отметить, что значения длин волн, при которых реализуется максимум декремента λ_m , составляют X мкм, что соответствует эксперименту [3]. В рассматриваемом (при $U = 1,8$ кВ) случае значение λ_m для электровзрывных покрытий на основе ниобия 1,59 мкм.

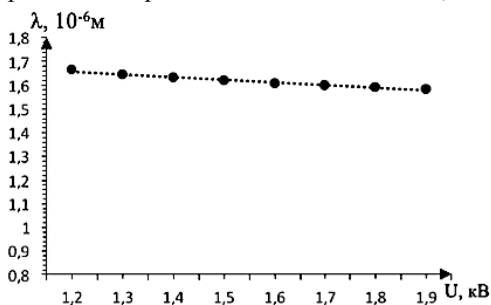


Рисунок 2 – зависимость длины волны для покрытия от зарядного напряжения

Была установлена прямая линейная зависимость длины волны от зарядного напряжения. Теоретически показано, что эффект перемешивания будет заметен при воздействии пары ниобий – титан.

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 18-32-00075 мол_а.

Литература:

1. Influence of thermomechanical processing on biomechanical compatibility and electrochemical behavior of new near beta alloy, Ti-20.6Nb-13.6Zr-0.5V / Mohsin Talib Mohammed, Zahid a Khan, Geetha Manivasagam, Arshad N Siddiquee // International Journal of Nanomedicine – 2015. Vol. 10. – P. 233-235.
2. Электровзрывное напыление износ- и электроэрозионностойких покрытий / Д. А. Романов, Е. А. Будовских, В. Е. Громов, Ю. Ф. Иванов. – Новокузнецк: Изд-во ООО «Полиграфист», 2014. – 203 с.
3. А.Я. Багаутдинов, Е.А. Будовских, Ю.Ф. физические основы электровзрывного легирования металлов и сплавов. Новокузнецк: СибГИУ, 2007. 297 с.

**Секция ТЕХНОЛОГИИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА,
МАШИННОГО ОБУЧЕНИЯ И ПРЕДСКАЗАТЕЛЬНОГО МОДЕЛИ-
РОВАНИЯ**

**РАЗРАБОТКА МОДУЛЕЙ К СИСТЕМЕ ОБУЧЕНИЯ
МАШИНИСТОВ ПУТЕУКЛАДЧИКА УК-25/9-18**

**Э.С. Бондарев, Д.В. Ермакова, А.В. Шаповалов, Н.А. Маслов
Сибирский государственный университет путей сообщения,
г. Новосибирск, bondareff.edik@yandex.ru**

В рамках работ, выполняемых по гранту ОАО «РЖД» для создания электронного учебника и тренажера «Путеукладчик УК-25/9-18», представлены модули системы обучения машиниста путеукладочного крана (разделы электронного учебника, в т.ч. видеоуроки с аудиосопровождением, конспекты лекций, словарь терминов, тестовые задания); модули тренажера-симулятора путеукладочного крана (3D модели элементов и крана; сценарии уровней игры-симулятора крана - техническое задание для программистов).

In the framework of the work performed under the grant of Russian Railways for the creation of electronic textbook and simulator "Laying crane UK-25/9-18" designed modules of the system training driver track-laying crane (sections of the electronic textbook, including videos with audio, teaching notes, a Glossary, test tasks); modules of the simulator of track-laying crane (3D models of elements and of the crane; script game levels - simulator crane - technical specification for programmers).

Цель работы: создание модулей системы обучения машинистов укладочного крана УК-25/9-18 (работа выполнена по гранту ОАО «РЖД» и гранту СГУПС на выполнение НИРС).

Задачи работы: разработка модулей электронного курса обучения машинистов кранов (разделы электронного учебника, в т.ч. сценарии видеоуроков, озвучивание видеоуроков; разработка тестовых заданий по разделам); разработка модулей тренажера-симулятора путеукладочного крана (3D моделирование элементов крана; сценарии уровней игры-симулятора крана); оценка эффективности практического применения системы обучения машинистов укладочного крана УК-25/9-18 на производстве.

Научная новизна работы заключается в создании эффективного инструмента (средства) совершенствования методики обучения рабочей

профессии «Машинист укладочного крана УК-25/9-18», не имеющего известных аналогов в этой области.

Методы исследований: компьютерное моделирование узлов, систем и крана в целом; анализ и классификация способов подготовки машинистов укладочного крана, анализ и классификация методов разработки учебного тренажера.

Существует два способа подготовки машинистов укладочного крана. Первый способ основан на применении реальной машины (более дорогостоящий, т.к. включает расходы на топливо, горюче-смазочные материалы, а также заработную плату бригаде машинистов), а второй предполагает использование тренажера. Примерный учебный план курса обучения машиниста крана составляет 960 часов. Практическая часть обучения в среднем занимает треть длительности обучения. При этом календарная длительность практической части зачастую оказывается слишком расщепленной в связи с невозможностью использования для обучения реального крана более чем на полтора-два часа в день. Использование электронного учебника и учебного тренажера (далее по тексту – система обучения) в практической подготовке машинистов повышает эффективность обучения и значительно снижает затраты на подготовку специалиста по этой рабочей профессии.

Система обучения для машинистов укладочного крана разработана с учетом требований учебных планов и программ для профессионального обучения по профессии машинист железнодорожно-строительных машин (ЖДСМ) ОАО «РЖД». Система обучения соответствует темам предмета специального курса «Управление, техническое обслуживание и эксплуатация ЖДСМ»: эксплуатация и управление краном (десять лабораторно-практических занятий). По теме «Управление железнодорожно-строительной машиной» система обучения обеспечивает: запуск двигателя, контроль работоспособности систем узлов и агрегатов, опробование тормозов; начало движения, разгон, торможение, остановку, реверсирование движения; перевод крана в рабочее и транспортное положения; управление краном, силовой установкой, отдельными рабочими узлами, механизмами и системами. Система обучения обеспечивает управление краном во время укладки железнодорожного пути с одного из двух постов управления (пост управления экипажной частью крана и пост управления крановым оборудованием). Управление краном включает управление: низковольтной схемой экипажной части; высоковольтной схемой экипажной части; тормозным оборудованием; низковольт-

ной схемой кранового оборудования; высоковольтной схемой кранового оборудования; гидравлическим оборудованием.

Рассмотрены варианты разработки учебного тренажера двумя методами: «вероятный»; «традиционный». Путь создания тренажера по «вероятному» методу состоит из трех этапов: закупка готового тренажера и базового программного обеспечения; разработка 3D модели машины и дополнительного программного обеспечения (ПО2); закупка стандартного программного обеспечения (ПО3) для моделирования машины. Из блок-схемы тренажера по «вероятному» методу видно, что программное обеспечение создается исходя из конфигурации пульта управления. Такое направление создания тренажера влечет за собой проблемы совместимости пульта управления с контроллером, контроллера с программным обеспечением (ПО2) 3D модели машины и (или) с программным обеспечением (ПО3) моделей систем машины, а также проблема совместимости ПО2 с ПО3. К недостаткам данного варианта создания тренажера можно отнести низкое качество визуализации рабочего процесса и высокую стоимость тренажера. «Традиционный» метод создания тренажера состоит также из трех этапов: закупка стандартного программного обеспечения (ПО) для моделирования систем машины; закупка стандартных контроллеров и командных органов; разработка 3D-модели машины и оборудования, а также дополнительного программного обеспечения (ПО2). В работе принят «традиционный» метод разработки учебного тренажера машиниста укладочного крана, как наиболее эффективный.

Результаты работы: модули электронного курса обучения машинистов путеукладочного крана (разделы электронного учебника, в т.ч. сценарии и конспекты лекций к видеурокам, видеуроки с аудиосопровождением, словарь терминов, тестовые задания); модули тренажера – симулятора путеукладочного крана (3D модели деталей и сборочных единиц крана (Рисунок 1); сценарии уровней игры – симулятора крана – техническое задание для программистов; игра-симулятор крана); оценка эффективности практического применения системы обучения машинистов укладочного крана УК-25/9-18 на производстве (рекомендации по устранению недостатков в подготовке машинистов); снижение затрат на подготовку квалифицированных специалистов для ОАО «РЖД».

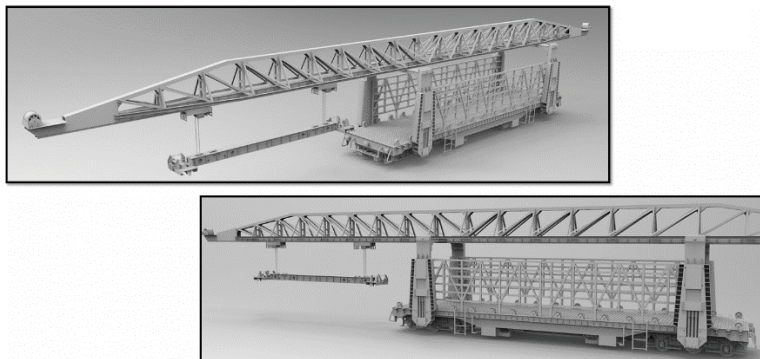


Рисунок 1 – 3D модель путеукладчика УК-25/9-18

В программных комплексах SolidWorks и Компас 3D разработаны трехмерные модели нескольких сотен деталей и десятки сборочных единиц. Написаны сценарии и конспекты лекций десяти видеоуроков, словарь терминов к электронному учебнику по эксплуатации и управлению укладочным краном. На основе сценариев к видеоурокам разработано 120 тестовых заданий к электронному учебнику по эксплуатации и управлению укладочным краном. На основе конспектов видеоуроков написаны сценарии десяти уровней к компьютерной игре-симулятору укладочного крана на основе конспектов видеоуроков, включающие в себя: текстовую формулировку действий и 120 заданий для игрока, графические пояснения для программистов (изображения деталей и узлов из видеоуроков). На основе полученных результатов программистами разработана игра-симулятор крана (с использованием межплатформенной среды разработки компьютерных игр – Unity), включающая десять уровней в двух режимах (обучение и тестирование).

В программе Automation Studio разработаны виртуальные модели гидравлической, пневматической и электрической систем крана, позволяющие выполнять поиск и выбирать пути устранения неисправностей узлов и систем.

Результаты научно-исследовательского проекта предполагается внедрить в СГУПС при подготовке специалистов по наземным транспортно-технологическим средствам (в т.ч. на базовой кафедре «Устройство и эксплуатация путей машин», ПМС-20, ст. Крахаль), при повышении квалификации кадров ОАО РЖД в ИТТиПК СГУПС, а также на сети Российских железных дорог.

Литература

1. Бондарев Э.С. Болбат О.Б. Разработка модуля «Привод колесной пары» для тренажера машиниста «Укладочный кран» // Научное сообщество студентов XXI столетия. Технические науки: сб. науч. тр. – Новосибирск: Изд-во АНС «СибАК», 2017. – С. 6-8.
2. Бондарев Э.С. Маслов Н.А. Разработка модуля «Привод грузовой лебедки» для тренажера машиниста «Укладочный кран» // Научное сообщество студентов XXI столетия. Технические науки: сб. науч. тр. – Новосибирск: Изд-во АНС «СибАК», 2018. – С. 79-83.
3. Бондарев Э.С. Маслов Н.А. Электронный курс для подготовки операторов укладочных кранов // Интеллектуальный потенциал Сибири: сб. науч. тр. – Новосибирск: Изд-во НГТУ, 2018. – С. 275-277.
4. Бондарев Э.С. Ермакова Д.В. Шаповалов А.В. Разработка модулей тренажера машиниста укладочного крана // Подъемно-транспортные, строительные, дорожные, путевые машины и робототехнические комплексы: сб. науч. тр. – Москва: Изд-во МАДИ, 2018. – С. 124-126.

СВЕРТОЧНЫЕ НЕЙРОННЫЕ СЕТИ ДЛЯ КЛАССИФИКАЦИИ ЧАСТИЧНО РАЗМЕЧЕННЫХ ИЗОБРАЖЕНИЙ

А.И. Ветров, А.А. Попов

**Новосибирский государственный технический университет,
г. Новосибирск, vetrov.andrei.95@gmail.com**

Сверточные нейронные сети имеют хорошие результаты в задаче классификации изображений, при этом, для достижения таких результатов требуются большие наборы размеченных данных. В случае, если размеченных данных слишком мало, можно воспользоваться методами обучения с частичным привлечением учителя. В данной работе были рассмотрены основные подходы в задаче обучения на частично размеченных данных. Были проведены исследования качества работы алгоритмов на различных наборах данных.

Convolutional neural networks have good results in the task of image classification, while to achieve such results, large sets of labeled data are required. In case the marked data is too small, you can use semi-supervised learning methods. In this paper, we reviewed the main approaches in the task of learning on partially marked data. Studies have been conducted on the quality of the algorithms on various data sets.

В некоторых задачах сверточные сети превосходят человека по качеству распознавания. Для того чтобы достичь таких успехов, сеть пришлось обучить на очень большом наборе размеченных изображений. К сожалению, получить такое количество данных является сложной задачей. Изображений может быть слишком много, что потребует достаточно времени для ручной разметки. Также, очень часто, для разметки изображений необходим эксперт, обладающий определенными знаниями в данной области (медицина, геология).

Одним из способов решения проблемы малого размера обучающей выборки являются методы обучения с частичным привлечением учителя. В отличие от методов обучения с учителем, которые используют только размеченные объекты, методы обучения с частичным привлечением учителя дополнительно используют неразмеченные объекты. Это позволяет понять структуру данных, что в свою очередь улучшает качество модели.

Опишем задачу обучения с частичным привлечением учителя более формально. В задаче обучения с учителем существует набор данных вида $(x, y) \in D$ из распределения $p(x, y)$, где x - изображения, а y - метка класса соответствующего изображения. Целью задачи является построить такую функцию $f(x, \theta)$, которая давала бы верную метку класса y для изображений из $p(x)$. В задаче обучения с частичным привлечением учителя дополнительно дается набор данных без меток $x \in D_{UL}$. Этот набор дополнительных данных будет использоваться для получения информации о виде распределения $p(x)$.

Выделяют три подхода для решения задачи классификация изображений методами обучения с частичным привлечением учителя:

1) Псевдо разметка (pseudo-labeling) [1]

Псевдо разметка - это эвристика, которая широко используется на практике. Суть данного метода заключается в том, что сеть, обученная на размеченной выборке, дает оценку вероятности принадлежности к каждому классу, элементам из неразмеченной выборки. Все элементы с вероятностью, превышающей заданный порог, используются при дальнейшем обучении сети, с этими метками классов.

2) Регуляризация согласованности (Consistency Regularization)

Этот класс методов описывает следующий подход: различное зашумление входного изображения $x \rightarrow \hat{x}$ не может изменять выход модели $f(x, \theta)$, то есть $f(x, \theta) = f(\hat{x}, \theta)$. Чаще всего при обучении сети,

к ошибке классификации добавляется ошибка вида $d(f(x, \theta), f(\hat{x}, \theta))$, где $d(\dots)$ - расстояние между двумя предсказаниями, обычно это среднеквадратическая ошибка или дивергенция Кульбака-Лейблера. Данный подход позволяет эффективно разделить кластеры соответствующие меткам класса. Рассмотрим несколько алгоритмов, которые используют данный подход.

а) π -модель [2]

Эта модель является простейшим вариантом реализации данного подхода. Она использует стандартные методы регуляризации нейронных сетей, такие как аугментации, зашумление и отсев нейронов, которые добавляют шум во входное изображение. Вычисленное расстояние между выходами модели с различными зашумлениями вычисляется по формуле $d(f(x, \theta), f(\hat{x}, \theta))$, где $d(\dots)$ - среднеквадратическая ошибка. Это расстояние добавляется к функции ошибки классификатора в виде масштабируемого гиперпараметра.

б) Средний учитель (mean teacher) [3]

Средний учитель похож на π -модель, но вместо изменения входного сигнала предлагается изменять параметры сети. $\hat{\theta}$ будет вычисляться как экспоненциальное скользящее среднее предыдущих параметров модели.

в) Виртуальное состязательное обучение (BCO) (virtual adversarial training) [4]

Вместо того, чтобы использовать внутреннюю случайность модели, виртуальное состязательное обучение аппроксимирует направление зашумления r_{adv} , которое наиболее существенно повлияет на выход функции предсказания. Возмущение может быть вычислено как:

$$r \sim N\left(0, \frac{\xi}{\sqrt{\dim(x)}} * I\right),$$

$$g = \nabla_r d(f(x, \theta), f(x + r, \theta)),$$

$$r_{adv} = \epsilon \frac{g}{\|g\|},$$

где ξ и ϵ являются скалярными гиперпараметрами. Затем аналогично предыдущим методам $d(f(x, \theta), f(x + r, \theta))$ добавляется к функции потерь в виде гиперпараметра.

3) Методы, основанные на минимизации энтропии (entropy-based) [5]

Суть данного метода заключается в том, чтобы заставить сеть делать максимально однозначные предсказания, независимо от фактического значения метки класса. Предполагая K возможных классов, можно ввести формулу минимизации энтропии:

$$-\sum_{k=1}^K f(x, \theta)_k \log f(x, \theta)_k$$

Этот метод не получил хороших результатов, но при этом его можно использовать для улучшения качества алгоритма BCO.

Исследование будет проводиться на двух наборах данных: CIFAR-10 и SVHN. CIFAR-10 – набор данных состоящий из 60000 изображений и содержащий в себе 10 классов объектов. SVHN – набор данных, содержащий 100000 изображений цифр с уличных знаков.

Проведем исследование качества классификации, взяв 4000 изображений для набора данных CIFAR-10 и 1000 изображений для SVHN. Метрикой качества будет процент неправильных ответов.

| Метод | CIFA R-10 | SV HN |
|---------------------------------|--------------|------------|
| Обучение с учителем | 20.26 % | 12. 83% |
| Псевдо разметка | 17.78 % | 7.6 2% |
| π -модель | 16.37 % | 7.1 9% |
| Средний учитель | 15.87 % | 5.6 5% |
| BCO | 13.86 % | 5.6 3% |
| BCO + минимизация эн- тропии | 13.13 % | 5.3 5% |

Как видно методы с частичным привлечением учителя, показывают лучший результат, чем методы обучения с учителем. При этом стоит заметить, что чем меньше размер обучающей выборки, тем больше отношение качества между методами обучения с учителем и методами с частичным привлечением учителя. При 4000 изображений улучшение в 1.5 раз, а при 1000 изображений, улучшение в 2.5 раза. Все это говорит о том, что в задаче классификации изображений методы обучения с частичным привлечением учителя могут найти множество применений.

Литература:

1. Lee D.H. Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks [Electronic resource] // ICML Workshop on Challenges in Representation Learning, –Atlanta, 2013. –URL: http://deeplearning.net/wp-content/uploads/2013/03/pseudo_label_final.pdf (accessed: 31.10.2018)
2. Samuli L, Timo A. Temporal ensembling for semi-supervised learning [Electronic resource] // Fifth International Conference on Learning Representations, –Toulon, 2017. –URL: <https://arxiv.org/abs/1610.02242> (accessed: 31.10.2018)
3. Szegedy C., Zaremba, W., Sutskever I., Bruna J., Erhan D., Goodfellow I., Fergus R. Intriguing properties of neural networks [Electronic resource] // In Second International Conference on Learning Representations, – Banff, 2014. –URL: <https://arxiv.org/pdf/1312.6199.pdf> (accessed: 31.10.2018)
4. Miyato T, Maeda S., Koyama M., Ishii S. Virtual adversarial training: a regularization method for supervised and semi-supervised learning [Electronic resource] // IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017. –URL: <https://arxiv.org/pdf/1704.03976.pdf> (accessed: 31.10.2018)
5. Grandvalet Y., Bengio Y. Semi-supervised learning by entropy minimization [Electronic resource] // Advances in Neural Information Processing Systems, –Vancouver 2005. –URL: <http://www.iro.umontreal.ca/~lisa/pointeurs/semi-supervised-entropy-nips2004.pdf> (accessed: 31.10.2018)

МОДЕЛИРОВАНИЕ МИМИКИ ЧЕЛОВЕКА

Г.И. Глухов, Т.А. Гулятьева

Новосибирский государственный технический университет,
г. Новосибирск, krksagl@gmail.com

Целью работы является разработка метода моделирования мимики человека. Необходимо учитывая звуковую дорожку речи человека синтезировать движение губ, соответствующее этой звуковой дорожке и создать видео его речи. Необходимо обработать и сопоставить звуковые функции для форм рта, учитывая контур губ в каждый момент времени, синтезировать новый контур и наложить его на исходный видеоряд так, чтобы он соответствовал входной звуковой дорожке.

The main purpose of scientific work is to develop a method or a program for modeling human facial expressions. Given audio of somebody person, it is necessary to synthesize a high quality video of this person speaking with accurate lip sync, composited into a target video clip. It is necessary to process and match the sound functions for the mouth forms, taking into account the contour of the lips at each moment in time, synthesize the new contour and superimpose it on the original video sequence so that it corresponds to the input sound track.

Современные методы решения такого типа задач базируются на построении рекуррентных нейронных сетей и их модификаций [1]. Рекуррентная нейронная сеть содержит обратные связи, позволяющие сохранять информацию, таким образом, возможно, предсказывать дальнейшие действия, опираясь на предыдущие. Это можно использовать для предсказания следующего кадра видео, опираясь на запись речи человека и предыдущие кадры.

Для реализации такого метода на начальном этапе требуется производить вычисление векторов целевых переменных в видеоканале и вектора признаков в звуковой дорожке. Для вычисления целевых переменных наиболее информативной составляющей является траектория движения губ говорящего. Для практической реализации операции отслеживания этой траектории в настоящее время широко используются активные контуры или активные контурные модели. В этом случае, как правило, требуется осуществить первичную подготовку анализируемого фрагмента изображения (выравнивание яркости, выделение интересующих фрагментов, бинаризация), а затем непосредственно провести адаптацию активного контура к полученной бинарной маске. Для выделения признаков в звуковом сигнале можно использовать мелкочастотные коэффициенты (MFCC), которые часто используются в

качестве характеристик звукового сигнала. Кепстр – энергетический спектр функции:

$$\ln|S(w)|^2,$$

определяемый выражением:

$$C_s(q) = \frac{1}{2\pi} \int \int_{-\infty}^{\infty} \ln |S(w)|^2 e^{i w q} dw.$$

Мел — единица высоты звука, основанная на восприятии этого звука нашими органами слуха, так как амплитудно-частотная характеристика человеческого уха не является прямой линией, а амплитуда не совсем точная мера громкости звука, поэтому вводятся эмпирически подобранные единицы громкости звука, например, фон. Фон - логарифмическая единица для оценки уровня громкости звука. Шкала фонов от шкалы децибелов отличается тем, что в ней значения громкости коррелируются с чувствительностью человеческого слуха на разных частотах. Подобные единицы измерения часто используют при решении задач распознавания, так как они позволяют приблизиться к механизмам человеческого восприятия, которое пока что лидирует среди известных систем распознавания речи. Итого MFCC - представление лог мощности спектра в мел частотной области. Описывают мощность огибающей спектра, которая (огибающая) характеризует модель речевого тракта. [2].

Поскольку признаки, которые необходимо выделить находятся в разных потоках видеозаписи, работа с ними ведется отдельно, а для их сопоставления используются временные промежутки.

Для обнаружения ключевых лицевых точек в реализованной программе используется библиотека dlib. При помощи обнаруженных лицевых ориентиров можно маркировать и извлекать области лица, такие как: рот, глаза, нос, брови и подбородок. Со всеми ориентирами лица связаны конкретные (x;y) координаты. Для извлечения этих точек используется модель, обученная на маркированном наборе данных iBUG 300-W dataset [3]. Эти сопоставления кодируются внутри словаря, используя который, можно извлечь индексы в массив ориентиров лица и затем извлечь различные черты лица. На данный момент для обучения нейронной сети используются только ключевые точки губ, поэтому извлекаются только они. Для получения wav-файла аудиопотока из видеозаписи для дальнейшей работы используется FFmpeg. Удобным форма-

том для работы со звуком в данном случае является WAV формат. WAV — это формат для хранения несжатого аудиопотока, широко используемый в медиа индустрии. Его особенность состоит в том, что для кодирования амплитуды выделяется фиксированное число бит. Это сказывается на размере выходного файла, но делает его очень удобным для чтения. Именно поэтому данный формат был выбран для преобразования в него потока звукоданных. MFCC-коэффициенты извлекаются при помощи библиотеки `python_speech_features`.

Для обучения используется нейронная сеть, в которой набор векторов mfcc-коэффициентов сопоставляется с вектором ключевых точек изображения идущим по времени после этого набора, и далее для каждого последующего вектора mfcc-коэффициентов в сопоставление ставится вектор точек.

Реализованный метод получает на вход последовательность характеристик звука записи речи человека, в виде MFCC-коэффициентов и на основе обученной нейронной сети, предсказывает положение губ, которое будет после произнесения человеком этой речи.

В дальнейшем планируется подключить к набору ключевых точек, не только рот, но и подбородок, а также нос. Остальные точки, такие как брови, скулы и глаза, будут подбираться в зависимости от положения точек, используемых для обучения, так как при речи, эти части человеческого лица менее подвижны и не играют важной роли при обучении. Также планируется попробовать использовать рекуррентные нейронные сети и LSTM-сети. После этого планируется наложить вектор полученных ключевых лицевых точек на текстуру лица для получение реального изображения. Также планируется уменьшить ошибку при помощи подбора оптимальных параметров для настроек нейронной сети.

Литература:

1. Synthesizing Obama: Learning Lip Sync from Audio / SUPASORN SUWAJANAKORN, STEVEN M. SEITZ, IRA KEMELMACHER-SHLIZERMAN ; University of Washington // ACM Transactions on Graphics / Vol. 36, No. 4, Article 95. Publication date: July 2017. – 2017. – Vol. 36, No. 4, Article 95. – С. 1-13.
2. Mel Frequency Cepstral Coefficient (MFCC) tutorial [Электронный ресурс]. –<http://practicalcryptography.com/miscellaneous/machine-learning/guide-mel-frequency-cepstral-coefficients-mfccs/>.

3. Facial point annotations [Электронный ресурс]. – <https://ibug.doc.ic.ac.uk/resources/facial-point-annotations/>.

НАБОР ДАННЫХ И МОДЕЛЬ ГЛУБОКОГО ОБУЧЕНИЯ ДЛЯ АНАЛИЗА ТЕКСТОВ ОТЗЫВОВ ПОЛЬЗОВАТЕЛЕЙ

Е.И. Грибков, Ю.П. Ехлаков
**Томский государственный университет систем
управления и радиоэлектроники,
г. Томск, drnemor@gmail.com**

В статье представлена модель разметки последовательности на основе структурного предсказания с использованием рекуррентной нейронной сети типа LSTM, позволяющая извлекать мнения пользователей из текстов отзывов. Для обучения и оценки качества работы модели создан размеченный набор текстов отзывов о мобильных телефонах из интернет-магазина Amazon. Обученная модель имеет $F1=61.9\%$ при извлечении аспектов, $F1=63.71\%$ - при извлечении оценочных высказываний.

The LSTM-based structured prediction model for sequence tagging capable for extraction of user opinions from reviews is described. The corpus of annotated user reviews about mobile phones from Amazon for model training and evaluation is created. Trained model shows $F1=61.9\%$ on aspects extraction task and $F1=63.71$.

Повсеместное распространение Интернета позволяет пользователям свободно обсуждать приобретаемые ими товары и услуги в сети. Для этого используются различные площадки: социальные сети, специализированные форумы, страницы интернет-магазинов. Анализ содержимого отзывов пользователей представляет большой практический интерес для коммерческих компаний, которые заинтересованы в повышении своей клиентоориентированности. Используя эти данные, руководство может формировать ассортимент, осуществлять индивидуальный подбор товаров и услуг и проводить различные кампании, направленные на повышение лояльности покупателя.

