

Metric time series classification using weighted dynamic warping relative to centroids of classes

A. V. Goncharov¹, V. V. Strijov²

1

Abstract

This paper discusses a problem of metric time series analysis and classification. The proposed classification model uses the matrix of distances between time series which is built with fixed distance function. The dimension of this distance matrix is very high and all related calculations are time-consuming. The problem of reducing the computational complexity is solved by selection reference objects and using them for describing classes. Model that uses dynamic time warping for building reference objects or centroids is chosen as a basic model. This paper introduces a function of weights for each centroid that influence on calculating the distance measure. Time series of different analytic functions and time series of human activity from accelerometer of mobile phone are used as the objects for classification. Properties and classification result of this model are investigated and compared with properties of basic model.

Keywords: *metric classification, weighted dynamic time warping, time series classification, centroid, distance function.*

References

- [1] *Kuznecov M. P., and N. P. Ivkin.* 2015. Algoritm klassifikacii vremennyh rjadov akselerometra po kombinirovannomu priznakovomu opisaniju [Time series classification algorithm using combined feature description]. Mashinnoe obuchenie i analiz dannyh [Machine Learning and Data Analysis], Vol. 1. No. 13. P. 1471 – 1483.
- [2] *Popova M. S., and V. V. Strijov.* 2015. Vybor optimal'noj modeli klassifikacii fizicheskoy aktivnosti po izmerenijam akselerometra [Selection of optimal physical activity classification model using measurements of accelerometer]. Informatika i ee primenija [Informatics and applications], Vol. 9. No. 1. P. 79–89.

¹1 — Moscow Institute of Physics and Technology, alex.goncharov@phystech.edu

²2 — Dorodnicyn Computing Centre, Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences, strijov@ccas.ru

- [3] *Ignatov A. D. and V. V. Strijov.* 2015. Human activity types recognition using quasiperiodic sets of time series. *Multimedia tools and applications*, P. 1–14.
- [4] *Faloutsos C., M. Ranganathan, and Y. Manolopoulos.* 1994. Fast Subsequence Matching in Time-Series Databases. *ACM International Conference on Management of Data (SIGMOD)*, P. 419–429.
- [5] *Berndt D. J., and J. Clifford.* 1994. Using dynamic time warping to find patterns in time series. *Workshop on Knowledge Discovery in Databases (KDD Workshop)*, P. 359–370.
- [6] *Keogh E. J., and C. A. Ratanamahatana.* 2005. Exact indexing of dynamic time warping. *Knowl. Inf. Syst.*, Vol. 7. No. 3. P. 358–386.
- [7] *Vlachos M., D. Gunopulos, and G. Kollios.* 2002. Discovering similar multidimensional trajectories. *IEEE International Conference on Data Engineering (ICDE)*, P. 673–684.
- [8] *Chen L., and R. T. Ng.* 2004. On the marriage of lp-norms and edit distance. *Very Large Data Bases (VLDB)*, P. 792–803
- [9] *Chen L., M. T. Özsu, and V. Oria.* 2005. Robust and fast similarity search for moving object trajectories. *ACM International Conference on Management of Data (SIGMOD)*, P. 491–502.
- [10] *Frentzos E., K. Gratsias, and Y. Theodoridis.* 2007. Index-based most similar trajectory search. *IEEE International Conference on Data Engineering (ICDE)*, P. 816–825.
- [11] *Morse M. D., and J. M. Patel.* 2007. An efficient and accurate method for evaluating time series similarity. *ACM International Conference on Management of Data (SIGMOD)*, P. 569–580.
- [12] *Chen Y., M. A. Nascimento, B. C. Ooi, and A. K. H. Tung.* 2007. SpADe: On Shape-based Pattern Detection in Streaming Time Series. *IEEE International Conference on Data Engineering (ICDE)*, P. 786–795.
- [13] *Keogh E. J., and M. J. Pazzani.* 1999. Scaling up Dynamic Time Warping to Massive Datasets. *Principles of Data Mining and Knowledge Discovery*. Prague: Springer Berlin Heidelberg, P. 1–11.
- [14] *Salvador S., and P. Chan.* 2004. Fastdtw: Toward accurate dynamic time warping in linear time and space. *Workshop on Mining Temporal and Sequential Data*, P. 11.
- [15] *Goncharov A. V., M. S. Popova, and V. V. Strijov.* 2015. Metricheskaja klassifikacija vremennyh rjadov s vyravnivaniem odnositel'no centroidov klassov [Metric time series classification using dynamic warping relative to centroids of classes]. *Sistemy i sredstva informatiki [Systems and Means of informatics]*, Vol. 25. No. 4. P. 52–64.

- [16] *Petitjean F., G. Forestier, G.I. Webb, A. E. Nicholson, Y. Chen, and E. Keogh.* 2014. Dynamic Time Warping Averaging of Time Series allows Faster and more Accurate Classification. IEEE International Conference on Data Engineering (ICDE), P. 470–479.
- [17] Data from accelerometer. Available at: http://sourceforge.net/p/mlalgorithms/TSLearning/data/preprocessed_large.csv (accessed November 15, 2015).

Метрическая классификация временных рядов со взвешенным выравниванием относительно центроидов классов*

А. В. Гончаров¹, В. В. Стрижов²

Аннотация: В работе рассматривается задача метрического анализа и классификации временных рядов. Метрические методы используют матрицу попарных расстояний, строящуюся при помощи фиксированной функции расстояния. Вычислительная сложность алгоритмов, использующих такую матрицу, по меньшей мере квадратична относительно числа временных рядов. Проблема снижения вычислительной сложности решается путем предварительного выделения эталонных объектов, центроидов классов, и последующего их использования для описания классов. В качестве базовой модели классификации выбрана модель, использующая динамическое выравнивание временных рядов для построения центроида. В работе предлагается ввести функцию весов центроида, влияющую на вычисление расстояния между объектами. Для анализа алгоритма построения центроида использованы как временные ряды элементарных функций, так и временные ряды физической активности человека с акселерометра мобильного телефона. Свойства построенной модели исследуются и сравниваются со свойствами модели, выбранной в качестве базовой.

Ключевые слова: взвешенное динамическое выравнивание; классификация временных рядов; центроид; функция расстояния

1 Введение

Рассматривается задача анализа и классификации временных рядов. Существуют различные способы ее решения: построение признакового пространства, использование

*Работа выполнена при финансовой поддержке РФФИ (проект 16-07-01163) и в рамках инициативы «МИТ-Сколтех».

