

РАЗДЕЛ 1. ВЫЧИСЛИТЕЛЬНЫЕ ТЕХНОЛОГИИ

ИДЕНТИФИКАЦИЯ СОБЫТИЙ АКУСТИЧЕСКОЙ ЭМИССИИ С ИСПОЛЬЗОВАНИЕМ МОДИФИЦИРОВАННЫХ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ

Ахмадеев Руслан Салаватович, студент гр. 21-Б.13

Математико-механический факультет

Санкт-Петербургский государственный университет,

Григорьева Анастасия Викторовна, к.т.н, доцент

Математико-механический факультет

Санкт-Петербургский государственный университет,

Максименко Маргарита Владимировна, к.т.н.,

ассистент кафедры высшей математики

Санкт-Петербургский горный университет императрицы Екатерины II

(г. Санкт-Петербург)

Аннотация. Акустическая эмиссия (АЭ) – один из наиболее актуальных методов контроля дефектов объектов промышленности. Для эффективной диагностики их состояния требуется классифицировать сигналы АЭ по предполагаемой природе дефекта, целесообразно использовать методы машинного обучения (МО). Основной результат работы – создание алгоритма классификации сигналов АЭ по типам дефектов на основе модифицированных методов машинного обучения без учителя. Полученная модель показала высокую точность на тестовых данных (80%) и большую устойчивость (90%), успешно опробована на данных реального объекта.

Ключевые слова: машинное обучение; акустическая эмиссия; классификация дефектов; тип повреждения.

Акустико-эмиссионный контроль является одним из наиболее распространенных и практичных методов неразрушающего контроля, используемых для диагностики дефектов у объектов промышленности. В частности, мы рассмотрим его применение на объектах ядерной энергетики.

Вкратце его суть можно описать так: на испытуемый объект устанавливаются датчики АЭ, объект начинают нагружать (давлением или температурой), датчики фиксируют показатели акустических сигналов, получаемые в процессе деформации предполагаемых дефектов.

В ходе применения диагностики АЭ получаем большой объем данных сигналов, который нужно уметь оперативно анализировать для контроля состояния важных (зачастую аварийно-опасных) объектов.

В связи с активным развитием методов машинного обучения, крайне логичным и целесообразным стало их внедрение в процессы анализа и диагностики объектов

различных отраслей промышленности, что подтверждает большое количество актуальных статей на эту тему [1,3,4].

Существующие коммерческие программные пакеты анализа данных АЭ, такие как Vallen System, Aline, не решают задачи отнесения зафиксированной группы сигналов к наиболее вероятному типу дефекта.

Цель работы – создание алгоритма классификации полученных сигналов по предполагаемому типу дефекта, основанного на стандартных методах МО без учителя с подкреплением учебными данными. Преимущество выбранного подхода перед алгоритмами МО с учителем: эти методы могут работать без предварительного обучения на данных, показывают высокую скорость, менее склонны к переобучению и более универсальны.

В качестве исходных данных принимается массив, содержащий АЭ сигналы с их различными характеристиками, такими как амплитуда, время, скорость, энергия, длительность, количество срабатываний датчика и т.д.

Методологическая основа работы – статья зарубежных авторов Identifying damage mechanisms of composites by acoustic emission and supervised machine learning [1].

В литературе можно встретить разные виды классификаций дефектов, в зависимости от их опасности, структуры и материалов, в которых они возникают. Наиболее универсальным и наглядным будет следующее деление источников АЭ на 5 типов дефектов [1]:

1. Трение, шумы, электромагнитные помехи.
2. Отслоения и расслоения.
3. Вытягивания и растяжения.
4. Трещины.
5. Разрывы и разломы.

У исходных данных был проведен отбор наиболее значимых признаков с использованием коэффициента ранговой корреляции Кендалла для более точной классификации сигналов [1].

Выделены наиболее значимые признаки сигналов – амплитуда, частота и энергия. Графики по аналогии со статьей [1] показывают амплитуду и частоту сигналов, но энергия тоже использовалась в классификации. К статье был приложен большой объем данных АЭ, полученных в ходе экспериментов с отмеченными типами дефектов.

Разработанный метод состоит из четырех этапов, которые описаны ниже.

Первый этап – предварительная обработка данных [2], позволяющая более точно и корректно работать методам МО. Во-первых, создание из имеющихся признаков данных новых. Такой параметр один — частота сигналов, рассчитываемая по формуле из [3]: $CNTS/D$. Во-вторых, фильтрация шумов. Обучающие данные были заранее очищены от выбросов, поэтому никакой предобработки для них не требуется, в то время как данные реального промышленного объекта отфильтрованы по частоте, и в рассмотрение брались только те сигналы, у которых частота меньше 0,6 Гц. Также проведена фильтрация по времени: изучаем только данные, полученные ночью, так как днем на объекте присутствовал антропогенный шум.

Второй этап – выбор и настройка базового алгоритма кластеризации сигналов. Для создания программной части алгоритма использовался язык Python с библиотеками sklearn, pandas и numpy. Опробованы некоторые стандартные функции библиотеки sklearn с относительно небольшим числом кластеров и гиперпараметров: K-Means, DBSCAN, hierarchical clustering. Рассмотрим подробнее, почему из перечисленных функций нами выбрана агломеративная кластеризация, и как мы ее модифицируем.

