

# Построение метода динамического выравнивания многомерных временных рядов, устойчивого к локальным колебаниям сигнала.\*

И. О. Автор<sup>1</sup>, И. О. Соавтор<sup>2</sup>, И. О. Фамилия<sup>1,2</sup>  
author@site.ru; co-author@site.ru; co-author@site.ru

<sup>1</sup>Организация, адрес; <sup>2</sup>Организация, адрес

Данная работа посвящена построению эффективного алгоритма динамического выравнивания многомерных временных рядов. Для решения данной задачи предлагается использовать функцию расстояния DTW между двумя многомерными временными рядами, согласно которому выравниваются две оси времени, при этом внутри функционала DTW выбирается расстояние между  $i$ -м и  $j$ -м измерениями такое, что оно устойчиво к локальным “сдвигам” сигнала. В качестве решения будет рассмотрено более продвинутое, основанное на DTW между парой измерений. Для проверки корректности используются как и реальные данные, например измерения активность мозга обезьян, так и искусственно сгенерированные, например движение сигнала в пространстве по часовой и против часовой стрелки.

**Ключевые слова:** многомерные временные ряды; DTW; динамическое выравнивание.

## 1 Введение

В данной работе исследуется проблема динамического выравнивания многомерных временных рядов, устойчивого к локальным колебаниям сигнала. Временной ряд - собранный в разные моменты времени статистический материал о значении каких-либо параметров (в простейшем случае одного исследуемого процесса). В данном случае рассматривается многомерный случай.

Базовое решение задачи с помощью метрики L2 расстояния между рядами не всегда оказывается эффективным. Таким примером являются 2 временных ряда, полученные при близком расположении датчиков с сигналами, которые могут зафиксировать один и тот же пик. Полученный пик окажет большое влияние на значение метрики L2.

Напротив, использование известного алгоритма DTW, но уже в многомерном случае позволит обойти проблему малого расстояния между датчиками.

Полученные алгоритмы тестировались на реальных данных и искусственно сгенерированных. Полученные результаты показали преимущество использования попарного DTW алгоритма.

## 2 Постановка задачи.

Рассматривается алгоритм решения задачи динамического выравнивания многомерных рядов.

$t = (t_1, t_2, \dots, t_n), r = (r_1, r_2, \dots, r_m)$ , где  $\forall i \in \{1, 2, \dots, n\} X_i, Y_i \in \mathbb{R}^K$  и  $\forall i \in \{1, 2, \dots, m\} X_i \in \mathbb{R}^K$  - 2 многомерных временных ряда размерности  $K$ .

Для этого рассмотрим задачу двухклассовой классификации 2 многомерных временных рядов, в которой в качестве метрики используется одна из нижеприведённых.

Метрика  $L_2$  между  $i$  сигналом 1 датчика и  $j$  сигналом 2 датчика:

$$d(i, j) = \sum_{k=1}^K (t(i, k) - r(j, k))^2$$

\*Работа выполнена при финансовой поддержке РФФИ, проекты № 00-00-00000 и 00-00-00001.

25 Расстояние DTW между парой сигналов определяется с помощью следующей рекур-  
 26 рентной формулы.

$$27 \quad D_{(i,j)} = d_{(i,j)} + \min(D_{(i-1,j)}, D_{(i,j-1)}, D_{(i-1,j-1)})$$

28 В качестве функционала ошибки будем использовать ассигасу.

29 Будут рассматриваться искусственно сгенерированные данные, которые жестко будут  
 30 разделены по положению пика на 2 класса.

31 Те ряды, которые имеют небольшое расстояние до пика, будут принадлежать одному  
 32 классу.

### 33 **3 Базовый эксперимент**

34 Суть эксперимента заключается в сравнении различных метрик для сравнения много-  
 35 мерных временных рядов, а именно: MDTW и L2.

36 Эксперимент проводится в 2 этапа.

37 **1.Искусственно сгенерированные данные**, на которых можно увидеть достоин-  
 38 ства и недостатки работы алгоритмов и предложить из улучшение.

39 **2.Реальные данные**, на которых можно протестировать полученные алгоритмы и  
 40 измерить качество.

41 Генерация многомерных временных рядов из 1 шага проводится с помощью функций  
 42 с отчётливым пиком:

$$43 \quad \sin(x)$$

$$44 \quad x^2$$

45 плотность нормального распределения

46 В дальнейшем следует сдвиг по временной шкале семества этих функций и некоторый  
 47 небольшой сдвиг максимума по шкале пространства.

$$f(x) \rightarrow f(x + a)$$

48 Также будем масштабировать значения

$$f(x) \rightarrow f(ax)$$

49 Суть 2 этапа заключается в проверке полученных на 1 шаге алгоритмов на реальных  
 50 данных и измерением точности.

51 В итоге будет выбран лучший алгоритм.

### 52 **Литература**

- 53 [1] Gineke ten Holt, Marcel Reinders, and Emile Hendriks. Multi-dimensional dynamic time warping  
 54 for gesture recognition. *Annual Conference of the Advanced School for Computing and Imaging*, 01  
 55 2007.
- 56 [2] Jörg P. Bachmann and Johann-Christoph Freytag. High dimensional time series generators. *CoRR*,  
 57 abs/1804.06352, 2018.
- 58 [3] Nurjahan Begum, Liudmila Ulanova, Hoang Anh Dau, Jun Wang, and Eamonn J. Keogh. A general  
 59 framework for density based time series clustering exploiting a novel admissible pruning strategy.  
 60 *CoRR*, abs/1612.00637, 2016.
- 61 [4] P. Sanguansat. Multiple multidimensional sequence alignment using generalized dynamic time  
 62 warping. *WSEAS Transactions on Mathematics*, 11:684–694, 08 2012.

63 *Поступила в редакцию*