Úloha: Řešení problému obchodního cestujícího

Vstup: množina uzlů U reprezentujících body.

Výstup: nalezení nejkratší Hamiltonovské kružnice mezi těmito uzly.

Nad množinou U nalezněte nejkratší cestu, která vychází z libovolného uzlu, každý z uzlů navštíví pouze jedenkrát, a vrací se do uzlu výchozího. Využijte níže uvedené metody konstrukčních heuristik:

- Nearest Neighbor,
- Best Insertion.

Výsledky porovnejte s výstupem poskytovaným nástrojem Network Analyst v SW ArcMap.

Otestování proveďte nad dvěma zvolenými datasety, které by měly obsahovat alespoň 100 uzlů. Jako vstup použijte existující geografická data (např. města v ČR s více než 10 000 obyvateli, evropská letiště, ...), ohodnocení hran bude představovat vzdálenost mezi uzly (popř. vzdálenost měřenou po silnici); pro tyto účely použijte vhodný GIS.

Výsledky s uvedením hodnot W, k, uspořádejte do přehledné tabulky (metodu Best Insertion nechte proběhnout alespoň 10x), a zhodnoťte je.

Pro implementaci obou konstrukčních heuristik použijte programovací jazyk Python, vizualizaci výstupů proveďte ve vhodné knihovně, např. `matplotlib`.

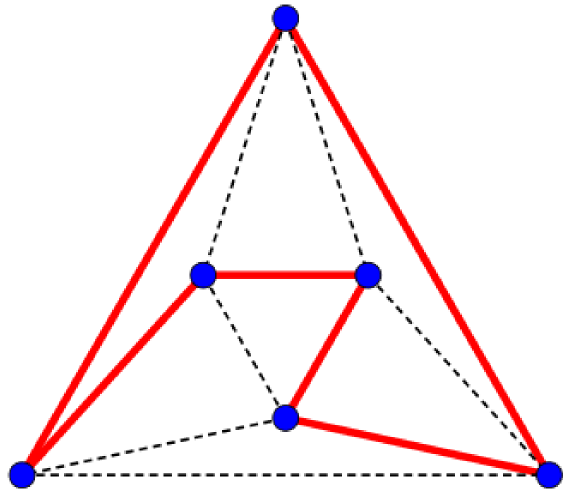
Čas zpracování: 3 týdny

Problém obchodního cestujícího (TSP)

Problém obchodního cestujícího vždy vzbuzoval velký zájem u matematiků a informatiků z důvodu snadné interpretace problematiky a zároveň složitého a dosud neexistujícího ideálního řešení. Znění problému lze vysvětlit jako: pokud si obchodní cestující přeje právě jednou navštívit každé místo na seznamu a vrátit se zpátky do původního místa, jakou cestou, aby splňovala podmínku nejnižšího celkového ohodnocení (nejkratší vzdálenost), se má obchodní cestující vydat?

TSP je představitelem rozsáhlé třídy problémů známých jako kombinatorické optimalizační problémy. TSP se zařazuje do kategorie úloh známých jako NP-úplné. U takových úloh neumíme exaktní řešení nalézt v dostatečně rychlém čase, jsme schopni získat pouze přibližné řešení, které však bude mít „dostatečnou“ kvalitu (bude akceptovatelné). Konkrétně když někdo dokáže nalézt efektivní (polynomiálně omezený) algoritmus pro problém obchodního cestujícího, mohou být poté nalezeny uspokojivé algoritmy pro všechny ostatní NP-úplné problémy. Do dnes však nebyl pro TSP polynomiálně omezený algoritmus vymyšlen.