¹Московский физико-технический институт, alex.goncharov@phystech.edu

²Вычислительный центр им. А. А. Дородницына Федерального исследовательского центра «Информатика и управление» Российской академии наук, strijov@ccas.ru

нейронных сетей, аппроксимация параметрическими функциями. Так, в [1] исследованы методы построения признакового описания временных рядов, в частности метод экспертного построения признаков и метод построения признакового описания на основе гипотезы порождения данных. Результаты [1] показывают, что построенное признаковое пространство адекватно описывает зависимую переменную. В [2] для решения задачи классификации использованы нейронные сети с небольшим числом связей между нейронами, обладающие свойством устойчивости к возмущениям данных. В [3] для классификации предложен алгоритм разбиения исходных временных рядов на периоды и их очистки от шумов. Предложены модификации алгоритма k ближайших соседей и нейронной сети для решения поставленной задачи. В вычислительном эксперименте, проведенном на реальных данных, оценена эффективность, а также проведено сравнение данных алгоритмов между собой. При этом показан высокий процент правильной классификации.

Построение матрицы попарных расстояний между всеми объектами в задаче метрической классификации является вычислительно трудоемкой задачей. Для снижения размерности задачи и вычислительных затрат предлагается решать задачу с предварительным выделением эталонных объектов, или же центроидов классов, и последующим их использованием для описания множества временных рядов.

Метрические методы используют различные функции расстояния для построения матрицы попарных расстояний: евклидово расстояние [4], метод динамического выравнивания временных рядов [5, 6], метод, основанный на нахождении наибольшей общей последовательности [7], Edit Distance with Real Penalty [8], Edit Distance on Real sequence [9], DISSIM [10], Sequence Weighted Alignment model [11], Spatial Assembling Distance [12] и другие. В качестве базового метода для построения функции расстояния в настоящей работе предлагается использовать динамическое выравнивание временных рядов (Dynamic Time Warping) [13]. Как показано в [14], этот метод находит наилучшее соответствие между двумя временными рядами, если они нелинейно деформированы друг относительно друга — растянуты, сжаты или смещены вдоль оси времени.

Базовой моделью классификации и анализа временных рядов в текущей работе принята модель, описанная в [15]. Там в качестве центроида выбирается объект выборки, являющийся ближайшим ко всем остальным объектам. Применяться же будет метод точного его вычисления. Это метод DBA, решающий задачу оптимизации для нахождения центроида. Алгоритм применения и доказательство корректности приведены в [16].

В настоящей работе вводится понятие вектора весов и матрицы весов центроида и описываются методы $vwDTW$ и $mwDTW$ (векторно- и матрично-взвешенный DTW) вычисления функции расстояния, основывающиеся на следующем предположении о форме временных рядов: в одном и том же классе находятся временные ряды, имеющие схожую форму с точностью до линейной или нелинейной деформации, локальных или глобальных сдвигов по оси времени. Предполагается, что в центроиде присутствуют характерные для всего класса участки, которым соответствуют большие веса вектора весов этого центроида. А функция расстояния, основанная на $vwDTW$

и использующая вектор весов центроида, точнее объединит объекты одного класса и разделит объекты разных классов, чем основанная на DTW. Поэтому в предлагаемой модели используется метод vwDTW как для вычисления центроида по методу DBA, так и для построения матрицы попарных расстояний. Также исследуются свойства, вид матрицы весов центроида и эффективность применения расстояния, вычисленного с ее помощью — mwDTW, в прикладных задачах.

Для дальнейшей классификации рядов по полученной матрице расстояний сниженной размерности применяется метод k ближайших соседей, как и в базовой модели. Процедура классификации выполняется в три шага. Первый — отбор эталонных объектов каждого класса. Второй — построение матрицы попарных расстояний сниженной размерности между временными рядами и эталонными объектами каждого класса. Третий — классификация временных рядов методом k ближайших соседей с помощью матрицы попарных расстояний.

Для проверки работоспособности такой модели проведен вычислительный эксперимент на реальных и синтетических данных. Эксперимент включает в себя анализ и классификацию данных при помощи построенной модели. Полученные результаты сравниваются с результатами применения базовой модели к тем же исходным данным.

2 Постановка задачи

Задана выборка $\mathcal{D} = \{(\mathbf{s}_i, y_i)\}_{i=1}^m$, состоящая из пар объект — ответ. Объектами служат временные ряды $\mathbf{s}_i \in \mathbb{R}^n$, а ответами являются метки класса — $y_i \in Y = \{1, \dots, E\}$, где $E \ll m$. Выборка разбита на обучающую \mathcal{D}_l и контрольную \mathcal{D}_t .

Определение 1. Модель классификации f — параметрическая функция объектов выборки, приближающая целевую зависимость y_i . В данной работе параметрами модели примем множество центроидов $\mathbf{C} = \{\mathbf{c}_e\}_{e=1}^E$ и множество векторов весов центроидов $\hat{\mathbf{W}} = \{\mathbf{w}_e\}_{e=1}^E$ или же матриц весов центроидов $\hat{\mathbf{W}} = \{\mathbf{W}_e\}_{e=1}^E$.

Определение 2. Функцией ошибки S модели f для задачи классификации будем считать

$$S(f, \mathcal{D}_t) = \frac{1}{|\mathcal{D}_t|} \sum_{i=1}^{|\mathcal{D}_t|} [f(\mathbf{s}_i) \neq y_i].$$

Требуется построить модель классификации $f : \mathbb{R}^n \rightarrow Y$, минимизирующую функцию ошибки S на контрольной выборке:

$$f_{\mathbf{C}, \hat{\mathbf{W}}} = \operatorname{argmin}_{\mathbf{C}, \hat{\mathbf{W}}} (S(f, \mathcal{D}_t)). \quad (1)$$

3 Вычисление значения функции расстояния

3.1 Общие понятия

В данной работе в качестве метрического расстояния между объектами предлагается использовать стоимость *взвешенного пути наименьшей стоимости* между этими объектами.

Даны два временных ряда: \mathbf{s}_1 и \mathbf{s}_2 . Будем считать, что $\mathbf{s}_1, \mathbf{s}_2 \in \mathbb{R}^n$. Пусть $\Omega^{n \times n}$ — это матрица, такая что ее элемент Ω_{ij} равен квадрату разности между i -м и j -м элементами последовательностей \mathbf{s}_1 и \mathbf{s}_2 :

$$\Omega_{ij} = (s_{1i} - s_{2j})^2.$$

Определение 3. Путем π между последовательностями \mathbf{s}_1 и \mathbf{s}_2 назовем упорядоченное множество пар индексов элементов матрицы Ω :

$$\pi = \{\pi_r\} = \{(i_r, j_r)\}, \quad r = 1, \dots, R, \quad i, j \in \{1, \dots, n\},$$

где R — длина пути, зависящая от выбора пути. Он должен удовлетворять следующим условиям.

Граничные условия. $\pi_1 = (1, 1)$ и $\pi_R = (n, n)$, т. е. начало и конец π находятся на диагонали в противоположных углах Ω .

Непрерывность. Пусть $\pi_r = (p_1, p_2)$ и $\pi_{r-1} = (q_1, q_2)$, $r = 2, \dots, R$. Тогда

$$p_1 - q_1 \leq 1, \quad p_2 - q_2 \leq 1.$$

Это ограничение нужно, чтобы в шаге пути π участвовали только соседние элементы матрицы (включая соседние по диагонали).