- KMeans

В классической реализации разбивает множество элементов векторного пространства на заранее известное число кластеров k . Действие алгоритма таково, что он стремится минимизировать среднеквадратичное отклонение на точках каждого кластера. Основная идея заключается в том, что на каждой итерации заново вычисляется центр масс для каждого кластера, полученного на предыдущем шаге, затем векторы разбиваются на кластеры вновь в соответствии с тем, какой из новых центров оказался ближе по выбранной метрике. Алгоритм завершается, когда на какой-то итерации не происходит изменения кластеров.

- DBScan

Алгоритм кластеризации DBSCAN основан на концепции образцов ядра, неосновных образцов и выбросов, а также на плотности кластеров:

1. Образцы ядра: образцы, представленные в области с высокой плотностью, имеют минимальное количество точек выборки с радиусом ϵ .

2. Образцы не ядра: образцы близкие к образцам ядра, но не являющиеся образцами ядра, а находящиеся очень близко к образцам ядра. Образцы не ядра лежат в радиусе ϵ от образцов ядра, но у них нет минимальных точек отбора проб.

3. Выбросы: образцы, которые не являются частью образцов ядра и образцов не ядра, и находящиеся далеко от всех образцов.

- Agglomerative clustering

Подвид иерархической кластеризации, работающий “снизу-вверх”. Изначально каждый элемент содержится в своем отдельном кластере. Затем на каждом шаге наиболее похожие элементы попарно объединяются в общий кластер. Степень “схожести” определяется расстояниями между кластерами, которое в свое очередь рассчитывается на основании выбранного нами метода (стандартные методы: “ward”, “single”, “complete”, “average”) и метрики расстояний между точками (например, евклидово или манхэттенское расстояние). С каждой итерацией порог расстояния между кластерами увеличивается, пока не достигнем нужного нам числа кластеров.

Программная реализация этих алгоритмов описана в [4].

После рассмотрения этих алгоритмов применительно к поставленной задаче был сделан вывод, что лучше всего по метрике ассигасу (процент совпадения полученных кластеров с типами дефектов из учебных данных) себя показала агломеративная кластеризация.

Вручную были настроены гиперпараметры модели, такие как метрика (в нашем случае остановились на cosine), способ определения межкластерного расстояния (в нашем случае average) [4].

Третий этап – модификация алгоритма. В ходе экспериментов было установлено, что целесообразно применение метода несколько раз для более точного определения дефектов, т.е. сначала мы делим все данные на 3 кластера (дефекты низкой, средней и высокой амплитуды). Затем делим крайние по амплитуде еще на 2. Получив тем самым искомые 5 кластеров дефектов. Тип дефекта для кластера в обучающих данных определяется преобладающей меткой.

Получили приемлемую для алгоритма без учителя точность: 87% для разрывов, 79% для расслоений, 73% для трения, 60% для трещин и 40% для растяжений, что сравнимо с результатами, полученными в статье с использованием метода KNN [1].



Рис. 1. Результаты алгоритма на обучающих данных

Четвертый этап – определить, какой кластер представляет собой какой тип дефекта. Самый простой алгоритм: объединить изучаемые данные с учебными, поделить на кластеры и каждому присвоить тип наиболее часто встречающегося в нем дефекта – не работает корректно, так как в размеченных данных дефекты имеют разные частоты, поэтому наиболее представленные могут вытеснить редкие, но не менее важные.

Для того чтобы оценить, к какому типу дефекта наиболее вероятно относится рассматриваемый кластер, введем так называемые коэффициенты надежности для каждого из дефектов, которые вычисляются по следующей формуле:

$$K[j] = \frac{100 * D[i][j]}{C[i]},$$

где $D[i]$ – доля точек из j – го кластера с i – ым типом дефекта, $C[i]$ – доля i – го типа дефекта в учебных данных.

Апробация алгоритма на данных реального объекта – бетонного кожуха реактора, нагружаемого термически в диапазоне от 250 °С до 400 °С. Показания фиксировались как с поверхности объекта, так и в глубине с помощью волноводов.

