

# Podstawy Biometrii

**Laboratorium #9:** Keystroke dynamics – szybkość pisania na klawiaturze

*inż. Maciej Szymkowski*

*Białystok, 07.05.2018*

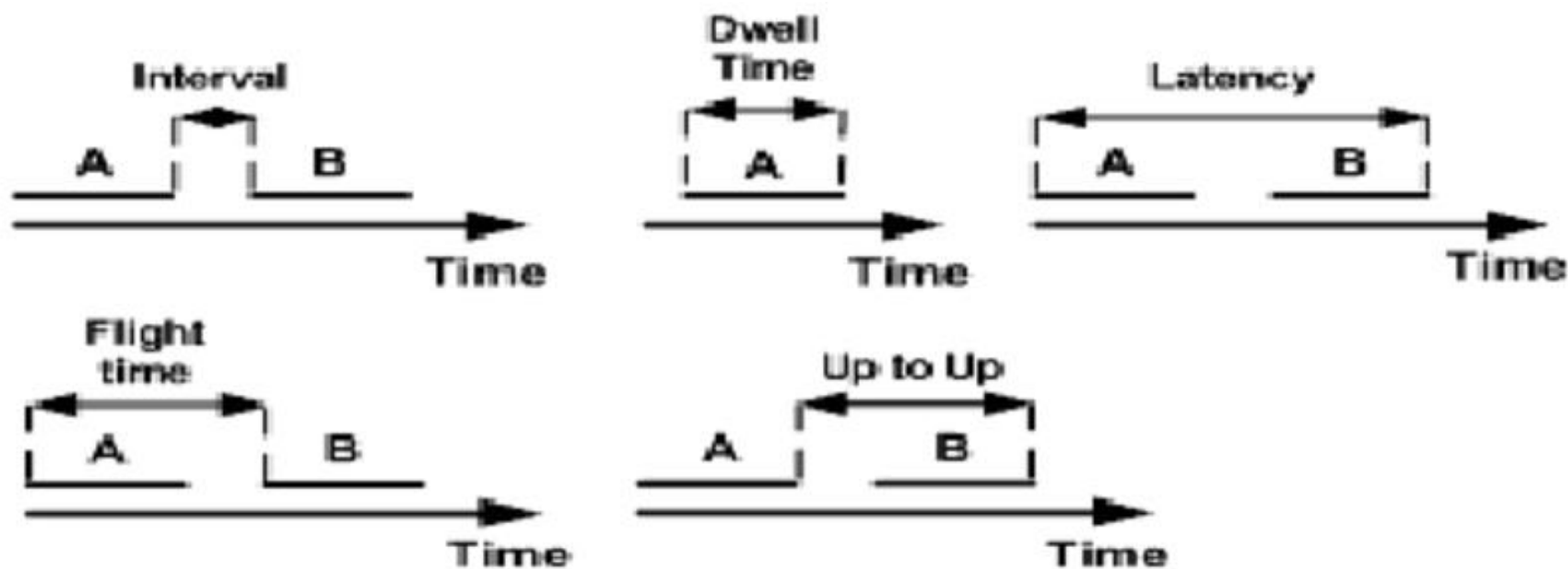
# Agenda

- ✓ Szybkość pisanie na klawiaturze jako cecha biometryczna
- ✓ Sposób pobierania keystroke dynamics
- ✓ Sposoby porównywania próbek
- ✓ Badanie jakości systemu bazującego na szybkości pisanie na klawiaturze.

# Szybkość pisania na klawiaturze jako cecha biometryczna

- ✓ Cecha behawioralna.
- ✓ Może być wykorzystywana jako cecha dodatkowa (wspierająca) w realnych systemach biometrycznych – nie jest cechą podstawową w tego typu systemach.
- ✓ Na jej postać duży wpływ może mieć wiek oraz stopień zaawansowania w używaniu systemów komputerowych.
- ✓ Jej podstawową wadą jest niskie różnicowanie pomiędzy osobami o podobnym wieku oraz o podobnej częstotliwości korzystania z systemów komputerowych (przykład: programiści w wieku 25 – 30 lat będą w bardzo podobnym czasie wprowadzać dane z użyciem klawiatury).
- ✓ W badaniach najczęściej wykorzystywane są dwa rodzaje czasu: dwell time oraz flight time.