Монотонность. Пусть $\pi_r = (p_1, p_2)$ и $\pi_{r-1} = (q_1, q_2)$, $r = 2, \dots, R$. Тогда выполняется хотя бы одно из условий

$$p_1 - q_1 \geq 1, \quad p_2 - q_2 \geq 1.$$

Это ограничение обусловлено природой рассматриваемых последовательностей и предназначено для монотонности функции выравнивания времени.

Физические ограничения. Как уже говорилось во введении, предполагается, что временные ряды одного класса имеют схожую форму и являются линейно или нелинейно деформированными или же смещены друг относительно друга. При этом считается, что подобного рода деформации и смещения являются малыми, локальными. В этом предположении выравнивающий путь в матрице слабо отклоняется от диагонали, то есть

$$\text{для каждого } \{i_r, j_r\} \in \pi \quad i_r - k \leq j_r \leq i_r + k,$$

где k определяется типом задачи и ее физическими ограничениями.

3.2 Векторно-взвешенный путь наименьшей стоимости

Дадим определение векторно-взвешенного пути наименьшей стоимости, vwDTW. Дан вектор весов $\mathbf{w} \in \mathbb{R}^n$.

Определение 4. *Стоимостью $Cost(\mathbf{s}_1, \mathbf{s}_2, \mathbf{w}, \pi)$ векторно-взвешенного пути π между последовательностями \mathbf{s}_1 и \mathbf{s}_2 с весом \mathbf{w} назовем*

$$Cost(\mathbf{s}_1, \mathbf{s}_2, \mathbf{w}, \pi) = \sum_{(i,j) \in \pi} w_j \Omega_{ij}. \quad (2)$$

Определение 5. *Векторно-взвешенным путем наименьшей стоимости (векторно-взвешенным выравнивающим путем) $\hat{\pi}$ между последовательностями \mathbf{s}_1 и \mathbf{s}_2 назовем взвешенный путь, имеющий наименьшую стоимость среди всех возможных векторно-взвешенных путей между последовательностями \mathbf{s}_1 и \mathbf{s}_2 :*

$$\hat{\pi} = \underset{\pi}{\operatorname{argmin}} Cost(\mathbf{s}_1, \mathbf{s}_2, \mathbf{w}, \pi). \quad (3)$$

Обозначим стоимость векторно-взвешенного выравнивающего пути между последовательностями \mathbf{s}_1 и \mathbf{s}_2 через $\rho(\mathbf{s}_1, \mathbf{s}_2, \mathbf{w}) = Cost(\mathbf{s}_1, \mathbf{s}_2, \mathbf{w}, \hat{\pi})$.

Для вычисления стоимости такого пути в данной работе используется модифицированный метод DTW — vwDTW (vector-weighted DTW). Согласно этому методу необходимо построить новую матрицу Γ , элементы которой определяются следующим образом:

$$\begin{aligned} \Gamma_{1j} &= w_j \Omega_{1j}, & \Gamma_{i1} &= w_1 \Omega_{i1}, & i, j &= 1, \dots, n, \\ \Gamma_{ij} &= w_j \Omega_{ij} + \min(\Gamma_{i,j-1}, \Gamma_{i-1,j}, \Gamma_{i-1,j-1}), & i, j &= 2, \dots, n. \end{aligned}$$

Элемент Γ_{ij} матрицы Γ равен стоимости векторно-взвешенного выравнивающего пути между последовательностями $\{s_{1a}\}_{a=1}^i$ и $\{s_{2a}\}_{a=1}^j$.

В качестве значения функции расстояния между двумя объектами выберем стоимость векторно-взвешенного выравнивающего пути между ними (3):

$$\rho(\mathbf{s}_1, \mathbf{s}_2, \mathbf{w}) = \Gamma_{nn}. \quad (4)$$

Заметим, что при единичном векторе весов vwDTW эквивалентен обычному DTW, описание которого приведено в [15].

3.3 Матрично-взвешенный путь наименьшей стоимости

Дадим определение матрично-взвешенного пути наименьшей стоимости, mwDTW. Дана матрица весов $\mathbf{W} \in \mathbb{R}^{n \times n}$.

Определение 6. *Стоимостью $Cost(\mathbf{s}_1, \mathbf{s}_2, \mathbf{W}, \pi)$ матрично-взвешенного пути π между последовательностями \mathbf{s}_1 и \mathbf{s}_2 с весом \mathbf{W} назовем*

$$Cost(\mathbf{s}_1, \mathbf{s}_2, \mathbf{W}, \pi) = \sum_{(i,j) \in \pi} W_{ij} \Omega_{ij}. \quad (5)$$

Определение 7. Матрично-взвешенным путем наименьшей стоимости (матрично-взвешенным выравнивающим путем) $\hat{\pi}$ между последовательностями \mathbf{s}_1 и \mathbf{s}_2 назовем взвешенный путь, имеющий наименьшую стоимость среди всех возможных матрично-взвешенных путей между последовательностями \mathbf{s}_1 и \mathbf{s}_2 :

$$\hat{\pi} = \underset{\pi}{\operatorname{argmin}} \operatorname{Cost}(\mathbf{s}_1, \mathbf{s}_2, \mathbf{W}, \pi). \quad (6)$$

Обозначим стоимость матрично-взвешенного выравнивающего пути между последовательностями \mathbf{s}_1 и \mathbf{s}_2 через $\rho(\mathbf{s}_1, \mathbf{s}_2, \mathbf{W}) = \operatorname{Cost}(\mathbf{s}_1, \mathbf{s}_2, \mathbf{W}, \hat{\pi})$.

Вычисление его стоимости происходит с помощью еще одного модифицированного метода DTW — mwDTW (matrix-weighted DTW). Согласно этому методу необходимо построить новую матрицу Γ , элементы которой определяются следующим образом:

$$\begin{aligned} \Gamma_{1j} &= W_{1j}\Omega_{1j}, & \Gamma_{i1} &= W_{i1}\Omega_{i1}, & i, j &= 1, \dots, n, \\ \Gamma_{ij} &= W_{ij}\Omega_{ij} + \min(\Gamma_{i,j-1}, \Gamma_{i-1,j}, \Gamma_{i-1,j-1}), & i, j &= 2, \dots, n. \end{aligned}$$

Элемент Γ_{ij} матрицы Γ равен стоимости матрично-взвешенного выравнивающего пути между последовательностями $\{s_{1a}\}_{a=1}^i$ и $\{s_{2a}\}_{a=1}^j$.

В качестве значения функции расстояния между двумя объектами выберем стоимость матрично-взвешенного выравнивающего пути между ними (6):

$$\rho(\mathbf{s}_1, \mathbf{s}_2, \mathbf{W}) = \Gamma_{nn}. \quad (7)$$

Заметим, что при использовании матрицы, состоящей из одних единиц, mwDTW переходит в обычный DTW, описание которого приведено в [15].

4 Вычисление параметров модели классификатора

4.1 Построение центроида

Пусть множество весов $\hat{\mathbf{W}}$ фиксировано. Построим множество центроидов \mathbf{C} .

