

SPECYFIKACJA ZADANIA

PROJEKT Z ALGORYTMÓW HEURYSTYCZNYCH

1. Temat projektu

Przygotować algorytm poszukujący optymalnej sieci kolejowej pomiędzy miastami z danego zbioru. Ponieważ do sieci kolejowej trzeba dostarczyć elektryczność, rozwiązanie powinno uwzględniać położenie elektrowni - im dalej są one od najbliższych torów, tym większy koszt budowy sieci.

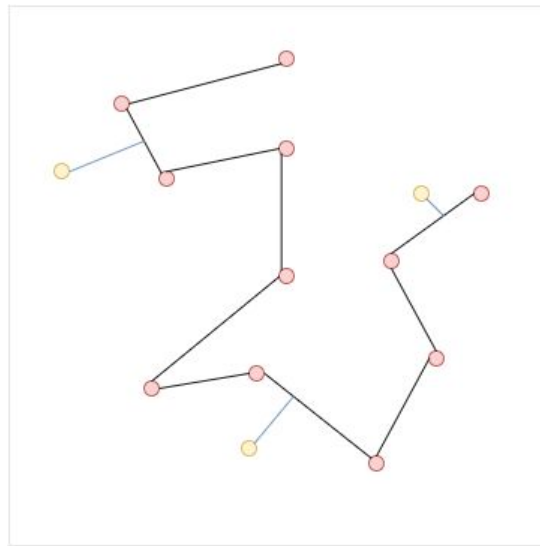
2. Opis zagadnienia, przyjęte założenia

Przyjęte przez nas założenia są następujące:

- miasta oraz elektrownie są reprezentowane jako punkty tzn. ich rozmiar nie ma znaczenia dla rozwiązania zadania
- położenie oraz liczba miast są z góry zdefiniowane dla danego wywołania programu
- położenie oraz liczba elektrowni są z góry zdefiniowane dla danego wywołania programu
- sieć kolejowa jest zbiorem odcinków pomiędzy miastami, umożliwiającym przemieszczenie się z danego punktu do innego dowolnego punktu
- koszt budowy całej infrastruktury będzie określony jako suma łącznej długości trakcji kolejowej pomnożonej przez koszt jednostki trakcji oraz sumarycznej długości połączeń pomiędzy elektrownią a najbliższymi torami pomnożoną przez koszt takiego połączenia
- poprzez odległość elektrowni od trakcji rozumiemy najmniejszą odległość pomiędzy punktem reprezentującym elektrownię i najbliższym punktem znajdującym się na odcinku reprezentującym połączenie pomiędzy dwoma miastami
- każdy punkt reprezentujący elektrownię jest połączony z najbliższym mu punktem na odcinku reprezentującym element trakcji
- zakładamy, że liczba elektrowni jest nie większa niż $n-1$ (liczba wszystkich krawędzi w sieci)

3. Definicja przestrzeni przeszukiwania

Elementem przestrzeni przeszukiwań jest drzewo rozpinające na zbiorze wierzchołków reprezentujących miasta wraz z połączeniami pomiędzy elektrowniami a krawędziami znajdującymi się w drzewie. Ściślej mówiąc oznacza to, że dla każdej elektrowni poprowadzony jest jeden odcinek od punktu ją reprezentującego do leżącego najbliżej niej punktu, znajdującego się na krawędzi drzewa. Poniżej na rysunku poglądowym przedstawiona jest przykładowa sieć kolejowa wygenerowana przez algorytm tj. pojedynczy osobnik populacji. Czerwone punkty reprezentują miasta zaś żółte elektrownie.



4. Propozycja postaci funkcji celu

Proponowana przez nas funkcja celu będzie miała postać:

$$F(x,y) = x * k1 + y * k2,$$

gdzie:

x - sumaryczna długość traktacji kolejowej,

$k1$ - koszt jednostki traktacji kolejowej,

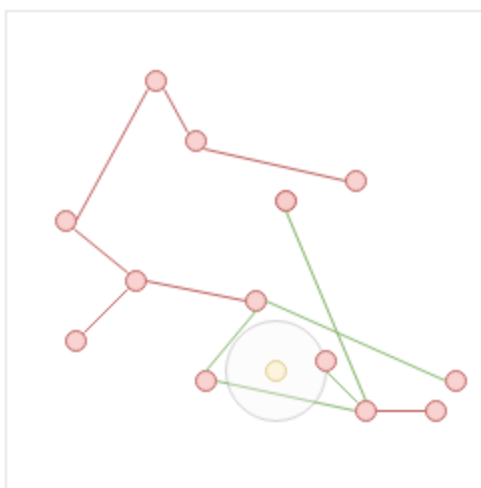
y - sumaryczna długość sieci elektroenergetycznych,

$k2$ - koszt jednostki sieci elektroenergetycznej.

