

Podzielenie grafu na części

Tsimafei Lukashevich, Paweł Solecki

01.03.2025

Cel projektu

Celem projektu jest stworzenie aplikacji dokonującej podziału grafu na określoną liczbę części przy minimalnej liczbie przeciętych krawędzi i zachowaniu równowagi w liczbie wierzchołków. Program działa w trybie wsadowym, umożliwiając konfigurację parametrów podziału. Dane wejściowe są przekazywane w formacie tekstowym, a wynik zapisywany w pliku tekstowym lub binarnym, co pozwala na ich ponowne wykorzystanie.

Dane wejściowe

Do poprawnego uruchomienia programu mamy zdefiniować graf za pomocą pliku tekstowego, czyli mamy podać ścieżkę do pliku z grafem, będącym argumentem wymaganym.

Program jest uruchamiany w terminalu, gdzie wszystkie parametry powinny być przekazywane przez linię poleceń.

Argumenty wywołania programu

Program akceptuje następujące argumenty wywołania:

- **-p / --parts** – liczba części podziału (np. -p 3, --parts 3), domyślnie 2.
- **-m / --margin** – maksymalny margines procentowy (np. -m 15, --margin 15), domyślnie 10%.
- **-o / --output** – format wyjściowy (-o txt lub --output txt dla pliku tekstowego, -o bin lub --output bin dla pliku binarnego), domyślnie tekstowy.
- **-f / --force** – pominięcie wbudowanego marginesu błędu, co pozwala na wykonanie podziału bez dodatkowych ograniczeń.
- **-v / --verbose** – szczegółowe logowanie przebiegu działania programu, w tym liczby iteracji i wyników pośrednich.

- `-F / --filename` – nazwa pliku wyjściowego (np. `--filename wynik.txt`), domyślnie wynik jest zapisywany do pliku domyślnego w wybranym formacie.

Przykładowe wywołania programu:

- `./program graf.txt --parts 3 --margin 20 --output txt`
 efektem będzie podzielenie grafu odczytanego z pliku `graf.txt` na 3 części tak, że liczba wierzchołków w powstałych częściach grafu nie będzie się różnić o więcej niż 20 procentowy oraz liczba przeciętych krawędzi będzie jak najmniejsza. Wynik będzie zapisany do pliku formatu `txt`.
- `./program graf.txt --parts 2 --force --verbose --filename wynik.txt`
 efektem będzie podzielenie grafu na 2 części bez ograniczeń marginesu błędu, z aktywnym szczegółowym logowaniem i zapisaniem wyniku do pliku `wynik.txt`.

Teoria

Podział grafu polega na podziale zbioru wierzchołków grafu na wzajemnie rozłączne grupy, przy jednoczesnym minimalizowaniu liczby krawędzi łączących te grupy. Celem podziału jest uproszczenie analizy danych lub zwiększenie efektywności obliczeń w systemach rozproszonych.

Wyróżnia się kilka podejść do podziału grafu:

- Podział losowy – przypisywanie wierzchołków do grup w sposób przypadkowy, co zapewnia równomierność, ale nie minimalizuje liczby przeciętych krawędzi.
- Algorytmy zachłanne – iteracyjne przypisywanie wierzchołków na podstawie lokalnych informacji, takich jak liczba połączeń do już przypisanych wierzchołków.
- Podział spektralny – oparty na analizie macierzy sąsiedztwa i jej wartości własnych, pozwalający na uzyskanie bardziej optymalnych podziałów kosztem większej złożoności obliczeniowej.
- Metody heurystyczne – np. algorytm Kernighana-Lina, który stopniowo poprawia istniejący podział poprzez zamianę wierzchołków między grupami.

Podczas oceny jakości podziału istotne są takie kryteria jak równomierny rozkład wierzchołków między grupami, minimalizacja liczby krawędzi między grupami oraz modularność, czyli miara spójności wierzchołków wewnątrz grup.

Komunikaty błędów

Program nie przyjmuje argumentów w trakcie działania więc mamy jedynie zadbać o poprawność danych wejściowych.

- Błędne definicja grafu: Linia 29: Niepoprawna definicja krawędzi. Wczytano: "1".
- Błędna liczba podziałów: Liczba podziałów musi być większa od 0 i mniejsza od liczby wierzchołków. Wczytano: "-1".
- Błędny margines procentowy: Margines procentowy musi być w zakresie 0-100. Wczytano: "123".
- Nieznany format wyjścia: Program obsługuje formaty txt i bin. Wczytano: "pdf".
- Masa ciała w danych wejściowych mniejsza/równa 0: Linia 31: Masa ciała musi być większa od 0 kg. Wczytano: "-2.1". Ciało zostało pominięte. Program wykrył ujemną masę (o wartości -2.1) w linii 31 danych wejściowych. Program ignoruje nieprawidłowe ciało.