4.1.1 Постановка задачи построения центроида

Определение 3. Пусть \mathcal{D}_e — множество элементов из \mathcal{D} , принадлежащих одному классу e из Y . Центроидом множества векторов $\mathcal{D}_e = \{\mathbf{s}_i | y_i = e\}_{i=1}^m$ по расстоянию ρ назовем вектор $\mathbf{c}_e \in \mathbb{R}^n$ такой, что

$$\mathbf{c}_e = \underset{\mathbf{c} \in \mathbb{R}^n}{\operatorname{argmin}} \sum_{\mathbf{s}_i \in \mathcal{D}_e} \rho(\mathbf{s}_i, \mathbf{c}),$$

где ρ — стоимость векторно- (матрично-) взвешенного пути наименьшей стоимости vwDTW (mvDTW).

Центроид найдем как решение оптимизационной задачи для vwDTW

$$\mathbf{c}_e = \operatorname{argmin}_{\mathbf{c} \in \mathbb{R}^n} \sum_{\mathbf{s}_i \in \mathcal{D}_e} \sum_{(t,t') \in \hat{\boldsymbol{\pi}}_i} \mathbf{w}_e(t) (\mathbf{s}_i(t') - \mathbf{c}(t))^2. \quad (8)$$

Или же для mwDTW:

$$\mathbf{c}_e = \operatorname{argmin}_{\mathbf{c} \in \mathbb{R}^n} \sum_{\mathbf{s}_i \in \mathcal{D}_e} \sum_{(t,t') \in \hat{\boldsymbol{\pi}}_i} \mathbf{W}_e(t,t') (\mathbf{s}_i(t') - \mathbf{c}(t))^2, \quad (9)$$

где $\hat{\boldsymbol{\pi}}_i$ — векторно- (матрично-) взвешенный выравнивающий путь между временными рядами \mathbf{s}_i и \mathbf{c} .

4.1.2 Решение задачи нахождения центроида методом DBA

Теорема 1 [16]. Пусть дано множество векторов $\mathcal{D}_e = \{\mathbf{s}_i | y_i = e\}_{i=1}^m$ одного класса, начальное приближение центроида \mathbf{c}_e и множество выравнивающих путей между каждым рядом и начальным приближением центроида $\{\hat{\boldsymbol{\pi}}_i\}_{i=1}^m$. Тогда локальный минимум задачи оптимизации (8) при единичном векторе весов \mathbf{v} в (4) (функция расстояния DTW) достигается при

$$\mathbf{c}_e(t) = \frac{1}{N} \sum_{\mathbf{s}_i \in \mathcal{D}_e} \sum_{t':(t,t') \in \hat{\boldsymbol{\pi}}_i} \mathbf{s}_i(t'), \quad (10)$$

$$N = \sum_{\mathbf{s}_i \in \mathcal{D}_e} \sum_{t':(t,t') \in \hat{\boldsymbol{\pi}}_i} 1.$$

Доказательство. Для поиска центроида и решения задачи оптимизации (8) воспользуемся необходимым условием экстремума. Запишем частные производные функционала по $\mathbf{c}_e(t)$, $t = 1, \dots, T$, и приравняем их к 0:

$$\frac{\partial F(\mathbf{c}_e, \mathcal{D}_e)}{\partial \mathbf{c}_e(t)} = \sum_{\mathbf{s}_i \in \mathcal{D}_e} \sum_{t':(t,t') \in \hat{\boldsymbol{\pi}}_i} 2(\mathbf{c}_e(t) - \mathbf{s}_i(t')) = 0. \quad (11)$$

Откуда и находим значение $\mathbf{c}_e(t)$, $t = 1, \dots, n$:

$$\mathbf{c}_e(t) = \frac{1}{N} \sum_{\mathbf{s}_i \in \mathcal{D}_e} \sum_{t':(t,t') \in \hat{\boldsymbol{\pi}}_i} \mathbf{s}_i(t'). \quad (12)$$

Данный метод вычисления центроида приведен в [16] и называется методом DBA. Там же приведено и его доказательство. При нахождении нового центроида множество выравнивающих рядов меняется, данную процедуру нужно проводить несколько раз, пока центроид не стабилизируется. При замене единичного вектора весов в функции расстояния vwDTW на произвольный справедливо следующее

Следствие 1. При использовании произвольного вектора весов центроида \mathbf{w} (замене DTW на $vwDTW$ с вектором весов \mathbf{w}) в задаче оптимизации (8) алгоритм DBA вычисления центроида находит локальный минимум при замене множества путей наименьшей стоимости $\{\tilde{\pi}_i\}_{i=1}^m$ на множество взвешенных путей наименьшей стоимости $\{\hat{\pi}_i\}_{i=1}^m$.

Доказательство данного следствия повторяет доказательство теоремы 1 при замене множества путей наименьшей стоимости $\{\hat{\pi}_i\}_{i=1}^m$ на множество взвешенных путей наименьшей стоимости $\{\pi_i\}_{i=1}^m$.

Для функции расстояния $mwDTW$ следствие сохраняет свою формулировку и доказательство при замене вектора весов центроида на его матрицу весов \mathbf{W} .

4.2 Оптимизация и ограничения вектора весов

Положим теперь множество центроидов \mathbf{C} фиксированным. Каждому центроиду \mathbf{c}_e из множества \mathbf{C} поставлен в соответствие вектор неотрицательных весов \mathbf{w}_e , принадлежащий множеству $\hat{\mathbf{W}}$. Значения данного вектора весов выделяют наиболее типичные для класса участки центроида, сопоставляя им большие веса. Вычислим этот вектор, решая задачу оптимизации

$$\mathbf{w}_e = \operatorname{argmin}_{\mathbf{w} \in \mathbb{R}^n} \sum_{\mathbf{s}_i \in \mathcal{D}_e} \sum_{(t, t') \in \pi_i} \mathbf{w}(t) (\mathbf{s}_i(t') - \mathbf{c}_e(t))^2. \quad (13)$$

При отсутствии ограничений на веса \mathbf{w}_e минимум (13) достигается при $\mathbf{w}_e = \mathbf{0}$. Для того чтобы избежать такого тривиального решения, введем ограничения на сумму элементов вектора весов

$$\sum_{t=1}^T \mathbf{w}_e(t) = T.$$

Предположим, что при решении задачи (13) нашлось t , для которого выполняется

$$\sum_{\mathbf{s}_i \in \mathbf{D}} \sum_{t': (t, t') \in \pi_i} (\mathbf{s}_i(t') - \mathbf{c}_e(t))^2 = 0.$$

Для таких t элемент решения задачи оптимизации (13) примет большие значения, которые обеспечат выполнение ограничений на сумму элементов. Это приведет к локальному скоплению больших значений вектора весов, что сильно ухудшит дальнейшую интерпретацию вектора весов, а также сделает метод чувствительным к малым изменениям входных данных. Поэтому введем ограничения на элементы вектора весов сверху:

$$\mathbf{w}_e(t) \leq \text{const}, t \in \{1, \dots, T\}.$$

Таким образом, исходная задача (13) примет следующий вид:

$$\mathbf{w}_e = \operatorname{argmin}_{\mathbf{w} \in \mathbb{R}^n} \sum_{\mathbf{s}_i \in \mathcal{D}_e} \sum_{(t, t') \in \pi_i} \mathbf{w}_e(t) (\mathbf{s}_i(t') - \mathbf{c}_e(t))^2, \quad (14)$$

$$\sum_{t=1}^T \mathbf{w}_e(t) = T, \quad 0 \leq \mathbf{w}_e(t) \leq \text{const}, \quad t \in \{1, \dots, T\},$$

где const — некоторая заданная константа.

