

<b>Tytuł</b>	<i>Performance analysis of smartphone-sensor behavior for human activity recognition</i>
<b>Autorzy</b>	Yufei Chen, Chao Shen
<b>Data publikacji</b>	1 marca 2017
<b>Źródło</b>	<a href="http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&amp;arnumber=7867876">http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&amp;arnumber=7867876</a>

Praca traktuje o rozpoznawaniu aktywności człowieka za pomocą smartfonu. Dane zbierane są podczas wykonywania przez uczestników badania typowych codziennych aktywności. Autorzy, na podstawie analizy innych publikacji, przedstawiają w sposób szczegółowy podejście, jakie można zastosować podczas HAR (ang. *Human Activity Recognition*, rozpoznawanie aktywności człowieka) – od akwizycji danych, poprzez wstępne przetworzenie sygnałów, konstrukcję wektorów cech sygnałów, implementację klasyfikatorów po sposób rozróżnienia pomiędzy aktywnościami.

### 1. Badane osoby

- **Ilość:** 10 osób płci żeńskiej
- **Przedział wagowy:** 55 kg do 72 kg
- **Przedział wzrostowy:** 165 cm do 178 cm
- **Wiek:** 20 lat do 23 lat

### 2. Analizowane aktywności

- Schodzenie po schodach (z 8-go na 1-sze piętro)
- Wchodzenie po schodach (z 1-go na 8-me piętro)
- Spacer (200 m)
- Jogging (200 m)
- Skakanie (15 razy)

Zlecono wykonanie czynności w sposób naturalny, z przerwami na odpoczynek.

### 3. Lokalizacja smartfonu

- W kieszeni spodni
- W kieszeni kurtki
- W dłoni
- Na ramieniu
- Przy biodrach (jako miejsce bliskie środka ciężkości człowieka)

Każdej osobie zlecono zarejestrowanie 5 wskazanych aktywności przy 5 wymienionych lokalizacjach smartfonu.

### 4. Zbieranie danych

- **Smartfon:** MEXZU MX3 z systemem Android 4.4.x wyposażony w 3-osiowy akcelerometr, żyroskop i czujnik pola magnetycznego
- **Sposób pomiaru:** ciągły pomiar za pomocą 3-osiowego akcelerometru i żyroskopu z częstotliwością próbkowania 20 Hz

## 5. Wstępne przetwarzanie sygnałów

- **Usunięcie szumów:** wybierając pomiędzy takimi filtrami cyfrowymi jak: *mean filter*, filtr dolno-przepustowy, *Gaussian filter*, filtr Kalmana i innymi zdecydowano się na taki, który będzie pochłaniał mało mocy obliczeniowej telefonu, z jednoczesnym krótkim czasem obliczeń, tj. *cubical smoothing algorithm* z 5-ma punktami aproksymacji.
- **Segmentacja danych – implementacja okna czasowego:** Długie szeregi czasowe potraktowano 1-sekundowym oknem czasowym (ang. *overlapped time window*), w wyniku czego uzyskano wiele krótkich szeregów czasowych (stopień nakładania (ang. *overlapping*)) ustawiono na 50%.
- **Segmentacja danych – segmentacja akcji:** podział długich szeregów czasowych na krótsze, zawierające cały cykl danej aktywności, jeszcze przed użyciem okna czasowego. Algorytm, który użyto to *cycle detection algorithm*, który polega na znajdowaniu w szeregu czasowym charakterystycznych punktów, jak np. przejścia przez 0, punkty szczytowe, punkty w dolinach.

## 6. Ekstrakcja cech z sygnałów

- W dziedzinie czasu:
  - Średnia
  - Odchylenie standardowe
  - Maksima i minima (zakres zmian sygnału)
  - Korelacja pomiędzy parami sygnałów z różnych osi czujnika
  - Rozstęp ćwiartkowy (różnica pomiędzy 3-cim i 4-tym kwartylem w statystyce opisowej, która może także wskazać na zmieniający się zakres sygnału)
  - DTW (ang. *Dynamic Time Warping distance*), mierzący podobieństwo dwóch ciągów czasowych zmieniających się w czasie (stosowane powszechnie w rozpoznawaniu mowy)