Существующие площадки решают проблему агрегации пользовательских мнений в одном месте, однако на сегодняшний день они не предлагают развитых средств для автоматизированного анализа отзывов пользователей.

Методы глубокого обучения широко используются в задачах обработки естественного языка, таких как частеречная разметка [1], анализ

синтаксической [2] структуры предложений, анализ тональности [3]. По сравнению с другими данный тип моделей не требует ручного задания входных признаков, вместе с тем достигая лучших показателей при решении задач.

В основе данной работы лежит понятие *мнения*. Мнение пользователя состоит из *аспекта* – характеристики, составной части или упоминания продукта, и *оценочного высказывания* – высказывания, характеризующее аспект мнения. Формально мнение определяется парой $O = \langle A(a_{begin}, a_{end}), D(d_{begin}, d_{end}) \rangle$, где $A(a_{begin}, a_{end})$ – аспект мнения, начинающийся на слове с порядковым номером a_{begin} и заканчивающийся на слове с номером a_{end} , $D(d_{begin}, d_{end})$ – оценочное высказывание, начинающийся на слове с порядковым номером d_{begin} и заканчивающийся на слове с номером d_{end} .

Для обучения и оценки качества модели необходимо сформировать размеченную выборку текстов отзывов с выделенными мнениями. За основу был взят набор отзывов пользователей интернет-магазина Amazon [4]. Всего было размечено 3232 отзыва о мобильных телефонах. Размеченная выборка содержит 9344 мнений, 1994 уникальных аспектов, 5124 уникальных оценочных выражений.

Задача извлечения пользовательских мнений может быть представлена как задача разметки последовательности, где для каждого элемента w_i входной последовательности $\{w_1, \dots, w_n\}$ необходимо определить метку класса (тег) y_i . Для этого мнения преобразуются из набора интервалов в последовательность тегов: O – отсутствие сущности; $B-Aspect$ – начало аспекта; $I-Aspect$ – продолжение аспекта; $B-Description$ – начало оценочного высказывания; $I-Description$ – продолжение оценочного высказывания. На рисунке 1 представлен пример задания мнения в вид последовательности тегов.

| | | | |
|-----------------------|---------------------------------------|-----------------|---|
| [Battery ₁ | life ₂] _{Aspect} | is ₃ | [good ₄] _{Description} |
| <i>B-Aspect</i> | <i>I-Aspect</i> | <i>O</i> | <i>B-Description</i> |

Рисунок 1 – Пример преобразование мнений в последовательность тегов

Для определения распределения вероятности тегов для каждого слова из входного предложения предлагается использовать следующее выражение:

$$P(\mathfrak{E}_t | \varphi(c_t))_i = \frac{\exp(W\varphi(c_t) + b)_i}{\sum_j \exp(W\varphi(c_t) + b)_j}$$

где c_t – конфигурация модели в момент времени t ,
 $\varphi(c_t)$ – функция, отображающая конфигурацию c_t в набор признаков,

W, b – параметры линейного преобразования.

Конфигурация определяется четверкой вида $c_t = (S, B, l, Y)$, состоящей из буфера B , содержащего в себе необработанные на текущий момент элементы входной последовательности; стека S , в который помещаются слова из найденных на текущий момент сущностей; метки последней найденной сущности l ; выходной последовательности меток Y . На каждом шаге t формируется вектор признаков $\varphi(c_t)$, который используется для определения метки текущего слова \mathfrak{E}_t и изменения текущей конфигурации в соответствии с этой меткой. Конфигурация изменяется в соответствии с набором правил, который позволяет соблюдать ограничения используемой системы разметки IOB.

Исходя из предложенной формы конфигурации c_t и правил её изменения, предлагается использовать $\varphi(c_t)$ следующего вида:

$$\varphi(c_t) = [B_i, \dots, B_{t+M_B}, S_1^{t-1}, \dots, S_{M_S}^{t-1}, Y_{t-1}^{t-1}, Y_{t-2}^{t-1}, E_{\mu}^t]$$

где B_i – i -тый элемент буфера B ,

S_j^t – j -тый элемент стека S на шаге t ,

E_k – k -тая строка матрицы E ,

Y_n^t – n -тый элемент последовательности меток на шаге t .

В качестве элементов буфера используются скрытые состояния, получаемые из последнего слоя многослойной двунаправленной LSTM [5]. Так, каждый элемент в буфере будет содержать информацию не только о слове на соответствующей позиции, но и о предшествующем и последующих контекстах.

Для обучения модели используется алгоритм оптимизации Adam [6] с параметрами $\text{lr} = 0.001$, $\beta_1 = 0.9$, $\beta_2 = 0.999$ и нормировкой градиента на уровне 3.0. Из-за небольшого объема данных оценки рассчитывались процедурой скользящего контроля с разбиением на 5 выборок

Результаты работы предложенной модели на примере обработки отзывов о мобильных телефонах из интернет-магазина Amazon приведены в таблице 1

Обученная модель была оценена по 3 критериям: точность (precision), полнота (recall) и F-меру (F1-measure) извлечения сущностей. Для аспектов Recall = 61.08%, Precision = 62.74%, F1 = 61.9%; для оценочных высказываний Recall = 62.49%, Precision = 64.98%, F1 = 63.71%.

Таблица 1 – Примеры извлеченных мнений.

| Товар | Мнения |
|-------------------|---|
| Sony Xperia XA | Phone is awesome; phone is easy to use; phone is perfect for those who need extra storage; battery life is mediocre. |
| Huawei P20 | Phone is a flagship performer; phone stopped receiving phone calls; phone is absolutely amazing for the price; the screen is fantastically large. |

Полученные результаты могут быть полезны специалистам в области обработки естественного языка, компьютерной лингвистике, а также бизнес-сообществу при реализации товаров и услуг и развитии их потребительских свойств.

Литература:

1. Natural language processing (almost) from scratch / Collobert R. [et al.] // Journal of Machine Learning Research. 2011. Vol. 12. P. 2493-2537.
2. Transition-based dependency parsing with stack long short-term memory / Dyer C. [et al.] // Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing. Beijing, China, 27-31 July 2015. Vol. 1. P. 334-343.

3. Zhang L., Wang S., Liu B. Deep learning for sentiment analysis: A survey // Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery. 2018. Vol. 8. No. 4. P. 1942-4787.
4. He R., McAuley J. Ups and downs: Modeling the visual evolution of fashion trends with one-class collaborative filtering // Proceedings of the 25th international conference on world wide web (WWW 2016). Montreal, Canada, 11-15 April 2016. P. 507-517.
5. Graves A., Jaitly N., Mohamed A. Hybrid speech recognition with deep bidirectional LSTM // Proceedings of the IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU 2013). Olomouc, Czech Republic, 8-12 December 2013. P. 273-278.
6. Kingma D. P., Ba J. Adam: A method for stochastic optimization // arXiv:1412.6980v9 [cs.LG]. 2017. [Электронный ресурс]: <https://arxiv.org/pdf/1412.6980.pdf> (дата обращения 10.10.2018).

КЛАСТЕРИЗАЦИЯ ГРАФОВ МЕТОДОМ ПОПАРНОГО ВЗВЕШЕННОГО ОБЪЕДИНЕНИЯ

А.В. Кочнев, В.М. Волкова

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

В данной работе проведен обзор проблем существующих алгоритмов кластеризации в графах. Кроме того, представлено описание разрабатываемого семейства алгоритмов кластеризации во взвешенных графах, основанного на различных метриках качества разбиений. Также предложены возможные модификации алгоритма для последующего анализа.

This paper reviews the existing problems of clustering algorithms in graphs. In addition, a description is given of the developed family of clustering algorithms in weighted graphs, based on different quality metrics of partitions. Also proposed are possible modifications of the algorithm for further analysis.

Зачастую при априорном анализе данных возникает необходимость выделения основных групп исследуемых объектов. В случае, когда множество исследуемых объектов не обладает атрибутами, идентифицирующими группы, разделение можно произвести, выделив кластера исследуемых объектов. Однако в случае отсутствия возможности отображения объектов в некоторое метрическое пространство, использование классических методов кластеризации вызывает затруднение. Также

довольно часто встречаются задачи, в которых известна лишь относительная связь между наблюдаемыми объектами. Именно для решения таких задач применяются методы кластеризации в графах.

На данный момент существует множество алгоритмов для кластеризации графов. Зачастую их применение для решения реальных задач вызывает затруднение, а результаты не вызывают доверия. Причинами этого для различных алгоритмов являются [1]:

- нелинейная зависимость сложности относительно объемов анализируемых данных;
- требование к наличию дополнительной информации, например количество кластеров;
- зависимость от порядка обхода анализируемых структур.

Для анализа качества решений задачи кластеризации используется два типа метрики: эффективности и качества разбиения.

Метрики качества говорят о том насколько предоставленное разбиение хорошо описывает структуру кластеров [2]. Примерами данных метрик являются модулярность Q (1) и различные виды разрезов (2) и (3):

$$Q = \frac{1}{2m} \sum_{i,j} \left(A_{i,j} - \frac{d_i d_j}{2m} \right) \delta(C_i, C_j), \quad (1)$$

$$\text{RatioCut}(\hat{C}_1, \dots, \hat{C}_{N_c}) = \sum_{k=1}^{N_c} \frac{\text{cut}(\hat{C}_k, V \setminus \hat{C}_k)}{|C_k|}, \quad (2)$$

$$\text{NormalizedCut}(\hat{C}_1, \dots, \hat{C}_{N_c}) = \sum_{k=1}^{N_c} \frac{\text{cut}(\hat{C}_k, V \setminus \hat{C}_k)}{\text{vol}(\hat{C}_k)}, \quad (3)$$

где C_i – некоторое разбиение (кластер), d_i – мощность i -ой вершины, $A_{i,j}$ – элемент матрицы смежности, N_c – количество разбиений, m – количество ребер. Разрез и объемы кластеров рассчитываются как:

$$\text{vol}(\hat{C}_i) = \sum_{i \in C_k} d_i,$$

$$\text{cut}(\hat{C}_1, \dots, \hat{C}_{N_c}) = \frac{1}{2} \sum_{k=1}^{N_c} \sum_{i \in c_{kj} \neq c_k} w_{i,j}.$$

Для анализа решения с известными истинными сообществами могут быть применены метрики эффективности NMI и F1-score [3].

В данной работе нами предлагается семейство алгоритмов иерархической кластеризации во взвешенных графах, основанный на метриках качества.

Логичным предположением о структуре кластеров взвешенных графов $G := (V, E)$, является то, что связи w_i (веса ребер) между элементами группы должна быть выше, чем между элементами разных групп. На основании данного предположения построим следующий алгоритм:

1. Положим, что каждая вершина является группой C_i .
2. Определим некоторый перцентиль от максимума массы ребра $w_{\text{trashhold}}$.
3. Парно объединим вершины в группы таким образом, что:
 - a. вершины объединяются в итерации единожды;
 - b. не объединяются вершины с ребрами легче, чем заданный перцентиль;
 - c. объединение продолжается, пока есть необъединенные пары вершин;
 - d. объединение происходит в порядке убывания весов ребер.
4. Зафиксируем значение метрики качества для полученных групп.
5. Перестроим граф по полученным группам, так что:
 - a. каждая группа станет вершиной;
 - b. ребра между вершинами группы и внешней вершиной объединяются, весом нового ребра будет суммарный вес объединенных.
 6. Если групп больше двух, пересчитаем значение перцентилья и перейдем к шагу 3.
7. Выберем разбиение с наибольшим значением заданной метрики качества.

Для данного алгоритма могут быть использованы различные процентные точки для отсеивания «шумных» ребер, а также различные метрики качества кластеризации. Однако данный алгоритм неприменим для графов без веса.

Для анализа качества работы предложенного семейства алгоритмов предлагается использовать тесты типа l-partition [4] и модель Ланси-

чинетти [5], а также классические наборы данных Zachary's Karate club, American college football [6] с добавлением весов.

Потенциальными преимуществами данных алгоритмов можем считать достаточно низкую сложность относительно количества обрабатываемых ребер, сравнимую с LPA [7], а также точность в плане сохранения самых крепких связей в кластере.

Литература:

1. *Силин И., Панов М.* Обзор и экспериментальное сравнение алгоритмов кластеризации графов // Информационные технологии и системы. – 2015. М.: Институт проблем передачи информации им. А.А. Харкевича РАН, 2015. – С. 1042-1059.
2. *Aaron Clauset, M. E. J. Newman, and Cristopher Moore.* Finding community structure in very large networks. *Phys. Rev. E*, 70:066111, Dec 2004.
3. *Jaewon Yang and Jure Leskovec.* Defining and evaluating network communities based on ground-truth. In *Proceedings of the ACM SIGKDD Workshop on Mining Data Semantics, MDS '12*, pages 3:1 3:8, New York, NY, USA, 2012. ACM.
4. *Condon A., Karp R.M.* Algorithms for Graph Partitioning on the Planted Partition Model. *Randomization, Approximation, and Combinatorial Optimization. Algorithms and Techniques 1999*, pages 116-140.
5. *Andrea Lancichinetti, Santo Fortunato, and Filippo Radicchi.* Benchmark graphs for testing community detection algorithms. *Phys. Rev. E*, 78:046110, Oct 2008.
6. *M. Girvan and M. E. J. Newman.* Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*, 99(12):7821-7826, 2002
7. *Usha Nandini Raghavan, Reka Albert, and Soundar Kumara.* Near linear time algorithm to detect community structures in large-scale networks. *Phys. Rev. E*, 76:036106, Sep 2007

ОБНАРУЖЕНИЕ НАРУШЕНИЙ ПАРКОВКИ АВТОМОБИЛЕЙ НА ОСНОВЕ АНАЛИЗА ИЗОБРАЖЕНИЙ

И.О. Кутырев, А.А. Попов

Новосибирский государственный технический университет,
г. Новосибирск, igkutyrev@gmail.com

В работе рассматриваются математические и алгоритмические аспекты создания автоматической системы обнаружения нарушений парковки. В ходе исследований проведен анализ наиболее популярных подходов к решению задач обнаружения объектов на изображении. Разработан сквозной алгоритм обнаружения нарушений парковок автомобилями. Проведено его исследование на различных наборах данных, а также предложена модификация для использования в режиме реального времени.

Related paper is dedicated to the mathematical and algorithmic aspects of creating an automatic system for detecting parking violations. During the research was performed the analysis of the most popular approaches for detecting objects in a single image. Developed complex algorithm for detecting car parking violations. Algorithm was tested on a various data sets, also, a real-time modification was purposed.

Системы автоматического контроля различных правонарушений в настоящее время активно внедряются в повседневную жизнь. Большинство из этих систем использует специальное оборудование для распознавания и обработки факта правонарушения. Однако, решить таким образом проблему проверки правильности парковки сложно: невозможно разместить специализированные камеры в каждом дворе и на каждой парковке. Подходом к решению данной проблемы является автоматическая подача заявки на проверку факта нарушения. Пользователь загружает фотографию, сделанную при помощи любой камеры, в специальную систему, после чего алгоритм проверяет, имеются ли на данной фотографии необходимые объекты, анализирует их взаимное расположение, и, по возможности, выносит решение о наличии или отсутствии правонарушения.

Целью данной работы является разработка алгоритма, позволяющего по фотографии припаркованного автомобиля сделать вывод о наличии на парковке нарушений трех видов:

- 1) парковка в зоне действия знака “стоянка запрещена”;
- 2) парковка в зоне действия знака “остановка запрещена”;
- 3) парковка на тротуаре.

При решении задачи не отслеживаются ограничения времени действия знаков и ограничения зоны действия знаков.

Обнаружение на фотографии знака “Парковка запрещена” производится при помощи алгоритма скользящего окна. Классическая реализация алгоритма подразумевает движение окна по изображению с шагом в один пиксель. В предлагаемой реализации число пикселей, на которое производится сдвиг, меняется в зависимости от результата сравнения областей. В случае, если с последнего положения окна мера совпадения областей уменьшилась, шаг увеличивается. Если значение меры совпадения областей увеличилось, окно возвращается к предыдущему положению и продолжает движение с меньшим шагом. Такой подход позволяет, с одной стороны, снизить количество вычислений, с другой, минимизировать потери точности за счет возврата на предыдущие положения окна. В таблице 1 приведены результаты сравнения двух алгоритмов. Ошибкой в данном случае называется среднее расстояние в пикселях от истинного положения знака до положения, определенного алгоритмом. Тестирование проводилось на кадрах записи с регистратора в разрешении 1980 на 1080 пикселей.

Таблица 1 – сравнение различных алгоритмов сопоставления с образцом

| Алгоритм | Доля верных распознаваний | Ошибка, пикселей | Среднее время обработки одного изображения, мс |
|---------------------------|---------------------------|------------------|--|
| Скользящего окна | 0.95 | 12.3 | 1048 |
| Быстрого скользящего окна | 0.9125 | 13.2 | 222 |

Для выделения на изображении автомобиля использовался метод Виолы-Джонса. Каскад предварительно обучался с использованием адаптивного бустинга на двух различных наборах данных: фотографии улиц города Бостон, США и фотографии с видеорегистратора автомобиля. Для увеличения объема выборки применялись различные виды аугментации: зеркальные отражения, накладка белых шумов, затемнение. Таблица 2 показывает сравнительные результаты исследований.

Таблица 2 – тестирование различных наборов данных для обучения

| Набор данных для обучения каскада | Доля верных распознаваний | Ложно-позитивные обнаружения в расчёте на одно изображение |
|-----------------------------------|---------------------------|--|
| На основе видеорегистратора | 0.75 | 0.09 |
| На основе фотографий улиц | 0.95 | 0.1 |

Выделение линии, разделяющей тротуар и проезжую часть производилось при помощи преобразования Хафа. Задача выделения границ проезжей части является достаточно сложной, и не до конца решена на текущем этапе исследований. Работы, исследующие поиск дорожного полотна на одиночных изображениях, используют “точку исчезновения” (vanishing-point). Так называется точка, в которой пересекаются левый и правый края дорожного полотна. Однако, полностью автоматизированная система не может использовать данный метод, так как поиск данной точки производится вручную. В результате исследований на наборе фотографий улиц достигнута точность составила 74%.

Предложенный алгоритм работает в три этапа. На первом этапе производится выделение линии, разделяющей тротуар и проезжую часть. Затем производится поиск автомобиля. Из всех найденных на изображении областей будем выделять одну, наибольшую по размеру. Это априорное предположение строится на том, что пользователь сделает фотографию, на которой автомобиль, совершивший нарушение, будет наибольшим из всех. На третьем этапе выделяется знак „парковка запрещена“ или „остановка запрещена“. Если область, соответствующую знаку, удастся обнаружить, производится поиск нарушения, связанного с парковкой под запрещающим знаком. Для этого определяется, какой из объектов расположен ближе к точке съемки – автомобиль или знак. В случае, если знак расположен ближе, чем автомобиль, нарушение считается доказанным. Если же область, соответствующая знаку, не обнаружена на изображении, проверяется, пересекает ли найденный автомобиль линию разделения тротуара и проезжей части. В случае, если автомобиль пересекает линию, нарушение считается доказанным. На рисунке 1 показан пример работы алгоритма.



Рисунок 1 – пример работы алгоритма

В качестве набора данных для тестирования алгоритма будем использовать 40 различных фотографий. 24 из них содержат нарушения парковки двух видов: парковка в зоне действия знака “остановка запрещена” или парковка на тротуаре. 16 не содержат нарушений парковки и являются фотографиями верно припаркованных машин, 6 из которых припаркованы под другими знаками. Результаты представлены в таблице 3.

Таблица 3 – тестирование алгоритма

| Доля верных определений присутствия нарушения | Доля ложноположительных определений присутствия нарушения | Доля верных определений отсутствия нарушения |
|---|---|--|
| 0.67 | 0.0 | 1.0 |

Исследование основных причин неверной работы алгоритма показывает, что в 24% случаев ошибка заключалась в неверном распознавании разделяющей линии, так как данная задача является наиболее сложной из исследуемых. В 8% случаев алгоритм неверно определял требуемый автомобиль на фотографии. Данные результаты показывают необходимость дальнейшего исследования проблемы.

Литература:

1. Marikhu R. et al. Police Eyes: Real world automated detection of traffic violations //Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), 2013 10th International Conference on. – IEEE, 2013. – С. 1-6.

2. Kong H., Audibert J. Y., Ponce J. General road detection from a single image //IEEE Transactions on Image Processing. – 2010. – Т. 19. – №. 8. – С. 2211-2220.

3. Negri P. et al. A cascade of boosted generative and discriminative classifiers for vehicle detection //EURASIP Journal on Advances in Signal Processing. – 2008. – Т. 2008. – С. 136.

АВТОМАТИЗАЦИЯ ПРОЦЕССА ПРОГНОЗИРОВАНИЯ ЗАГРУЗКИ ОТЕЛЯ

М.И. Лебедева, М.А. Бакаев

**Новосибирский государственный технический университет,
г. Новосибирск, 1332698@mail.ru**

На основе данных о загрузке отеля за 2014-2017 г. была предложена методика прогнозирования загрузки на будущий период с применением нейронной сети. Также был построен и оценен качественно график загрузки.

The occupancy forecasting method for the future was proposed based on the hotel load data for 2013-2017 using a neural network. In addition, the load schedule was built and qualitatively evaluated.

Введение. Необходимость автоматизации бизнес-процессов обусловлена высокой скоростью развития современного бизнеса. Автоматизация функционирующих систем необходима для поддержания эффективной работы всего предприятия. Прогнозирование спроса является важным показателем в сфере гостиничного бизнеса. Прогнозные значения позволяют оценить уровень загрузки, оптимизировать затраты компании, а также выстроить гибкую модель по формированию цены на проживание. Поэтому автоматизация процесса загрузки отеля так важна в условиях жесткой конкуренции и экономически-нестабильной ситуации, прогнозировать такой важный показатель для отеля как загрузка весьма нелегкая задача.

Исследование. Нейронная сеть - это последовательность нейронов, соединенных между собой синапсами. В свою очередь синапс - это связь между двумя нейронами. А нейрон определяется вычислительной единицей, которая получает информацию, производит над ней различные простые вычисления и передает ее дальше. Нейронные сети необходимы для решения задач прогнозирования, классификации и/или

управления [1]. Из преимуществ нейронных сетей следует выделить простоту в использовании и богатые возможности. Нейронные сети учатся на примерах. Пользователь нейронной сети выбирает данные, а затем запускает алгоритм обучения, который автоматически определяет структуру данных.

Нейронная сеть представляет собой инструмент, благодаря которому можно моделировать разные зависимости, к примеру, линейные или нелинейные модели. Возможность моделирования нелинейных моделей является главной особенностью нейронных сетей. Способность нейронной сети к обобщению и выделению скрытых зависимостей между входными и выходными данными делает ее способной к прогнозированию. Обученная нейронная сеть способна предсказывать данные на основе существующих факторов и предыдущих значений. Для прогнозирования будущих значений необходимо подготовить данные для обучения и проверки работы сети, подобрать соответствующую топологию, основные характеристики и параметры обучения нейронной сети.

Данные для анализа были взяты из системы управления отелем Fidelio [2]. Fidelio Version 8 - это полностью интегрированный, гибкий программный пакет, предназначенный для максимальной эффективности работы отеля. Система содержит все функции повседневной деятельности отеля, включая все аспекты управления и обслуживания отеля. Данные по загрузке были выгружены за 2014-2017 годы. Для построения модели нейронной сети использовался программный продукт StatisticaSoft. Было определено какая сеть даст наиболее точный результат. Наиболее хорошие результаты получаются при использовании типа сети - MLP (мультиперсептрон) с топологией сети 1-7-1, о чем иллюстрирует рисунок 1.

| Summary of active networks (Spreadsheet1) | | | | | | | | | |
|---|------------|----------------|------------|----------------|------------|--------------------|----------------|-------------------|-------------------|
| Index | Net. name | Training perf. | Test perf. | Training error | Test error | Training algorithm | Error function | Hidden activation | Output activation |
| 1 | MLP 1-14-1 | 0.909244 | 0.967076 | 0.004911 | 0.003482 | BFGS 7 | SOS | Logistic | Logistic |
| 2 | MLP 1-15-1 | 0.924016 | 0.967063 | 0.004136 | 0.002440 | BFGS 81 | SOS | Tanh | Tanh |
| 3 | MLP 1-18-1 | 0.909464 | 0.967189 | 0.004904 | 0.003545 | BFGS 6 | SOS | Tanh | Logistic |
| 4 | MLP 1-7-1 | 0.922398 | 0.969553 | 0.004217 | 0.002422 | BFGS 39 | SOS | Tanh | Tanh |
| 5 | MLP 1-7-1 | 0.928234 | 0.968274 | 0.003954 | 0.002261 | BFGS 103 | SOS | Exponential | Logistic |

Рисунок 4 - Сети MLP с различной топологией

Видно, что эффективность обучающего и тестового набора наиболее близки к 1. Чем ближе эти параметры к 1, тем качественнее модель. На рисунке 2 представлен график загрузки отеля за 2014-2017 годы и прогноз на 2018 год.

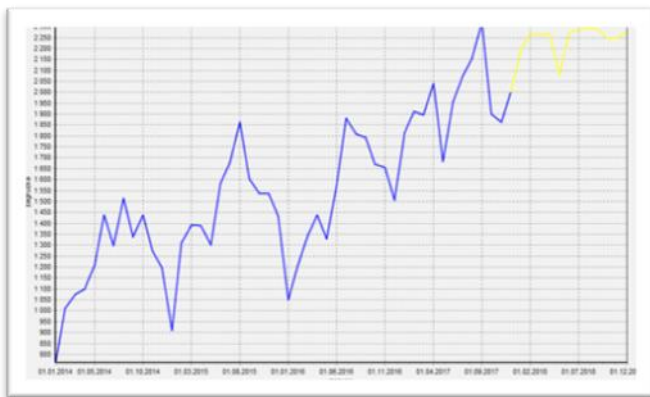


Рисунок 5 - График загрузки отеля за 2014-2018

Реализация. Для поддержания конкурентных преимуществ компании необходима высокая скорость реакции на внутренние и внешние изменения [3]. Поэтому компании крайне важно обладать информацией загрузки гостиницы на будущий период. С помощью информации о загрузке на будущий период можно перепроектировать бизнес-процессы компании, такие как продажа профильных услуг отеля, предоставление проживания, предоставление питания, а также процессы маркетинга по привлечению новых и удержанию старых клиентов, добавив туда информацию о прогнозе загрузки отеля, что позволит специалистам размещать различного рода информацию об акциях, специальных предложениях, новых программах лояльности в случае прогноза низкой загрузки. А в случае высокой, корректировать ценовую политику, производить больший закуп, подготавливать большее количество номеров, а также привлекать большее количество персонала, с целью максимизации прибыли и удовлетворенности клиентов.

Литература:

1. Хайкин С. Нейронные сети: полный курс. 2-е издание. М.: Вильямс, 2006. 1104 с.
2. Официальный сайт: <http://www.micros.rs/en/fidelio-suite-8.html>
3. Миняев Е.В. Методы совершенствования бизнес-процессов //Е.В.Миняев // Вестник ОмГУ. Серия: Экономика. -2010. -№1

РЕШЕНИЕ ЗАДАЧ КЛАССИФИКАЦИИ С ИСПОЛЬЗОВАНИЕМ ТРАНСДУКТИВНОГО МЕТОДА ОПОРНЫХ ВЕКТОРОВ

А.Г. Морковкин, А.А. Попов

Новосибирский государственный технический университет,
г. Новосибирск, ag.morkovkin@gmail.com

В данной работе приводятся результаты сравнения двух алгоритмов классификации при использовании частично размеченных выборок: метода опорных векторов и трансдуктивного метода опорных векторов. Было разработано соответствующее приложение и произведены исследования алгоритмов на сгенерированных выборках и на реальных прикладных задачах с использованием различных метрик качества. На выборках с относительно небольшой размеченной частью трансдуктивный метод опорных векторов показывал результаты лучше, чем классический метод опорных векторов.