Celem zadania jest minimalizacja funkcji celu czyli dobór jak najkrótszych tras pomiędzy miastami, pozwalających na dojazd z dowolnego punktu do innego dowolnego punktu oraz wyznaczenie optymalnych połączeń z elektrowniami.

Dystans pomiędzy dwoma miastami będzie liczony analitycznie na zasadzie określenia odległości jednego punktu od drugiego. Następnie uzyskane odległości będą sumowane i mnożone przez odpowiednio dobrany współczynnik kosztu k_1 tworząc łączny koszt trakcji kolejowej.

Koszt połączeń pomiędzy drzewem a elektrowniami będzie określany podczas generacji populacji początkowej. Dla każdej wylosowanej krawędzi będzie sprawdzane, czy jest szansa, że punkt na niej leżący znajduje się bliżej niż najbliższe miasto dla elektrowni. Na przykład jeżeli oba punkty tworzące krawędź znajdują się w skrajnie odległej części przestrzeni względem elektrowni należy ją odrzucić. Przykładowy dobór krawędzi do rozpatrzenia pokazano na poniższym rysunku (czerwone krawędzie zostaną odrzucone, zielone będą sprawdzane). Następnie po odrzuceniu części krawędzi będzie liczona odległość od danej elektrowni, a dalej dobierane takie połączenie, żeby jego długość była najmniejsza. Trzeba zaznaczyć, że w szczególnych przypadkach elektrownie mogą leżeć na krawędziach drzewa (lub mieć takie same współrzędne jak miasta) co oznaczałoby, że koszt budowy sieci łączącej elektrownie z torami nie będzie wpływał na funkcję celu gdyż wyniesie zero. Po wygenerowaniu populacji początkowej, odległości od elektrowni będą wyliczane tylko wtedy kiedy w procesie mutacji zostanie usunięta krawędź, której odległość od elektrowni była najmniejsza. W związku z tym konieczne będzie ponowne wyznaczenie takiej odległości oraz wyznaczenie kosztu sieci energetycznej. Proces mutacji został opisany w dalszej części dokumentacji.



5. Metody optymalizacji (metaheurystyki)

Do rozwiązania problemu zaprojektowania optymalnej sieci kolejowej wybraliśmy algorytm ewolucyjny, w ramach którego wykonywane będą niżej opisane operacje.

- Generacja populacji początkowej

Podczas generacji zostanie utworzona początkowa pula osobników składająca się z drzew rozpinających na zbiorze wierzchołków, które reprezentują położenia miast wraz z “dowiązanyymi” połączeniami pomiędzy

wybranymi krawędziami drzewa a wierzchołkami reprezentującymi elektrownie. Algorytm generacji polega na utworzeniu dwóch zbiorów - pierwszy będzie zawierał wszystkie wierzchołki, które jeszcze nie zostały wykorzystane, drugi z kolei te, które dowiązano już do drzewa. Kolejne kroki algorytmu będą następujące:

1. Wybieramy jeden spośród wierzchołków jeszcze niewykorzystanych.
2. Jeżeli zbiór wierzchołków wykorzystanych jest:
 - a. pusty - to dołączamy do niego wybrany wierzchołek,
 - b. niepusty - to do wynikowego drzewa dołączamy krawędź, która składa się z wybranego wierzchołka oraz losowo wybranego wierzchołka wykorzystanego.
3. Jeżeli zbiór wierzchołków niewykorzystanych jest pusty, to algorytm kończy działanie. W przeciwnym przypadku powtórz kroki 1-2.

Dla każdego wygenerowanego drzewa obliczana jest wartości funkcji celu (dla każdej elektrowni dołączany jest dodatkowo odcinek od punktu, który reprezentuje jej położenie do najbliższego punktu należącego do drzewa). Dla każdej utworzonej krawędzi łączącej dwa miasta obliczamy odległość między nią a wszystkimi punktami reprezentującymi elektrownie. Jeżeli dla danej krawędzi odległość jest minimalną odległością elektrowni od rozpatrywanej sieci kolejowej, to zapisujemy dla niej tę krawędź i odległość między nimi.

Liczebność populacji początkowej będzie określana jako parametr wprowadzany przy wywołaniu programu.

- **Selekcja**

Selekcja osobników do operacji krzyżowania będzie odbywać się według zasad proporcjonalnego typu selekcji, tzn. prawdopodobieństwo wybrania osobnika będzie wprost proporcjonalne do przypisanej mu wartości funkcji celu.

- **Krzyżowanie**

Dla samego operatora krzyżowania istotne jest, aby powstałe w wyniku jego zastosowania rozwiązanie składało się wyłącznie z krawędzi, które pochodzą od jego rodziców. Należy zwrócić również uwagę, aby powstałe rozwiązanie w dalszym ciągu było drzewem rozpinającym.