**Powyższe wielkości są wielkościami statystycznymi, które nie uwzględniają w większości przypadków istotnych cech sygnałów. Z tego względu najbardziej użytecznym z nich jest DTW.**

- W dziedzinie częstotliwości:
  - Współczynniki transformaty Fouriera (FFT coefficients) – wybrano jedynie współczynniki od 2-go do 9-go ze względu na czasochłonny proces obliczeniowy FFT (pierwszy współczynnik reprezentuje średnią z dziedziny czasu, z tego względu pominięto go)
- W dziedzinie falkowej (Wavelet-domain features):
  - Energia falkowa (wavelet energy) – funkcja matematyczna stosowana do podziału danej funkcji lub ciągłego sygnału w dziedzinie czasu na składowe o różnej skali („... into different scale components...”). Na podstawie analizy różnych publikacji zastosowano trójrzędową falkę *Daubechies* na pięciu poziomach, w celu dekompozycji danych przyspieszenia w pionowym kierunku i wyekstrahowano współczynniki z 4-go i 5-go poziomu w celu obliczenia energii falki, która wykazała w przybliżeniu największą energię przy 0.625 do 2.5 Hz.

## 7. Wybór cech sygnałów

Ekstrakcja cech pozwoliła utworzyć 89-elementowe wektory. W celu zmniejszenia czasu obliczeń postanowiono odrzucić cechy o najmniejszym znaczeniu (nie przyczyniających się do detekcji aktywności), zastosowano *test Kołmogorova-Smirnova* dla dwu prób, by sprawdzić czy dwa zbiory cech są znacząco różne (z hipotezą zerową, że dwa zbiory danych są z tego samego rozkładu i z hipotezą alternatywną, że pochodzą z różnych rozkładów). Odrzucono te cechy pomiędzy którymi test wykazał istotność większą niż ustalone  $p\text{-value} = 0.05$ . W celu wiarygodnego porównywania cech znormalizowano je za pomocą normalizacji  $z\text{-score}$ .

## 8. Implementacja klasyfikatorów

Rozpoznawanie aktywności człowieka jest „wieloetykietowym” problemem klasyfikacji (więcej niż jedna aktywność jest uwzględniana podczas klasyfikacji) – z tego względu konieczne jest użycie wieloklasowych klasyfikatorów. Zdecydowano się na następujące algorytmy:

- a) K-Nearest Neighbours
- b) Random Forests
- c) Support Vector Machines

## 9. Ocena skuteczności metodologii

Przyjęto trzy procedury treningu i testowania, i na ich podstawie oceniono skuteczność klasyfikatorów.

- a) **Model spersonalizowany - jeden do jednego:** zarówno dane treningowe jak i testowe pochodzą od tej samej osoby. Rozważa się taki model, gdyż ze względu na płeć, wiek, wagę czy wzrost, każdego człowieka charakteryzują unikalne sygnały.
- b) **Model uogólniony – wszyscy do jednego:** dane treningowe pochodzą od wszystkich osób, a testowe od jednej osoby. Ten model z kolei jest brany pod uwagę, ponieważ z reguły do trenowania klasyfikatora potrzebny jest duży zbiór danych, zwłaszcza gdy tak jak tutaj analizuje się długie wektory. Zebranie dużej ilości danych treningowych byłoby żmudne i męczące, jeśli dokonywałoby się to na jednej osobie. Poza tym model uogólniony redukuje anormalną interferencję danych.
- c) **Model uogólniony – reszta do jednego:** dane testowe pochodzą od tej samej osoby, natomiast treningowe od reszty badanych. Jedyna różnica pomiędzy tym modelem, a powyższym jest taka, że nie uwzględnia się w danych treningowych danych treningowych testowanej osoby. Model ten jest konstruowany do celów sprawdzenia jak bardzo dane od jednej osoby wpływają na uogólniony wynik, to znaczy, czy uogólniona informacja może ustalić aktywność konkretnej osoby z wystarczającą dokładnością.

## 10. Procedura treningu i testowania

Do oceny skuteczności klasyfikatorów potrzebne jest utworzenie zbiorów danych treningowych oraz testowych. Zrobiono to poprzez podział zebranych danych na mniejsze fragmenty. Do walidacji zastosowano tzw. *10-krotną walidację krzyżową* (ang. *10-fold cross validation*).

