

# Algorytmy OnLine

## Lista 1

Adrian Herda

March 2025

### 1 Treść zadania

Dla listy jednokierunkowej zdefiniujmy operacje  $Access(i)$ , która przegląda listę od początku i sprawdza, czy element  $i$  jest na liście, a w przypadku braku elementu na liście wstawia go na końcu listy. Kosztem operacji  $Access(i)$  będzie liczba przejranych elementów listy. Zbadaj średni koszt  $n$  operacji  $Access(X)$ , gdzie  $X$  jest zmienna losowa przyjmująca wartości ze zbioru  $\{1, \dots, 100\}$ .

Rozważ następujące rozkłady zmiennej losowej  $X$ :

- jednostajny  $Pr[X = i] = \frac{1}{100}$ ,
- harmoniczny  $Pr[X = i] = \frac{1}{i \cdot H_{100}}$ , gdzie  $H_{100}$  jest setna liczba harmoniczna,
- dwuharmoniczny  $Pr[X = i] = \frac{1}{i^2 \cdot \hat{H}_{100}}$ , gdzie  $\hat{H}_{100} = \sum_{i=1}^{100} \frac{1}{i^2}$  jest n-ta liczba dwuharmoniczna,
- geometryczny  $Pr[X = i] = \frac{1}{2^i}$ , dla  $i < 100$ , oraz  $Pr[X = 100] = \frac{1}{2^{99}}$ .

Zastosuj następujące metody samoorganizacji listy:

- brak samoorganizacji,
- *move-to-front* – szukany element jest przenoszony na początek listy,
- *transpose* – szukany element jest przesuwany o jedną pozycję do przodu,
- *count* – z każdym elementem wiążemy licznik, który jest inkrementowany przy szukaniu tego elementu, a lista jest uporządkowana malejąco względem liczników.

Przeprowadź eksperymenty dla  $n$  ze zbioru  $\{100, 500, 1000, 5000, 10000, 50000, 100000\}$  i przygotuj krótkie sprawozdanie ilustrujące uzyskane wyniki.

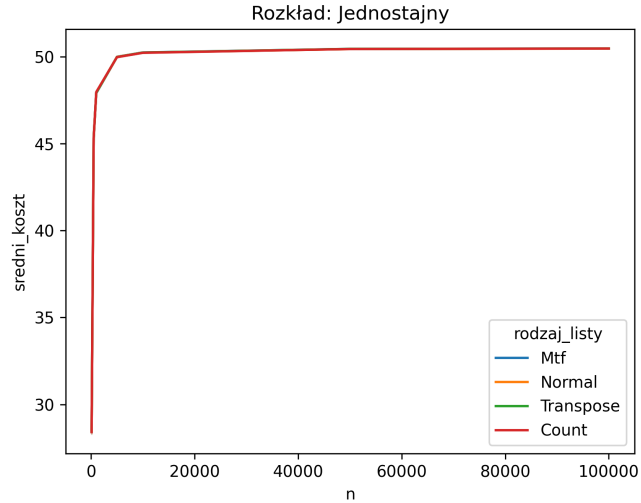


Figure 1: Średni koszt operacji dla list przy rozkładzie jednostajnym

## 2 Wyniki

### 2.1 Rozkład jednostajny

Analiza uzyskanych wyników pokazuje, że niezależnie od zastosowanej metody samoorganizacji listy, różnice w kosztach operacji są niewielkie. Oznacza to, że w badanym zakresie wartości  $n$  i dla przyjętych rozkładów zmiennej losowej  $X$ , żadna z metod nie zapewnia znaczącej przewagi nad pozostałymi.

### 2.2 Rozkład harmoniczny

Analiza wykresu ujawnia wyraźne różnice między metodami organizacji listy. Metody transpose i count charakteryzują się zbliżonymi wynikami, podczas gdy strategie normal i move-to-front (MTF) wykazują podobne wartości. Natomiast między tymi dwiema grupami obserwujemy zauważalną różnicę w kosztach operacji. Oznacza to, że wybór pierwszej grupy metod organizacji listy ma duży wpływ na efektywność operacji  $Access(X)$ , przy czym metody *transpose* i *count* oferują odmienną charakterystykę działania niż *normal* i *MTF*.