Działanie operatora rekombinacji będzie polegać na tym, że utworzony zostanie zbiór krawędzi, który będzie sumą zbiorów krawędzi rodziców (krzyżowanie zawsze dotyczy dokładnie dwóch osobników). Następnie z tak powstałego grafu zostanie wygenerowane drzewo rozpinające poprzez wybór

krawędzi o minimalnych odległościach od poszczególnych elektrowni, a pozostałe krawędzie dobierane będą w sposób losowy.

- **Mutacja**

Operator mutacji powinien modyfikować wygenerowane rozwiązanie w taki sposób, że wprowadzona zostanie do niego informacja, która nie była zawarta w rozwiązaniach jego rodziców. Nie może być to jednak zmiana zbyt drastyczna. Wydaje się, że odpowiednią modyfikacją może być wymiana jednej spośród krawędzi, które składają się na rozpatrywane drzewo rozpinające. Wymiana taka przeprowadzana jest w następujący sposób - wybiera się taką krawędź między miastami, która nie należy do rozpatrywanego drzewa i dołącza się ją do zbioru krawędzi tego drzewa. Taka operacja spowoduje utworzenie cyklu w drzewie reprezentującym rozwiązanie, dlatego w następnym kroku usuwa się losową krawędź z tego cyklu, ale inną niż ta dodana. Jeżeli usunięta krawędź posiadała połączenie z elektrownią, to należy ponownie obliczyć optymalne połączenie tej elektrowni z rozpatrywaną siecią.

- **Sukcesja**

W każdej kolejnej iteracji bieżąca populacja zastępowana jest przez zbiór osobników składający się z n najlepszych rozwiązań z aktualnej populacji oraz wszystkich nowych osobników powstałych w wyniku operacji krzyżowania i mutacji.

Warunek zatrzymania algorytmu zostaje osiągnięty w momencie, gdy przez określoną liczbę iteracji nie zostało znalezione rozwiązanie, dla którego wartość funkcji celu jest większa od wartości funkcji celu dla osobnika dotychczas najlepszego lub zostanie przekroczona dopuszczalna liczba iteracji (zadana jako parametr).

6. Przewidywane wyniki pracy

Wynik działania algorytmu realizującego metaheurystykę jest uzależniony od wielu czynników, w szczególności od doboru wartości parametrów oraz wprowadzonej losowości. Dlatego, ostateczna efektywność będzie zależeć od zadanych wartości m.in. takich jak:

- liczebność populacji - parametr określa ile osobników zostanie wygenerowanych na samym początku działania algorytmu i będzie stanowiła podstawę do dalszej ewolucji rozwiązań (kolejne populacje będą miały ten sam, stały rozmiar)
- prawdopodobieństwo mutacji osobników - każdy nowy osobnik otrzymany w wyniku operacji krzyżowania będzie w następnym kroku poddawany operacji mutacji z prawdopodobieństwem równym wartości tego parametru

- liczba najlepszych osobników z bieżącej populacji, które znajdą się również w populacji następnej
- minimalna liczba iteracji

Na ostateczne wyniki znaczący wpływ będzie miał również dobór współczynników kosztu w funkcji celu. Z racjonalnego punktu widzenia koszt jednostki sieci elektroenergetycznej powinien być mniejszy niż koszt jednostki sieci kolejowej. Przewiduje się, że coraz lepsze wyniki funkcji celu będziemy otrzymywać wraz ze zwiększaniem liczebności populacji początkowej oraz całkowitej liczby iteracji algorytmu. W celu uzyskania jak najlepszych wyników w jak najkrótszym czasie przewiduje się strojenie wartości poszczególnych parametrów w trakcie testowania, tak aby dobrać dla nich optymalne wartości zaproponowanego rozwiązania.

Algorytm będzie testowany dla różnego rodzaju konfiguracji rozmieszczenia zarówno punktów reprezentujących miasta jak i punktów odpowiadających położeniu elektrowni m.in.:

- losowe położenie miast i elektrowni
- równomierne rozmieszczenie miast
- rozmieszczenie miast oraz elektrowni w skupiskach
- rozmieszczenie miast na krańcach przestrzeni

Głównym wskaźnikiem, który będzie reprezentował skuteczność algorytmu będzie wartość funkcji celu w zależności od liczby wykonanych iteracji w ramach danego wykonania programu. Analizując wykres przedstawiający tę zależność, będzie można ocenić jakość wygenerowanej populacji początkowej, a następnie prześledzić przebieg procesu poszukiwania najlepszego rozwiązania. Prezentowane będą również optymalne sieci uzyskiwane podczas przebiegu algorytmu.