

ОРГАНИЗАЦИЯ ПЕРВИЧНОГО ПОИСКА ВИБРАЦИОННЫХ СИГНАЛОВ В БОЛЬШИХ ХРАНИЛИЩАХ ДАННЫХ

Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь

Базаревский В.Э., Базаревский Вл.Э.

Бранцевич П.Ю. – кандидат техн. наук, доцент

В современных системах поддержки принятия решений на основе вибрационной диагностики одной из главных задач является обеспечение эффективного поиска вибросигналов. В данных тезисах рассматриваются несколько способов организации данных для эффективного их анализа и поиска.

Одной из задач, подлежащей решению в рамках создания системы вибрационной диагностики конструкций с вращательным движением, является поиск вибросигналов, аналогичных наблюдаемым в данный момент по ряду характеристик. Поиск таких прецедентов позволяет делать своего рода предсказания о возможных вариантах развития наблюдаемой ситуации на основе исторических данных об исследуемой или аналогичной ей конструкциях. Логично, что чем больше прецедентов будет храниться в базе, тем более точными будут такие предсказания.

Рассмотрим несколько способов организации хранилищ сигналов в контексте алгоритмической сложности [1] поиска сигналов по ним. Примем, что N – число хранимых сигналов, M – число атомарных характеристик, по которым ведется сравнение.

Таблица 1. Алгоритмическая сложность способов организаци данных

Способ организации данных	Сложность поиска
Векторы признаков, построенные на основе сигналов	$O(M \log N)$
Сигналы в бинарном виде	$O(M N)$
Преобразованные сигналы в виде пространственных структур	$O(M \log N)$

Каждый из указанных подходов имеет свои преимущества и может использоваться совместно с другими. Одним из преимуществ использования пространственных структур для организации поиска сигнальных данных является возможность использования пространственных индексов баз данных наряду с широкими возможностями параметризации поиска, тогда как поиск на основе предопределенных векторов признаков сильно ограничен и слабо расширяем в процессе использования и поддержки хранилища сигналов.

Исходя из таблицы 1, наиболее слабым местом в производительности поиска сигналов является M – число атомарных характеристик, по которым ведется сравнение. Существует несколько способов уменьшения M без существенных потерь в качестве поиска, наиболее эффективными из которых являются методы применения алгоритмов хеширования, чувствительных к расстоянию для многомерных векторов (Locality sensitive hashing) [2]. Не смотря что данные методы являются вероятностными и не обеспечивают 100% точности поиска, для большого числа сигналов они обеспечивают приемлемое качество поиска. Такой подход возможен только для минимизации M для поиска по векторам признаков.

Для поиска на базе пространственных структур такой подход невозможен, однако M может быть минимизировано за счёт приведения вибросигналов от временного к частотному представлению (спектру) последующими фильтрацией и построением огибающей по полученной функции. Так как сложность сравнения пространственных структур зависит от числа вершин, описывающих их, построение огибающей по отфильтрованному от шума сигналу значительно уменьшит число вершин результирующей фигуры, тем самым значительно уменьшив M без потери полезных характеристик сигнала.

Список использованных источников:

1. Du, Ding-Zhu. Theory of Computational Complexity / Du, Ding-Zhu; Ko, Ker-I // John Wiley & Sons, 2000. – 512 p.
2. Anand Rajaraman. Mining of Massive Datasets / Anand Rajaraman, Jeffrey David Ullman // Cambridge University Press – Cambridge, 2011. – 326 p.