**Dla modelu (a):** losowy podział danych od tej samej osoby na 10 części o tej samej długości, a następnie ustawienie każdego podzbioru jako danych testowych, a pozostałych podzbiorów jako danych treningowych. Na końcu obliczenie średniej dokładności.

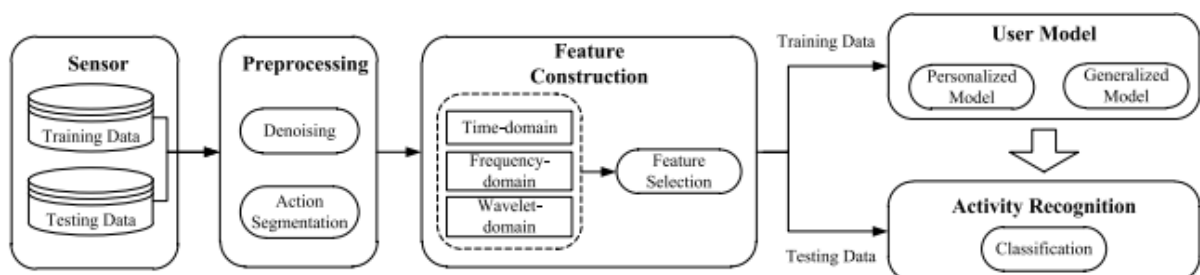
**Dla modelu (b):** podział danych od każdej osoby na 10 podzbiorów, a następnie dla i-tego podzbioru utworzyć zbiór testowy, a pozostałych podzbiorów do utworzenia zbioru treningowego.  
**Dla modelu (c):** podobnie jak wyżej, z tym, że w zbiorze treningowym nie uwzględniać i-tego podzbioru osoby, dla której wykonuje się test.

Badanie skuteczności przeprowadzono uwzględniając następujące czynniki:

- 1) Lokalizacja smartfonu
  - 2) Przestrzeń użytkownika
  - 3) Kombinacja czujników
  - 4) Stopień niestabilności danych
- 1) Skuteczność oceny w zależności od miejsca ulokowania telefonu jest uwarunkowana aktywnością. Spośród rozważanych aktywności najlepiej wypadło umieszczenie telefonu w kieszeni spodni przy biodrach, gdyż sygnał jest najbardziej stabilny – ruch w kierunku przód-tył, wyraźna okresowość sygnału, kierunek pionowy. Najskuteczniejszy klasyfikator to *Random Forests*.
  - 2) Głównie ze względu na tę analizę rozważa się opisane wcześniej 3 modele, gdyż każda osoba posiada charakterystyczne dla niej cechy biometryczne (np. osoby starsze, często mniej ruchliwe i wysportowane wykonują krótsze kroki niż osoby w średnim wieku, te z kolei dłuższe niż małe dzieci). W modelu (c) zaobserwowano najmniejszą skuteczność klasyfikatorów. Biorąc pod uwagę tę analizę najlepiej zastosować model (b).
  - 3) Ta analiza z kolei wykazała, że z reguły jest lepiej łączyć różne czujniki niż dokonywać analizy danych z różnych czujników osobno. Jeden z nich często „uzupełnia informacje”, których brak może wystąpić w danych z drugiego czujnika.

## 11. Podsumowanie

- 1) Odrzucenie modelu *reszta-do-jednego*, zważywszy na fakt, iż każda osoba jest unikalna i posiada swoisty „schemat” poszczególnych ruchów.
- 2) Najskuteczniejszy model to *wszyscy-do-jednego*, gdyż pojedyncza osoba zwiększa zakres dokładności, a dane uogólnione powodują odrzucenie wartości niestabilnych, nienormalnych ze zbioru treningowego.
- 3) „Kombinacja” danych z więcej niż jednego czujnika daje więcej informacji, co może się przyczynić do uwzględnienia większej ilości cech sygnału, a w związku z tym lepsze rozróżnienie HAR.
- 4) Dobrą praktyką, której tu nie zastosowano, byłoby ujednolicenie układu współrzędnych – geograficznego, do którego odnosi się lokalizację człowieka i własnego układu smartfonu, która zmienia się mniej lub więcej zależnie od zmian orientacji telefonu.



Rysunek 1 Schemat postępowania w rozpoznawaniu sygnałów