This paper presents the results of comparing two classification algorithms using partially labeled samples: the support vector machine and the transductive support vector machine. An appropriate application was developed and algorithms were investigated on the generated samples and on real experimental problems using various quality metrics. On samples with a relatively small labeled part, the transductive support vector machine showed better results than the classical support vector machine.

Классификация данных является одной из основных задач машинного обучения. Задача классификации состоит в том, чтобы по некоторым входным характеристикам объекта определить принадлежность этого объекта к одному из классов.

В настоящее время с задачей классификации сталкиваются во многих областях. Можно привести в пример несколько прикладных задач: медицинская диагностика (поставить диагноз пациенту по набору медицинских характеристик), кредитный скоринг (принять решение о выдаче/отказе кредита по анкете заемщика), геологоразведка (определить наличие полезных ископаемых по данным зондирования почв), синтез химических соединений (спрогнозировать свойства получаемого соединения по параметрам химических элементов).

Одним из самых популярных методов классификации является метод опорных векторов, разработанный Владимиром Вапником [1] в 1995 году.

Задача классификации относится к разделу обучения с учителем (supervised learning). Обучение с учителем позволяет восстанавливать значения функций, заданной на подмножестве области определения. Для

методов обучения с учителем необходимо, чтобы для всей выборки входных данных ставился в соответствие некоторый отклик. Зачастую в задачах классификации имеется большая выборка, на которой нет возможности составить данное соответствие для всей выборки. Это может случиться из нескольких причин: чтобы классифицировать всю выборку вручную требуется много времени, для классификации необходимо привлечение эксперта из этой области, что не всегда возможно.

Для того чтобы избежать этих проблем можно вместе с обучающей выборкой использовать тестовую выборку. Неразмеченные данные использованные совместно с небольшим количеством размеченных данных, могут обеспечить прирост качества обучения (Рисунок 1). Алгоритмы, основанные на обработке частично-размеченной выборке, называются методами частичного обучения. Частичное обучение занимает промежуточное положение между обучением с учителем и обучением без учителя.

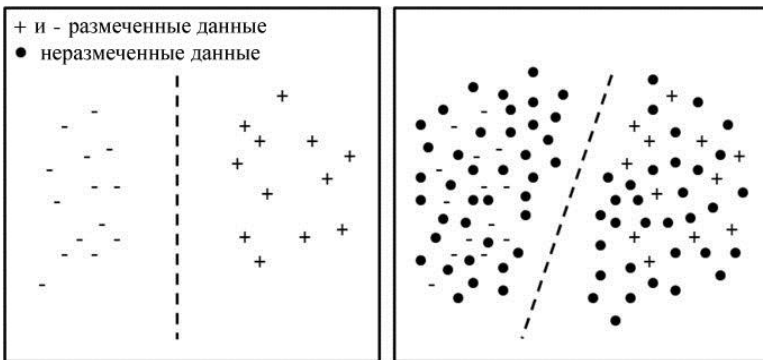


Рисунок 1 – Слева: разделяющая поверхность, построенная с использованием только размеченных данных. Справа: разделяющая поверхность, построенная с использованием как размеченных данных, так и неразмеченных данных.

В ходе работы были проведены исследования, направленные на сравнение алгоритмов метода опорных векторов (SVM) [2] и трансдуктивного метода опорных векторов (TSVM) [3]. Исследования проводились на сгенерированных и реальных выборках. Для тестирования использовались открытые наборы данных: Flame, Spiral, Adult, Titanic, Reuters, Diabetes.

Исследования проводились следующим образом. В выборке размечались разное количество данных (1, 5, 10, 20 процентов). Для оценки качества алгоритмов использовались две метрики качества: доля правильно помеченных элементов (accuracy) и F-мера.

Проведено исследование качества классификации на сгенерированной выборке, состоящей из 240 элементов и 2 классов. При 20 процентах размеченных элементов, оба алгоритма правильно поместили более 95 процентов элементов, что говорит о высоком качестве классификации (Таблица 1). Однако при малом количестве размеченных элементов, метод опорных векторов показывает результаты значительно хуже трансдуктивного метода опорных векторов.

Таблица 1 – Результаты на выборке «Flame»

| Размеченные данные, % | SVM | | TSVM | |
|-----------------------|----------|---------|----------|---------|
| | Accuracy | F-мера | Accuracy | F-мера |
| 1 | 0.62917 | 0.45321 | 0.92083 | 0.91851 |
| 5 | 0.75417 | 0.69726 | 0.9125 | 0.90637 |
| 10 | 0.90417 | 0.86351 | 0.97083 | 0.96834 |
| 20 | 0.96667 | 0.96274 | 0.9875 | 0.98039 |

В исследовании, направленном на оценку качества многоклассовой классификации на сгенерированной выборке, состоящей из 312 элементов и 3 классов, при небольшом количестве размеченных элементов метод трансдуктивных векторов показывает себя лучше метода опорных векторов (Таблица 2). При увеличении количества размеченных элементов метод опорных векторов также показывает высокое качество классификации.

Таблица 2 – Результаты на выборке «Spiral»

| Размеченные данные, % | SVM | | TSVM | |
|-----------------------|----------|---------|----------|---------|
| | Accuracy | F-мера | Accuracy | F-мера |
| 1 | 0.34615 | 0.31686 | 0.65385 | 0.65537 |
| 5 | 0.68589 | 0.68372 | 0.90064 | 0.89816 |
| 10 | 0.83654 | 0.83506 | 0.96154 | 0.96208 |
| 20 | 0.91026 | 0.91165 | 0.98718 | 0.98736 |

Также были проведены исследования на реальных данных. В данных исследованиях трансдуктивный метод опорных векторов показал себя лучше метода опорных векторов.

Из полученных результатов было выяснено, что трансдуктивный метод опорных векторов более эффективен при малом количестве размеченных объектов, но в то же время его обучение занимает больше времени, чем метод опорных векторов.

Литература:

1. Cortes C., Vapnik V. Support-vector networks //Machine learning. – 1995. – Т. 20. – №. 3. – С. 273-297.
2. Vapnik V. Pattern recognition using generalized portrait method //Automation and remote control. – 1963. – Т. 24. – С. 774-780.
3. Joachims T. Transductive inference for text classification using support vector machines //ICML. – 1999. – Т. 99. – С. 200-209.

АЛГОРИТМ ВЫЯВЛЕНИЯ КРУПНОМАСШТАБНЫХ АСТРОФИЗИЧЕСКИХ ОБЪЕКТОВ ПО КОСМИЧЕСКОЙ ПАУТИНЕ

В.О. Тальников, В.М. Волкова

**Новосибирский государственный технический университет,
г. Новосибирск, talnikov.2013@stud.nstu.ru**

Рассматривается задача распознавания космологических структур в автоматическом режиме трёх типов: скопление, стена, цепочка. В процессе работы был реализован алгоритм FRiS-Tax, состоящий из двух последовательных этапов FRiS-Cluster и FRiS-Class. Также реализовано выделение космологических структур на основе метода главных компонент.

The problem of recognition of cosmological structures in the automatic mode of three types is considered: cluster, wall, chain. In the process of work, the FRiS-Tax algorithm was implemented, consisting of two consecutive stages of FRiS-Cluster and FRiS-Class. Also, the separation of cosmological structures based on the principal component analysis is realized.

Предметом изучения современной астрофизики является исследование физических процессов во Вселенной, того как эти процессы повлия-

яли на самоорганизацию и эволюцию астрофизических объектов, а также на дальнейшую их динамику и взаимодействие. Несмотря на большие успехи в области наблюдений звездного неба зачастую с помощью телескопов можно увидеть лишь статическую картинку. По существу, математическое моделирование – это основной, если не единственный, подход к теоретическому исследованию астрофизических процессов и объектов.

Исходная задача при проведении данной работы заключалась в проведении такого анализа данных, представляющих собой результат моделирования эволюции астрофизических объектов, который бы позволил в автоматическом режиме обнаруживать скопление галактик нескольких типичных форм: скопление, цепочка, стена. Для решения этой задачи использовались методы машинного обучения, а именно – таксономии, классификации «без учителя».

Исходными данными для распознавания являлся массив объектов – точек, каждая из которых задана своей яркостью и значениями трех пространственных координат. Экспертом выбирался порог, с использованием которого фиксируется набор точек, яркость которых превышает порог. По нему осуществляется поиск кластеров (сгущений) объектов.

Для распознавания космологических структур был выбран алгоритм FRiS-Tax, так как на начальном этапе принадлежность объектов выборки к тому или иному образу неизвестна. Таксоны могут быть сложной формы. Во входных данных только выборка объектов.

Алгоритм FRiS-Tax, предложенный Загоруйко и Борисовой [1], основанный на понятии компактности кластера – отношении количества граничных точек к общему количеству точек образа – позволяет выделять группы объектов сложной геометрической формы.

Программа распознавания космологических структур основана на алгоритме FRiS-Tax. FRiS-Tax состоит из двух процедур. Процедура FRiS-Cluster находит центры локальных сгустков объектов и формирует линейно разделимые кластеры. Процедура FRiS-Class объединяет кластеры в классы произвольной сложной формы – таксоны. С помощью МГК определяется к какому типу астрофизических структур относятся выделенные таксоны.

В алгоритмах активно используется расстояние, использующее идею конкурентного сходства. Для оценки конкурентного сходства используется FRiS-функция расчета расстояния от объекта z до конкурирующих образов A и B , удовлетворяющая ряду фундаментальных свойств [2]. Нелинейная модификация FRiS-функции может быть записана в виде:

$$F(z, A | B) = 1 - 2R_A(R_B + R_A) + t \cdot \sin(4\pi R_A(R_B + R_A)),$$

где $R_i = r_i / d_i$, d_i – дисперсия класса i , $i = \overline{A, B}$, r_i – евклидово расстояние от объекта z до объекта i . Данная функция близка к восприятию сходства человеком. Опытным путем было выявлено, что оптимальным коэффициентом нелинейности будет $t = 0.14$ (подробно процедура представлена в [2]).

Также предложена и реализована процедура классификации: отнесения выделенной структуры к одному из типов космологических структур. Был сформулирован ряд решающих правил, использующих соотношения размерностей таксонов. Размерности же получаются как проекции объектов (точек в трехмерном пространстве) на оси, выделяемые с помощью метода главных компонент (МГК). Для этого на собственные вектора, получаемые в МГК, проецируется каждый объект рассматриваемого таксона. Вычисляется норма отрезка от проекции точки до центра масс таксона. Для оценки размерностей кластера выбирается наибольшая норма по каждой оси ($V_{i_1}, V_{i_2}, V_{i_3}$). На основании вычисленных трех размерностей каждого кластера, он может быть отнесен к одному из типов структуры по следующим правилам:

- $\log_{10}(V_{i_1}) \approx \log_{10}(V_{i_2}) \approx \log_{10}(V_{i_3})$ для структуры типа «скопление»;
- $\log_{10}(V_{i_1}) \approx \log_{10}(V_{i_2}) < \log_{10}(V_{i_3})$ для структуры типа «цепочка»;
- $\log_{10}(V_{i_1}) \approx \log_{10}(V_{i_2}) > \log_{10}(V_{i_3})$ для структуры типа «стена».

Алгоритм реализован в распараллеленном виде с помощью MPI. Идея заключается в том, чтобы для алгоритма FRiS-Cluster разделить входные данные (массив точек) на некоторое число процессов с тем, чтобы каждый процесс обработал свой блок.

При такой схеме распараллеливания очевидно, что каждый процесс теряет информацию о объектах, находящихся в других процессах, и столпы выбираются по не полным данным. Эту потерю информативности компенсирует алгоритм FRiS-Class (подробное алгоритмическое описание процедуры представлено в [3]), который объединяет рядом стоящие столпы в один таксон, тем самым восстанавливая информативность. Такой подход позволяет достигать такого же показателя качества, что и при последовательной работе алгоритма, но за меньшее время.

Для ускорения работы программы было принято решение поменять связь алгоритмов FRiS-Cluster и FRiS-Class. Идея заключается в том, чтобы запускать FRiS-Class не на каждой итерации, а один раз после всех итераций алгоритма FRiS-Cluster, но на объектах с максимальным локальным показателем качества, который высчитывается в конце каждой итерации. Также принято решение после каждого нахождения очередного локального максимума отсчитывать определённое число итераций. Если за отведённое число итераций не будет найден следующий локальный максимум, то вычисления прекращаются. Этот набор модификаций позволил ускорить программу в 20 раз без потери качества на всех проведенных тестах. Следующим шагом было изменение евклидовой метрики на манхэттенскую, что позволило увеличить скорость работы алгоритма в 1.5 раза:

- евклидова метрика: $r = \sqrt{(A.x - B.x)^2 + (A.y - B.y)^2 + (A.z - B.z)^2}$,
- манхэттенская метрика: $r = |A.x - B.x| + |A.y - B.y| + |A.z - B.z|$,

где A, B – объекты, x, y, z – координаты 3-х мерного пространства.

Реализован алгоритм **FRiS-Cluster** – первый этап алгоритма **FRiS-Tax**. По результатам работы алгоритма видно, что он достаточно оптимально находит центры локальных сгустков объектов, выявляет все кластеры простой формы, обладающие максимальной компактностью и помехоустойчивостью. Реализован **FRiS-Class** – второй этап алгоритма **FRiS-Tax** для анализа полученных кластеров на предмет того, не являются ли некоторые из них частью одного кластера и последующего объединения их в классы сложной формы – таксоны. Реализовано автоматическое распознавание выделяемых таксонов как системы, состоящей из объектов трех видов космологических структур: скопление, стена или цепочка.

Данная работа является частью совместного проекта с группой ученой из ИВМиМГ СО РАН [4].

Литература:

1. Загоруйко Н.Г., Борисова И.А. Когнитивный анализ данных – Новосибирск: Академическое издание издательство «Гео» 2013 – 186 с.
2. Математическое моделирование и классификация крупномасштабных космологических структур на суперЭВМ / В.О. Тальников, В.М. Волкова, И.М. Куликов, И.Г. Черных, А.Ф. Сапетина // Параллельные вычислительные технологии (ПаВТ) 2017: короткие статьи и опи-

сание плакатов XI междунар. науч. конф. [Казань, 3–7 апр. 2017 г.]. – Челябинск: Издательский центр ЮУрГУ, 2017. – С. 525.

3. Тальников В.О., Волкова В.М. Распознавание крупномасштабных астрофизических объектов по космической паутине = Recognizing the large-scale astrophysical objects cosmic web // Обработка информации и математическое моделирование : материалы Рос. науч.-техн. конф. [Новосибирск, 25–26 апр. 2017 г.]. – Новосибирск : СибГУТИ, 2017. – С. 14-18.

4. AstroPhi: A code for complex simulation of dynamics of astrophysical objects using hybrid supercomputers / I.M. Kulikov, I.G. Chernykh, A.V. Snytnikov, B.M. Glinskiy, A.V. Tutukov // Computer Physics Communications. – 2015. – V. 186. – P. 71-80. doi: 10.1016/j.cpc.2014.09.004

СТАТИСТИЧЕСКИЙ АЛГОРИТМ ОБНАРУЖЕНИЯ СИГНАЛОВ В НЕЙРОКОМПЬЮТЕРНЫХ ИНТЕРФЕЙСАХ

Ю.А. Часовских, С.Е. Радченко, А.Г. Вострецов
Новосибирский Государственный Технический Университет
г. Новосибирск, r1505@mail.ru

Предложено математическое описание сигнала нейрокомпьютерного интерфейса, учитывающее статистический характер отсчетов. Для преодоления априорной неопределенности амплитуды сигнала предложен равномерно наиболее мощный алгоритм обнаружения. Результаты моделирования подтвердили целесообразность его применения во встраиваемых системах.

The mathematical model for neurocomputer interface signal has been supposed basing on the stochastic nature of its' sample elements. The a priori uncertainty of the signal amplitude was overcome by the developing of uniformly the most powerful detection algorithm. Numerical modeling results proved its' efficiency and the expediency to be applied in the embedded electronics.

Введение. Нейрокомпьютерные интерфейсы находят применение при решении широкого круга задач, среди которых стоит выделить задачи управления различными устройствами. При этом необходимо правильно выделять и трактовать характерные сигналы, соответствующие движениям оператора. Проблемы, с которыми при этом сталкиваются разработчики, вызваны наличием обилия каналов, индивидуальными особенностями операторов, плохой повторяемостью позиционирования

датчиков. Это, среди прочего, приводит к априорной неопределенности амплитуд характерных сигналов и параметров шумовой составляющей. В данной работе предлагается подход, позволяющий устойчиво обнаруживать характерные сигналы в условиях параметрической априорной неопределенности.

Разработка алгоритма.

Один из возможных характерных сигналов при работе с нейроинтерфейсами приведен на Рис.1. Для разработки алгоритма и его статистического исследования необходима аналитическая модель сигнала. Поскольку в исследуемом процессе присутствует шумовая составляющая, обусловленная влиянием разного рода артефактов, модель сигнала будет представлять собой аддитивную смесь полезной составляющей и гауссовского шума с нулевым средним и неизвестной дисперсией:

$$x_i = s \cdot f_i + \eta_i, \quad i \in 1..N,$$

где s - неизвестная амплитуда сигнала f_i , f_i - отсчеты огибающей, η_i - отсчеты гауссовского шума.

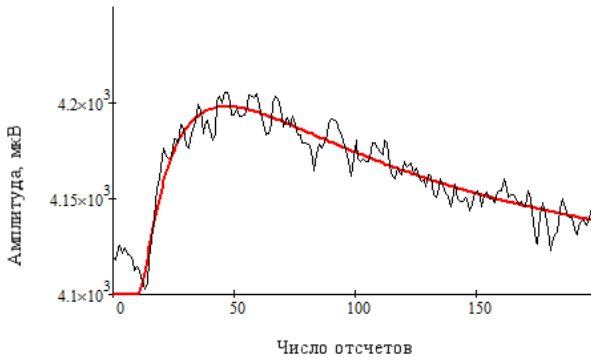


Рисунок 1. Сигнал нейроинтерфейса и его огибающая

Форма такого сигнала может быть представлена при помощи лог-нормальной кривой [1]:

$$f_i = \frac{1}{\sqrt{2\pi i \sigma'}} \exp\left(-\frac{(\ln(i) - \mu)^2}{2\sigma'^2}\right) + k,$$

где $\sigma' = 1$, $\mu = 0$, k - параметр смещения, определяемый при аппроксимации.

Реализация сигнала, полученного в соответствии с предложенной моделью, приведена на Рис.2.

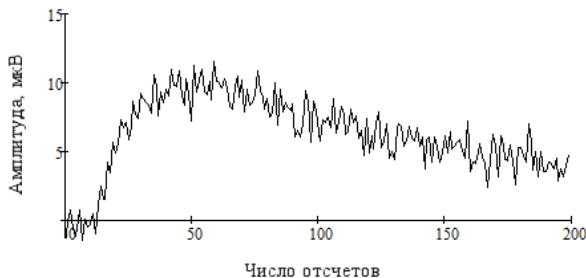


Рисунок 2. Реализация модели сигнала нейроинтерфейса

Совместная плотность распределения вероятностей наблюдаемого процесса получена в виде произведения частных плотностей распределения отсчетов выборки:

$$\omega(\mathbf{x}) = \prod_{i=1}^N \omega(x_i) = \left(\frac{1}{\sqrt{2\pi}\sigma} \right)^N \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^N (x_i - sf_i)^2 \right),$$

где σ - неизвестный энергетический параметр шума.

Задача обнаружения, таким образом, заключается в проверке статистических гипотез [2]:

$$\begin{cases} H_0 : s = 0 \text{ (сигнал отсутствует),} \\ H_1 : s > 0 \text{ (сигнал присутствует);} \end{cases}$$

Решающая функция равномерно наиболее мощного алгоритма получена в виде:

$$\varphi(x) = \begin{cases} 1, & T(x) > C, \\ 0, & \text{иначе;} \end{cases}$$

где $T(x) = \sum_{i=1}^N f_i x_i$, C - пороговая константа.

Результаты моделирования.

Для записи электроэнцефалограмм использовался нейроинтерфейс Eros Emotiv и лицензионное программное обеспечение от компании Emotiv Systems – Emotiv Xavier Pure EEG. Тесты проводились в несколько сессий. Первая сессия включала в себя движение глаз вверх с задержкой в таком положении около 5 секунд. Аналогично проводились испытания

«влево», «вниз», «вправо». Фрагменты исследования представлены на Рис.3.

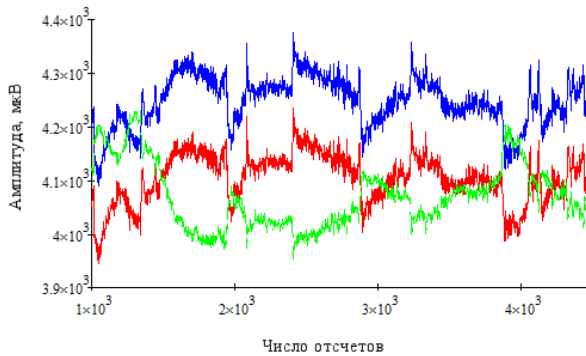


Рисунок 3. Фрагмент электроокулограммы при записи сигнала "глаза вверх"

После визуализации полученных данных были выделены участки с наиболее явным проявлением сигналов, соответствующих движениям глаз вверх(они и послужили основой для создания модели сигнала).

Имитационное моделирование алгоритма выполнено при помощи программы MathCAD. На Рис.4 представлена зависимость мощности алгоритма при различных уровнях вероятности ложного обнаружения.

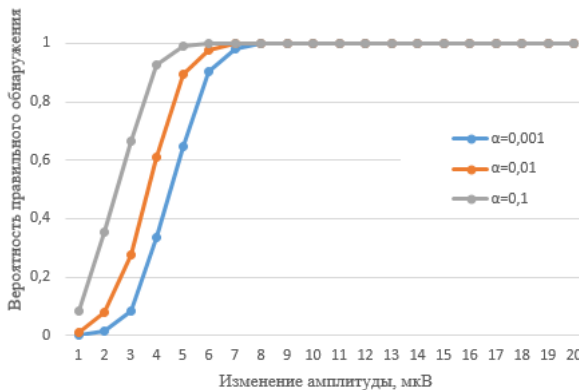


Рисунок 4. Зависимость вероятности правильного обнаружения от изменения амплитуды сигнала при различных значениях вероятности ложной тревоги

Видно, что при снижении заданного уровня вероятности ложной тревоги не приводит к существенному уменьшению вероятности правильного обнаружения. Алгоритм устойчив к изменениям уровня сигнала, а также не требует существенных вычислительных ресурсов, что обеспечивает его широкую применимость во встраиваемых системах.

Литература

1. Тихонов В. И. Статистическая радиотехника. - М., 1982. - 624 с.
2. Богданович В.А., Вострецов А.Г. Теория устойчивого обнаружения, различения и оценивания сигналов. – М.: Физматлит, 2004. – 320 с.

МЕТОДЫ КЛАСТЕРИЗАЦИИ ДЛЯ РЕШЕНИЯ ЗАДАЧ МАРШРУТИЗАЦИИ ТРАНСПОРТА С ДОПОЛНИТЕЛЬНЫМИ УСЛОВИЯМИ

А.А. Чепур, В.С. Карманов
Новосибирский Государственный Технический Университет,
г. Новосибирск, tchepur@mail.ru

Рассматриваются методы кластерного анализа для решения задач маршрутизации транспорта с учетом ограничений на его грузоподъемность и время прибытия.

Methods of cluster analysis are considered for solving transport routing problems taking into account restrictions on its carrying capacity and arrival time.

Цель работы

Применение методов кластеризации и создание программного модуля для решения задачи маршрутизации с учетом ограничений на грузоподъемность транспортных средств и время прибытия.

Введение

Задача маршрутизации транспорта является одной из задач комбинаторной оптимизации, в которых для парка транспортных средств должен быть определен маршрут до определенного количества потребителей. Оптимизация маршрутов, позволяющая снизить расходы на транспортировку, являются актуальной задачей для компаний, занимающихся транспортными грузоперевозками.

Постановка задачи

Имеется некоторое множество точек $V = \{v_0, v_1, \dots, v_n\}$, где v_0 – точка, в которой происходит загрузка транспортного средства (склад, депо), она является начальной и конечной точкой маршрута, $\{v_1, \dots, v_n\}$ – множество точек в которые нужно осуществить доставку груза [1,2]. Каждая точка из множества $\{v_1, \dots, v_n\}$ обладает несколькими свойствами: *weight* – вес доставляемого груза, t^{begin} – время после которого нужно осуществить доставку, t^{end} – время до которого нужно осуществить доставку.

Решением задачи является множество маршрутов (подмножеств множества V) с заданием порядка обхода на каждом из них, с учетом общего веса маршрута, который не должен превышать заданного норматива, и учета попадания моментов приезда и отъезда в заданное временное окно.

Методы решения задачи

Для решения задач маршрутизации с учетом дополнительных условий существует достаточно большое количество различных методов. Многие из них являются эвристическими методами, так как точные алгоритмы не всегда дают решение за приемлемое время при большой размерности задачи. В текущей работе рассматривается метод иерархической агломеративной кластеризации [3,4]. Подобные методы направлены на создание иерархии вложенных кластеров, где каждый новый кластер строится путем объединения более мелких кластеров на основе матрицы мер сходства.

Критерием для определения меры схожести кластеров является расстояние между точками по трем измерениям: долготе, широте и времени прибытия в эту точку. Варианты расчета расстояния в объединенном пространстве «географические координаты – время» задаются формулами (1-4).

$$D_{ij}^1 = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (t_i^e - t_j^e)^2}, \quad (1)$$

$$D_{ij}^2 = \alpha_1 \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} + (1 - \alpha_1) |t_i^e - t_j^e|, \quad (2)$$

$$D_{ij}^3 = |x_i - x_j| + |y_i - y_j| + |t_i^c - t_j^c|, \quad (3)$$

$$D_{ij}^4 = \alpha_1 \rho((x_i, y_i), (x_j, y_j)) + (1 - \alpha_1) |t_i^c - t_j^c|, \quad (4)$$

где x_i, y_i – соответственно широта и долгота точки, из которой выезжает транспортное средство; x_j, y_j – широта и долгота точки, в которую должно приехать транспортное средство; t_i^c, t_j^c – центры временного окна i, j точек; α_1 – вес; $\rho((x_i, y_i), (x_j, y_j))$ – расстояние по графу дорог, получаемое обращением к одному из геоинформационных картографических сервисов.

Среднее значение временного интервала, в течение которого нужно осуществить доставку, вычисляется как

$$t_i^c = \frac{t_i^{begin} + t_i^{end}}{2}. \quad (5)$$

В методе одиночной связи на основании матрицы сходства определяются два наиболее схожих или близких объекта, которые образуют первый кластер. На следующем шаге выбирается объект, который имеет наибольшее сходство хотя бы с одним из объектов, включенных в кластер.

Результат кластеризации входных данных (Рис. 1) можно представить в виде дендрограммы (Рис. 2). Вариант построенного маршрута приведен на Рис. 3.

