

## МЕТОД ОЦІНКИ МЕТРИК ПОДІБНОСТІ ЧАСОВИХ РЯДІВ

Ставиченко С.Г.<sup>1</sup>, Філатова Г.Є.<sup>2</sup>

<sup>1</sup> аспірант кафедри КІП, НТУ «ХПІ», Харків, Україна

<sup>2</sup> професор кафедри КІП, д-р техн. наук, НТУ «ХПІ», Харків, Україна

Метрики подібності часових рядів знаходять широке застосування при їх аналізі, наприклад, під час пошуку за зразком, кластеризації, виявленні аномалій, а також як функція втрат у задачах машинного навчання. Характеристики та поведінка функції подібності мають суттєвий вплив на кінцеву ефективність методу.

У цій роботі запропоновано метод оцінки такої характеристики як стійкість функції відстані до часових спотворень ряду. Суть методу полягає у плавному спотворенні часового ряду в часі, з подальшою оцінкою кореляції ступеня спотворення з результатами функції відстані, розрахованими для оригінального та спотвореного ряду.

Нехай  $X(t_i)$  - це часовий ряд, де  $t_i = t_0 + i \cdot \Delta t$  - часові мітки ряду з постійним кроком  $\Delta t$ . На базі  $t_i$  розраховуємо новий ряд часових міток із гармонічно спотвореним кроком:

$$\tau_i = \lambda \cdot \sin\left(t_0 + \frac{2\pi t_i}{P}\right) + t_i,$$

де  $P$  - період спотворення,  $\lambda$  - параметр, що характеризує ступінь спотворення.

Тоді спотворений ряд у початковій часовій сітці  $t_i$  можна отримати методом інтерполяції значення сигналу для двох найближчих часових міток  $\tau_i$ :

$$X^{t_i, \lambda} = X(\tau_k) + \frac{X(\tau_{k+1}) - X(\tau_k)}{\tau_{k+1} - \tau_k} \cdot (t_i - \tau_k),$$

де  $t_k \leq t_i \leq t_{k+1}$

Запропонована метрика оцінки  $R$  функції відстані  $D$  розраховується як кореляційний коефіцієнт між значеннями функції відстані для початкового та спотвореного ряду з різним значеннями параметру спотворення  $\lambda$  та самим параметром  $\lambda$ :

$$\begin{aligned} \lambda_k &= \lambda_0 + k \cdot \Delta\lambda; \\ D_k &= D\left([X(t_i) \mid t \in (0, n]], [\widehat{X}(t_i, \lambda) \mid t \in (0, n]]\right); \\ R_{\lambda D} &= r(\lambda, D), \end{aligned}$$

де  $r$  - коефіцієнт кореляції Пірсона.

Метод був застосований для оцінки функцій відстані між дискретними сигналами ЕКГ. Були розглянуті середньоквадратична (MSE) та середньомодульна (MAE) відстань, коефіцієнт кореляції Пірсона, індекс структурної подібності (SSIM) та відстань, отримана алгоритмом динамічної трансформації часової шкали (DTW Distance). Отримані результати показали, що SSIM має найкращий показник кореляції (0.94) зі ступенем спотворення в порівнянні з іншими метриками. DTW відстань має найбільш плавну зміну значення, але лише при не значних спотвореннях, що обумовлено наявністю різких піків сигналу.