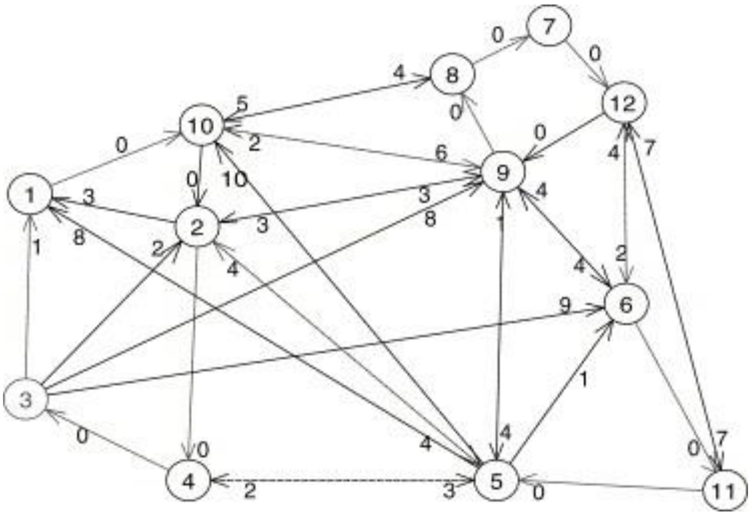
Prvním krokem k vyřešení rozsáhlých TSP musí být dobrá matematická formulace problému. V případě problému obchodního cestujícího je matematickou strukturou graf, kde je každá zastávka označená uzlem. Každé spojení dvou různých uzlů se poté označuje jako hrana. Každé hraně přísluší její ohodnocení často vyjádřené vzdáleností mezi uzly. V případě, že každé hraně přísluší směr, v němž lze hrana projít, jedná se o orientovaný graf. Pokud na směru nezáleží, mluvíme o grafu neorientovaném. V momentě, kdy se obchodní cestující může dostat z libovolného uzlu do kteréhokoliv jiného, graf se považuje za uzavřený. Okružní cesta poté koresponduje s podmnožinou hran v určitém pořadí, v teorii grafů je definována pod pojmem Hamiltonovská kružnice. Délka této kružnice je potom dána součtem ohodnocení jí příslušejících hran.



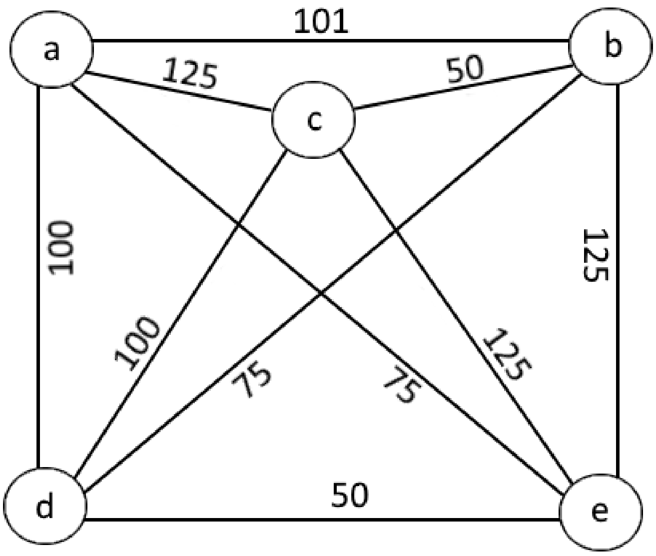
Obr. č. 1: Příklad Hamiltonovské kružnice (zdroj: https://commons.wikimedia.org/wiki/File:01_Octahedral_graph-Hamilton_circle.svg)

Záleží na směru, ve kterém je hrana grafu projita obchodním cestujícím. Rozlišuje se totiž symetrický (STSP) a asymetrický problém obchodního cestujícího (ATSP). Pro formulaci asymetrického TSP na celkovém počtu uzlů je využito binární proměnné. Hodnota 1 je přiřazena hraně vyskytující se ve výsledné cestě, nulou je potom označena hrana nepříslušící do Hamiltonovské kružnice. Každý uzel grafu musí mít právě jednu hranu směřující do něho a z něj. Tato omezení však stále nestačí, neboť by byl umožněn vznik nesouvislých smyček. Z tohoto důvodu je třeba doplnění formulace asymetrického problému obchodního cestujícího, které musí zamezit možnému vzniku těchto tzv. podcest. U asymetrického problému je povolena odlišná hodnota ohodnocení hrany v opačném směru.

Při řešení symetrického problému je směr projití hrany nepodstatný. Proto se ohodnocení hrany v obou směrech musí rovnat. Dá se tedy vycházet z toho, že mezi dvěma uzly existuje právě jedna hrana. Ve výsledné Hamiltonovské kružnici musí každý uzel tvořit právě dvě hrany. Stejně jako u asymetrického řešení, musí existovat omezení eliminující vznik podcest. Symetrický problém je v podstatě speciální případ problému asymetrického. V praxi se však ukázalo, že algoritmy vytvořené pro asymetrický problém špatně fungují pro problém symetrický. Když si uvědomíme, že případ ATSP lze jednoduše převést na případ symetrický s dvojnásobným počtem uzlů, jakýkoliv algoritmus vytvořený pro STSP může být použit pro vyřešení ATSP.