# Szybkość pisania na klawiaturze jako cecha biometryczna



Źródło: [https://www.researchgate.net/publication/221247068\\_Identity\\_theft\\_computers\\_and\\_behavioral\\_biometrics/figures?lo=1&utm\\_source=google&utm\\_medium=organic](https://www.researchgate.net/publication/221247068_Identity_theft_computers_and_behavioral_biometrics/figures?lo=1&utm_source=google&utm_medium=organic)

# Pobieranie keystroke dynamics

```
test.addKeyListener(new KeyListener() {  
    int lastCode = -1;  
    long timeLast = -1;  
    @Override  
    public void keyTyped(KeyEvent e) {}  
  
    @Override  
    public void keyPressed(KeyEvent e) {  
        lastCode = e.getKeyCode();  
        timeLast = System.currentTimeMillis();  
    }  
});
```

```
@Override  
public void keyReleased(KeyEvent e) {  
    if(lastCode != -1) {  
        long measuredTime = System.currentTimeMillis() - timeLast;  
        System.out.println(String.format("Code: %d, time %d", lastCode,  
measuredTime));  
    }  
    }  
});
```

# Sposoby porównywania próbek

- ✓ Niech  $X_1 = [x_1, x_2, \dots, x_{26}]$  oraz  $Y_1 = [y_1, y_2, \dots, y_{26}]$
- ✓ Odległość pomiędzy próbkami może być mierzona z wykorzystaniem następujących metryk:

| Metryka     | Wzór                                                 |
|-------------|------------------------------------------------------|
| Euklidesowa | $d(X_1, Y_1) = \sqrt{\sum_{i=1}^{26} (x_i - y_i)^2}$ |
| Manhattan   | $d(X_1, Y_1) = \sum_{i=1}^{26}  x_i - y_i $          |
| Czebyszewa  | $d(X_1, Y_1) = \max_i  x_i - y_i $                   |

# Metoda $k$ -Najbliższych Sąsiadów

- ✓ Metoda ta jest jednym z najprostszych i najbardziej powszechnych klasyfikatorów. Wykorzystywana jest bardzo często w przypadku danych, które mają postać liczbową i nie niosą ze sobą żadnych dodatkowych informacji.
- ✓ Procedura ta może zostać opisana w sposób następujący:
  - 1) Jako wejście do procedury należy wprowadzić dane (w formie wektorów liczbowych) oraz określić wartość parametru  $k$ .
  - 2) Obliczana jest odległość pomiędzy wybraną próbką oraz każdą próbką z bazy (bez aktualnie badanej próbki (metoda leave-one-out)).
  - 3) Następnie sortowane rosnąco są rezultaty względem odległości od klasy.
  - 4) Wybiera się  $k$  pierwszych rekordów.
  - 5) Klasa najliczniej reprezentowana w zbiorze jest uznawana za klasę badanego wektora. Gdy dwie klasy są równolicznie reprezentowane przyjmuje się, że ta klasa jest klasą wektora jeżeli jej sumaryczne odległości są najmniejsze.

# Badanie jakości systemu

✓ Jakość systemu może zostać zbadana na wiele różnych sposobów, aczkolwiek Państwa będzie obowiązywał bardzo prosty schemat:

- 1) Obliczamy z wykorzystaniem metody  $k$ -Najbliższych Sąsiadów, klasę dla każdego z obiektów z bazy danych.
- 2) Jeżeli klasa przypisana obiektowi jest zgodna z jego rzeczywistą klasą, wtedy dodajemy 1 do licznika poprawnych klasyfikacji.
- 3) Po zliczeniu wszystkich poprawnych klasyfikacji, stopień poprawności klasyfikacji będzie obliczany na podstawie wzoru:  $\frac{\text{klasyfikacje}_{\text{poprawne}}}{\text{wszystkie obiekty}} \cdot 100\%$



# Zadanie do realizacji

- ✓ Państwa zadaniem jest przygotowanie pełnoprawnego systemu biometrycznego opartego o szybkość pisanie na klawiaturze.
- ✓ Oceniane będą następujące elementy:
  - ✓ [1pkt] Możliwość zbierania próbek (oraz poprawność wykonania tego elementu) + zapis wektorów do bazy danych (średni czas dla każdej z liter + klasa obiektu).
  - ✓ [2pkt] Badanie jakości systemu z wykorzystaniem metody k-Najbliższych Sąsiadów i metryk: Euklidesa, Manhattan oraz Czebyszewa.
  - ✓ [0,5pkt] Jakość przygotowanego sprawozdania oraz opracowane wnioski.
  - ✓ [0,5pkt] Możliwość wykonania identyfikacji oraz weryfikacji.
- ✓ Wynikiem Państwa pracy musi być sprawozdanie w którym pokrótce opiszą Państwo przygotowany system (sposób pobierania próbek), zamieszczą wykresy dotyczące jakości (przy różnej liczbie  $k$ ), dokonają analizy porównawczej z programem, który jest dostępny na platformie CEZ2 oraz zaprezentują wnioski z wykonanych eksperymentów.
- ✓ Grupa PS3 – termin oddania: 21.05.2018.

DZIĘKUJĘ ZA UWAGĘ