4.3 Оптимизация матрицы весов

4.3.1 Ограничения матрицы весов

Как и при оптимизации вектора весов положим множество центроидов \mathbf{C} фиксированным. Каждому центроиду \mathbf{c}_e некоторого класса e из множества \mathbf{C} сопоставлена матрица неотрицательных весов \mathbf{W}_e , принадлежащая множеству $\hat{\mathbf{W}}$.

Используя те же соображения, что и для случая использования расстояния vwDTW , определим задачу нахождения матрицы весов центроида как задачу оптимизации с ограничениями при использовании расстояния mwDTW :

$$\mathbf{W}_e = \underset{\mathbf{W} \in \mathbb{R}^{n \times n}}{\text{argmin}} \sum_{\mathbf{s}_i \in \mathcal{D}_e} \sum_{t, t' \in \pi_i} \mathbf{W}_e(t, t') (\mathbf{s}_i(t') - \mathbf{c}_e(t))^2, \quad (15)$$

$$\sum_{t=1}^T \sum_{t'=1}^T \mathbf{W}_e(t, t') = T^2, \quad 0 \leq \mathbf{W}_e(t) \leq \text{const}, \quad t \in \{1, \dots, T\},$$

где const — некоторая заданная константа.

4.3.2 Сглаживание полученной матрицы

Полученная матрица не является устойчивой к изменению входных данных: при использовании других временных рядов выравнивающие пути будут иметь другой вид, а значит, и решение задачи оптимизации будет другое. Более устойчивой матрица получится после процедуры сглаживания.

Будем говорить, что элемент матрицы $\mathbf{W}_e(t, t')$ содержится в множестве Φ , если существует временной ряд $\mathbf{s}_i \in \mathcal{D}_e$ такой, что путь наименьшей матрично-взвешенной стоимости проходит через элемент $\Omega(t, t')$ в матрице Ω , построенной для временного ряда \mathbf{s}_i и центроида.

При решении задачи оптимизации изменяться в меньшую сторону будут элементы $\{\mathbf{W}_e(t, t')\} \in \Phi$. При этом $\mathbf{W}_e(t, t') \notin \Phi$ достигнет своей верхней границы для выполнения ограничений, накладываемых на сумму элементов матрицы.

Выберем произвольный элемент матрицы весов $\mathbf{W}_e(t, t') \notin \Phi$. Вероятность того, что при добавлении нового временного ряда (например из тестовой выборки) выполнится $\mathbf{W}_e(t, t') \in \Phi$, выше, если среди ближайших к $\mathbf{W}_e(t, t')$ элементов в строке матрицы весов многие содержатся в множестве Φ . Тогда значение такого элемента должно быть похожим на значения соседних. Добьемся этого, выполнив сглаживание матрицы весов центроида:

$$\widetilde{\mathbf{W}}_e(t, t') = \frac{1}{2\delta} \sum_{k=-\delta}^{\delta} \mathbf{W}_e(t, t' + k),$$

где δ — величина окна сглаживания, а $\widetilde{\mathbf{W}}_e$ — искомая матрица весов центроида.

4.4 Задача оптимизации параметров модели

Задача оптимизации параметров модели сведена к комбинации задач оптимизации (8) и (14), (15) для vwDTW:

$$\mathbf{w}_e, \mathbf{c}_e = \operatorname{argmin}_{\mathbf{c}, \mathbf{w} \in \mathbb{R}^n} \sum_{\mathbf{s}_i \in \mathcal{D}_e(t, t') \in \pi_i} (\mathbf{w}(t)(\mathbf{s}_i(t') - \mathbf{c}(t))^2), \quad e = 1, \dots, E, \quad (16)$$

$$\sum_{t=1}^T \mathbf{w}_e(t) = T, \quad 0 \leq \mathbf{w}_e(t) \leq \text{const}, \quad t \in \{1, \dots, T\}.$$

Или же для mwDTW:

$$\mathbf{W}_e, \mathbf{c}_e = \operatorname{argmin}_{\mathbf{W} \in \mathbb{R}^n \times \mathbb{R}^n} \sum_{\mathbf{s}_i \in \mathcal{D}_e(t, t') \in \pi_i} (\mathbf{W}_e(t, t')(\mathbf{s}_i(t') - \mathbf{c}_e(t))^2), \quad e = 1, \dots, E, \quad (17)$$

$$\sum_{t=1}^T \sum_{t'=1}^T \mathbf{W}_e(t, t') = T^2, \quad 0 \leq \mathbf{W}_e(t) \leq \text{const}, \quad t \in \{1, \dots, T\}.$$

Эту задачу будем решать, вычисляя сначала множество центроидов \mathbf{C} при фиксированном начальном приближении множества весов центроидов $\hat{\mathbf{W}}$, а затем вычисляя множество весов центроидов $\hat{\mathbf{W}}$ при фиксированном множестве центроидов \mathbf{C} . Таким образом, алгоритм вычисления параметров модели будет иметь следующий вид:

Шаг 1. Начальное приближение вектора весов центроида:

$$\mathbf{w}_e = \mathbf{1}, \quad e = 1, \dots, E.$$

Или же начальное приближение матрицы весов центроида:

$$\mathbf{W}_e = \mathbf{1}, \quad e = 1, \dots, E.$$

Шаг 2. Начальное приближение центроида класса — произвольный элемент класса:

$$\mathbf{c}_e = \mathbf{s}_j \in \mathcal{D}_e, \quad e = 1, \dots, E.$$

Шаг 4. Вычисление центроида при фиксированном векторе (матрице) весов центроида как решение задачи оптимизации (8).

Шаг 3. Вычисление вектора (матрицы) весов центроида при фиксированном центроиде как решение задачи оптимизации (14), (15).

5 Вычислительный эксперимент

Для проверки свойств введенной функции расстояния, а также выбранной модели был проведен вычислительный эксперимент на реальных и синтетических данных. Свойства функций расстояния vwDTW и mwDTW для наглядности продемонстрированы на синтетических данных, представляющих собой смещенные и линейно деформированные временные ряды аналитических функций: $\sin x$, \sqrt{x} , x^2 , — длиной 100 точек. В выборке находилось 100 временных рядов каждого класса: 50 в обучающей и 50 в контрольной. Обозначим классы как 1 — $\sin x$, 2 — \sqrt{x} , 3 — x^2 соответственно.