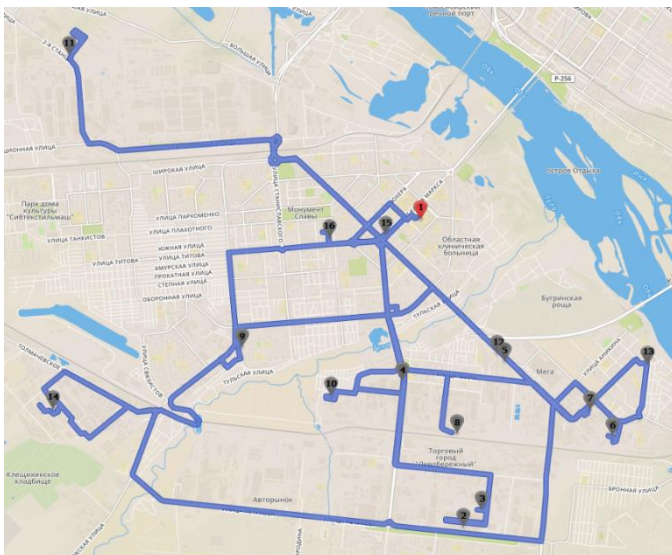


Рисунок 3 – Пример построенного маршрута.

Литература

1. Трофимов Д., Федуков А. Задача маршрутизации транспорта [Электронный ресурс] – 2006. – URL: <http://rain.ifmo.ru/cat/view.php/theory/unsorted/vrp-2006>
2. М. М. Заводченко, В. С. Карманов, М. Б. Олимпиади Применение методов кластерного анализа при решении задач маршрутизации транспорта // Интеллектуальные информационные системы : тр. Всерос. конф. с междунар. участием (Воронеж, 12-13 дек. 2017 г.). - Воронеж : ВГТУ, 2017. - Ч. 1. – С. 162-165. - 350 экз. - ISBN 978 - 5 - 7731 - 0565 - 7
3. Чубукова И.А. Data Mining [Электронный ресурс] / URL: <https://www.intuit.ru/studies/courses/6/6/lecture/182>
4. Айвазян С.А., Бухштабер В.М., Енюков И.С., Мешалкин Л.Д. Прикладная статистика: Классификация и снижение размерности - М.: Финансы и статистика, 1989. – С. 254-259

ВЫЯВЛЕНИЕ ТИПОВЫХ ДВИГАТЕЛЬНЫХ ШАБЛОНОВ ПО ДАННЫМ ИНЕРЦИАЛЬНЫХ ДАТЧИКОВ МОБИЛЬНЫХ УСТРОЙСТВ

Д.Ю. Шкатула, В.С. Карманов
Новосибирский Государственный Технический Университет,
г. Новосибирск, proxor@proxorgames.ru

Рассмотрен один из методов навигации внутри помещений, основанный на показаниях инерциальных датчиков мобильных устройств. Разработана и подготовлена информационная инфраструктура, посредством которой на сервер по транспортному протоколу TCP поступают показания датчиков. Развернутая на сервере нейронная сеть на основе больших данных анализирует и выявляет типовые двигательные шаблоны человека, которые в дальнейшем предполагается использовать при решении задачи позиционирования.

One of the methods of indoor navigation based on the readings of inertial sensors of mobile devices is considered. The information infrastructure has been developed and prepared, through which sensors are read by the TCP protocol on the server. A neural network deployed on a server, on the basis of big data, analyzes and identifies typical human movement patterns that are to be used in the future for solving positioning tasks.

В настоящий момент существует множество технологий навигации, например, GPS и ГЛОНАСС, основанных на спутниковой связи. Одна из проблем заключается в недоступности этих сетей в помещениях с экранирующими материалами, под землёй или из-за метеоусловий. В таком случае хорошим решением может стать использование инерциальной навигации.

Инерциальная навигация - метод навигации, основанный на свойствах инерции тел, являющийся автономным, то есть не требующим наличия внешних ориентиров или поступающих извне сигналов. Сущность инерциальной навигации состоит в определении ускорения объекта и его угловых скоростей с помощью установленных на движущемся объекте приборов и устройств, а по этим данным — местоположения (координат) этого объекта, его курса, скорости, пройденного пути и др., а также в определении параметров, необходимых для стабилизации объекта и автоматического управления его движением. Это осуществляется с помощью датчиков линейного ускорения – акселерометров; гироскопических устройств – гироскопов; вычислительных устройств – ЭВМ.

В рамках данного исследования в качестве аппаратной и программной платформы будут использоваться мобильные устройства на базе

операционных систем iOS и Android. Такие мобильные устройства являются полным комплексом, так как являются ЭВМ, а также содержат инерциальные датчики – акселерометры и гироскопы.

Каждое мобильное устройство имеет свои локальные оси, основанные на первоначальном положении устройства (*рис. 1*). Одной из основных задач стоит перевод локальных координат устройства в мировые, что можно реализовать при помощи матрицы перехода. Тем самым можно получить азимут, направление движения, а также вычислить точку отсчёта. Также одной из основных задач является подготовка инфраструктуры для сбора данных в режиме реального времени.

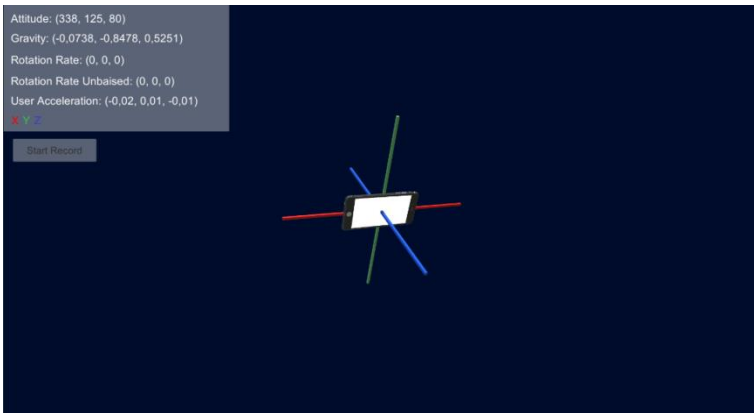


Рисунок 1 Мобильное приложение для снятия данных с устройства (iOS).

На *Рисунке 1* проиллюстрирована работа мобильного приложения, которое снимает показания с датчиков устройства в реальном времени. Трёхмерная модель устройства полностью копирует поведение реального устройства, то есть его положение в пространстве относительно наблюдателя в стартовой позиции. Также на рисунке обозначена локальная система координат устройства. В верхнем левом углу имеется список снимаемых показаний (округлённых для удобства просмотра):

- Attitude – угол поворота устройства относительно осей;
- Gravity – положение устройства в пространстве, наклон;
- Rotation Rate – скорость вращения (рад/с);
- User Acceleration – линейное ускорение (мм/с²).

Также в ходе работы программы можно начать трансляцию данных на сервер при помощи транспортного протокола ТСР, тем самым пополнив базу данных записей. На основе большого количества данных пришедших на сервер, специальное ПО анализирует эти данные и проводит сравнительную характеристику с типовыми двигательными шаблонам.

В качестве типовых двигательных шаблонов могут рассматриваться следующие:

- Прямолинейное движение (шаг, бег);
- Движение по лестнице;
- Движение на эскалаторе, в лифте;
- Поворот, разворот (на месте, в движении);
- Наклон;
- Сгибание-разгибание руки в локте;
- Подъём-опускание руки вперёд, в сторону;
- Разговор по сотовому телефону;
- Общение по сотовому телефону в текстовом режиме (мессенджеры);
- Съёмка (фото или видео) на камеру сотового телефона;
- Съёмка селфи.

Дополнительно решаются задачи выявления других типовых шаблонов и их комбинаций. Например, при последовательном выполнении шаблонов решается задача распознавания момента перехода от одного к другому, при параллельном (одновременном) выполнении шаблонов – задача их распознавания и разделения.

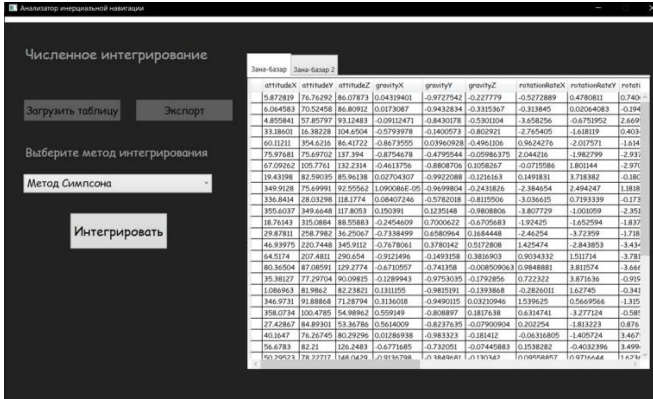


Рисунок 2. Программа для анализа данных, пришедших на сервер

На *Рисунке 2* проиллюстрирована программа, которая считывает файлы, пришедшие на сервер. Далее производится обработка полученных данных (фильтрация шумов датчиков, численное интегрирование данных линейного ускорения, перевод в неподвижную систему координат и пр.)

Также имеется возможность экспорта данных в таблицу Excel, для дальнейшей обработки и построения графиков (*Рисунок 3*):

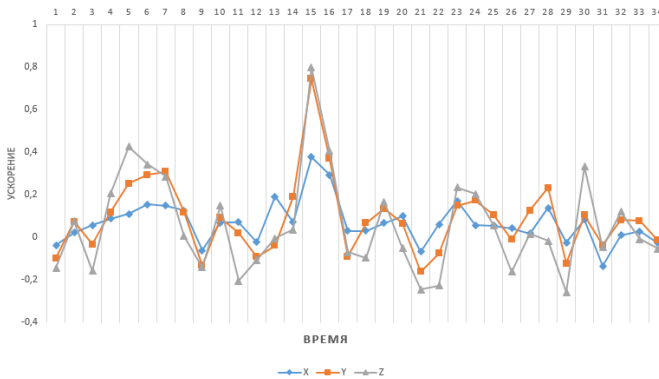


Рисунок 3. График линейного ускорения для одной записи (фрагмент).

Выявление и распознавание типовых двигательных шаблонов в дальнейшем планируется применять при решении задачи позиционирования, т.е. определения положения по данным инерциальных датчиков.

Литература

1. *Sebastian O.H. Madgwick*. An efficient orientation filter for inertial and inertial/magnetic sensor arrays [Электронный ресурс] / URL: http://x-io.co.uk/res/doc/madgwick_internal_report.pdf
2. *Shala U., Rodriguez A*. Indoor Positioning using Sensor-fusion in Android Devices [Электронный ресурс] / Kristianstad University, Department Computer Science. — Sweden, 2011. — URL: <https://www.diva-portal.org/smash/get/diva2:475619/FULLTEXT02.pdf>
3. *Османец А.А.* ОПРЕДЕЛЕНИЕ МЕСТОПОЛОЖЕНИЯ ТЕЛЕФОНА ПО ДАННЫМ СЕНСОРОВ // Машинное обучение и анализ данных. 2014. Т. 1. № 9. С. 1232-1245.

ИССЛЕДОВАНИЕ РАБОТЫ АЛГОРИТМА ИДЕНТИФИКАЦИИ НОТ ДЛЯ ВЫБОРА МЕТОДА ОПРЕДЕЛЕНИЯ ГРАНИЦ НОТЫ

А.Ю. Якимук, К.М. Томская
**Томский государственный университет систем
управления и радиоэлектроники,**
г. Томск, ya@keva.tusur.ru

В данной статье рассматривается задача определения влияния выбора метода скалирования для идентификации нот на получаемые результаты. Определение наличия погрешности в распознанных нотах, зависящее от выбранной методики усреднения, способно снизить число ошибок первого и второго рода.

This article discusses the problem of determining the influence of the choice of scaling method for identifying notes on the results obtained. The determination of the presence of errors in the recognized notes, depending on the chosen averaging technique, can reduce the number of type I and type II errors.

Введение

В общепринятом подходе соотнесения частот с музыкальными нотами существует следующая закономерность: для каждой ноты указано точное значение частоты. Если придерживаться такого подхода, то

можно сделать вывод о том, что все промежуточные частоты не относятся ни к одной ноте. Однако, колебания частот во время исполнения ноты располагаются в некотором интервале в районе указанных в таблицах значений. Грамотно построенная шкала интервалов частот способна снизить число ошибок, получаемых при идентификации исполненных диктором нот. Исследование данного вопроса является одним из основных при создании программы по распознаванию нот[1].

1. Построение интервальных шкал

Проведение эксперимента осуществлялось с помощью программного комплекса Amadeus [2], дополненного возможностью выбора шкалы в начальном меню настройки (рис.1). Выбираемый файл конфигурации содержал значения для начала и конца интервалов каждой ноты.

При построении файлов конфигураций с интервальными шкалами были рассмотрены следующие методы усреднения: средняя гармоническая, средняя геометрическая, средняя арифметическая, средняя квадратическая и средняя кубическая. Кроме того, в связи с природой рассматриваемых величин, было добавлено разделение границ нот по формуле вычисления частот звуков.

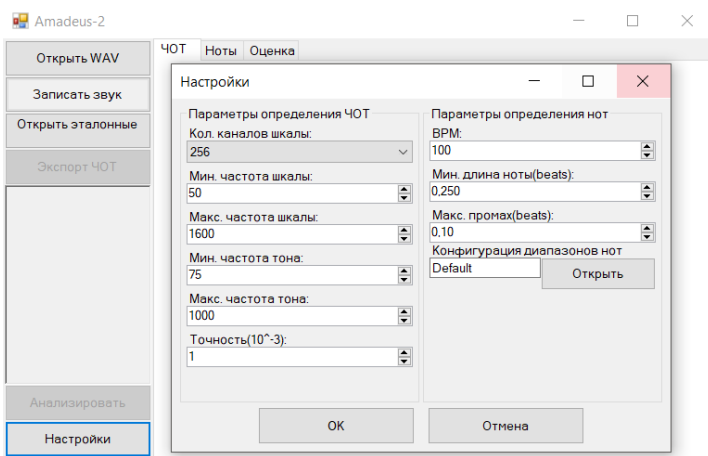


Рисунок 1 – Внесение возможности выбора шкалы идентификации нот

Для этого была модифицирована формула, применяемая в равномерном темпированном строю для получения значения следующей ноты, путем понижения степени, в которую возводился коэффициент, в 2 раза. Итоговые формулы определения верхней и нижних границ частот для i – й ноты приобрела следующий вид:

$$f_{i_{\text{Н}}} = f_{i-1} * 2^{\frac{1}{24}} \quad (1)$$

$$f_{i_{\text{В}}} = f_i * 2^{\frac{1}{24}} \quad (2)$$

Обозначим данный вид шкалы идентификации нот как логарифмический. На рисунке 2 представлен фрагмент полученной таблицы шкал идентификации нот в диапазоне от ля 1-й октавы до соль-диез 2-й октавы. Полученные интервалы задействованы в программном комплексе таким образом, что конец интервала одной ноты является началом для другой. При этом минимальное значение не включается в интервал в каждом случае. При таком подходе на диапазоне частот не остается пустых значений, которые не отнесены ни к одной из нот, а границы интервалов расположены строго между нотами, что не позволит получить ситуацию, при которой в одном интервале находится более одной ноты. Вопрос определения чисто исполненных нот решается использованием в программе коэффициента, регулирующего полноту использования интервала идентификации нот. Настройка данного параметра осуществляется пользователем в зависимости от способностей диктора.

| Частота звучания ноты | Среднее | | | | | | | | | | | |
|-----------------------------|---------------|----------|-----------------|----------|----------------|----------|----------------|----------|----------------|----------|------------|----------|
| | Гармоническое | | Логарифмическое | | Геометрическое | | Арифметическое | | Квадратическое | | Кубическое | |
| | Ниж.гр. | Верх.гр. | Ниж.гр. | Верх.гр. | Ниж.гр. | Верх.гр. | Ниж.гр. | Верх.гр. | Ниж.гр. | Верх.гр. | Ниж.гр. | Верх.гр. |
| 698,456 | 678,290 | 718,623 | 678,571 | 718,921 | 678,572 | 718,923 | 678,856 | 719,223 | 679,138 | 719,522 | 679,421 | 719,822 |
| 739,989 | 718,623 | 761,355 | 718,921 | 761,671 | 718,923 | 761,672 | 719,223 | 761,990 | 719,522 | 762,308 | 719,822 | 762,625 |
| 783,991 | 761,355 | 806,627 | 761,671 | 806,962 | 761,672 | 806,963 | 761,990 | 807,300 | 762,308 | 807,636 | 762,625 | 807,972 |
| 830,609 | 806,627 | 854,591 | 806,962 | 854,946 | 806,963 | 854,948 | 807,300 | 855,305 | 807,636 | 855,661 | 807,972 | 856,017 |

Рисунок 2 – Шкалы идентификации нот

Сравнение полученных шкал (рис. 3) показало, что отклонение от идеального значения ноты для каждого из методов усреднения данных составляет от 2,72 % у кубического метода до 2,88% у гармонического

для нижней границы и от 2,88% у гармонического до 3,05 у кубического для верхней границы.

| Частота звучания ноты | Среднее | | | | | | | |
|-----------------------|----------------|----------|-----------|------------|-----------------|----------|-----------|------------|
| | Гармоническое | | | | Логарифмическое | | | |
| | Ниж.гр. | Верх.гр. | Δ ниж.гр. | Δ верх.гр. | Ниж.гр. | Верх.гр. | Δ ниж.гр. | Δ верх.гр. |
| 65,406 | 63,518 | 67,295 | -1,8885 | 1,8885 | 63,544 | 67,323 | -1,8622 | 1,9164 |
| 739,989 | 718,623 | 761,355 | -21,3661 | 21,3658 | 718,921 | 761,671 | -21,0682 | 21,6817 |
| 1975,530 | 1918,495 | 2032,569 | -57,0355 | 57,0392 | 1919,295 | 2033,413 | -56,2355 | 57,8830 |
| Частота звучания ноты | Среднее | | | | | | | |
| | Геометрическое | | | | Арифметическое | | | |
| | Ниж.гр. | Верх.гр. | Δ ниж.гр. | Δ верх.гр. | Ниж.гр. | Верх.гр. | Δ ниж.гр. | Δ верх.гр. |
| 65,4064 | 63,544 | 67,323 | -1,8620 | 1,9166 | 63,571 | 67,351 | -1,8355 | 1,9447 |
| 739,989 | 718,923 | 761,672 | -21,0664 | 21,6833 | 719,223 | 761,990 | -20,7665 | 22,0010 |
| 1975,530 | 1919,295 | 2033,417 | -56,2354 | 57,8869 | 1920,095 | 2034,265 | -55,4350 | 58,7350 |
| Частота звучания ноты | Среднее | | | | | | | |
| | Квадратическое | | | | Кубическое | | | |
| | Ниж.гр. | Верх.гр. | Δ ниж.гр. | Δ верх.гр. | Ниж.гр. | Верх.гр. | Δ ниж.гр. | Δ верх.гр. |
| 65,4064 | 63,597 | 67,379 | -1,8090 | 1,9727 | 63,62 | 67,41 | -1,7825 | 2,0008 |
| 739,989 | 719,522 | 762,308 | -20,4668 | 22,3186 | 719,82 | 762,62 | -20,1674 | 22,6357 |
| 1975,530 | 1920,895 | 2035,113 | -54,6349 | 59,5827 | 1921,69 | 2035,96 | -53,8359 | 60,4294 |

Рисунок 3 – Сравнение отклонений в полученных шкалах

В среднем ширина интервала относительно значения исследуемой ноты составляет: для гармонического метода – 5,774%, для логарифмического и геометрического – 5,777%, для арифметического – 5,779%, для квадратического – 5,782%, а для кубического – 5,784%.

2. Экспериментальная проверка интервальных шкал

Для исследования полученных шкал дополнительно было сделано 29 контрольных аудиозаписей для распознавания исполненных дикторами нот. Каждая аудиозапись содержала отдельную спетую ноту.

Для анализа было отобрано 3 эксперта с обязательным наличием музыкального образования. Все полученные от экспертов оценки исполненных дикторами нот были проверены с помощью коэффициента конкордации Кендалла, показавшем высокую степень согласованности [3]. Далее все аудиозаписи были проанализированы с помощью программы Amadeus с поочередным использованием всех шкал.

В ходе вычислительного эксперимента было замечено, что программа вывела идентичные результаты вне зависимости от выбора шкалы. Причиной для подобного результата может служить то, что среди запи-

сей самой высокой исполненной нотой является фа-диез второй октавы, звучащая на высоте 739,98 Гц. На данном уровне частот разница между границами интервалов разных шкал достаточно мала: разница между средним гармоническим и средним кубическим не превышает 1,5 Гц, что составляет 0,17% от значения частоты, на которой прозвучала исследуемая нота.

Заключение

Итогом данного исследования стал вывод о том, что алгоритм идентификации нот работает корректно вне зависимости от того, какая методика усреднения данных была использована при формировании шкалы интервальных значений звучания ноты. Полученные результаты свидетельствуют о том, что на диапазоне частот, на котором осуществляется пение, разница значений границ не существенна, и для упрощения алгоритма можно использовать среднее арифметическое.

Литература

1. Конев А.А., Онищенко А.А., Костюченко Е.Ю., Якимук А.Ю. Автоматическое распознавание музыкальных нот // Научный вестник НГТУ том 60, №3, 2015, с. 32–47
2. Якимук А.Ю., Конев А.А., Осипов А.О. Программный комплекс для автоматизации моделирования сегментации речевых сигналов и вокальных исполнений // Вестник Иркутского государственного технического университета. 2017. Т. 21. № 10. С. 53–64.
3. Томская К.М. Определение метода шкалирования для идентификации нот с помощью частот основного тона // Российская наука в современном мире: сборник статей XVI международной научно-практической конференции. 2018. С. 88-91.

**ЧИСЛЕННЫЙ МЕТОД ОЦЕНКИ ПАРАМЕТРОВ
МАТЕМАТИЧЕСКИХ МОДЕЛЕЙ В ФОРМЕ
ЛОГИСТИЧЕСКИХ ФУНКЦИЙ**

Е.А. Афанасьева, В.Е. Зотеев
Самарский государственный технически университет,
г. Самара, elenaafa6277@gmail.com

Разработан новый численный метод оценки параметров математических моделей в форме логистических функций. В основе данного метода лежит построение разностных уравнений, описывающих результаты наблюдений, и нахождение среднеквадратичной оценки их коэффициентов. Алгоритм построения численного метода и его апробация рассмотрены на примере математической модели, построенной с помощью функции Верхульста.

A new numerical method for estimating the parameters of mathematical models in the form of the function of Verhulst has developed. The basis of this method is the construction of difference equations describing the results of observations, and the RMS estimate of their coefficients. The algorithm for constructing a numerical method and its approbation are considered on the example of a mathematical model complied using of the function of Verhulst.

При построении математических моделей различных социально-экономических процессов широкое применение получили логистические функции (тренды). Основной проблемой при моделировании экономических процессов является нахождение достоверной оценки параметров логистического тренда по данным, полученным в результате наблюдений. Эффективное решение поставленной задачи можно реализовать с помощью современных информационных и компьютерных технологий, применяя известные методы нелинейного оценивания параметров [1]. Однако при большом разбросе данных относительно построенной модели на основе логистической функции известные подходы не позволяют получить необходимую точность вычислений, при этом требуя выбор начального приближения вектора оценки параметров. Данная процедура сопровождается большими трудностями.

Разработан новый численный метод, основанный на построении разностных уравнений, описывающих результаты наблюдений. В свою

очередь, коэффициенты уравнений связаны с параметрами логистической функции известными соотношениями.

Рассмотрим алгоритм численного метода на примере нахождения оценки параметров логистической функции Верхульста, имеющей вид:

$$\mathfrak{F}(t) = A_0 / (1 + A_1 \exp(-\alpha t)), \quad (1)$$

где A_0 , A_1 и α – параметры, которые необходимо идентифицировать, основываясь на данных наблюдений. Используя подход, представленный в [2], необходимо составить рекуррентное соотношение, связывающие два последовательно идущих значения дискретной функции $\mathfrak{F}_k(\tau k) = A_0 / (1 + A_1 \exp(-\alpha \tau k))$, где τ – период дискретизации:

$$\mathfrak{F}_{k-1} \mathfrak{F}_k = \mu_1 \mathfrak{F}_k + \mu_2 \mathfrak{F}_{k-1}, \quad k = 1, 2, \dots, N-1. \quad (2)$$

Учитывая естественный разброс данных ε_k в результатах наблюдений $y_k = \mathfrak{F}_k + \varepsilon_k$, $k = 0, 1, 2, \dots, N$, где N – объем выборки результатов наблюдений, на основе рекуррентного соотношения (2) становится возможным построить систему линейных разностных уравнений, которая описывает результаты наблюдений:

$$\begin{cases} y_0 = \mu_3 + \varepsilon_0; \\ y_k y_{k-1} = \mu_1 y_k + \mu_2 y_{k-1} + \eta_k; \\ \eta_k = \varepsilon_{k-1} (y_k - \mu_2) + \varepsilon_k (y_{k-1} - \mu_1), k = 1, 2, 3, \dots, N-1. \end{cases} \quad (3)$$

Коэффициенты системы уравнений (3) связаны с параметрами функции Верхульста (1) следующими соотношениями:

$$\mu_1 = \frac{A_0}{1 - e^{\alpha \tau}}, \quad \mu_2 = -\mu_1 e^{\alpha \tau}, \quad \mu_3 = \frac{A_0}{1 + A_1}. \quad (4)$$

Для нахождения среднеквадратичной оценки коэффициентов системы разностных уравнений (3) необходимо учесть условие минимизации функционала $\|\varepsilon\|^2 = \|y - \mathfrak{F}\|^2 = \sum_{k=0}^{N-1} (y_k - \mathfrak{F}_k)^2 \rightarrow \min$. Далее воспользуемся матричной формой записи и представим систему разностных уравнений (3) в виде обобщенной регрессионной модели:

$$\begin{cases} b = F\mu + \eta; \\ \eta = P_\mu \varepsilon. \end{cases} \quad (5)$$

Элементы модели (5) формируются на основе данных наблюдений y_k . Например, матрица P_μ линейного преобразования вектора естественного разброса данных ε формируется следующими соотношениями: $p_{11} = 1$, $p_{ii} = y_{i-2} - \mu_1$, $p_{i,i-1} = y_{i-1} - \mu_2$, $i = \overline{2, N}$; остальные элементы матрицы равны P_μ нулю.

На следующем шаге представим модель (5) в виде $P_\mu^{-1}b = P_\mu^{-1}F\mu + \varepsilon$. Понятно, что условие минимизации среднеквадратичного отклонения построенной модели от данных наблюдений $\|y - \mathfrak{F}\|^2$ сводится к минимизации функционала $\|\varepsilon\|^2 = \|P_\mu^{-1}b - P_\mu^{-1}F\mu\|^2 \rightarrow \min$. Из этого условия следует, что для вычисления оценок коэффициентов обобщенной регрессионной модели можно воспользоваться итерационной процедурой уточнения вектора среднеквадратичных оценок \mathfrak{F} коэффициентов разностного уравнения:

$$\mathfrak{F}^{(i+1)} = \left(F^T \Omega_{\mathfrak{F}^{(i)}}^{-1} F \right)^{-1} F^T \Omega_{\mathfrak{F}^{(i)}}^{-1} b, \quad \Omega_{\mathfrak{F}^{(i)}} = P_{\mathfrak{F}^{(i)}} P_{\mathfrak{F}^{(i)}}^T, \quad (6)$$

где $i = 0, 1, 2, \dots$ – номер итерации. Начальное приближение вектора среднеквадратичных оценок $\mathfrak{F}^{(0)}$ можно найти из условия минимизации

ции невязки $\|r\|^2 \rightarrow \min$ следующим соотношением

$$\mathcal{K}^{(0)} = (F^T F)^{-1} F^T b.$$

Для нахождения оценок параметров логистической функции Верхулста (1) воспользуемся среднеквадратичными оценками коэффициентов разностных уравнений. Из (4) получаем соотношения:

$$\mathcal{C} = -\frac{1}{\tau} \ln \left(-\frac{\mu_1}{\mu_2} \right), \quad \mathcal{A}_0 = \mu_1 + \mu_2, \quad \mathcal{A}_1 = \frac{\mu_1 + \mu_2}{\mu_3} - 1. \quad (7)$$

Были проведены численно-аналитические исследования разработанного численного метода с помощью компьютерного моделирования, которые показали помехозащищенность оценок параметров логистической функции Верхулста и высокую адекватность построенной модели по данным наблюдений.

