

# Dokumentacja wstępna

## ***Optymalizacja zbioru publikacji pod kątem maksymalizacji punktów: algorytm zachłanny, referencyjny***

Ewelina Chmielewska

Mateusz Nosek

### **Cel projektu:**

Celem projektu jest zaimplementowanie w języku Python algorytmu zachłannego będącego referencyjnym, którego celem będzie maksymalizacja liczby punktów uzyskiwanych przez uczelnię za zgłaszane publikacje oraz sprawdzenie jego wydajności i poprawności na zbiorze testowym.

### **Opis problemu:**

Za publikacje – artykuły, monografie (autorstwo, redakcję, rozdział), artykuły recenzyjne, przekłady oraz edycję źródeł (autorstwo, redakcję, rozdział) – przyznaje się określoną ilość punktów. Punkty te są następnie dzielone pomiędzy autorów, zgłaszających swój wkład w dzieło. Liczba punktów, jaką można przyznać danemu autorowi ogółem, albo za konkretne dzieło jest jednak ograniczona i zależy między innymi od czasu pracy i udziału w dyscyplinie danego autora, a także od typu stworzonego przez niego dzieła. Autorzy mogą upoważniać jednostkę do swoich publikacji, lecz nałożone na nich osobno ograniczenia pozostają w mocy. Liczba dzieł zgłaszanych przez jednostkę również podlega pewnym limitom, zależnym od liczby zatrudnianych przez nią pracowników oraz ich działalności. Deklarowanie przez pracowników dyscyplin, upoważnianie jednostki do swoich publikacji oraz deklarowanie przynależności do N w zatrudniających podmiotach zwiększają ogólne ograniczenie liczby dzieł.

Naszym zadaniem jest zaimplementowanie algorytmu zachłannego referencyjnego, który wybierając dzieła do publikacji przez jednostkę będzie przestrzegał ograniczeń wynikających ze specyfikacji problemu oraz będzie maksymalizował punkty uzyskiwane przez jednostkę. Projekt w uproszczeniu można sprowadzić do dwóch problemów plecakowych – jeden z nich wynika z ograniczeń narzuconych na autora publikacji, zaś drugi z ograniczeń narzuconych na jednostkę.

### **Reprezentacja rozwiązania**

Rozwiązanie zostanie zaprezentowane w formie pary wektorów. W każdym z nich kolejne komórki odpowiadać będą kolejnym autorom. W pierwszym wektorze, w pojedynczej komórce znajdować się będzie wektor publikacji danego autora, natomiast w drugim (dla tego samego autora) umieszczony zostanie wektor zer i jedynek, odpowiadający wliczeniu publikacji do zbioru lub nie.

### **Oznaczenia:**

0 – dana publikacja nie została wliczona

1 – dana publikacja została wliczona

**Funkcja celu:**

$$f = \sum_i \left( \sum_j a_{i,j} w_{i,j} - 250 \max \left( 0, \sum_j a_{i,j} u_{i,j} - 4d_i \right) \right) - 250 \max \left( 0, \sum_{i,j} a_{i,j} u_{i,j} - 3 \sum_i d_i \right)$$

$a_{i,j}$  - przedstawienie do ewaluacji wkładu autora i w dzieło j; może mieć wartość 0 albo 1

$w_{i,j}$  - kwant punktowej wartości przeliczeniowej autora i w dziele j

$u_{i,j}$  - udział jednostkowy autora i w dziele j; ułamek z zakresu (0,1]

$d_i$  - średnia wartość iloczynu wymiaru czasu pracy i udziału w dyscyplinie autora i.

**Funkcja naprawy:**

$$q = X \rightarrow R$$

$$q = \sum_i \left( \sum_j a_{i,j} w_{i,j} \right) - \sum_{k,l} w_{k,l}$$

$a_{i,j}$  - przedstawienie do ewaluacji wkładu autora i w dzieło j; może mieć wartość 0 albo 1

$w_{i,j}$  - kwant punktowej wartości przeliczeniowej autora i w dziele j

$w_{k,l}$  - kwant punktowej wartości przeliczeniowej autora k w dziele l, które zostały usunięte z rozwiązania z powodu przekroczenia ograniczeń (naprawa).

**Algorytm:**

1. Wyliczenie wartości pomocniczych

Dla każdego autora:

- a) Posortowanie listy publikacji ze względu na wartość publikacji
- b) Zaczynając od najwyższej ocenianej publikacji: dodawanie kolejnych publikacji, sprawdzanie spełnienia ograniczeń nałożonych na autora
- c) Zapamiętanie listy zaakceptowanych publikacji – np. odpowiednie ich oznaczenie
- d) Zwrócenie: SP, LP

2. Posortowanie listy autorów ze względu na SP/LP

3. Obliczenie początkowej wartości dla funkcji heurystycznej H: suma SP dla kolejnych X autorów

4. Tworzenie listy publikacji zaakceptowanych przez jednostkę

Dla każdej publikacji:

- a) Sprawdzenie ograniczeń dla jednostki w związku z dodaniem danej publikacji
- b) Obliczenie wartości funkcji celu dla przypadków: akceptacja/ odrzucenie publikacji
- c) Wyliczenie wartości funkcji heurystycznej: brak zmian/ dodanie średniej wartości publikacji autora znajdującego się X pozycji niżej. W obu przypadkach: odjęcie wartości aktualnie rozważanej publikacji
- d) Zaakceptowanie lub odrzucenie publikacji

X – liczba autorów brana pod uwagę przy wyznaczaniu wartości funkcji heurystycznej

SP – suma wartości publikacji danego autora, które przyjęte zostały jako „do rozważenia” (spełniają ograniczenia dotyczące publikacji, jakie nałożono na autora)

LP – liczba publikacji danego autora, które przyjęte zostały jako „do rozważenia”