Пример такой выборки и результаты нахождения параметров модели по обучающей выборке показаны на рис. 1. По оси абсцисс отложены значения времени, а по оси ординат — значения временного ряда. На левом верхнем графике приведены примеры аналитических функций, используемых в создании синтетической выборки временных рядов. На правом верхнем графике изображены центроид (нижний временной ряд) и вектор весов (верхний временной ряд) для класса 1. Аналогично на нижних графиках показаны результаты для классов 2 и 3.

Векторы весов описывают наиболее информативные участки центроида. Так, для центроида класса 1 ($\sin x$) наиболее информативными оказались минимумы и максимумы, в отличие от точек перегиба.

Для сравнения свойств полученной функции расстояния с функцией расстояния DTW были посчитаны расстояния до центроидов для всех временных рядов контрольной выборки с помощью DTW и vwDTW , после чего производилась классификация. Каждому временному ряду контрольной выборки ставилась в соответствие метка того класса, расстояние до центроида которого было минимальным. Результат классификации с помощью функции расстояния vwDTW — 97%, а для функции расстояния DTW — 84%, что на 15% меньше. Построенная в работе функция расстояния лучше разделила временные ряды разных классов, сгруппировала их вокруг соответствующих центроидов.

Также были построены матрицы весов центроида на синтетических данных. Вследствие хорошей интерпретируемости, в работе приведено визуальное отображение матрицы весов центроида только для класса 1. На рис. 2 сверху слева изображены примеры временных рядов. Сверху справа — множество выравнивающих путей по матрице Ω . Нижний левый рисунок — матрица весов центроида до сглаживания. Правый левый рисунок — после сглаживания. Синие цвета соответствуют маленьким значениям элементов матрицы, а красные — большим. Хорошо просматривается периодичность в матрице, напоминающая периодичность вектора весов. Предполагается, что функция mwDTW будет лучше разделять классы, которые сильно различаются между собой, так как матрица весов центроида учитывает и среднее отклонение выравнивающего пути от диагонали в матрице Ω . При отклонениях пути сильнее типичного для данного класса, элементам пути будут приписываться большие веса, что видно из структуры матрицы весов центроида.

Использование функции mwDTW для классификации синтетических временных рядов улучшает классификацию по сравнению с DTW . Этот результат аналогичен

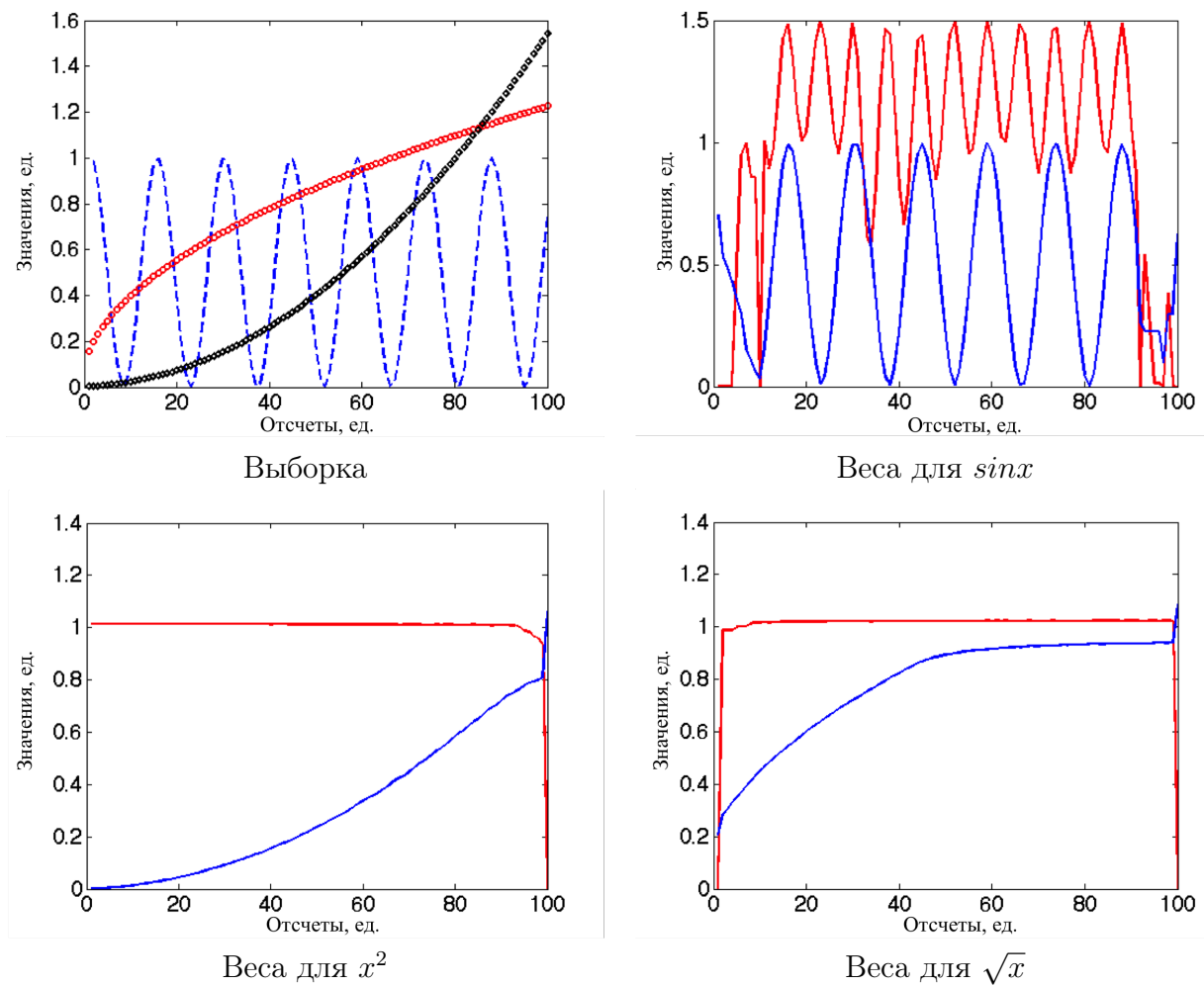


Рис. 1: Примеры синтетических временных рядов аналитических функций (выборка), результаты построения центроида и вектора весов для синтетических данных

случаю использования vwDTW — 97%.

Для демонстрации работы предложенной модели на реальных данных и ее сравнения с базовой моделью классификации были использованы данные акселерометра мобильного телефона. Вследствие большой вычислительной сложности они сравнивались на данных, представляющих собой 600 временных рядов длиной 200 точек, каждый из которых представляет собой абсолютные значения ускорения мобильного телефона, объединяя три временных ряда: временной ряд ускорения по оси X (200 измерений), оси Y (200 измерений) и оси Z (200 измерений). Выделено шесть типов физической активности: ходьба, бег, сидение, стояние, подъем, спуск. Временные ряды записывались акселерометром, который находился в кармане у человека, выполняющего один из типов физической активности, после чего разделялись на 10-секундные сегменты. Примеры таких временных рядов приведены на рис. 3.