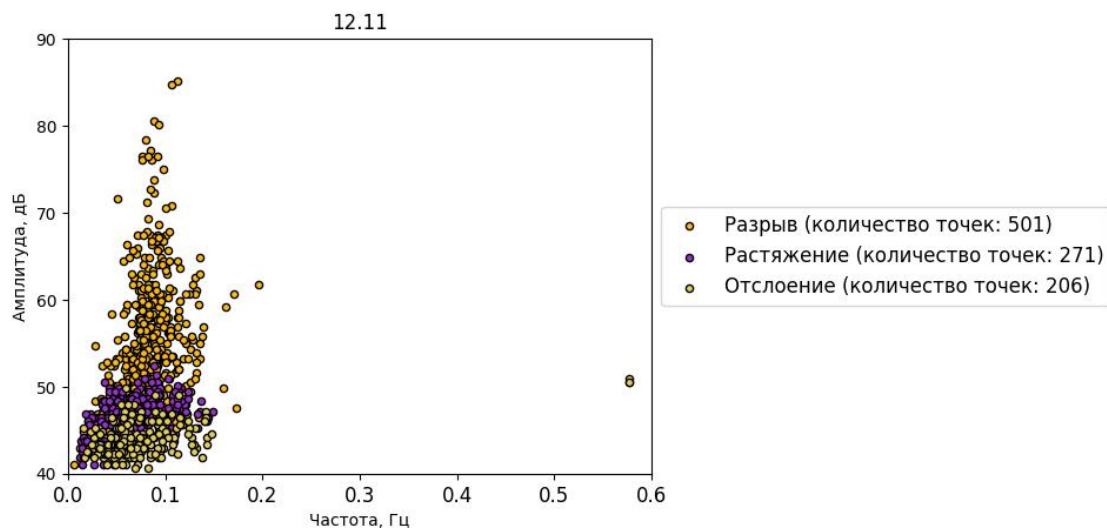


Рис. 2. Результат алгоритма на реальных данных

Метод имеет высокую устойчивость: при применении его на дефектах одного типа несколько раз около 90% сигналов попадет в один и тот же тип кластера.

Выводы: метод показал результаты высокой точности на учебных данных, и был успешно опробован на данных реального объекта.

Дальнейшее усовершенствование метода может быть связано с его внедрением в процессы локации дефектов, изучением применимости к объектам другого типа (в ближайших планах – тонкостенные резервуары высокого давления), корректировкой гиперпараметров и коэффициентов. Из апробации метода в том виде, который представлен на данный момент, запланировано проведение тестов на данных калибровки [6] рассмотренного выше объекта.

Список литературы

1. Identifying damage mechanisms of composites by acoustic emission and supervised machine learning (Renato S.M. Almeida, Marcelo D. Magalhães, Md Nurul Karim, Kamen Tushtev, Kurosch Rezwan).
2. Балагурин, П. Обработка данных акустических сигналов для применения к ним коэффициента корреляции Пирсона и их кластеризация / П. Балагурин, П. Макарихин, А. В. Григорьева // Fourth Conference on Software Engineering and Information anagement (SEIM-2019). 2019. С. 41–47.
3. Сайт библиотеки sklearn: сайт. – 2024. – URL: <https://scikit-learn.org/stable/modules/clustering.html#clustering>.
4. Analysis of acoustic emission data for structural health monitoring applications (Manindra Kaphle).
5. Cluster analysis of acoustic emission signals and infrared thermography for defect evolution analysis of glass/epoxy composites (Yanan Zhang, Bo Zhou, Fangai Yu, Changzheng Chen)
6. Григорьева, А. В. Метод обработки данных акустико-эмиссионного контроля для определения скорости и локации каждого сигнала / А. В. Григорьева,

ЧИСЛЕННОЕ МОДЕЛИРОВАНИЕ ЗАДАЧ ТЕЧЕНИЯ И ПЕРЕНОСА В ПЕРФОРИРОВАННЫХ ОБЛАСТЯХ С НЕОДНОРОДНЫМИ ГРАНИЧНЫМИ УСЛОВИЯМИ

*Герасимова Лена Аркадьевна, студент гр. Б-ПМИ-20,
Алексеев Валентин Николаевич, к.ф.-м.н, доцент
Институт математики и информатики
Северо-Восточный федеральный университет им. М.К. Аммосова
(г. Якутск)*

Аннотация. В данной работе исследуется задача моделирования течения и переноса жидкости в пористых средах с учетом воздействия граничного условия, известного как граничное условие Робина. Используется смешанный метод конечных элементов для аппроксимации уравнения Стокса, которое описывает движение жидкости. Также изучается уравнение переноса концентрации вещества, решаемое последовательно с учетом различных значений коэффициента диффузии с использованием метода SUPG. Описаны результаты численного моделирования в трёхмерной модели.

Ключевые слова: метод конечных элементов; несжимаемая жидкость; задача Стокса; уравнение переноса; граничное условие Робина; численное моделирование; перфорированная область; SUPG.

Проблема течения и переноса жидкости в пористых средах привлекает все больше внимания исследователей и инженеров в силу ее важности для различных областей науки и техники. Пористые среды, такие как грунты и горные породы, играют решающую роль во многих процессах, включая геологические формации, нефтегазовую эксплуатацию, подземное хранение энергии и различные инженерные приложения [1-5].

Одним из ключевых аспектов моделирования течения и переноса в пористых средах является учет граничных условий, особенно в случае взаимодействия с окружающими средами. Граничные условия третьего рода, такие как граничное условие Робина, описывающее реакцию на стенках пористой среды, играют важную роль в точном описании этих процессов. Исследование и разработка численных методов для моделирования таких условий имеет критическое значение для точного прогнозирования поведения системы в реальных условиях [6].

Сосредоточимся на численном моделировании потока и переноса жидкости в пористых средах с использованием граничного условия Робина. Используем смешанный