Для апробации численного метода был сделан прогноз числа мобильных телефонов на 1 тыс. человек взрослого населения. Экспериментальная выборка данных была взята из социологических опросов Левада-Центр (<http://www.levada.ru/>) за период с 2000 по 2007 год в РФ. Для моделирования была использована логистическая функция Верхулста, построение моделей осуществлялось двумя классическими методами (метод Ньютона и метод на основе логарифмирования) и разработанным численным методом на основе разностных уравнений. После был сделан прогноз роста доли взрослых пользователей мобильным телефоном на март 2014 года с помощью трех методов, результаты прогнозирования были сравнены с данными Левада-Центр.

Полученные результаты:

- Метод на основе логарифмирования– 75%;
- Метод Ньютона– 78%;
- Разработанный метод– 79%;
- Данные Левада-Центр– 81%;

Ясно, что результат, полученный разработанным методом (79%), наиболее близок к статистическим данным (81%).

Таким образом, в работе был разработан новый численный метод оценки параметров математических моделей в форме логистических функций, приведены формулы, позволяющие вычислить параметры

тренда через коэффициенты разностных уравнений, показана методика оценки погрешности результатов вычислений и проведены численно-аналитические исследования. Разработанный метод был апробирован на примере прогноза будущего роста показателя числа телефонов на 1 тыс. человек взрослого населения в РФ. Сравнение полученного результата с имеющимися статистическими данными показало состоятельность нового численного метода.

Литература:

1. Демиденко Е.З. Линейная и нелинейная регрессии. М.: Финансы и статистика, 1981. – 303 с.
2. Зотеев В.Е. Параметрическая идентификация диссипативных механических систем на основе разностных уравнений / Под ред. Радченко В.П. – М.: Машиностроение, 2009. – 344 с.

ИССЛЕДОВАНИЕ МОДЕЛЕЙ ВЫБОРА МАРШРУТА НА ПРИМЕРЕ ТРАНСПОРТНОЙ СЕТИ АКАДЕМГОРДКА

М.Д. Верига, С.П. Сарычев
Сибирский государственный университет путей сообщения
г. Новосибирск, mari.veriga.96@mail.ru

В данной работе рассмотрена и построена модель транспортной сети Академгородка при помощи средств программного комплекса PTV Visum. В ходе исследования были смоделированы различные ситуации движения маршрутов на сети, а также рассмотрен фактор информированности пассажиров и его влияние на выбор маршрута. Иными словами был оценен эффект от повышения фактора информированности пассажиров на среднее время, затрачиваемое на ту или иную поездку.

A model of the transportation network of Akademgorodok was reviewed and constructed using the tools of the PTV Visum software package. During the study, various situations of movement of routes on the network were modeled, and the factor of passenger awareness and its influence on the choice of route was considered. The effect of raising the passenger awareness factor on the average time spent on a particular trip was evaluated.

В условиях роста уровня автомобилизации существенно повышается значение сбалансированного использования ресурсов для эффективной организации дорожного движения. При этом, на участках высокой ин-

тенсивности движения основной ресурс – земельный – является дефицитным и одним из наиболее финансово емких. Современным инструментом сбалансированного использования ресурсов и является транспортное планирование.

В настоящее время хорошо развитая транспортная сеть Новосибирска отчасти компенсирует разбросанность города и необходимость перемещений на большие расстояния по его территории. Но к сожалению, сложившаяся дорожная ситуация не позволяет говорить о том, что все компоненты транспортной системы города работают согласованно. Особенную роль в этой системе играет пассажирский транспорт, т.к. его элементы в большей степени поддаются управлению, координации, систематизации, а так же упорядоченности транспортных потоков.

Современное развитие интеллектуальных систем позволяет повысить эффективность работы пассажирского транспорта и минимизировать расходы, связанные с ним.

Средствами программного комплекса PTV Visum была создана математическая модель транспортных потоков сети Академгородка.

При построении модели соблюдался следующий алгоритм:

- создание модели транспортного предложения;
- создание модели транспортного спроса;
- создание модели воздействия (создание и распределение движения, выбор нужного режима, перераспределение на транспортную сеть);
- подведение итогов по созданной модели;
- проведение нескольких экспериментов (4) и вычисление нового показателя (среднего времени поездки).

Фрагмент транспортной модели представлен на рисунке 1.

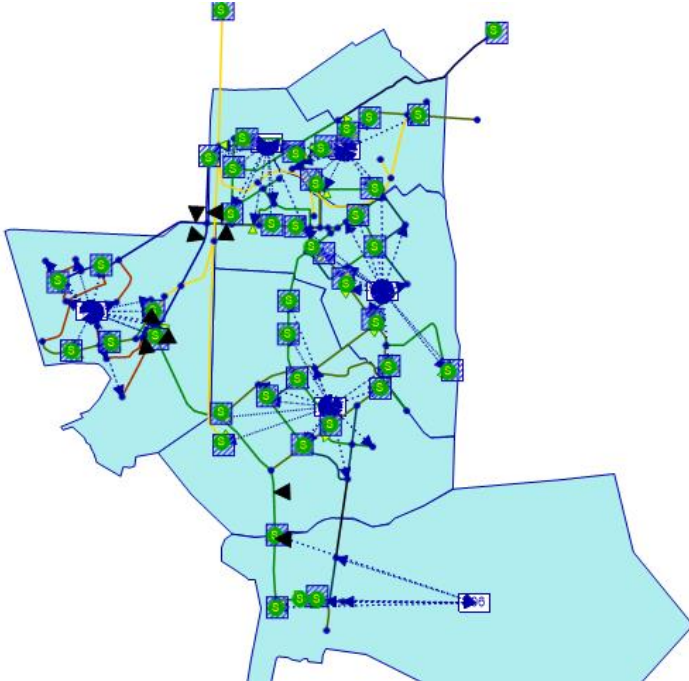


Рисунок 1 – Модель Академгородка в ПК PTV Visum

В перераспределении по интервалам принципиально считается, что пассажиры знают интервалы и время движения маршрутов, и от того, какой еще информацией помимо этого они владеют, зависит их поведение при выборе маршрута на посадку или пересадку.

При проведении эксперимента были использованы следующие модели выбора маршрута:

- Отсутствие информации и экспоненциально распределенные интервалы;
- Отсутствие информации и постоянные интервалы;
- Информация о прошедшем времени ожидания;
- Информация о следующем времени отправления для маршрутов, останавливающихся на остановке.

Последний случай имеет место, например, тогда, когда на остановках будут установлены динамические информационные системы для

пассажиров. В этом случае пассажиры могут на основе информации на табло узнать, какой из отправляющихся маршрутов в актуальной ситуации потратит меньше всего времени, чтобы доставить их к цели (так называемый маршрут с наименьшим остаточным временем поездки).

Таким образом, в ходе исследования были смоделированы различные ситуации движения маршрутов на сети, а также рассмотрен фактор информированности пассажиров и его влияние на выбор маршрута. Иными словами оценен эффект от повышения фактора информированности пассажиров на среднее время, затрачиваемое на ту или иную поездку.

Полученные в работе результаты позволяют сформировать рекомендации по изменению функционирования транспортной модели Академгородка.

В дальнейшем планируется решение задачи оптимизации маршрутной сети как Академгородка, так и всего Новосибирска.

Литература:

1. Лозе Д. Моделирование транспортного предложения и спроса на транспорт для пассажирского и служебного транспорта – обзор теории моделирования // Организация и безопасность дорожного движения в крупных городах: Сборник докладов седьмой международной научно-практической конференции. СПб гос. архит.строит. ун-т. СПб, 2006 – 544 с.
2. Мягков В.Н., Пальчиков Н.С., Федоров В.П. Математическое обеспечение градостроительного проектирования. Л.: Наука, 1989.

ОБНАРУЖЕНИЕ ДЕФЕКТОВ В ПРОДУКЦИИ МЕТОДОМ МАШИННОГО ОБУЧЕНИЯ

Д.М. Гойал, В.С. Тимофеев

**Новосибирский государственный технический университет,
г. Новосибирск, gojal.2014@corp.nstu.ru**

В данной статье рассматривается задача контроля качества продукции. Для решения проблемы предлагается осуществлять обнаружение дефектов путем анализа изображений продукции методом машинного обучения. В результате будет разработано программное обеспечение, позволяющее предпри-

ятию ускорить процесс выявления причин появления бракованной продукции для создания более совершенных технологий изготовления.

This article deals with the issue of the control of product quality. To solve the problem it is proposed to detect defects by analyzing images of the product using the method of machine learning. As a result, the creation of software will allow the company to speed up the process of identifying the causes of the manufacturing defect to create advanced production techniques.

В наше время контроль качества продукции играет важную роль и является неотъемлемой частью производственного процесса. Он, в первую очередь, направлен на проверку надежности в процессе изготовления продукции, ее потребления или эксплуатации. Процесс контроля заключается в установлении соответствия действительных значений физических величин установленным предельным значениям [1]. Контроль должен отвечать на вопрос находится ли анализируемая физическая величина в поле допуска или выходит за его пределы.

В частности, на предприятиях, занимающихся производством преформ из полиэтилентерефталата (ПЭТ), очень важно осуществлять контроль качества. Преформа – это первоначальная заготовка для последующего создания ПЭТ-контейнеров (бутылок), отличающихся по типу горловины, весу, цвету и используемым добавкам, специально созданная под требования потребителей для различных сегментов. Так, например, на предприятии ООО "Эталон" в качестве соблюдения и регулирования норм производят измерительный контроль, осуществляемый с помощью различных средств измерений: фотометра, контрольных весов, толщинометра, лампы с фильтром поляризованного света, позволяющей обнаружить дефекты, не поддающиеся определению невооруженным глазом.

При производстве преформ, задача контроля качества осложнена огромным количеством произведенных изделий, а также трудностью диагностирования визуальных дефектов. Вследствие этого значительно повышаются затраты на обеспечение контроля качества (увеличивается штат сотрудников для просмотра выпущенных преформ, так как автоматически определять визуальные дефекты в силу их большого количества не представлялось возможным). Для решения этой проблемы на производстве был разработан комплекс «ЛИКА», задача которого сводится к мгновенному сканированию преформы и анализу полученных изображений (снимков). Однако на данном этапе отсутствует программное обеспечение, осуществляющее их точный анализ. Поэтому применяется «грубая» оценка полученных данных. Для проведения

данной оценки был введен такой параметр, как порог отбраковки, который устанавливается вручную для различных преформ. Если обнаружено значение, отклоняющееся от среднего на величину порога отбраковки, то преформа считается бракованной. Решение задачи таким способом дает неточный результат, а также не позволяет автоматически определить тип дефекта, поэтому было принято решение разработать программное обеспечение, позволяющее автоматизировать диагностику данных со снимков для точного определения наличия дефекта.

Для сканирования преформы используется прибор с зарядовой связью, называемый ПЗС-матрица, размерности 512 на 512. Полученный массив данных может быть представлен, как изображение с разрешением 0,262 Мп. Использование такого набора данных позволит диагностировать только два дефекта: «пузырь» и «непроплав», так как только эти дефекты влияют на проходимость света через преформу.

Задача классификации дефектов сводится к нахождению их визуальных отличий. На рисунке 1 представлена преформа с сочетанием дефектов «пузырь» и «непроплав». Для «пузыря» характерна плохая проходимость света в преформе на границе дефекта, в то время как для «непроплава» это явление наблюдается по всей площади. Такого рода брак особенно критичен, поскольку приводит к взрыву контейнера при его дальнейшем использовании.

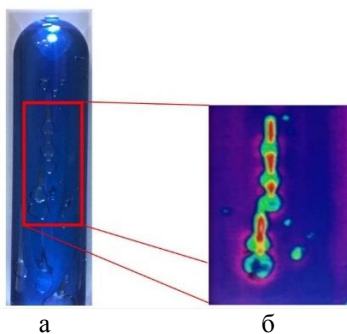


Рисунок 1 – преформа с дефектом пузырь, непроплав и их результат сканирования: а – фото преформы; б – результат сканирования

Для реализации данной задачи предполагается использовать сверточную нейронную сеть, входящую в состав технологий глубокого обучения. Данный метод позволяет значительно уменьшить количество настраиваемых весов, сделать модель устойчивой к повороту и сдвигу

распознаваемого изображения [2]. Для создания обучающей выборки был предусмотрен автоматический сбор информации (снимки преформ с комплекса «ЛИКА»). В ходе обучения нейронной сети будут найдены оптимальные значения весов для ядра свертки, что позволит использовать разработанную модель для классификации известных изображений из обучающей выборки, а также других изображений преформ.

Созданное программное обеспечение позволит не только определить наличие брака, но и классифицировать дефекты. Предприятие сможет использовать полученные результаты для формирования бизнес-отчетов, анализ которых поможет ускорить процесс выявления причин появления бракованной продукции для создания более совершенных технологий изготовления.

Литература:

1. Зайцев Г.Н. Управление качеством. Технологические методы управления качеством изделий: Учебное пособие. Стандарт третьего поколения. – СПб.: Питер, 2014. – 272с.: ил. – (Серия «Учебное пособие»).
2. Lee H., Grosse R., Ranganath R., Ng A.Y. Convolutional Deep Belief Networks for Scalable Unsupervised Learning of Hierarchical Representations. Proceedings of the 26th Annual International Conference on Machine Learning. 2009.

МАТЕМАТИЧЕСКАЯ КОМПЬЮТЕРНАЯ МОДЕЛИ ДЛЯ РЕШЕНИЯ ОПЕРАТИВНОЙ ПРОБЛЕМЫ РЕГИОНАЛЬНЫХ ГРУЗОПЕРЕВОЗОК

Е.А. Загирова, А.В. Панюков
Южно-Уральский государственный университет,
г. Челябинск, zagirovaea@mail.ru

В работе рассмотрена задача распределения транспортных средств разных видов между пунктами доставки (регионов) с разными видами товаров. Была построена математическая модель распределения транспортных средств между пунктами доставки с различными видами товаров и разработана программа автоматизации данного процесса.

The paper considers the different types vehicles distribution problem between delivery points (regions) with various types of goods. A mathematical model was built for the distribution of vehicles between delivery points with different types of goods and a program was developed to automate this process.

На текущий момент распределение и планирование перевозок больших объемов товаров требует значительных временных затрат и трудовых ресурсов или автоматизации. В данной работе предложен один из вариантов автоматизации деятельности отдела логистики для планирования перевозок товара между регионами [1].

На основе анализа рынка были выявлены следующие требования к информационной системе:

- возможность учёта потоков товаров как в пределах, так и за пределами системы рассматриваемого региона;
- возможность учёта многономенклатурных перевозок;
- возможность учёта перевозок различными видами транспорта;
- удобный пользовательский интерфейс.

Исходными данными задачи моделирования региональных товарных потоков является следующая информация:

1. О транспортных средствах: количество видов транспортных средств, вместительность транспортных средства по весу (в килограммах).
2. О товаре: количество видов товара, количество товара, требуемого к поставке (в килограммах), вес одной единицы каждого вида товара.
3. О неустойке: количество денежных средств, требуемых к выплате при недоставке товара по маршруту. Для построения оптимального плана перевозок каждого вида товара по каждому маршруту необходимо минимизировать целевую функцию по неустойке.

Для решения данной задачи будем использовать методы распределительной задачи линейного программирования, а именно методы решения задачи о назначениях. Согласно ее специфике, что каждым грузом необходимо будет оперировать как одним целым. То есть будут иметь место булевы переменные: груз либо доставляется, либо нет. Каждое транспортное средство также является неделимым.

Пусть:

- J – множество всех регионов, $i, j \in J$;
- r – вид товара, $r \in R$;
- k – вид транспорта, $k \in K$;

- d – маршрут из i в j , $d = \{(i, j): i, j \in J, i \neq j\}$, $d \in D$

К показателям искомого плана перевозок будут относиться следующие величины:

- B_{ij}^k – количество доступного транспорта вида $k \in K$ для перевозки от региона $i \in J$ до региона $j \in J$;
- E_{ij}^r – объем заказа на перевозку товара $r \in R$ от региона $i \in J$ до региона $j \in J$;
- e_{ij}^r – неудовлетворенная часть спроса на перевозку товара $r \in R$ от региона $i \in J$ до региона $j \in J$;
- c_{ij}^r – размер неустойки при невыполнении перевозки продукта $r \in R$ от региона $i \in J$ до региона $j \in J$;
- λ^{rk} – обобщенный удельный объем, требуемый для перевозки единицы товара $r \in R$ на транспорте вида $k \in K$;
- α_r – удельный объем одной единицы товара вида $r \in R$;
- β_k – удельный объем, требуемый для перевозки единицы товара на транспорте вида $k \in K$;
- z_{ij}^r – есть объем продукта, перенесенный с маршрута на транзитный маршрут.

Тогда многоиндексная транспортная задача [2] формулируется следующим образом: требуется определить такие величины x_{ij}^{rk} – количество продукта r , который был перевезен на транспорте k из региона i в j , чтобы выполнялись следующие соотношения:

$$\sum_{r \in R} c_{ij}^r e_{ij}^r \rightarrow \min,$$

$$r \in R, i, j \in J;$$

$$\sum_{r \in R} \lambda^{rk} x_{ij}^{rk} \leq B_{ij}^k, k \in K,$$

$$\sum_{k \in K} x_{ij}^{rk} + e_{ij}^r = E_{ij}^r, r \in R$$

| | | | | | | |
|---------|-------------|-------------|-------------|-------------|-------------|-------------|
| Маршрут | 6 | | | | | |
| | Транспорт 1 | Транспорт 2 | Транспорт 3 | Транспорт 4 | Транспорт 5 | Транспорт 6 |
| Товар 1 | 0 | 0 | 1500 | 0 | 0 | 0 |
| Товар 2 | 7000 | 0 | 0 | 0 | 0 | 0 |
| Товар 3 | 0 | 0 | 500 | | 0 | 0 |

Рисунок 2 – Форма загрузки решения

Далее планируется реализация алгоритмов решения данной задачи с возможностью транзита товара через промежуточные пункты.

Литература:

1. *Панюков А.В., Пивоварова Ю.В.* Развитие транспортной логистики в челябинской области: проблемы и перспективы применения информационных технологий // Вестник Южно-Уральского государственного университета. Серия: Экономика и менеджмент. – 2017. – Т. 11, № 1. – С. 7-11.
2. *Раскин, Л.Г., Кириченко И.О.* Многоиндексные задачи линейного программирования – М.: Радио и связь, 1989. – 240 с.
3. *Панюков А.В., Телегин В.А.* Техника программной реализации потоковых алгоритмов // Вестник Южно-Уральского государственного университета. Серия: Математическое моделирование и программирование. – 2008. – № 27 (127). – С. 78-99.

ЭКСПЕРТНАЯ СИСТЕМА ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ ПРИ ТОРГОВЛЕ НА ФОНДОВЫХ РЫНКАХ

А. А. Капленко, А. В. Фаддеенков
Новосибирский Государственный Технический Университет
г. Новосибирск, gregorydoodle79@gmail.com

В данной работе описано и исследовано влияние коэффициентов Шарпа и Сортино на экспертную систему поддержки принятия решений при торговле на фондовых рынках. Разработаны советники, реализующие конкретные торговые стратегии. Проведены анализ эффективности торговых стратегий и исследования параметров рынка, определены их оптимальные значения.

In this work the influence of the Sharpe and Sortino ratios on the expert decision-making support system when trading on stock market is described and investigated. Advisers that implement specific trading strategies were developed. Analysis of the effectiveness of trading strategies and investigations of market parameters were conducted, their optimal values were determined.

Введение

В настоящее время при анализе эффективности финансовых инвестиций используются различные коэффициенты, отражающие соотношения доходности и риска анализируемого финансового актива. Данные подходы можно использовать и при анализе эффективности торговых стратегий на рынке Форекс.

Эффективность торговых стратегий оценивается с использованием различных коэффициентов. При этом чаще всего рассматривается эффективность в сравнении с некоторым «безрисковым активом» или «доходностью рынка». Данные понятия используются как базисы для сравнения. В нашем случае под доходностью безрискового актива можно понимать доходность банковского депозита или доходность по государственным ценным бумагам ГКО, ОФЗ.

Цель работы

Цель: провести анализ эффективности торговых стратегий. Выявить оптимальные параметры распределения баланса трейдера для торговых советников на основе показателей коэффициентов Шарпа и Сортино.

Задачи:

- проанализировать модель рынка;
- рассмотреть существующие торговые стратегии;
- разработать торговые советники;
- рассчитать коэффициенты Шарпа и Сортино для каждого разработанного советника;
- распределить денежные средства между советниками на основе показателей коэффициентов эффективности, изменяя внутренние параметры торговых стратегий.

Торговые стратегии

На рисунке 1 приведены результаты работы торговых советников. Было разработано 4 торговых стратегий, работающих на временном периоде 2017.01.01 – 2018.01.01 гг. с таймфреймом в Н1.

Средний уровень риска для каждой стратегии – 2-5%.

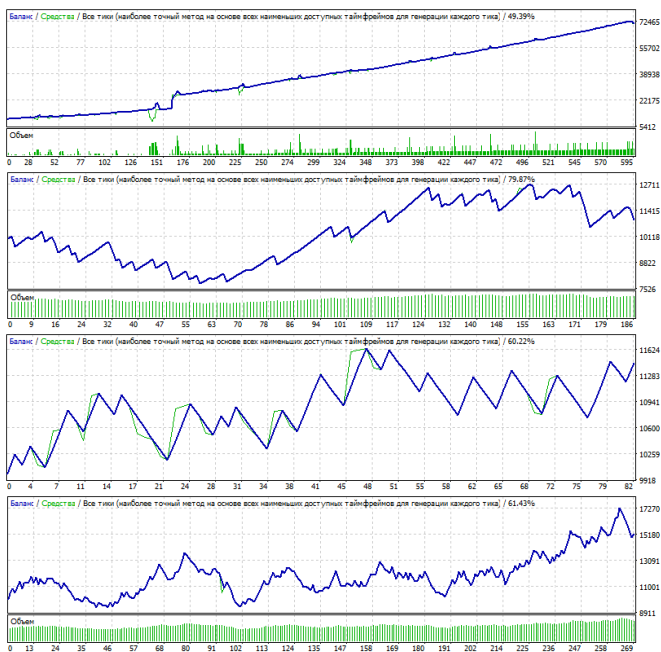


Рисунок 1 – Результаты работы советников

Коэффициент Шарпа

Коэффициент Шарпа показывает, какую доходность приносит актив на единицу риска. Чем больше значение коэффициента по рассматриваемому активу, тем больше получит инвестор за принятый на себя риск. Вычисляется коэффициент Шарпа на заданном временном горизонте инвестиций по следующей формуле:

$$SR = \frac{R_p - R_{FR}}{\sigma_p},$$

где R_p – доходность паттерна за анализируемый период, R_{FR} – доходность безрискового актива за анализируемый период, σ_p – стандартное отклонение доходности паттерна за анализируемый период.

Задача выбора паттерна p^* , оптимального с точки зрения коэффициента Шарпа, имеет следующий вид:

$$p^* = \underset{p \in P}{\text{Arg max}}(SR) .$$

Коэффициент Сортино

Коэффициент Сортино показывает доходность инструмента, взвешенную по риску. Чем выше этот показатель, тем инструмент более эффективен с точки зрения сочетания доходности и риска.

В расчетах вместо показателя стандартной дисперсии, которая описывает все возможные отклонения как в положительную, так и в отрицательную сторону, используется полудисперсия, учитывающая только отрицательную волатильность.

Расчет полудисперсии ведется по следующей формуле:

$$\sigma_p^- = \sqrt{\frac{\sum_{t=1}^T \min\{(R_p(t) - R_{FR}(t), 0)\}^2}{T}} .$$

На основе этого показателя коэффициент Сортино рассчитывается по формуле:

$$SOR = \frac{R_p - R_{FR}}{\sigma_p^-} ,$$

где R_p – доходность паттерна за анализируемый период, R_{FR} – доходность безрискового актива за анализируемый период.

Задача выбора паттерна p^* , оптимального с точки зрения коэффициента Сортино, имеет следующий вид:

$$p^* = \underset{p \in P}{\text{Arg max}}(SOR)$$

Распределение баланса трейдера в зависимости от показателей коэффициентов Шарпа и Сортино

В таблицах 1-2 приведены показатели коэффициентов Шарпа и Сортино для разработанных торговых советников.

Таблица 1 – Показания коэффициентов торговых советников с уровнем риска в 2%.

| Советники | Коэффициент Шарпа (SR) | Коэффициент Сортино (SOR) |
|-----------|------------------------|---------------------------|
| MA | 0,852 | 1,3809 |
| SMA + WPR | 0,425 | 0,3902 |
| RSI | 0,679 | 1,178 |

| | | |
|---------------|-------|-------|
| Parabolic SAR | 0,795 | 1,132 |
|---------------|-------|-------|

Таблица 2 – Показания коэффициентов торговых советников с уровнем риска в 5%.

| Советники | Коэффициент Шарпа (<i>SR</i>) | Коэффициент Сортино (<i>SOR</i>) |
|---------------|---------------------------------|------------------------------------|
| MA | 0,604 | 1,012 |
| SMA + WPR | 0,2201 | 0,117 |
| RSI | 0,526 | 0,8107 |
| Parabolic SAR | 0,725 | 0,896 |

С учетом полученных данных можно провести анализ эффективности той или иной торговой стратегии. Предположим, что изначальное распределение депозита [25%, 25%, 25%, 25%]. Исходя из определений коэффициентов эффективности, советник с максимальными показателями коэффициентов будет приносить больше доходов. Отсюда следует, что советник, созданный на основе торговой стратегии MA, будет получать ~50-70% от общего депозита, что в итоге принесёт трейдеру больше общей прибыли. Тогда распределение будет следующим: [64%, 5%, 13%, 18%].

Заключение

В ходе работы была предложена модель оптимизации распределения баланса трейдера в зависимости от показателей коэффициентов Шарпа и Сортино. Разработаны торговые стратегии.

В дальнейшем будет разработано ПО, которое сможет автоматизировать процесс перераспределения денежных средств согласно коэффициентам, что позволит избежать убытков. Программы для автоматизации распределения ресурсов между торговыми советниками на основе показаний коэффициентов Шарпа и Сортино нам не известны, поэтому дальнейшие разработки целесообразны.

Литература

1. Полтева Т. В. Оценка эффективности управления инвестиционным портфелем // Карельский научный журнал. – 2016. № 4 (17). – С. 195–198.

2. Расчет коэффициентов эффективности управления [Электронный ресурс]. Режим доступа: <http://www.beintrend.ru/2010-11-08-10-38-49> (дата обращения: 02.10.18).

ТЕХНОЛОГИЯ АЙТРЕКИНГА И ЕЕ ПРИМЕНЕНИЕ В МАРКЕТИНГЕ

А.Г. Климкина, Н.В. Федорова
Сибирский государственный университет науки и технологий
имени академика М.Ф. Решетнева
г. Красноярск, rukario2187@gmail.com

Дано определение айтрекинга, описаны области применения технологии и метод сбора статистической информации. Приведены реальные выводы и практики, полученные в результате применения айтрекинга. Определена польза и практичность использования описанной технологии в маркетинге.