Данные разделялись на обучающую и контрольную выборку. В обучающую выборку входило по 70 временных рядов каждого вида физической активности, а в контрольную — по 30 временных рядов. Производилась классификация методом k ближайших соседей, использующим построенную матрицу попарных расстояний. Осуществлялся контроль качества при помощи кросс-валидации. В табл. 1 приведены результаты классификации при использовании новой модели, использующей функцию vwDTW , и базовой модели классификации.

Качество классификации базовой модели ухудшилось в сравнении с вычислительным экспериментом в работе [15], так как теперь используются абсолютные значения ускорения, а не последовательно соединенные временные ряды ускорения вдоль трех координат. Использование функции расстояния vwDTW улучшает классификацию для всех классов физической активности, повышая общий уровень классификации на 9%.

Таблица 1: Сравнение эффективности предложенной (vwDTW) и базовой модели классификации на данных [17]

Модель	Бег	Ходьба	Вверх	Вниз	Сидение	Стояние	Общее
vwDTW .							
Точность по критерию скользящего контроля, %	97	95	79	75	95	95	89
DTW [15].							
Точность по критерию скользящего контроля, %	95	92	60	60	85	90	80

Полученные параметры модели для реальных данных изображены на рис. 4. Для класса бег, например, хорошо просматривается периодичность временного ряда как для центроида, так и для вектора весов центроида.

Для сравнения модели, использующей mvDTW , с базовой моделью классификации были использованы данные, представляющие собой 600 временных рядов длиной 100 точек. Временные ряды такой длины выбраны для разумного ограничения времени работы методов численной оптимизации. В остальном этот эксперимент повторяет тот, что выполнен для сравнения базовой модели и модели, использующей vwDTW . Результаты классификации приведены в табл. 2. Видно, что длина временных рядов, а следовательно, и количество периодов временного ряда, существенно влияют на качество классификации. Качество базовой модели сильно упало по сравнению с временными рядами длиной 200 точек. При этом использование mvDTW улучшает классификацию на 13%.

Таблица 2: Сравнение эффективности предложенной (mwDTW) и базовой модели классификации и алгоритма разделяющей классификации на данных [17]

Модель	Бег	Ходьба	Вверх	Вниз	Сидение	Стояние	Общее
mwDTW.							
Точность по критерию скользящего контроля, %	95	95	78	76	90	90	87
DTW [15].							
Точность по критерию скользящего контроля, %	88	83	55	60	82	80	74

6 Заключение

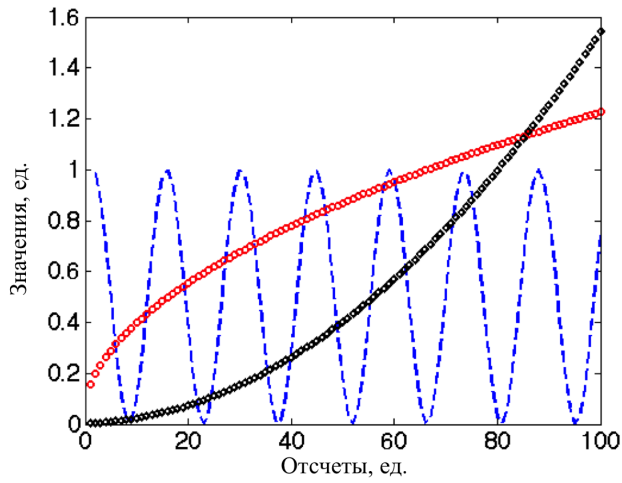
В работе описан новый подход к работе с центроидами временных рядов, построена модель, использующая веса и матрицы центроидов, и показаны ее преимущества перед моделью, описанной в работе [15]. Продемонстрировано влияние структуры, длины и физического смысла временных рядов на результаты классификации. В последующих работах данный подход будет совершенствоваться. Предполагается ускорить данный алгоритм за счет эффективного нахождения пути наименьшей стоимости, его аппроксимации. Такой подход применим и ко всем типам временных рядов, где возможна нелинейная деформация как по оси времени, так и по оси значений временного ряда.

Список литературы

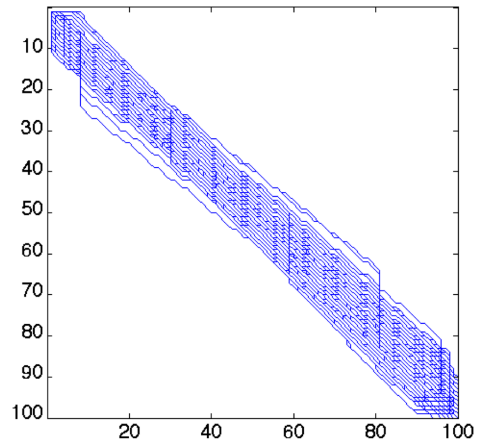
- [1] Кузнецов М. П., Ивкин Н. П. Алгоритм классификации временных рядов акселерометра по комбинированному признаковому описанию // Машинное обучение и анализ данных, 2015, Т. 1. Вып. 13. С. 1471–1483.
- [2] Попова М. С., Стрижов В. В. Выбор оптимальной модели классификации физической активности по измерениям акселерометра // Информатика и ее применения, 2015, Т. 9. Вып. 1. С. 79–89.
- [3] Ignatov A. D., Strijov V. V. Human activity types recognition using quasiperiodic sets of time series collected from a single tri-axial accelerometer // Multimedia tools and applications, 2015, Springer US, P. 1–14.
- [4] Faloutsos C., Ranganathan M., Manolopoulos Y. Fast Subsequence Matching in Time-Series Databases // SIGMOD International Conference on Management of Data, 1994, Minneapolis, ACM, P. 419–429.