### 2.3 Rozkład podwójnie harmoniczny

Wykres dla rozkładu podwójnie harmonicznego wykazuje podobne cechy jak w poprzednim przypadku, jednak zauważalna jest już różnica między metodami normal i move-to-front (MTF), szczególnie dla mniejszych wartości  $n$ . Różnica ta stopniowo maleje wraz ze wzrostem liczby operacji. Dodatkowo metoda

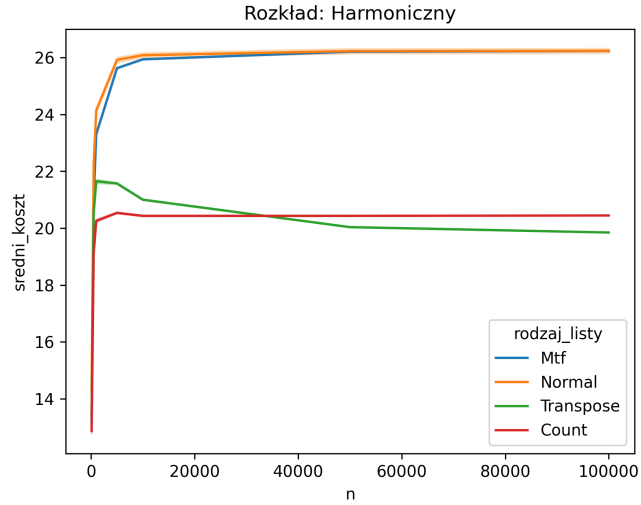


Figure 2: Średni koszt operacji dla list przy rozkładzie harmonicznym

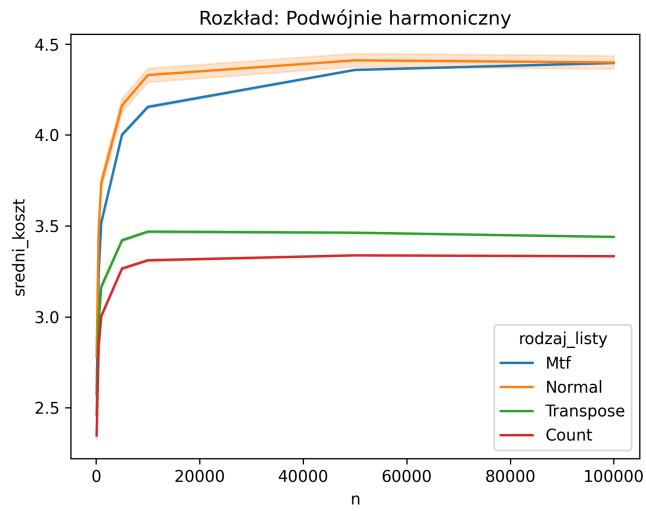


Figure 3: Średni koszt operacji dla list przy rozkładzie podwójnie harmonicznym

count osiąga minimalnie lepszy średni koszt niż transpose, co sugeruje, że dynamiczne dostosowywanie kolejności elementów na podstawie liczby odwołań przynosi niewielką, ale zauważalną poprawę efektywności.

## 2.4 Rozkład geometryczny

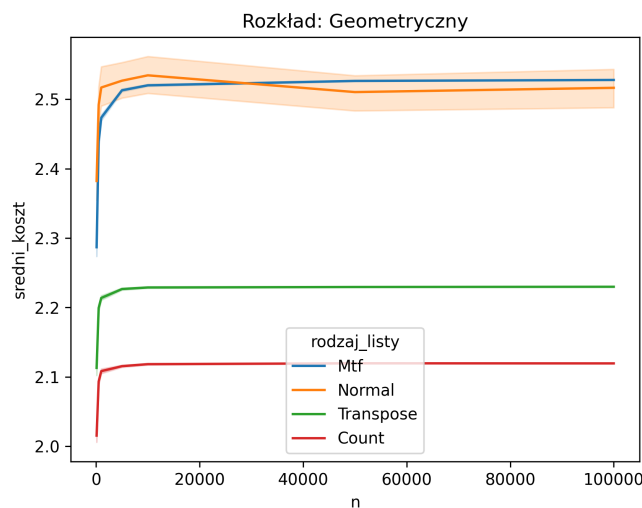


Figure 4: Średni koszt operacji dla list przy rozkładzie geometrycznym

Różnica widoczna na poprzednim wykresie pomiędzy operacjami count a transpose stopniowo się poszerza, podczas gdy różnica pomiędzy normal a mtf ulega zmniejszeniu. Dodatkowo, metoda normal wykazuje znaczne wahania wyników, co wyróżnia ją na tle innych metod, które prezentują bardziej stabilne rezultaty. Takie zachowanie metody normal może wskazywać na jej większą wrażliwość na zmiany w danych wejściowych lub zastosowanej metodzie obliczeniowej, co może wymagać dalszej analizy i optymalizacji w kontekście porównania z innymi podejściami.

## 3 Wnioski

Wnioski z analizy wskazują, że metody count i transpose sprawdzają się lepiej w przypadkach, gdy rozkład jest niesymetryczny, wykazując lepszą wydajność w takich warunkach. Z kolei w przypadku rozkładu jednostajnego, który charakteryzuje się symetryczną strukturą, nie zaobserwowano istotnych różnic pomiędzy metodami. W kontekście ogólnej efektywności, metoda count okazała się najlepsza, prezentując stabilność i precyzję wyników, szczególnie w sytuacjach, gdzie rozkład danych odbiega od symetrii.