The definition of is given, the fields of application of the technology and the method of collecting statistical information are described. The real conclusions and practices obtained as a result of the application of eye tracking are given. The benefits and practicality of using the described technology in marketing are determined.

Куда мы смотрим в первую очередь, открывая страницу в браузере? На что обращаем внимание, когда видим рекламу, прежде всего? Куда падает наш взор при выборе товара со стеллажа? Если опереться на значимые и достоверные статистические данные, то ответ будет прост. А помочь с этим может айтрекинг.

Айтрекинг — технология, позволяющая наблюдать и записывать движения глаз: расширение зрачка, его перемещение. [1] Она используется в разных дисциплинах: психология, когнитивные науки, медицина, маркетинг. [2] Она относится к массовому статистическому наблюдению и является как количественным, так и качественным методом сбора и анализа информации.

Особо важны «точки фиксации» — области, в которых взгляд пользователя задерживается на продолжительное время, чтобы обработать воспринимаемое изображение и «саккада» — процесс движения взгляда пользователя между точками фиксации. [1] Например, если глаз человека двигается слишком быстро, то пользователь не

воспринимает то, что видит, а воспроизведение саккад показывает, в какой последовательности человек считывал информацию.

Общий метод определения фокуса взгляда пользователя предполагает сравнение позиции ближнего инфракрасного света (отраженного глазом) с позицией зрачка. Эта информация в сочетании с информацией о положении головы наблюдаемого может быть использована для определения точки, на которой сфокусирован взгляд пользователя, посредством чего определяются соответствующие координаты на экране. [1]

Все полученные данные проектируются и группируются и, вследствие чего, их можно визуализировать и интерпретировать, чтобы сделать нужные выводы. Таким образом, определяется:

- то, что пользователь видит, и то, что он не видит;
- время достижения любой заданной точки фиксации (это может быть связано с тем, как легко или трудно было найти элемент);
- длина любой заданной фиксации (насколько привлекателен был элемент);
- количество фиксаций на элемент (насколько он был понятен).

[1]

С 1980 года айтрекинг начали применять в маркетинге для изучения поведения человека при просмотре рекламных объявлений. [3] Сегодня эта технология проникла во все сферы маркетинга и является одним из популярных методов исследования у ведущих маркетинговых и дизайнерских агентств. Она широко применяется при разработке дизайна упаковки, сайта, рекламных баннеров.

На основе айтрекинга сейчас строятся основные концепции веб-дизайна. Так, например, в ходе экспериментов было выявлено, что люди почти всегда просматривают текстовые материалы, следуя очертаниям английской буквы «F», и уделяют больше внимания информации, расположенной у левого края экрана. [4] Важно, что для языков, читаемых справа налево, результат прямо противоположный. Именно поэтому такие сайты, как «ВКонтакте», «Amazon», «Википедия» и другие размещают всю важную информацию в левом верхнем углу.

Помимо этого, были сделаны выводы о том, что любые видеоролики на странице поиска привлекают большее внимание, особенно, если находятся среди первых трех результатов. [4]

Интересен вывод о практичности использования визуальных сигналов для направления внимания пользователя в нужную область сай-

та. Люди подсознательно следуют взглядам других, а с самого рождения нас приучали идти по стрелкам, указывающим направление. Если направить взгляд человека на фотографии в нужный текст, то пользователь подсознательно будет смотреть туда же. [4] Таким образом, маркетолог может заставить пользователя обратить внимание на нужную ему (маркетологу) информацию.

Кроме того, в ходе исследований было открыто явление «баннерной слепоты». Пользователи сайтов неосознанно игнорируют рекламные блоки, а также объекты, похожие на них.

За все время существования айтрекинг дал много ценных находок веб-дизайнерам и маркетологам. Его применение позволяет понять, по каким причинам пользователи переходят с одной страницы на другую, как ведут себя на конкретном сайте. Использование айтрекинга является частью процесса оптимизации конверсии.

Таким образом, мы начинаем глубже и нагляднее понимать мотивацию человека и принцип его восприятия информации. Несмотря на то, что традиционные аналитические инструменты, такие как фиксация кликов и скроллов, по-прежнему являются эффективными, их сочетание с айтрекингом открывает огромные возможности для совершенствования юзабилити веб-интерфейсов.

Стоит помнить, что реальную пользу айтрекинг может принести только в том случае, если сбором и анализом данных будет заниматься специалист со знаниями и опытом в этой области. [1]

Айтрекинг – недешевая технология для сбора нужных статистических данных. Поэтому прежде чем применять данную технологию, маркетологу следует оценить, использует ли он весь потенциал менее дорогих методов. Если ответ будет положительным, то айтрекинг – это следующий естественный шаг. Если отрицательный, то использование таких недорогих методов, как наблюдение или интервью пользователей, позволит сделать хороший скачок вперед.

Литература

1. Айтрекинг: для чего он нужен и когда использовать? [Электронный ресурс]. URL: <https://lpgenerator.ru/blog/2016/04/12/ajtreking-dlya-chego-nuzhen-i-kogda-ispolzovat/> (дата обращения 05.11.2018);

2. Окулография [Электронный ресурс]. URL: <https://ru.wikipedia.org/wiki/%D0%9E%D0%BA%D1%83%D0%BB%D0%BE%D0%B3%D1%80%D0%B0%D1%84%D0%B8%D1%8F> (дата обращения 06.11.2018);

3. История айттрекинга: краткий обзор [Электронный ресурс]. URL: <https://lpgenerator.ru/blog/2015/06/29/istoriya-ajtrekinga-kratkij-obzor/> (дата обращения 05.11.2018);

4. 7 уроков из практики айттрекинга [Электронный ресурс]. URL: <https://lpgenerator.ru/blog/2013/06/30/7-urokov-iz-praktiki-ajtrekinga/> (дата обращения 05.11.2018).

ИССЛЕДОВАНИЕ СТАТИСТИЧЕСКИХ ХАРАКТЕРИСТИК ПОМЕХ ИНЕРЦИАЛЬНЫХ ДАТЧИКОВ НА МОБИЛЬНЫХ УСТРОЙСТВАХ

Д.А. Коржуева, В.С. Карманов

**Новосибирский Государственный Технический Университет,
г. Новосибирск, korzhueva.daria@gmail.com**

Одним из современных методов навигации является инерциальная навигация, основывающаяся на определении углов поворота, наклона объекта, ускорения и его численном интегрировании. От точности показаний такой навигации зависит работа многих систем, поэтому ставится задача по обнаружению и минимизации воздействия шумов на характеристики инерциальной навигационной системы.

One of the modern methods of navigation is inertial navigation, based on the definition of the angles of rotation, tilt of the object, acceleration and its numerical integration. The operation of many systems depends on the accuracy of the indications of such navigation; therefore, the task is to detect and minimize the effect of noise on the characteristics of the inertial navigation system.

С ростом научно-технического прогресса возникла необходимость в определении местоположения различных объектов реального мира, поэтому была разработана система, позволяющая решить такую задачу. Такая система называется спутниковая система навигации (GPS). Она позволяет определять географические координаты наземных, водных и воздушных объектов и широко применяется в различных государственных структурах.

Спутниковая система навигации несовершенна, например, когда объект находится в закрытом помещении, так как многочисленные препятствия будут значительно искажать сигнал или вовсе препятствовать его прохождению. По этой причине существует так называемая система внутреннего позиционирования (indoor positioning system, IPS), которая

не имеет проблему аналогичную GPS. Такая система включает в себя следующие технологии позиционирования: звук, инерциальную навигацию, инфракрасное излучение, камеру, тактильные системы, Bluetooth, Wi-Fi и другие.

Рассматривалась одна из технологий, называемая инерциальной навигацией. Инерциальная навигация (инерциальная навигационная система, ИНС) - один из современных методов навигации, основывающийся на показаниях таких датчиков как акселерометр и гироскоп, а также на численной обработке показаний этих датчиков. Преимущество такой навигации заключается в её полной автономности, что позволяет использовать её для объектов, находящихся в закрытых помещениях. ИНС можно использовать для определения местоположения в торговых комплексах, музеях, вокзалах, станциях метро, в медицинском обслуживании, подземном строительстве и других сферах. Однако, такой метод имеет и существенный недостаток – шум, искажающий показания датчиков ИНС. Для сокращения влияния шума используются различные методы фильтрации.

Цель работы – исследование статистических характеристик шума для минимизация его влияния на показания датчиков ИНС.

В качестве объекта для работы с датчиками было выбрано мобильное устройство, располагающее следующими сенсорами: акселерометр, гироскоп. Было написано приложение на языке Java с использованием Android-библиотек, записывающее показания сенсоров в файл csv. Показания сенсоров считываются с периодом $T = 200 \pm 0,001$ мс.



Рисунок 6. План здания (второй этаж первого корпуса НГТУ).

Для получения данных было пройдено несколько треков, на каждом из которых записывались показания датчиков устройства. Первый пройденный трек начинается от точки 1 до точки 2, второй трек следует в обратном направлении (рис. 1). В последующих исследованиях будут рассматриваться показания с этих двух треков.

Ускорение связано со скоростью, а скорость в свою очередь с расстоянием по формулам

$$v(t) = v_0 + \int_{t_0}^t a(\tau) d\tau, \quad s(t) = s_0 + \int_{t_0}^t v(\tau) d\tau \quad (1)$$

где t_0 – время начала движения, с;

t_1 – время окончания движения, с;

t – переменная времени, с; $t_0 \leq t \leq t_1$;

v_0 – начальная скорость, м/с;

s_0 – начальное расстояние, м.

Количество показаний, которое необходимо обработать, достаточно велико, поэтому было решено использовать для этих целей свободную

систему математических вычислений – GNU Octave. К тому же эта система имеет встроенные функции численного интегрирования.

Акселерометр выводит значения ускорения по трём осям x , y и z (рис. 2), поэтому зависимость расстояния от ускорения рассматривались для каждой из этих осей. Гироскоп показывает угловую скорость вращения относительно трёх осей (рис. 3).

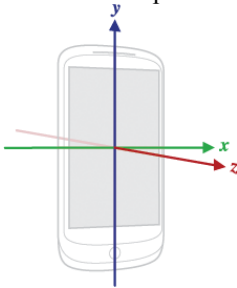


Рисунок 7. Оси, определяющие положение устройства

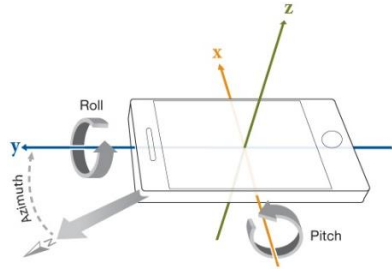


Рисунок 8. Определение поворота вокруг осей

В ходе численного интегрирования ускорений по осям x , y и z было выявлено, что графики пути будут иметь примерно одинаковый вид, поэтому имело смысл взять для рассмотрения показания по одной из осей. При замере трека телефон находился в плоскости Oxy и перемещался вдоль оси y , по этой причине была выбрана ось y .

Как показали графики (рис. 4-5), траектория движения, полученная экспериментально, далека от фактической прямолинейной траектории, также расстояние, на котором могут быть получены показания датчиков, достаточно велико, а значит точность таких данных крайне низкая.

Также была устранена линейная компонента, вносящая вклад в поведение графиков, однако, траектория по каждому из замеров получилась различной, что уже говорит о том, что помимо линейной компоненты имеются ещё некоторые факторы, искажающие данные, полученные датчиками ИНС.

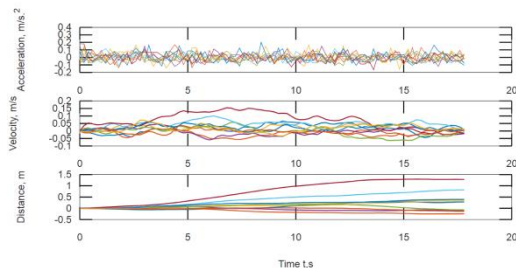


Рисунок 9. Устройство находилось в неподвижном положении

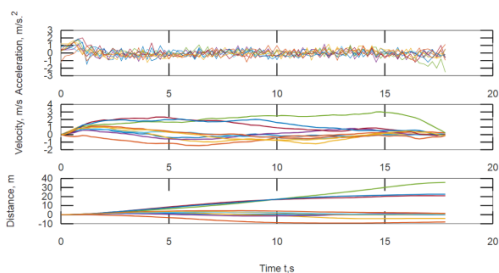


Рисунок 10. Устройство перемещалось по первому треку

Выводы – при использовании ИНС имеется проблема, влияющая на точность измерения положения тела, называемая шумом приборов. Показания такой системы нуждаются в фильтрации шума одним или несколькими из наиболее эффективных методов.

Литература

1. Кузовков Н. Т., Салычев О. Инерциальная навигация и оптимальная фильтрация — 1982.
2. Mautz R. Indoor positioning technologies [Электронный ресурс] / ETH Zurich. – Germany, 2012. – URL: <https://www.research-collection.ethz.ch/bitstream/handle/20.500.11850/54888/eth-5659-01.pdf>
3. Shala U., Rodriguez A. Indoor Positioning using Sensor-fusion in Android Devices [Электронный ресурс] / Kristianstad University, Department Computer Science. — Sweden, 2011. — URL: <https://www.diva-portal.org/smash/get/diva2:475619/FULLTEXT02.pdf>
4. Бутаков Н.А., Намиот Д.Е. Инерциальная навигация мобильных устройств [Электронный ресурс] / Московский государственный

университет имени М.В. Ломоносова. – Москва, 2014. – URL:
http://linkstore.ru/articles/inertial_navigation.pdf

ПОСТРОЕНИЕ РЕГРЕССИОННЫХ МОДЕЛЕЙ МЕТОДОМ УСЕЧЕННОГО МАКСИМАЛЬНОГО ПРАВДОПОДОБИЯ

К.А. Молочная, А.В. Фаддеев

**Новосибирский государственный технический университет,
г. Новосибирск, afama051@gmail.com**

Объектом исследования является метод усеченного максимального правдоподобия.

Цель работы – разработка и реализация алгоритма оценивания параметров регрессионной модели с использованием MTL-метода.

В процессе работы проводились экспериментальные исследования метода усеченного максимального правдоподобия в сравнении с методом наименьших квадратов.

В результате исследования было установлено, что метод усеченного максимального правдоподобия работает точнее на данных с засорениями, чем метод наименьших квадратов.

The object of study is the method of truncated maximum likelihood.

The aim of the work is to develop and implement an algorithm for estimating the parameters of a regression model using the MTL method.

In the course of the work, experimental studies of the truncated maximum likelihood method were carried out in comparison with the least squares method.

As a result of the study, it was found that the truncated maximum likelihood method works more accurately on clogged data than the least squares method.

Регрессионный анализ – это статистический метод исследования того, каким образом влияют независимые переменные на зависимую переменную.

Задачами регрессионного анализа являются установление формы зависимости между переменными, оценка функции регрессии, оценка неизвестных значений (прогноз значений) зависимой переменной.

Наиболее известный метод регрессионного анализа – метод наименьших квадратов или сокращенно МНК. Данный метод по результатам наблюдений позволяет найти оценку параметров регрессионного уравнения. Но возможен и другой подход к решению данной задачи. В случае линейного регрессионного уравнения с ошибкой регрессии, рас-

пределенной по нормальному закону с нулевым математическим ожиданием, можно производить оценку методом максимального правдоподобия (ММП). При таких условиях результаты работы этих двух методов совпадут.

Однако у ММП есть свои несовершенства. Для получения статистически верных результатов необходимо, чтобы выполнялся ряд условий, например, отсутствие свойства несмещенности оценок или ограничения на свойства входных данных.

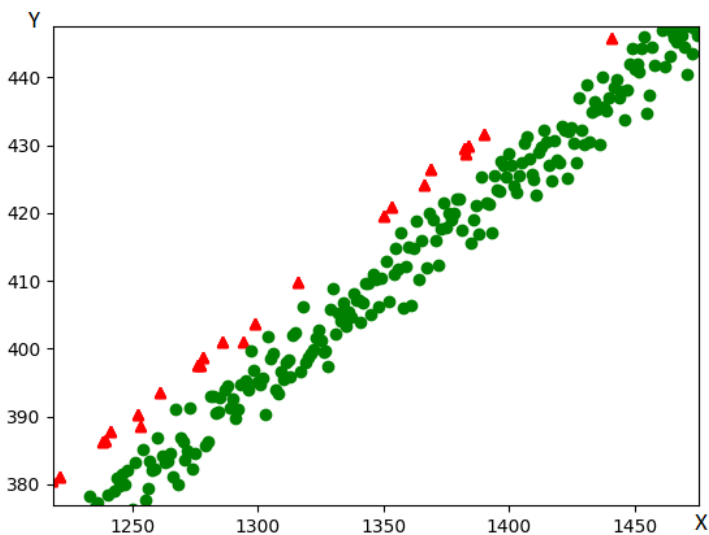
Как известно, в действительности, практически невозможно достичь выполнения всех условий и получить идеальную ситуацию. Одним из вариантов места ошибки может быть значение независимых переменных. Они могут, подобно зависимым переменным содержать разного рода ошибки, которые наделяют исходные данные свойством неоднозначности. Так же при непосредственном проведении наблюдений есть вероятность изменения условий проведения наблюдений, что может привести к неоднородности данных или даже к получению резко выделяющихся из основной массы наблюдений, иначе называемых выбросами.

В эконометрике определение наличия выбросов является отдельной задачей, для которой используют специальные, устойчивые к выбросам методы оценивания параметров. Одним из таких является метод усеченного максимального правдоподобия (МТЛ-метод). Данный метод менее чувствителен к выбросам, чем ММП. В случае, если ошибки наблюдений принадлежат нескольким нормальным распределениям с разными параметрами, а не к одному распределению, то этот метод не сообразится, а также точно оценит параметры.

Цель данной работы – разработка алгоритма оценивания параметров регрессионной модели с использованием МТЛ-метода.

Суть данного метода состоит в том, что параметры оцениваются по наблюдениям, которые вносят наибольший вклад в функцию правдоподобия. Т.е. некоторое количество наблюдений с минимальным вкладом удаляются, считаясь лишними, а оценка проводится по оставшимся (усеченным) наблюдениям.

Ниже представлен график, точки которого – это часть наблюдений. Треугольниками отмечены те, что согласно методу, вносят наименьший вклад или, иначе говоря, являются выбросами. Точки – остатки выборки, по которому в дальнейшем будут оцениваться параметры уравнения регрессии.



Литература:

1. Фаддеенков, А.В. Оценивание параметров регрессионных моделей методом усеченного максимального правдоподобия / А.В. Фаддеенков, Е.А. Хайленко // Научный вестник НГТУ. – Новосибирск, 2016. – том 65, №4. – С. 135-145.
2. Тимофеев, В.С. Эконометрика: учебник / В.С. Тимофеев, А.В. Фаддеенков, В.Ю. Щеголкин. – Новосибирск : Изд-во НГТУ, 2015. – 346 с.

ПОИСК КОНФЛИКТОВ МЕЖДУ ТРЕБОВАНИЯМИ К ПРОГРАММНОМУ ПРОДУКТУ

М.Ш. Мургазина, Т.В. Авдеенко
Новосибирский государственный технический университет,
г.Новосибирск, murtazina@corp.nstu.ru

В докладе рассматриваются вопросы обнаружения конфликтов между требованиями к программному продукту. Предлагается подход к формированию модели поиска конфликтов для онтологии требований программного продукта.

The report discusses the detection of conflicts between software requirements. An approach to the construction of a conflict search model for the software product requirements ontology is proposed.

Выявление взаимосвязей требований является одной из задач анализа требований. В центре внимания многих исследований по оценке взаимосвязей требований находится выявление конфликтов между парами требований [1].

В рамках данной работы предлагается подход к поиску конфликтов между требованиями. В каждом требовании выделяется три части: актор (Actor), действие (Action) и объект (Object), на который направлено действие. Например, дано два требования:

1. Пользователь может редактировать свои комментарии на стене.
2. Пользователь не может редактировать чужие комментарии на стене.

Каждое требование представляется в онтологии в виде экземпляров классов. В каждом требовании выделяются экземпляры классов и отношения между экземплярами классов актор, действие и объект (Рисунок 1). Это может быть сделано вручную или в полуавтоматическом режиме при помощи методов автоматической обработки текста.

Все требования могут быть разделены на две группы: требования, которые пересекаются по акторам, действиям и/или объектам, и требования, не пересекающиеся ни по одному из этих элементов. При этом пересекающиеся требования могут содержать конфликты.

Для определения типа отношения между требованиями может быть составлена система правил вида:

- Если Тип_отношения (Актор1, Актор2)
- И Тип_отношения (Действие1, Действие2)
- И Тип_отношения (Объект1, Объект2)

Тогда Тип_отношения_между_требованиями $\in \{ \text{Конфликтуют}, \text{Пересекаются}, \text{Не_пересекаются} \}$

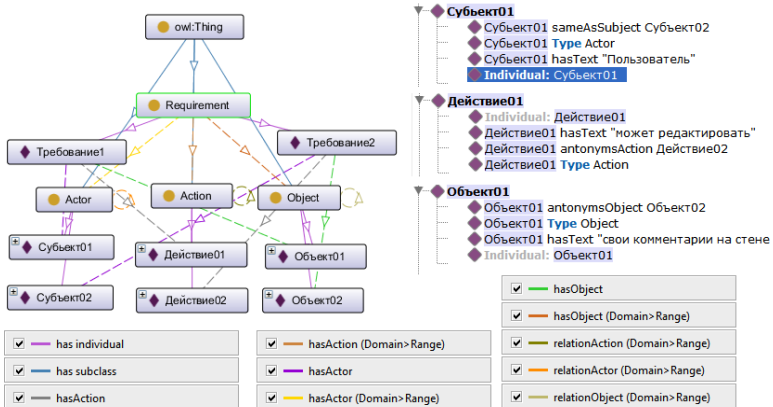


Рисунок 1 – Фрагмент онтологии

В онтологии заранее определяются возможные отношения между акторами, действиями и объектами пары требований. На рисунке 2 представлены свойства объектов анализируемой онтологии.

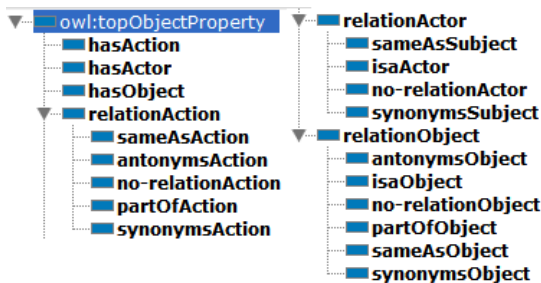


Рисунок 2 – Свойства объектов онтологии

Для реализации предлагаемого подхода был разработан конструктор модели на языке python. Данное приложение имеет графический интерфейс и позволяет пользователю в несколько «кликов» построить систему правил для анализа отношений между элементами требований. На первом шаге пользователь выбирает файл OWL онтологии, для которой

строится модель и место, куда модель будет записана. На втором – устанавливает соответствие классов онтологии элементам требований. На третьем – выбирает, какие свойства объектов должны быть включены в анализ. На последнем шаге для сгенерированных правил пользователь выбирает тип отношения между требованиями. На рисунке 3 представлен последний шаг построения модели. Приведенная выше пара требований будет обработана правилом:

IF sameAsSubject(Actor1, Actor2) and antonymsAction(Action1, Action2) and antonymsObject(Object1, Object2) Then Interaction(Requirement1, Requirement2).

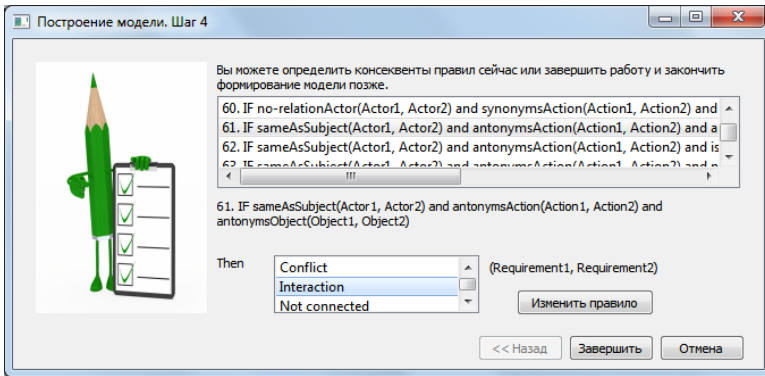


Рисунок 3- Экранная форма конструктора модели

Для пары требований из примера устанавливается тип отношения «Пересекаются», так как оба требования относятся к одному и тому же актору. Конфликтными, например, будут являться те требования, где для одного и того же актора назначены противоположные действия с одним и тем же объектом. Действия и объекты сопоставляются между собой при помощи лингвистической онтологии. В общем случае может быть использован WordNet.

По построенной модели далее производится сопоставление требований, внесенных в онтологию. Для реализации анализа данные из онтологии переводятся в факты на языке Prolog, а правила модели в правила на языке Prolog. Для реализации программного продукта использованы python-библиотека ruSwiP и среда разработки SWI-Prolog 7.6.4. Для работы с онтологией использована среда Protégé 5.2.

Литература:

1. Salado A., Nilchiani R. The Concept of Order of Conflict in Requirements Engineering // IEEE Systems Journal. - 2014. - DOI: 10.1109/JSYST.2014.2315597.

ВЫЯВЛЕНИЕ АНОМАЛЬНЫХ НАБЛЮДЕНИЙ В ДАННЫХ БОЛЬШИХ РАЗМЕРНОСТЕЙ НА ОСНОВЕ ГЕОМЕТРИЧЕСКОГО АВОД-ПОДХОДА

Н.С. Олейник, В.Ю. Щеколдин
Новосибирский государственный технический университет,
г. Новосибирск, nikitaolejnik@mail.ru

В работе рассматривается задача выявления наблюдений в многомерных выборках, которые существенно отличаются от других. Рассматриваются метод решения задачи на основе анализа расстояния между наблюдениями и метод, основанный на нахождении максимальных углов между наблюдениями, представленными точками многомерного пространства. Разработана алгоритмическая реализация этих методов с использованием критерия Шовене, метода максимального относительного отклонения, а также использования аппарата кумулятивных кривых.

The paper deals with the task of identifying observations from a large sample that would be different from others. Method of finding emissions, based on the distance between observations and the method based on finding the maximum angles between observations are analyzed. The proposed implementation of these methods using the Chauvene criterion, the method of maximum relative deviation, as well as the use of cumulative curves is realized.

1. Введение

В современном мире во многих сферах деятельности приходится работать с большими объемами данных, которые необходимо собирать, обрабатывать, анализировать, получать и интерпретировать результаты, а также правильно использовать их, в частности, разрабатывать конкретные полезные для практики рекомендации по контролю и управлению изучаемых процессов.

На этапе предварительной обработки данных часто возникает проблема выявления специфических элементов в выборке, называемых выбросами. Одно из наиболее популярных определений понятия «выброс» или «аномального наблюдения» было предложено в [2] Д. Хоукин-

сом (1980), который говорил, что «выброс – это наблюдение, которое так сильно отличается от других наблюдений, что может возникнуть предположение, что оно появилось в выборке принципиально другим способом». Появление выбросов может быть вызвано как ошибками при получении данных, так и определяться особенностями изучаемой генеральной совокупности, что приводит к изменениям в структуре соответствующих распределений.

Как правило, при обработке данных выбросы исключаются из выборки, что может позволить получить статистически более корректные результаты. Однако с практической точки зрения выброс может не являться «ошибкой», а наоборот, указывать на какую-то важную особенность анализируемого явления или процесса, что приводит к необходимости разработки так называемых робастных подходов.