Obr. č. 2: Asymetrický problém obchodního cestujícího nad orientovaným grafem (převzato z Turkensteen 2006 s. 64)



Obr. č. 3: Symetrický problém obchodního cestujícího nad neorientovaným grafem (Zdroj: https://www.quora.com/What-is-the-full-c-code-of-the-traveling-salesman-problem)

Přesné postupy řešení takových problémů vyžadují algoritmy, které generují jak dolní, tak horní hranici skutečné minimální hodnoty v konkrétním případě. Jakákoliv okružní cesta, která vede skrz každý uzel právě jednou, je realizovatelným řešením s určitou hodnotou, která nemůže být menší než hodnota cesty optimální Hamiltonovské kružnice. Abychom zjistili blízkost horní hranice k optimální hodnotě, musíme znát dolní hranici optimální hodnoty. Pokud se horní a spodní hranice shodují, je dosaženo důkazu optimálního řešení. Pokud se neshodují, tak je odhad relativní chyby horní hranice určen pomocí rozdílu horní a spodní hranice vyděleným spodní hranicí. Proto je třeba horní a spodní hranice k nalezení prokazatelně optimálních řešení obtížných kombinatorických úloh.

Algoritmům, které zhotovují možná řešení, vykazují tedy horní hranici pro optimální hodnotu, se říká heuristické. Tyto řešené strategie produkují výsledky často bez spolehlivé garance toho, jak daleko mohou být od optimálního výsledku. Heuristické algoritmy, které hledají možná řešení jediného pokusu, jsou nazývány jako konstruktivní heuristiky, zatímco algoritmům, které se iterativně pokouší o vylepšení počátečního řešení, se říká vylepšující heuristiky.

Podle jednoznačnosti řešení dělíme algoritmy (nejen TSP) na deterministické a nedeterministické. Deterministický algoritmus vždy za stejných vstupních podmínek vytvoří stejné výsledky. Každý aktuální i následující krok vykonávání algoritmu je vždy jednoznačně definován, což je rozdíl oproti nedeterministickým algoritmům, kde následující krok nemusí být vždy jednoznačně určen.

Jak už název problému napovídá, jedno z možných využití je právě pro obchodní cestující, nebo spíše pro dodavatelské firmy. Dobře naplánovaný rozvoz může takové firmě ušetřit nezanedbatelné množství peněz. Další možné uplatnění lze nalézt ve výrobě. Je-li například potřeba do nějaké desky/plochy vytvořit velké množství otvorů na přesně daná místa, tak je výhodné zjistit ideální pořadí vytváření děr, aby se děrovací přístroj nepřesunoval zbytečně z jednoho konce desky na druhý.

Řešení problému

Pro řešení TSP byly vybrány algoritmy Nearest Neighbor a Best Insertion.

Nearest Neighbor

Jednoduchý princip zakládající se na hledání nejbližšího souseda euklidovskou vzdáleností. V prvním kroku se zvolí výchozí uzel, přidá se do Hamiltonovské cesty a označí se jako zpracovaný. V následujícím kroku se pro něj nalezne z množiny nezpracovaných uzlů nejbližší soused, který se přidá do téže cesty. Proces je opakován, dokud nezbyde žádný nezpracovaný uzel.

<u>Pseudokód:</u>

Označ všechny uzly jako nezpracované.

Vyber náhodný uzel u.

Přidej uzel u do Hamiltonovské cesty.

Označ uzel u jako zpracovaný.

Nastav hodnotu délky Hamiltonovské cesty na w = 0.

Dokud existuje alespoň jeden nezpracovaný uzel:

    Najdi nejbližšího nezpracovaného souseda uzlu u a nahraď ho za uzel u.

    Zvyš hodnotu délky Hamiltonovské cesty o vzdálenost k nejbližšímu sousedovi.

