

УДК 004.021

ВЫЯВЛЕНИЕ ГЕОМЕТРИЧЕСКИХ ПАТТЕРНОВ ВО ВРЕМЕННЫХ РЯДАХ С ИСПОЛЬЗОВАНИЕМ ЭВОЛЮЦИОННЫХ АЛГОРИТМОВ

Мережников М.В. (Университет ИТМО), **Калюжная А.В.** (Национальный центр когнитивных разработок), **Хватов А.А.** (Национальный центр когнитивных разработок)
Научный руководитель – кандидат технических наук, доцент Института Дизайна и Урбанистики Калюжная А.В.
(Национальный центр когнитивных разработок)

Выявление геометрических паттернов во временных рядах позволяет визуализировать свойства рядов в удобной для восприятия форме. Решение данной задачи подразумевает построение модели, обученной на исходных данных, ключевым аспектом которой является ее интерпретируемость. Использование генетического алгоритма позволяет определить структуру описывающей модели в форме, удобной для визуальной интерпретации.

Применение управляемых данными алгоритмов для выявления моделей обосновано частым отсутствием достаточного количества знаний для аналитического описания исследуемых явлений, для которых, однако, есть возможность проведения рядов наблюдений и получения набора данных. Зачастую в основе подобных алгоритмов лежат методы машинного обучения. Одним из случаев применения данного подхода является использование управляемых данными алгоритмов для построения моделей, описывающих поведение временных рядов на основе их реализаций и позволяющих оценить наличие геометрических паттернов с целью удобной визуальной интерпретации свойств рядов.

Целью данной работы является создание управляемой данными системы вывода аналитического вида временных рядов, позволяющей определить геометрические паттерны во временном ряде. Подробный анализ эффективности построения моделей достигается за счет валидации алгоритма на искусственных входных данных с заранее известными характеристиками, а также реальных данных, свойства которых уже были исследованы. В качестве входных данных алгоритм использует нормированный по амплитуде ряд временных отсчетов исследуемого процесса. Для исследования робастных свойств алгоритма в исходные данные добавляются шумы различной мощности.

Исследуемый метод основан на генетическом алгоритме, который подбирает наиболее подходящую структуру выражения на основе определенного набора базовых функций, выступающих в роли геометрических паттернов. В набор входят удобные для визуальной интерпретации тригонометрические функции и импульсы. Параметры функций оптимизируются при помощи алгоритма дифференциальной эволюции, а значимость каждого слагаемого в итоговом выражении оценивается при помощи использования L1 – регуляризации при линейной регрессии на нормированных данных, что не дает модели включать в себя плохо интерпретируемые составляющие ради малого увеличения точности описания входных данных. Возможность последовательного решения задачи позволяет закреплять выявленные на предыдущих этапах эволюции паттерны, уменьшая размерность задачи на последующих шагах.

Полученный алгоритм успешно выявляет структуру искусственно заданных временных рядов, состоящих из композиции некоторого числа базовых функций. Похожие результаты при добавлении шумов с высокой дисперсией говорят о робастности алгоритма. Результаты работы с реальными рядами хуже в силу наличия в них конструкций, которые сложно описать с высокой точностью набором базовых функций.