Другая проблема состоит в множественности как самих данных, так и их размерности. При этом существует понятие «проклятие размерности»: ситуация, когда из-за увеличения размерности пространства экспоненциально возрастает объём данных, что влечёт за собой существенное повышение вычислительной сложности применяемых алгоритмов, а также сопровождается необходимостью хранения существенных объёмов данных.

2. Постановка задачи

Пусть имеется база данных $D \in R^d$, где d – размерность пространства наблюдений. Для дальнейшей обработки используем скалярное произведение $v: R^d \times R^d \rightarrow R$ и связанную с ним норму $\mu: R^d \rightarrow R_0^+$. Необходимо проанализировать базу данных на предмет наличия в ней аномальных наблюдений в соответствии с индуцированной нормой. При больших значениях величины d применение стандартных одномерных методов выявления выбросов оказывается, как правило, малоэффективно, поскольку они не позволяют учитывать геометрию анализируемого пространства. Поэтому за последнее десятилетие широкое распространение получили подходы, основанные на использовании геометрических свойств взаимного расположения элементов выборки в пространстве R^d . Согласно [1, 3] наиболее популярными методами такого типа являются DiBOD, DeBOD, ABOD.

3. Анализ существующих методов

DiBOD – Distance-based outlier detection: метод заключается в том, что необходимо определить число элементов выборки, находящихся в ε -окрестности каждого наблюдения. При этом выбросами признаются

те наблюдения, для которых число «соседних» элементов выборки существенно меньше среднего их числа. К преимуществам метода DiBOD относятся относительно небольшое число вычислительных операций и простота реализации алгоритма. В тоже время метод слабо инвариантен к многомерным задачам (начиная с $d = 3$), что сокращает возможности его применения.

DeBOD – Depth-based outlier detection: метод базируется на понятии «глубины пространства», введенном П. Рузьё (P. Rousseeuw) [4]. Выбросами в этом случае будут являться точки, которые соответствуют наименьшей «глубине». Данный метод способен обрабатывать большой объём данных за короткое время, но он также не является инвариантным для многомерных задач и в определенных ситуациях может существенно «заблуждаться» относительно правильного выделения аномальных наблюдений.

ABOD – Angle-based outlier detection: метод заключается в том, что для каждого наблюдения необходимо определить угол в многомерном пространстве, под которым из данной точки видны все остальные наблюдения выборки. Выбросами в этом случае будут являться точки пространства, которым соответствуют наименьшие значения таких углов. Для определения значений углов, начиная с которых можно считать наблюдение выбросом, можно использовать аппарат кумулятивных кривых, изложенный в [5].

Наиболее эффективен подход ABOD для решения задач большой размерности, при этом, как показано в [3], эффективность существенно снижается при уменьшении величины d . Поскольку предполагается решение практических задач с существенно большим значением d , то в качестве базового было решено рассматривать именно метод ABOD.

4. Разработка структуры тестов

Схема I. Исходные данные (в пространстве с $d = 10$, объем исходной выборки меняется от $N = 100$ до $N = 1000$) представляются в виде эллипсоида фиксированного радиуса ρ , в котором равномерно размещается 90% наблюдений, а оставшиеся 10% располагаются в кольце с радиусами от 1.5ρ до 2ρ . Таким образом, в результате работы алгоритма должно выявляться в среднем 10% выбросов.

Схема II. Исходные данные представляются в виде многомерного эллипсоида, у которого каждая полуось, начиная от главной, больше предыдущей в отношении 3:2. В нем равномерно размещается 90%

наблюдений, а оставшиеся 10% располагаются в области между двумя эллипсоидами с полуосями $1.5\rho_i / 2\rho_{i+1}$ и $2\rho_i / 2.5\rho_{i+1}$, где ρ_i – соответствующие полуоси исходного эллипсоида, нумерованные, начиная от главной. В результате работы программы также ожидается, что методы будут выявлять в среднем 10% выбросов.

Для обеспечения статистической устойчивости результатов моделирования проводится многократное воспроизведение алгоритма с числом повторений $n=1000$. Полученные результаты изучаются на предмет возможности описания каким-либо законом распределения, а затем усредняются.

Выводы.

В работе рассмотрены три метода, позволяющие проводить выявление выбросов в многомерных выборках. В качестве базового предлагается выбрать ABOD как самый подходящий для решения существенно многомерных задач. Рассмотренные алгоритмы, а также выбранные для сравнения методы одномерного поиска выбросов такие как критерий Шовене и максимального относительного отклонения, реализованы в виде программного комплекса на языке C++.

При разработке программного обеспечения были решены разнообразные оптимизационные проблемы, связанные с нахождением эффективных способов оценивания параметров для нелинейных моделей при решении задачи определения числа выбросов с привлечением аппарата кумулятивных кривых. Планируется развитие программной системы для обнаружения аномальных наблюдений на платформе Python/C#. Эффективность работы изложенных методов будет протестирована на разработанной схеме экспериментов для данных в многомерном пространстве. Полученные результаты будут использоваться для разработки рекомендаций по применению данных методов при решении практических задач.

Литература.

1. Angiulli F., Pizzuti C. Fast outlier detection in high dimensional spaces. // In Tech. Report, n. 25, ISI-CNR, 2002. – pp. 15-27.
2. Hawkins D. Identification of Outliers. – Chapman and Hall, 1980.
3. Kriegel H., Schubert M., Zimek A. Angle-based outlier detection in high-dimensional data. // In Proc. of the 14th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD'08), Las Vegas, NV, 2008. – pp. 444-452.

4. Rousseeuw P., Aelst S., Driessen K., Agullo J. Robust multivariate regression. – *Technometrics*, Vol. 46, 2004. – pp. 293-305.
5. Shehekoldin V. Developing the risk classification based on ABC-analysis of possible damage and its probability // *International Forum: Proceedings of 11th International Forum on Strategic Technology (IFOST-2016)*. – Novosibirsk, 2016. – pp. 317-319.

СТАТИСТИЧЕСКОЕ ИССЛЕДОВАНИЕ РЕГИОНАЛЬНЫХ ОСОБЕННОСТЕЙ ПРЕСТУПНОСТИ В РОССИЙСКОЙ ФЕДЕРАЦИИ

А.В. Попкова, Е.В. Макаридина
Новосибирский государственный университет
экономики и управления «НИНХ»,
г. Новосибирск, anaporkova57564@gmail.com

В работе рассмотрены особенности преступности в России с помощью статистических методов анализа. В результате осуществлена типологизация субъектов РФ по уровню преступности.

The paper discusses the characteristics of crime in Russia using statistical methods of analysis. As a result, a typology of the subjects of the Russian Federation in terms of crime was made.

Предметом исследования правовой статистики является количественная сторона тех явлений, которые входят в сферу деятельности прокуратуры, МВД, ФСБ, ФНС, судов, исправительно-трудовых учреждений и других правоохранительных органов. Следует учитывать, что объектом анализа становятся не все совершённые преступления, а лишь те из них, которые были обнаружены, и по поводу которых велось расследование и уголовное судопроизводство, либо применены меры административного воздействия. Следует учитывать эту особенность при анализе региональных различий преступности, так как наблюдается разная степень латентности преступности в отдельных субъектах [2]. Также существует проблема непопулярности использования статистических методов анализа в криминологии «потому что у нас нет нового поколения криминологов. То есть, по технологиям анализа мы остались на месте 1989 года» [1].

Наиболее разработанной, как в организационном, так и в научном отношении частью является уголовно-правовая статистика, изучающая преступность, жертв преступлений и реакции со стороны уголовного законодательства [5]. Статистические методы позволяют изучить преступность в самых разных направлениях: охарактеризовать динамику объема, структуру по различным признакам (видам преступлений, территории, по степени общественной опасности и т.д.), зависимость преступности от важнейших социально-экономических и политических факторов, выявить причинно-следственные связи, тенденции и закономерности развития, оценить эффективность работы правоохранительных органов и степень влияния принимаемых нормативно-правовых актов и т.д.

При характеристике преступности в РФ, важно принимать во внимание территориальные масштабы. Практически в каждом субъекте сложились собственные специфические экономические, социально-культурными этно-демографические условия. Соответственно, региональная преступность также имеет свои особенности.

На основании данных Росстата проведен анализ структуры преступности на период 2017 года в разрезе федеральных округов по различным критериям (рис. 1-2).



Рисунок 1 – Структура числа зарегистрированных преступлений в РФ по федеральным округам

Почти четвертая часть от всего объема преступлений приходится на ЦФО, чуть менее пятой части зарегистрировано в СФО; в то время как минимальный вклад у СКФО: в нем преступлений совершено почти в 7 раз меньше, чем в ЦФО и в 5 раз меньше, чем в СФО – 3,4%. Помимо общего количества преступлений криминогенную обстановку в регионе необходимо рассматривать в привязке к раскрытию преступлений и выявлению лиц, преступления совершивших. В РФ на каждую тысячу пре-

ступлений выявлено 470 преступников, в ЦФО - 380, что является минимальным значением данного показателя, а в СКФО значение величины было наибольшим – 540 выявленных лиц, совершивших преступление, на каждую тысячу преступлений. Отметим, что доля преступлений, совершенных группой лиц, в общем количестве предварительно расследованных преступлений, в данном округе наибольшая – 0,48%, что, вероятно, частично обуславливает число выявленных преступников [3].

Анализ структуры преступлений по видам показывает, что максимальную долю среди рассмотренных видов преступлений имеют кражи (69%). В Сибирском ФО наблюдается наибольшая доля краж относительно других видов преступлений, по сравнению с этим же показателем в рассмотренных округах, имеющая значение, равное 0,726. Минимальную среди рассмотренных видов преступлений имели изнасилование и покушения на изнасилование - 0,003 в РФ.

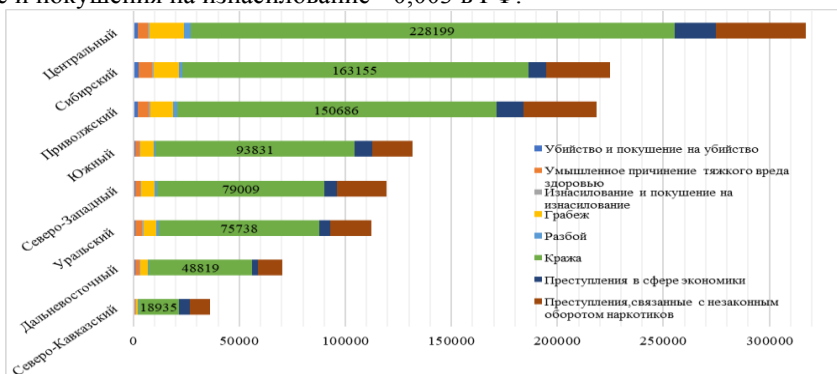


Рисунок 2 – Структура зарегистрированных преступлений по видам

Наиболее часто для характеристики преступности в России используют общий коэффициент преступности, показывающим количество преступлений на 100 тысяч человек населения по региону. Типологическая группировка федеральных округов по данному показателю позволяет сделать следующие выводы: единственным округом с низким уровнем преступности, как и предполагалось по уже проведенному анализу, является СКФО. В группу со средним уровнем преступности вошли: Центральный, Северо-Западный, Южный и Приволжский ФО; высокий уровень преступности наблюдался в Уральском, Сибирском и Дальневосточном округах. При проведении аналогичной группировки по субъектам РФ, отметим, что высокий уровень преступности отмеча-

ется в Республике Бурятия, Республике Тыва, Забайкальском крае и Сахалинской области. низкий – в Московской области, г. Санкт – Петербурге, Республике Саха (Якутия), и других, в частности, НСО имеет средний уровень преступности.

Группировка по субъектам соотносится с результатами группировки по округам: все регионы, относящиеся к СКФО попали в группу с низким уровнем преступности, а, например, Дальневосточный округ, как и Уральский, таких регионов вообще не содержит.

В качестве интегрального показателя, характеризующего общую криминогенную обстановку в регионах, можем использовать многомерную среднюю. В интегратор включим показатели: количество нераскрытых преступлений в отчетном периоде, количество преступлений, совершенных в состоянии алкогольного опьянения, количество преступлений, совершенных несовершеннолетними и при их соучастии, количество преступлений, совершенных лицами, ранее совершавшими преступления [4]. После нормирования этих показателей и типологизации округов согласно уровню преступности, получаем, что СКФО снова попадает в первый интервал, в который включаются округа, имеющие «лучшую» криминогенную обстановку, однако, ДФО, который по результатам предыдущей группировки являлся округом с высоким уровнем преступности, также попал в данную группу. Это объясняется тем, что абсолютные значения включенных показателей у ДФО являются наименьшими после СКФО, и высокий коэффициент преступности в ДФО не сопровождается высокими уровнями рецидивной преступности, преступности несовершеннолетних, низкой раскрываемостью преступлений. В итоге наихудшая криминогенная обстановка наблюдается в ПФО, СФО и ЦФО.

Таким образом, применение статистических методов при изучении преступности позволяет провести сравнительный анализ региональной преступности, детально рассмотреть особенности данного явления по каждому отдельному субъекту.

Литература

1. *Кузнецов Д.* Преступность в России: что мы в действительности о ней знаем?.2016. URL: https://life.ru/t/найка/910726/priestupnost_v_rossii_chno_my_v_dieistvitielnosti_o_niei_znaiem (дата обращения: 25.10.2018)

2. *Мирзоев А.А.* Методологические рекомендации и анализ недостатков использования статистики в изучении уголовно-правовых дисциплин // Современный ученый. – 2017. - №3. – С. 237-240
3. Официальный сайт Федеральной службы государственной статистики URL: <http://www.gks.ru> (дата обращения: 23.10.2018)
4. Официальный сайт Информационно-аналитический портал правовой статистики Генеральной прокуратуры Российской Федерации URL: <http://crimestat.ru> (дата обращения: 28.10.2018)
5. *Тихонова Е.А., Шиперова Е.С., Терехов А.М.* Данные правовой статистики в уголовно-правовой сфере как индикатор эффективности управления правовой системой в Российской Федерации // *Фундаментальные и прикладные исследования: от теории к практике: материалы II международной научно-практической конференции, приуроченной ко Дню российской науки.* – Воронеж: Изд-во «АМиСта», 2018. – С.24-29.

РАЗРАБОТКА ПРОГРАММНОГО КОМПЛЕКСА ДЛЯ ЭМУЛЯЦИИ ЛОГИСТИЧЕСКОЙ СИСТЕМЫ И ИССЛЕДОВАНИЯ ЕЁ ХАРАКТЕРИСТИК

Б.С. Соловьев, В.С. Карманов, В.Ю. Щеколдин
Новосибирский Государственный Технический Университет,
г. Новосибирск

Для исследования характеристик логистической системы, разработан программный комплекс, эмулирующий сеть поставок. Изучены алгоритмы оценки объёма заказов. Проведены исследования эффективности информационного взаимодействия между участниками сети поставок в различных вариантах топологий.

To study the characteristics of the logistics system, a software package, that emulates a supply chain, has been developed. The algorithms for estimating the volume of orders have been researched. A study of the effectiveness of information interaction between the participants of the supply chain in different variants of topologies has been conducted.

Введение

На кафедре ТПИ НГТУ разработан программный комплекс, позволяющий моделировать торговые взаимодействия в сетях поставок различной топологии, от линейной цепи до деревьев разных форм. Каждый элемент сети поставок использует один из алгоритмов оценивания зака-

за, по которому он определяет объём заказа для своего поставщика. Программный комплекс предназначен для исследования эффективности используемых алгоритмов, топологии сети на конкретных данных спроса или на случайно сгенерированном спросе. На результатах моделирования сети поставок также можно увидеть "эффект хлыста". "Эффект хлыста" (эффект кнута, эффект бычьего хвоста, эффект производности спроса, эффект Форрестера; англ. bullwhip effect, demand amplification, whip-saw, whiplash effect, Forrester effect) – феномен в цепях поставок, который заключается в усилении амплитуды колебаний спроса (объема заказов) по мере удаления от реального источника спроса в цепи поставок [1, 2]. Эффект хлыста имеет крайне негативное воздействие на эффективность управления цепями поставок, что приводит к чрезмерному увеличению страховых запасов, ненужному увеличению производственных затрат и накладных расходов, потенциальному ухудшению качества товара, и, что еще хуже, ухудшению качества обслуживания клиентов и упущенным продажам, увеличению расходов на логистику и т.д. После окончания моделирования программный комплекс выдаёт статистику в текстовом и графическом виде (графики, гистограммы).

Цель работы

Разработать инструмент моделирования сети поставок, для оценки эффективности алгоритмов формирования заказов, определения оптимальных способов взаимодействия участников сети.

Постановка задачи

Имеется сеть поставок, состоящая из нескольких участников. Участники являются элементами топологии сети (дерево как частный случай). На первом уровне находится производитель, на втором - дистрибьютор, на третьем - оптовый торговец и на четвёртом розничный торговец. Участники взаимодействуют друг с другом каждую единицу времени (недели). Спрос для розничных торговцев моделируется установленным распределением с параметрами. У оптовых поставщиков спрос задаётся розничными торговцами. Соответственно у дистрибьюторов - оптовые поставщики, а у производителя - дистрибьюторы. Производитель решает, сколько продукции произвести.

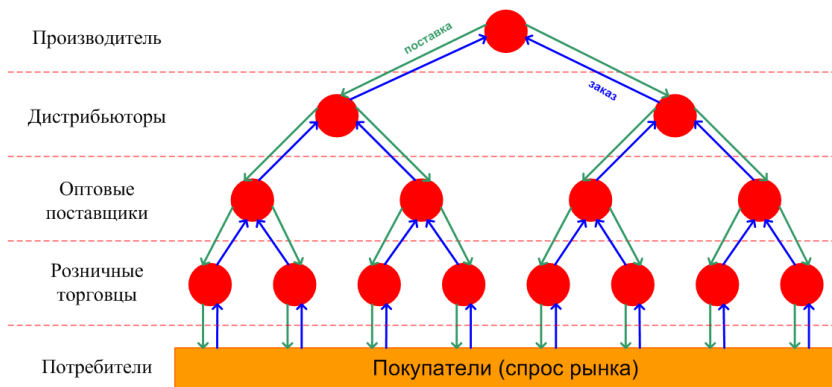


Рисунок 1 - Пример топологии

На рисунке 1 показан пример топологии торговой сети в виде бинарного дерева.

Каждый участник выполняет следующие шаги:

- принять заказы от своих клиентов (с нижнего уровня);
- получить товар от своего поставщика (с верхнего уровня);
- отправить товар своему клиенту дальше по цепочке (на нижний уровень);
- сделать новый заказ своему поставщику (на верхний уровень);
- заплатить за хранение товара.

Выбор объема заказа в каждом раунде является единственным решением, которое принимают участники сети поставок. За остатки на складе участник платит за хранение. За неудовлетворение спроса участник платит, за каждую единицу товара для покрытия спроса, повышенную цену.

Исходные данные:

- d – количество периодов, данные которых используются для вычисления заказов;
- l – срок поставки, недели;
- t – цена хранения одной единицы товара на складе, денежные единицы;
- ec – наценка при перекупке, %.

Целевая функция:

$$\bar{m} = \frac{1}{weeks \cdot 4} \left[\sum_{i=1}^{weeks} (m_i^p + m_i^d + m_i^w + m_i^r) \right] \rightarrow max,$$

где $weeks$ – количество периодов наблюдения (моделирования) сети, $m_i^p, m_i^d, m_i^w, m_i^r$ – состояния капиталов участников сети.

Уравнение управления сетью поставок:

$$s_i^p = s_{i-1}^p + u_i^p - x_i^d,$$

где s – количество продукции на складе (складской остаток), u – объём поступившего товара, x – объём заказанного товара. Описать это можно так: уровень запаса на конец периода i равен сумме уровня запаса на конец предшествующего периода ($i - 1$) и объема поступлений p_i за вычетом отгруженных товаров x_i за период i [3].

Требуется:

Подобрать алгоритмы и стратегии, чтобы минимизировать затраты в сети поставок и, таким образом, получить максимальную прибыль для всех участников сети.

Описание алгоритмов и результаты

Были разобраны несколько алгоритмов, по которым участники сети делают закупки. Все они зависят от «истории закупок», длина которой d является параметром. Использовались такие алгоритмы как алгоритм последнего заказа, скользящее арифметическое среднее порядка d , взвешенное скользящее среднее порядка d на основе арифметической и геометрической прогрессии, балансная оценка на основе анализа предыдущих заказов, запасов и поставок.

На рисунке 2 приведён пример работы прототипа [4] на выборках из нескольких. Красным цветом обозначен производитель, синим дистрибьютор, зелёным – оптовый поставщик, жёлтым – розничный продавец, фиолетовым – средние значения по сети. Сплошной линией изображается усреднённое значение изучаемого случайного процесса, пунктирной – экстремальные значения, штрихпунктирной – размеры доверительного интервала вычисляемого через одно значение среднеквадратического отклонения.



Рисунок 2 - Результаты моделирования на нескольких реализациях

Заключение

В ходе работы было проведено моделирование рынка, рассмотрены алгоритмы оценивания заказов. Разработан программный комплекс, моделирующий торговые взаимодействия согласно выявленным алгоритмам. Данное программное обеспечение может быть использовано для обоснования выбора стратегии при управлении сетями поставок в распределительной логистике с целью минимизации убытков.

Литература

1. Lee H. L., Padmanabhan V., Whang S. Information distortion in a supply chain: The bullwhip effect // Management Science – 1997. – № 43 (4) – p. 546.
2. Щавелева Н. А., Гусев Д. А. Способы устранения эффекта хлыста в цепях поставок // «Научное сообщество студентов XXI столетия. Экономические науки»: материалы IX студенческой международной заочной научно-практической конференции. Новосибирск, 28 марта 2013 г. – Новосибирск: Изд. «СибАК» – 2013. – С. 256-266
3. Чуйкова Ю. С. Динамическая модель пополнения товарного запаса предприятия с учётом заданных ограничений // Вестник Самарского государственного экономического университета – 2008. – № 11 (61) – с. 125-128.
4. Б. С. Соловьев, В. С. Карманов МОДЕЛИРОВАНИЕ ЭФФЕКТА ХЛЫСТА В СКЛАДСКОЙ ЛОГИСТИКЕ // НАУКА. ТЕХНОЛОГИИ. ИННОВАЦИИ // Сборник научных трудов в 10ч. / под ред. проф. Б.Ю.Лемешко, проф. А.Попова, проф. М.Э.Рояка, доц. В.С.Тимофеева. – Новосибирск: Изд-во НГТУ, 2017. –Часть 2. – 274 с

ПРИМЕНЕНИЕ АЛГОРИТМОВ ПЛАНИРОВАНИЯ ЭКСПЕРИМЕНТА К ПОСТРОЕНИЮ ОЦЕНОЧНЫХ ПОДМНОЖЕСТВ МЕТОДА MCD

И.О. Усова, Е.А. Хайленко

Новосибирский государственный технический университет,
г. Новосибирск, kemirinaou@gmail.com

В работе рассмотрена задача устойчивого оценивания параметров моделей с ошибками в объясняющих переменных. Рассмотрены основные понятия методов планирования эксперимента. Проведены исследования, в которых отображается изменение условия D-оптимального плана при применении различных спектров плана, а также в условиях разного рода выбросов

The problem of stable estimation of errors in variables models is considered in the work. The basic concepts of experiment planning methods are considered. Research, displaying change in the condition of the D-optimal design when applying different spectra of the design and in the conditions of various kinds of emissions, have been conducted.

На практике довольно часто встречаются задачи поиска оценок неизвестных параметров моделей с ошибками в объясняющих переменных (EIV) [3], и для того, чтобы при оценивании параметров EIV-моделей учитывать появление такого рода ошибок необходимо использовать специальные методы оценивания параметров. Кроме того, в выборке могут присутствовать грубые ошибки (выбросы) и классические методы будут давать некорректные оценки при обработке таких данных.

В данной работе будет рассматриваться метод наименьшего определителя ковариационной матрицы (MCD), который дает возможность оценивать параметры не по всем наблюдениям, а только по тем, которые позволяют достичь наименьшего определителя ковариационной матрицы, отсеивая наблюдения имеющих грубые ошибки.

Для того, чтобы добиться большей точности результатов при оценивании параметров регрессионных моделей можно использовать методы планирования эксперимента. Планирование эксперимента позволит получить максимально информативные точки, по которым будет производиться оценивание.

Одним из наиболее известных и часто используемых является критерий D-оптимальности.

План ξ^* называется D-оптимальным, если план максимизирует определитель информационной матрицы $M(\xi) : \xi^* = \text{Arg max} |M(\xi)|$

или минимизирует определитель дисперсионной матрицы $D(\xi)$:
 $\xi^* = \text{Arg min } |D(\xi)|$.

Информационная матрица $M(\xi)$ определяется [1]:

$$M(\xi) = \sum_{i=1}^n \frac{\omega_i}{\sigma_1(x_i, \theta)} \left(\frac{\partial m(x_i, \theta)}{\partial \theta} \right)^T \left(\frac{\partial m(x_i, \theta)}{\partial \theta} \right),$$

где $\sigma_1(x_i, \theta) = \left(1, \frac{\partial m(x_i, \theta)}{\partial x_i} \right) \Sigma_{\delta \varepsilon} \left(1, \frac{\partial m(x_i, \theta)}{\partial x_i} \right)^T$.

Как известно, эллипсоид рассеивания оценок параметров D-оптимального плана имеет минимальный объем.

В качестве условия D-оптимальности будет использоваться выражение вида: $d_{ML}(x, \xi^*) = f^T(x)M^{-1}(\xi^*)f(x) - m\sigma_1(x, \theta) = 0$

В данной работе перед применением планирования эксперимента будет производиться оценивание параметров с помощью метода максимального правдоподобия. Данный метод позволит получить истинные значения x_{ij}^* известной зашумленной переменной \bar{x} , с помощью которых будет найдена ошибка наблюдений $\bar{\varepsilon}$.

Метод максимального правдоподобия подразумевает, минимизацию функционала вида:

$$F = \frac{1}{2} \sum_{i=1}^n \left[\frac{(x_i - x_i^*)^2}{\sigma^2(x_i)} - \frac{(y_i - (\beta_0 + \sum_{j=1}^p \beta_j x_{ij}^*))^2}{\sigma^2(y_i)} \right].$$

Алгоритм метода максимального правдоподобия описан в работе [2].

Исследования проводились на примере простой линейной регрессии:
 $y = \beta_0 + \beta_1 x_1$, $x_j = x_j^* + \delta$.

Для данной модели D-оптимальным будет являться равновесный двухточечный план вида $\xi_1^* = \begin{Bmatrix} -1 & 1 \\ 0.5 & 0.5 \end{Bmatrix}$ При проведении исследований

рассмотрим также двухточечный неравновесный план

$$\xi_2^* = \begin{Bmatrix} -1 & 1 \\ 0.33 & 0.67 \end{Bmatrix} \text{ и четырёхточечный равновесный план}$$

$$\xi_3^* = \begin{Bmatrix} -1 & 1 & -3 & 3 \\ 0.25 & 0.25 & 0.25 & 0.25 \end{Bmatrix}$$

Условие D-оптимальности плана эксперимента

| | Без выбросов | Горизонтальные выбросы | Вертикальные выбросы | Смешанные выбросы |
|----------------|--------------------------------|--------------------------------|--------------------------------|--------------------------------|
| План ξ_1^* | [0, 0] | [0, 0] | [0, 0] | [0, 0] |
| План ξ_2^* | [0.014, 0.009] | [0.624, 0.075] | [0.189, 0.069] | [0.164, -0.242] |
| План ξ_3^* | [-0.051, -0.056, 0.036, 0.041] | [-0.852, -0.567, 0.074, 0.057] | [-0.436, -0.277, 0.047, 0.032] | [-0.035, 0.053, -0.355, 0.107] |

По таблице видно, что при использовании оптимального плана эксперимента ξ_1^* , условие D-оптимальности выполняется. В случае, когда выбирается неравновесный план ξ_2^* или план с большим количеством точек в спектре ξ_3^* , условие оптимальности не выполняются. Также можно заметить, что при отсутствии выбросов отклонение от оптимальности есть, но оно не так велико, как отклонение при наличии в данных выбросов. Поэтому в дальнейшей работе будем находить наблюдения, которые войдут в оценочное подмножество метода MCD исходя из их информативности.