- [5] *Berndt D. J., Clifford J.* Using dynamic time warping to find patterns in time series // Workshop on Knowledge Discovery in Databases, at the 12th Int'l Conference on Artificial Intelligence, 1994, Seattle, WA, P. 359–370.
- [6] *Keogh E. J., Ratanamahatana C. A.* Exact indexing of dynamic time warping // Knowl. Inf. Syst., 2005, Vol. 7. No. 3. P. 358–386.
- [7] *Vlachos M., Gunopulos D., Kollios G.* Discovering similar multidimensional trajectories // IEEE International Conference on Data Engineering, 2002, San Jose, IEEE Computer Society, P. 673–684.
- [8] *Chen L., Ng R. T.* On the marriage of lp-norms and edit distance // Very Large Data Bases (VLDB), 2004, Toronto, Morgan Kaufmann, P. 792–803
- [9] *Chen L., Özsu M. T., Oria V.* Robust and fast similarity search for moving object trajectories // ACM International Conference on Management of Data (SIGMOD), 2005, Baltimore, ACM, P. 491–502.
- [10] *Frentzos E., Gratsias K., Theodoridis Y.* Index-based most similar trajectory search // IEEE International Conference on Data Engineering (ICDE), 2007, Istanbul, IEEE Computer Society, P. 816–825.
- [11] *Morse M. D., Patel J. M.* An efficient and accurate method for evaluating time series similarity // ACM International Conference on Management of Data (SIGMOD), 2007, Beijing, ACM, P. 569–580.
- [12] *Chen Y., Nascimento M. A., Ooi B. C., Tung A. K. H.* SpADe: On Shape-based Pattern Detection in Streaming Time Series // IEEE International Conference on Data Engineering (ICDE), 2007, Istanbul, IEEE Computer Society, P. 786–795.
- [13] *Keogh E. J., Pazzani M. J.* Scaling up Dynamic Time Warping to Massive Datasets // Principles of Data Mining and Knowledge Discovery, 1999, Prague, Springer Berlin Heidelberg, P. 1–11.
- [14] *Salvador S., Chan P.* Fastdtw: Toward accurate dynamic time warping in linear time and space // Workshop on Mining Temporal and Sequential Data, 2004, Seattle, P. 70–80.
- [15] *Гончаров А. В., Попова М. С., Стрижов В. В.* Метрическая классификация временных рядов с выравниванием относительно центроидов классов // Системы и средства информатики, 2015, Т. 25. Вып. 4. С. 52–64.
- [16] *Petitjean F., Forestier G., Webb G. I., Nicholson A. E., Chen Y., Keogh E.* Dynamic Time Warping Averaging of Time Series allows Faster and more Accurate Classification // IEEE International Conference on Data Engineering (ICDE), 2014, Chicago, IEEE Computer Society, P. 470–479.

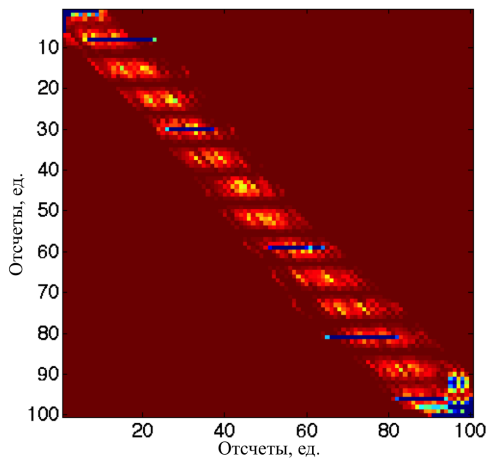
- [17] Data from accelerometer. Available at: http://sourceforge.net/p/mlalgorithms/TSLearning/data/preprocessed_large.csv (accessed November 15, 2015).



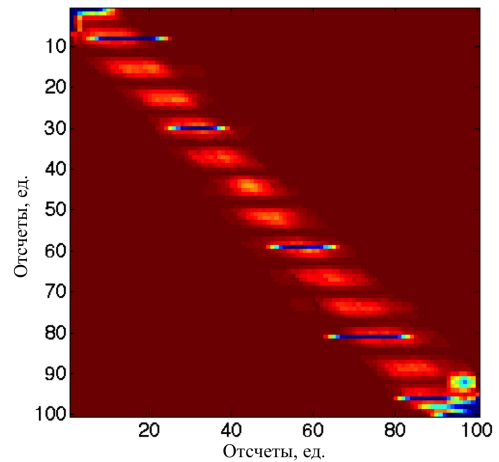
Выборка



Множество выравнивающих путей



Начальная матрица весов



Сглаженная матрица весов

Рис. 2: Примеры синтетических временных рядов аналитических функций (выборка), результат построения матрицы весов центроида класса $\sin(x)$

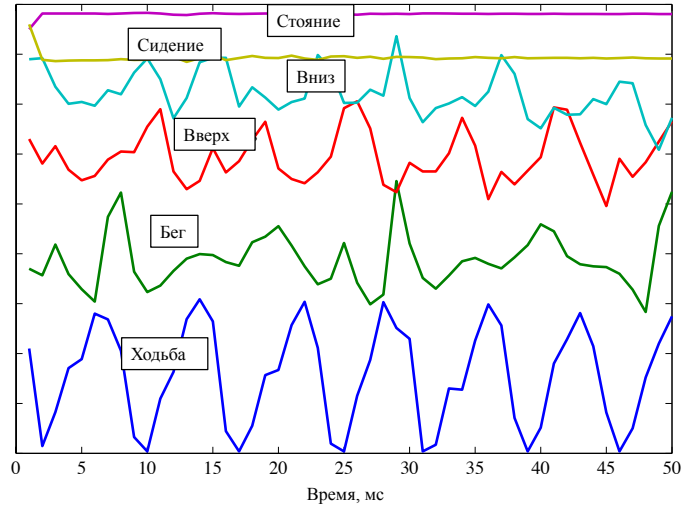


Рис. 3: Примеры временных рядов измерений акселерометра для разных видов физической активности

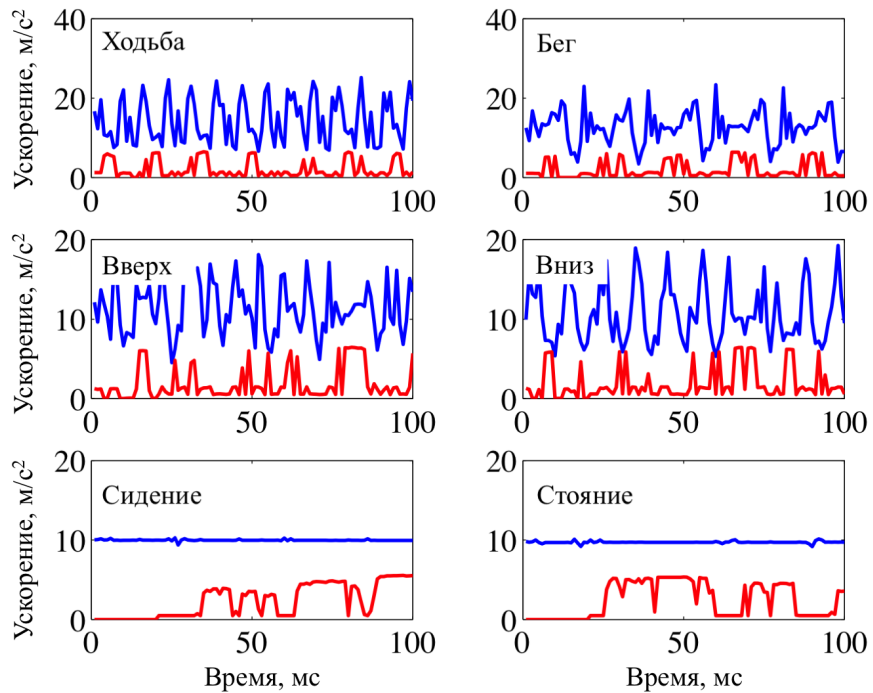


Рис. 4: Примеры временных рядов измерений акселерометра для разных видов физической активности