

# Разработка параллельного алгоритма поиска похожих подпоследовательностей временного ряда для сопроцессора Intel Xeon Phi\*

А.В. Мовчан, М.Л. Цымблер

Южно-Уральский государственный университет

Временные ряды широко распространены во всех сферах человеческой деятельности. Например, в медицине создается большое количество временных рядов, таких как электрокардиограммы, электроэнцефалограммы, данные по экспрессии генов и др. Для обнаружения знаний используются методы интеллектуального анализа. К интеллектуальному анализу временных рядов относятся такие задачи как кластеризация, классификация, прогнозирование, поиск аномалий. В основе многих алгоритмов интеллектуального анализа временных рядов лежит поиск похожих подпоследовательностей. Существует множество работ, посвященных данной задаче. Некоторые из них рассматривают возможность параллельной обработки временного ряда, но отсутствуют работы, посвященные параллельной реализации данной задачи для сопроцессоров Intel Xeon Phi.

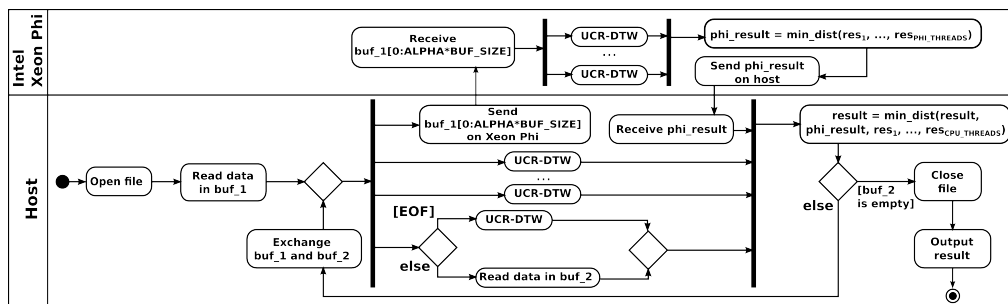


Рис. 1. Схема работы параллельного алгоритма

В данной работе рассматривается реализация алгоритма поиска похожих подпоследовательностей временного ряда для сопроцессора Intel Xeon Phi (см. Рис. 1). Схожесть определяется с помощью алгоритма динамической трансформации времени (Dynamic Time Warping). Этот алгоритм вычисляет оптимальную последовательность трансформации времени между двумя временными рядами, что позволяет сравнивать временные ряды, которые различаются скоростью. Для ускорения расчета применяется алгоритм UCR-DTW, использующий оптимизации, описанные в работе [1]. Для обработки временного ряда используются ресурсы центрального процессора (CPU) и сопроцессора. Часть данных буфера, равная константе ALPHA ( $0 < ALPHA \leq 1$ ), пересылается на сопроцессор для обработки, а другая часть обрабатывается CPU. Одновременно с обработкой данных один из потоков CPU загружает из файла следующую часть данных в другой буфер. Обработка на CPU и сопроцессоре осуществляется параллельно, а после выбираются лучшие результаты.

## Литература

1. Rakthanmanon T., Campana B., Mueen A., Batista G., Westover B., Zhu Q. Searching and mining trillions of time series subsequences under dynamic time warping // The Eighteen ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, August 12 – August 16, 2012, New York, New York, USA, ACM. 2012. P. 262–270.

\*Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 12-07-00443-а.