

ПРИМЕНЕНИЕ МОДИФИЦИРОВАННЫХ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ АНАЛИЗА ДАННЫХ АКУСТИЧЕСКОЙ ЭМИССИИ

Григорьева А.В., Максименко М.В., Макарихин П.А.

СПбГУ, Санкт-Петербург, Россия

Nasty001@mail.ru

Основной задачей данной работы являлась разработка интеллектуальной системы обработки данных АЭ, которая позволит с высокой точностью определять местоположение источников АЭ на корпусе объекта. Для этих целей использованы алгоритмы машинного обучения, характеристики которых настроены авторами с учетом особенностей акустико-эмиссионных сигналов и благодаря этому автоматически подстраиваются под исследуемый объект. Это позволяет адаптировать метод для каждого объекта без необходимости перенастройки алгоритма вручную. Объектами исследования на данном этапе являются тонкостенные стальные сосуды высокого давления.

Первым необходимым шагом являлось нахождение (с некоторой погрешностью) местоположения каждого из источников АЭ, зафиксированных во время эксперимента. Этот метод разработан и подробно описан авторами в [1].

Вторым шагом является кластеризация. На данный момент существует много методов кластеризации: *k*-средних, *affinity propagation*[2], агломеративная кластеризация, различные нейронные сети и многое другое. Однако, из-за специфики данных АЭ важно выполнение нескольких требований к алгоритму: число кластеров заранее неизвестно, результат должен быть одинаков для одного набора данных (то есть не должно быть зависимости от случайных величин, как, например, в алгоритме *k*-средних) и устойчивость к выбросам и шумам. Алгоритм DBSCAN [3] отвечает всем этим требованиям. Суть разработанного метода кластеризации заключается в автоматическом вычислении диапазонов характеристик для данного конкретного объекта. Ему необходимо задать в качестве начальных данных одну из опасных зон на поверхности, в которой с большой вероятностью имеется дефект. Она является образцом для поиска аналогичных зон на том же объекте. Таким образом, для работы метода необходимы координаты прямоугольника, в котором находится зона-образец, и количество высокоамплитудных импульсов в этом прямоугольнике. Число определяется опытным путем и варьируется от 3 до 10 в зависимости от того, насколько поврежден объект, исходя из [4]. Удачный выбор такой зоны может повлиять на качество кластеризации, так как будет более точно рассчитаны гиперпараметры алгоритма DBSCAN: максимальное расстояние между элементами для объединения их в один кластер, допустимая разница скоростей (скорость высчитывается для каждого источника индивидуальная на шаге 1), разница усредненных амплитуд, и другие 15 параметров метрики. Выбор параметров основывался на следующих статьях: [4, 5, 6]. Дополнительной сложностью является то, что разные параметры обладают разной размерностью и значимостью, а следовательно применение евклидовой метрики для вычисления расстояния между элементами кластеров будет некорректным. Для решения этой проблемы была разработана и применена следующая метрика:

$$f(i, j) = w_1 * l + w_2 * v + w_3 * a$$

i, j – точки из набора данных; w_1, w_2, w_3 – весовые коэффициенты расстояния, скорости и амплитуды соответственно;

$$l = \sqrt{\left(\sum_{k=0}^n \frac{x_{ik}}{n} - \sum_{k=0}^m \frac{x_{jk}}{m}\right)^2 + \left(\sum_{k=0}^n \frac{y_{ik}}{n} - \sum_{k=0}^m \frac{y_{jk}}{m}\right)^2}$$

$$v = \left| \sum_{k=0}^n \frac{v_{ik}}{n} - \sum_{k=0}^m \frac{v_{jk}}{m} \right|$$

$$a = \min \left(\sum_{k=0}^n \frac{a_{ik}}{n} - \sum_{k=0}^m \frac{a_{jk}}{m} \right)$$

где x , y , v , a – усредненные координаты, скорость, амплитуда точек в наборах потенциальных расположений источников сигналов, n , m – количество точек в этих наборах.

Для каждого из параметров (расстояния, скорости и амплитуды) были использованы соответствующие пороговые значения для расчета глобального эpsilon, используемого в алгоритме DBSCAN:

$$EPS = w_1 * eps_l + w_2 * eps_v + w_3 * eps_a$$

Разработанная система прошла проверку на имеющихся данных АЭ-контроля трёх крупных производственных объектов. Полученные результаты хорошо согласуются с результатами, полученными при исследовании этих объектов другими методами неразрушающего контроля после проведения зачистки поверхности. Применение разработанного метода кластеризации сигналов значительно сокращает площади объекта, рекомендованные к проверке специалистом.

1. Григорьева А. В., Максименко М. В. Метод обработки данных акустико-эмиссионного контроля для определения скорости и локации каждого сигнала // Компьютерные исследования и моделирование. – 2022. – Т. 14. – №. 5. – С. 1029-1040.
2. Frey Brendan J, Dueck Delbert. Clustering by passing messages between data points // science. — 2007. — Vol. 315, no. 5814. — P. 972–976.
3. A density-based algorithm for discovering clusters in large spatial databases with noise. / Martin Ester, Hans-Peter Kriegel, Jörg Sander et al. — 1996. — Vol. 96, no. 34. — P. 226–231.
4. Гомера Виктор Петрович, Смирнов Анатолий Джумамуратович, Нефедьев Евгений Юрьевич. Повышение достоверности АЭ-контроля оборудования, содержащего внутренние устройства // В мире неразрушающего контроля. — 2014. — no. 3. — P. 20–24.
5. Yaroslavkina E. Investigation of methods for determining the integral characteristics of periodic signals separated in time. — 2018. — P. 1–4
6. Evaluating the Probability of Detecting Acoustic Emission Signals Using the Amplitude Discrimination Method with a Low Signal-toNoise Ratio / IA Rastegaev, IS Yasnikov, II Rastegaeva et al. // Russian Journal of Nondestructive Testing. — 2022. — Vol. 58, no. 3. — P. 157–166