Литература:

1. Konstantinou Maria, Dette Holger, Locally optimal designs for errors-in-variables models [Электронный ресурс] / Maria Konstantinou, Holger Dette. – Режим доступа: [https:// www.ruhr-uni-](https://www.ruhr-uni-)

bochum.de/imperia/md/content/mathematik3/publications/mem-unblinded.pdf.

2. Грешилов А. А. Математические методы построения прогнозов : научное издание / А. А. Грешилов, В. А. Стакун, А. А. Стакун. – М. : Радио и связь, 1997. – 112 с..

3. Rousseeuw Peter J., A Fast Algorithm for the Minimum Covariance Determinant Estimator [Электронный ресурс] / Peter J. Rousseeuw, Katrien VanDriessen. – Режим доступа : <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.45.5870&re=rep1&t type=pdf>.

О ПРОБЛЕМАХ ПОДБОРА МАКРОЭКОНОМИЧЕСКИХ ПОКАЗАТЕЛЕЙ ДЛЯ ОПЕРЕЖАЮЩИХ ИНДИКАТОРОВ

И.К. Холбоева, А.Ю. Тимофеева

**Новосибирский Государственный Технический Университет,
г. Новосибирск, kholboevaira@mail.ru**

В данной работе рассматриваются методы подбора опережающих индикаторов для дальнейшего прогнозирования продаж. Рассмотрены макроэкономические показатели, которые могут выступать в качестве опережающих индикаторов. Для анализа предлагаемых индикаторов рассчитаны коэффициенты корреляции между сводными индексами (опережающим, синхронным и запаздывающим) и выбранными макроэкономическими индикаторами. Определено, при каких лагах достигается наибольшая корреляция.

This paper discusses methods for selecting leading indicators for further sales forecasting. Macroeconomic indicators that can act as leading indicators are considered. For the analysis of the proposed indicators, the correlation coefficients between the composite indices (leading, roughly coincident and lagging) and the selected macroeconomic indicators were calculated. It was determined at which lags the highest correlation is achieved.

Временные ряды макроэкономических показателей содержат большое количество наблюдений (порядка сотни), в то время как число таких показателей в десятки раз больше. Необходимы специальные методы обработки предикторов, так как применение обычной регрессионной модели невозможно.

Методологии можно разделить на две широкие категории, характеризующиеся количеством учитываемых переменных. В исследованиях

Stock, Watson (1991), Camba-Mendez (2001) использовалось небольшое количество выбранных предикторов. Затем, в предположении перекрестно-некоррелированных ошибок, оцениваются факторные модели с использованием фильтра Калмана. Но основная работа Stock, Watson (2002) предполагает прогнозирование с использованием большого набора предикторов, где информация обобщается с использованием оценки главных компонент (Principal Component Analysis, PCA) [1]. С использованием PCA только небольшое количество ортогональных переменных остается в качестве входных данных.

Альтернативный подход, который может справиться с отбором переменных, когда их количество очень велико, - регуляризация (shrinkage), часто реализуется с использованием регрессии с ограничением на параметры (регрессия LASSO). LASSO успешно применяется для оперативного прогнозирования продаж с внешней информацией. Huang (2014) и Ma (2015) решали проблему выбора переменных, изменяя влияние каждой переменной по отдельности. При прогнозировании макроэкономических переменных LASSO было сочтено полезным при выборе соответствующих предикторов. Модель LASSO становится все более привлекательной в свете проблем с большими данными (Tibshirani, 2011). В отличие от PCA, этот подход дает прямое представление о конечной модели, что ценно для прогнозирования продаж [2].

Трудной проблемой в обоих вышеуказанных методах является выбор набора предикторов и переменных. Voivin, Ng (2006) предполагают, что кросс-корреляция регрессоров в больших наборах данных может привести к неточным прогнозам, и, следовательно, меньший набор, скорее всего, обеспечит меньшую среднюю ошибку прогноза. Ключом к этой проблеме является использование методов выбора переменных. Kapetanios (2007) и Kapitanios, Marcellino и Papailias (2012) используют нестандартную оптимизацию информационных критериев для определения подходящих инструментов и прогнозируют макроэкономические переменные с обнадеживающими результатами. В исследованиях Alvarez (2012) также решает эту проблему и сравнивает динамические факторные модели малого и большого масштаба с использованием набора данных Стока и Уотсона для США. Однако их подход ограничен этим конкретным классом моделей [1].

Для дальнейшего анализа были определены макроэкономические показатели, представленные в таблице 1. Эти показатели могут быть использованы в качестве опережающих индикаторов. Сводный циклический индикатор содержит значения трех обобщенных индикаторов

экономической конъюнктуры – сводного опережающего индекса (СОИ), сводного синхронного индекса (ССИ) и сводного запаздывающего индекса (СЗИ).

Таблица 2 – Макроэкономические показатели

| № | Показатель | Источник | Период, гг. | Периодичность |
|---|---|-------------------------------|-------------|---------------|
| 1 | Сводный циклический индикатор | Институт «Центр развития» ВШЭ | 1997-2018 | Месячная |
| 2 | Денежный агрегат М2 | Центральный Банк РФ | 1993-2018 | Месячная |
| 3 | Индекс реального курса рубля к доллару США | Центральный Банк РФ | 2004-2018 | Месячная |
| 4 | Базовый ИПЦ | Росстат | 1991-2018 | Месячная |
| 5 | Международные резервы | Центральный Банк РФ | 1998-2018 | Месячная |
| 6 | Оборот розничной торговли | Росстат | 2000-2017 | Месячная |
| 7 | Индекс цен на акции РТС | Московская Биржа | 1995-2018 | Месячная |
| 8 | Отношение М2 к объему золотовалютных резервов | Центральный Банк РФ | 1998-2018 | Месячная |
| 9 | Отношение М2 к денежной базе | Центральный Банк РФ | 2002-2018 | Месячная |

Для анализа приведенных выше опережающих индикаторов был проведен корреляционный анализ. Коэффициенты корреляции r показывают степень зависимости индикаторов от сводных циклических индикаторов. В качестве примеров были рассмотрены данные показателей за период 2005-2011 гг. Из всех данных была удалена сезонная компонента, и агрегат М2 взят как временной ряд первых разностей.

Как видно из таблицы 2 наиболее связанными индикаторами являются денежный агрегат М2 и курс рубля к доллару США.

Таблица 3 - Коэффициенты корреляции опережающих индикаторов

| Индикаторы | СОИ | | ССИ | | СЗИ | |
|--|--------|-----|--------|-----|--------|-----|
| | г | Лаг | г | Лаг | г | Лаг |
| Денежный агрегат М2 | 0,530 | 0 | 0,498 | -3 | 0,490 | -9 |
| Индекс реального курса рубля к доллару США | 0,419 | -4 | 0,457 | -8 | 0,375 | -8 |
| Базовый ИПЦ | -0,414 | -6 | -0,347 | -8 | 0,195 | 0 |
| Международные резервы | -0,699 | -10 | -0,658 | -13 | -0,703 | -16 |
| Цены на акции РТС | 0,482 | -8 | 0,447 | -10 | 0,043 | 0 |

На рисунке 1 показаны графики сводного опережающего индекса, объема денежного агрегата М2 и отношение реального курса рубля к доллару США. Все показатели на графике приведены к одному отрезку [0,1]. Черной вертикальной линией показано примерное начало финансово-экономического кризиса России 2008-2010 гг. Как видно из графика, СОИ начал реагировать на кризис раньше остальных показателей.

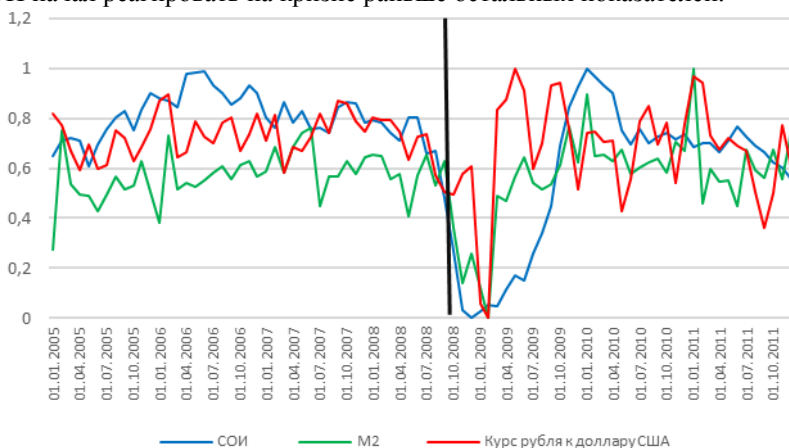


Рисунок 11 – Графики опережающих индикаторов

Денежный агрегат М2 реагирует на кризис, не смотря на небольшой коэффициент корреляции с СОИ, так же, как и показатель отношения курса рубля к доллару США. Отношение курса рубля к доллару США реагирует с большим опозданием, чем остальные, но гораздо быстрее реабилитируется. У индикатора «Международные резервы» примерно такой же коэффициент r , но реагирует с большим лагом. В целом, опираясь на кризис 2008-2010 годов, можно сделать вывод, что все индикаторы способны в разной степени заранее реагировать на кризис, но с небольшими задержками. В дальнейшем с целью автоматизации процесса прогнозирования продаж с использованием макроэкономической динамики необходимо разработать процедуру автоматического отбора опережающих индикаторов.

Литература:

1. Camba-Mendez G, Kapetanios G, Papailias F, Weale M. An automatic leading indicator, variable reduction and variable selection methods using small and large datasets // ECB Working Paper. - 2015. - №1773.

2. Sagaert Y, Kourentzes N Tactical sales forecasting using a very large set of macroeconomic indicators // European Journal of Operational Research. – 2018. – 264(2). – pp. 558-569.

ТИПОЛОГИЯ СУБЪЕКТОВ РФ ПО УРОВНЮ ИНВЕСТИЦИОННОЙ ПРИВЛЕКАТЕЛЬНОСТИ

К.С. Юшина, Л.К.Серга

**Новосибирский государственный университет
экономики и управления «НИНХ»,
г. Новосибирск, k.s.yushina@nsuem.ru**

В работе проведена оценка уровня инвестиционной привлекательности регионов РФ и выполнена их типология исходя из данного фактора. Главной проблемой является определение способа дифференциации инвестиционно привлекательных субъектов. В ходе работы разработан один из вариантов оценки инвестиционной привлекательности регионов РФ. Типология проведена регионов путем кластерного и портфельного анализа.

In this work there is assess the level of investment attractiveness of the regions of the Russian Federation. The main goal is to conduct a typology based on this factor.

The main problem is to determine how to differentiate investment attractive regions. During the work, one of the method for assessing the investment attractiveness of the regions of the Russian Federation was developed. A typology of regions was carried out through cluster and portfolio analysis

В условиях необходимости привлечения инвестиций в экономику страны оценка уровня инвестиционной привлекательности регионов становится весьма важной задачей. Так как необходимо установить, какие регионы можно отнести к инвестиционно привлекательным, а какие – нет.

Уровень инвестиционной привлекательности каждого субъекта РФ предлагаем определять как среднее значение из инвестиционного потенциала и инвестиционного риска. Подробнее об этом раскрывается в статье «Статистический анализ инвестиционной привлекательности регионов РФ» [8].

В результате анализа рассчитанного уровня инвестиционной привлекательности получаем, что самыми привлекательными для инвесторов (значения превышают 0,2) являются г. Москва, Республика Саха (Якутия), Тюменская область, Красноярский край. Это связано, в первую очередь, с высоким инвестиционным потенциалом, а также с богатыми природными ресурсами в большинстве выделенных регионов. Самыми непривлекательными оказались Республика Тыва, Республика Марий-Эл, г. Севастополь и Республика Крым. Можно также заметить, что уровень инвестиционного риска внутри этих регионов весьма высокий, что, несомненно, сказывается на привлекательности регионов.

Для проведения типологии субъектов по уровню инвестиционной привлекательности в координатах: инвестиционный потенциал – инвестиционный риск построена модифицированная матрица GE/McKinsey, которая позволяет выделить 4 типа регионов: 1) «Лидеры», 2) «Зона роста», 3) «Аутсайдеры», 4) «Перспективные» (Рисунок 1)

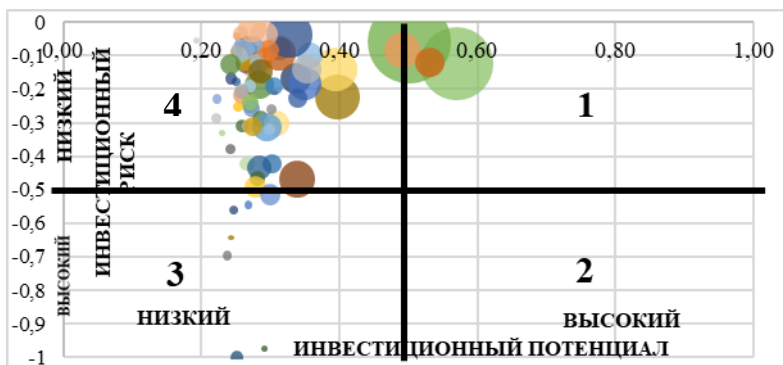


Рисунок 1 - Матрица «Инвестиционный потенциал– Инвестиционный риск» для субъектов РФ в 2016 г.

В результате, регионы РФ можно разбить на 4 группы:

1) «Лидеры» - г. Москва, Тюменская область и Республика Саха (Якутия). Данные регионы являются самым привлекательным вариантом инвестирования для потенциального инвестора, поскольку их инвестиционный потенциал высок, а инвестиционные риски незначительны.

2) «Зона роста» - Данные регионы имеют возможность минимизировать свои риски при сохранении высокого инвестиционного потенциала и перейти в группу «Лидеры». Среди субъектов РФ в 2016 году таких не оказалось.

3) «Аутсайдеры» - Ивановская область, Республика Хакасия, Республика Тыва, Республика Марий Эл, г. Севастополь, Республика Крым и Новосибирская область. Данным субъектам, чтобы стать «Перспективными», нужно в первую очередь снижать риски.

4) «Перспективные» - самая многочисленная группа, куда вошли 72 региона, не вошедшие в остальные типы. Чтобы не попасть в зону «Аутсайдеров», следует направить все действия, на то, чтобы низкий уровень инвестиционного риска соответствовал высокому уровню потенциала.

Однако значения нескольких из вышеперечисленных субъектов РФ находятся в критических зонах, то есть на переходе от одного типа к другому. Например, Красноярский край и Тюменская область расположены между «Перспективными» и «Лидерами» - для укрепления лидерских позиций, нужно поддерживать и увеличивать инвестиционный потенциал в регионе. Новосибирская и Кемеровская области граничат

между «Перспективными» и «Аутсайдерами», поэтому органам управления этих регионов нужно, прежде всего, разрабатывать программы по минимизации инвестиционных рисков региона.

Для рассмотрения места выбранных регионов по факторам, потенциально влияющих на инвестиционную привлекательность, проведен кластерный анализ регионов Российской Федерации. При построении матрицы расстояний использовалась Эвклидова метрика, а для объединения субъектов – метод Варда.

Кластеризация регионов по методу иерархической кластеризации представлена на рисунке 2.

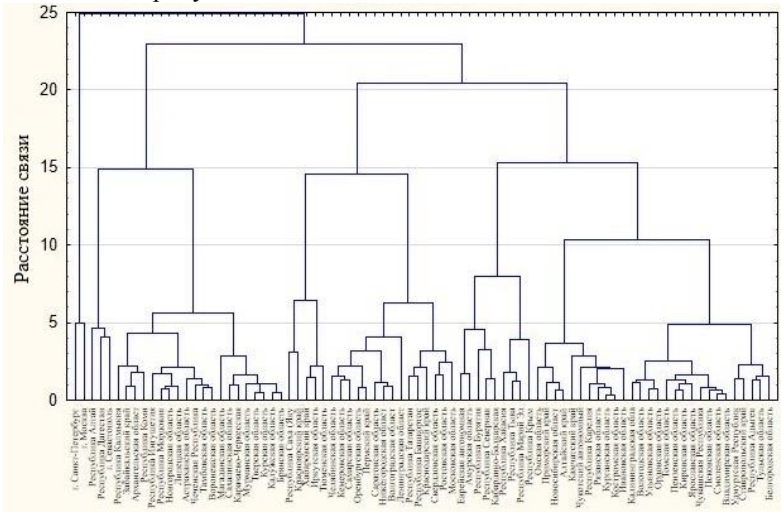


Рисунок 2 – Кластеризация регионов по методу иерархической кластеризации за 2016 год

Построенный дендрит показывает, что оптимальным является разбиение на 3 кластера. Для индентификации кластеров проведена кластеризация методом К-средних. Состав кластеров получился таким же. В 1 кластер вошли 41 регион со средним уровнем инвестиционной привлекательности с низким уровнем риска и низким уровнем потенциала. Во 2 кластер с низкой привлекательностью вошли 39 регионов при высоком риске и среднем потенциале. А в 3 – 2 региона: г. Москва и г. Санкт-Петербург, с высоким уровнем инвестиционной привлекательности (низкий риск - высокий потенциал).

Таким образом, совместное использование кластерного и портфельного анализа позволило выявить наиболее проблемные регионы, которые нуждаются в изменениях в первую очередь. Многие регионы являются «перспективными», поэтому нужно пересмотреть региональную политику каждого субъекта, обратив внимание на улучшение, в первую очередь, социально-экономической, ресурсно-сырьевой, производственной сфер деятельности, с целью повышения уровня инвестиционного потенциала. В регионах, находящихся в критических точках, следует обратить внимание на те факторы, которые способствуют повышению уровня инвестиционной привлекательности. Поэтому следует разрабатывать государственные программы и проводить региональные мероприятия, направленные на совершенствование инвестиционного законодательства субъектов РФ, способствование и оказание помощи инвесторам со стороны региональных властей путем льгот, установление инвестиционной доступности и открытости, активное привлечение иностранных инвестиций, а также развитие инвестиционной инфраструктуры.

Литература:

1. Вуколов Э.А. Основы статистического анализа. Практикум по статистическим методам и исследованию операций с использованием пакетов STATISTIC A и EXCEL Учебное пособие. Москва, изд-во Форум, 2008 год. – 454 с.
2. Глинский В.В. Портфельный анализ: сущность и применение // Финансы и бизнес. – 2008. – № 4. – С. 105–109.
3. Регионы России. Социально-экономические показатели. 2017: Р32 Стат. сб. / Росстат. - М., 2017. - 1402 с.
4. Серга Л.К., Никифорова М.И., Румынская Е.С., Хван М.С. Прикладное использование методов портфельного анализа // Вестник НГУ-ЭУ. – 2012. – № 3. – С. 146–158
5. Серга Л.К. О подходах к решению задачи идентификации нечетных совокупностей // Вестник НГУЭУ – 2013. - №3. – С. 83-91.
6. Юшина К.С. Статистический анализ инвестиционной привлекательности регионов РФ // Труды 9-й Международной научно-практической конференции студентов и аспирантов «Статистические методы анализа экономики и общества» (15-18 мая 2018 г.) – Национальный исследовательский университет «Высшая школа экономики», 2018, стр. 286-289.

7. Glinskiy V., Serga L., Chemezova E., Zaykov K. Clusterization economy asa way build sustainable development of the region. Procedia CIRP. Vol. 40. "13th Global Conference on Sustainable Manufacturing - Decoupling Growth from Resource Use". Elsevier Science Publishing Company, Inc. pp. 324-328.

СОДЕРЖАНИЕ

НАУЧНОЕ НАПРАВЛЕНИЕ – ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ И ОБРАБОТКИ ДАННЫХ

Секция МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ, АНАЛИЗ И ОБРАБОТКА ДАННЫХ

| | | |
|--|--------------------------|----|
| Абдулов А.А., Воловикова З. А., Мустакимов М. Ф., Балахчи А. Г., Веснин А.М. | Иркутск | 3 |
| Астанина М.С., Шерemet М.А. | Томск | 5 |
| Бардашов Д.С., Быков А.Г., Лосев Д.В. | Томск | 8 |
| Витчикова Л.Ю., Черникова О.С. | Новосибирск | 12 |
| Гелевера И.О., Куделин С.П., Инатович Ю.В. | Екатеринбург | 17 |
| Гераськин Р.А., Игнатъева А.А., Спектор А.А. | Новосибирск | 21 |
| Демьяненко А.В., Рабинович Е.В. | Новосибирск | 24 |
| Демьяненко А.В., Рабинович Е.В. | Новосибирск | 26 |
| Дружинин В.Г., Морозов В.А. | Санкт-Петербург | 29 |
| Дяченко Э.П., Дяченко Н.П., Иванова М.И. | Московская обл., д. Веря | 34 |
| Замашикова А.А., Щеколдин В.Ю. | Новосибирск | 39 |
| Кулабухова С.О., Чубич В.М. | Новосибирск | 43 |
| Мальцева Е.А., Карманов В.С. | Новосибирск | 48 |
| Мартюгин А.В. | Набережные Челны | 52 |
| Маслов К.А., Токарева О.С. | Томск | 56 |
| Овчинников Н.Н., Двойнишников С.В. | Новосибирск | 60 |
| Осинцева Е.А., Чимитова Е.В. | Новосибирск | 63 |
| Переверзина Т.В. | Новосибирск | 67 |
| Поверин Д.В., Постовалов С.Н. | Новосибирск | 70 |
| Пыжлаков К.С., Бакаев М.А. | Новосибирск | 75 |
| Шабля Ю.В., Репкин А.С., Кручинин Д.В. | Новосибирск | 78 |
| Садовникова Я.С., Мандровский К.П. | Москва | 83 |
| Саклаков В.М. | Томск | 87 |
| Семенов Д.О., Двойнишников С.В. | Новосибирск | 90 |

| | | |
|---|-------------|-----|
| Соломатин В.П., Гришаев В.Ю., Марупова К.Р., Саклаков В.М. | Томск | 93 |
| Филипенко Н.Ю., Рабинович Е.В. | Новосибирск | 97 |
| Халин Д.С., Семёнова М.А., | Новосибирск | 100 |
| Четвертакова Ю.С., Черникова О.С. | Новосибирск | 105 |
| Шапорева Г.О., Бизяев А.А., Яковицкая Г.Е. | Новосибирск | 109 |
| Шевченко Д.А., Осипов А.Л. | Новосибирск | 110 |
| Шерстобитова А.О., Емельянова Т.В. | Томск | 114 |
| Ширяев Б.В., Ющенко А.Ю., Безрук А.В. | Томск | 118 |

**Секция МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ, АНАЛИЗ
И ОБРАБОТКА ДАННЫХ**

| | | |
|--|--------------|-----|
| Гриф А.М., Персова М.Г. | Новосибирск | 121 |
| Гумалевский Р.В., Соловейчик Ю.Г. | Новосибирск | 125 |
| Гусельникова О.О., Бердников В.С. | Новосибирск | 129 |
| Данилов Н.И., Бердников В.С. | Новосибирск | 133 |
| Карасенко И.И., Соловейчик Ю.Г. | Новосибирск | 134 |
| Кузорова А.Е., Суханов Д.Я. | Томск | 138 |
| Кунц А.В., Ступаков И.М. | Новосибирск | 142 |
| Набагова А.А., Гольцев В.А., Метелкин Н.А. | Екатеринбург | 144 |
| Патрушев И.И., Персова М.Г. | Новосибирск | 148 |
| Пельменев К.Г., Остертак Д.И., Драгунов В.П. | Новосибирск | 153 |
| Росляков С., Суханов Д.Я. | Томск | 157 |
| Толстобров И.А., Ступаков И.М. | Новосибирск | 160 |
| Филяков А.Д., Романов Д.А., Невский С.А., Соснин К.В. | Новокузнецк | 163 |

**Секция ТЕХНОЛОГИИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА,
МАШИННОГО ОБУЧЕНИЯ И ПРЕДСКАЗАТЕЛЬНОГО МОДЕ-
ЛИРОВАНИЯ**

| | | |
|--|-------------|-----|
| Бондарев Э.С., Ермакова Д.В., Шаповалов А.В., Маслов Н.А. | Новосибирск | 169 |
| Ветров А.И., Попов А.А. | Новосибирск | 172 |
| Глухов Г.И., Гульяева Т.А. | Новосибирск | 177 |

| | | |
|--|-------------|-----|
| Грибков Е.И., Ехлаков Ю.П. | Томск | 180 |
| Кочнев А.В., Волкова В.М. | Новосибирск | 184 |
| Кутырев И.О., Попов А.А. | Новосибирск | 188 |
| Лебедева М.И., Бакаев М.А. | Новосибирск | 192 |
| Морковкин А.Г., Попов А.А. | Новосибирск | 195 |
| Тальников В.О., Волкова В.М. | Новосибирск | 198 |
| Часовских Ю.А., Радченко С.Е., Вострецов А.Г. | Новосибирск | 202 |
| Чепур А.А., Карманов В.С. | Новосибирск | 206 |
| Шкатула Д.Ю., Карманов В.С. | | 211 |
| Якимук А.Ю., Томская К.М. | Томск | 215 |

**Секция ЭКОНОМИКО-МАТЕМАТИЧЕСКИЕ МЕТОДЫ,
СТАТИСТИКА И ЭКОНОМЕТРИЯ**

| | | |
|--|-------------|-----|
| Афанасьева Е.А., Зотеев В.Е. | Самара | 220 |
| Верига М.Д., Сарычев С.П. | Новосибирск | 224 |
| Гойал Д.М., Тимофеев В.С. | Новосибирск | 227 |
| Загирова Е.А., Панюков А.В. | Новосибирск | 230 |
| Капленко А.А., Фаддеенков А.В. | Новосибирск | 234 |
| Климкина А.Г., Федорова Н.В. | Красноярск | 239 |
| Коржуева Д.А., Карманов В.С. | Новосибирск | 242 |
| Молочная К.А., Фаддеенков А.В. | Новосибирск | 247 |
| Муртазина М.Ш., Авдеенко Т.В. | Новосибирск | 250 |
| Олейник Н.С., Щеколдин В.Ю. | Новосибирск | 253 |
| Попкова А.В., Макаридина Е.В. | Новосибирск | 257 |
| Соловьев Б.С., Карманов В.С., Щеколдин В.Ю. | Новосибирск | 261 |
| Усова И.О., Хайленко Е.А. | Новосибирск | 266 |
| Холбоева И.К., Тимофеева А.Ю. | Новосибирск | 269 |
| Юшина К.С., Серга Л.К. | Новосибирск | 273 |

НАУКА. ТЕХНОЛОГИИ. ИННОВАЦИИ
Сборник научных трудов в 9 частях

г. Новосибирск, 03-07 декабря 2018 г.

Часть 2

Под ред. Гадюкиной А.В.

Подписано в печать 03.12.2018. Формат 60×84/16. Бумага офсетная.

Уч.-изд. л. 16,5. Печ. л. 17,75. Тираж 100 экз. Заказ № 2

Отпечатано в типографии

Новосибирского государственного технического университета
630073, г. Новосибирск, пр. К. Маркса, 20