    Přidej uzel u do Hamiltonovské cesty a označ ho jako zpracovaný.

Vytvoř Hamiltonovskou kružnici přidáním prvního uzlu cesty na její konec.

Zvyš hodnotu délky Hamiltonovské kružnice o vzdálenost dvou posledních uzlů.

Best Insertion

Metoda Best Insertion v prvním kroku vytvoří iniciální Hamiltonovskou kružnici z třech náhodných uzlů, které označí jako zpracované. Dále je z množiny nezpracovaných uzlů náhodně vybrán uzel, pro který se hledá nejlepší místo přiřazení do kružnice. Nejlepší místo je definováno nejnižším přírůstkem ohodnocení celé Hamiltonovské kružnice. Pro každou dvojici uzlů tvořících hranu Hamiltonovské cesty je využito trojúhelníkové nerovnosti, v níž se nový přírůstek vypočítá relaxací existující hrany. Po nalezení onoho minima je uzel přiřazen do optimálního místa v kružnici a označen jako zpracovaný. Tímto způsobem jsou do kružnice přiřazeny i zbývající nezpracované uzly.

Pseudokód:

Označ všechny uzly jako nezpracované.

Vyber 3 náhodné uzly, přidej je do Hamiltonovské kružnice a označ je jako zpracované.

Vypočítej délku w Hamiltonovské kružnice.

Dokud existuje alespoň jeden nezpracovaný uzel:

    Vyber náhodný nezpracovaný uzel u a označ ho jako zpracovaný.

    Pro každou sousedící dvojici uzlů Hamiltonovské kružnice:

        Vypočítej přírůstek délky kružnice.

    Vlož uzel mezi uzly kružnice, kde vyšel přírůstek nejnižší.

    Zvyš délku kružnice w o tento přírůstek.

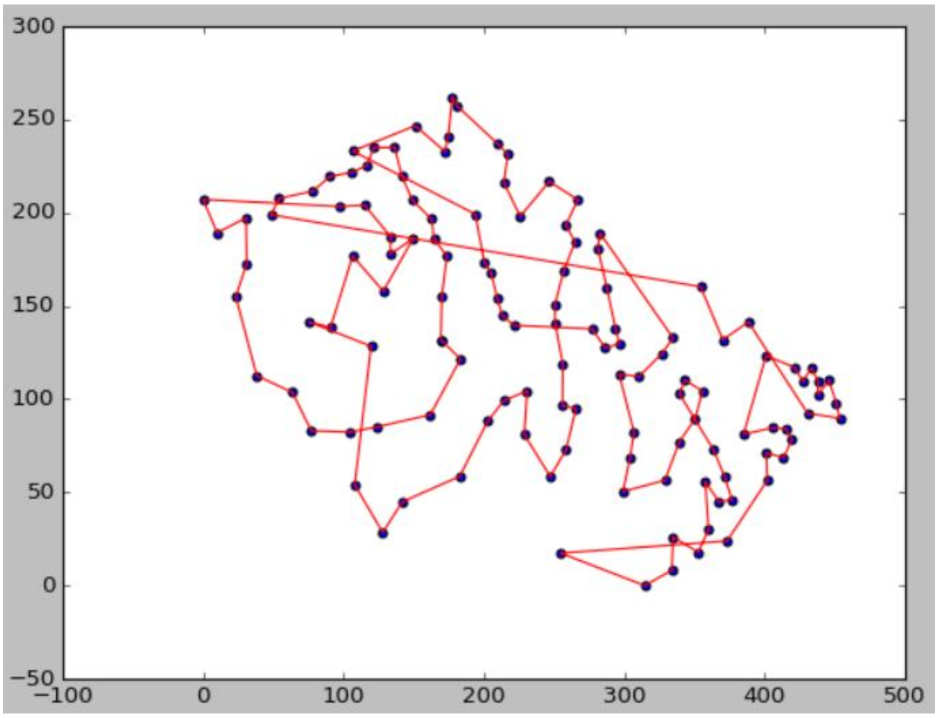
Experiments and results

Pro testování algoritmů byly použity 2 vstupní datasety. Prvním je shapefile výběru obcí Libereckého kraje, druhým je potom shapefile obcí s rozšířenou působností s počtem obyvatel větším než 10 000. Níže jsou v tabulce uvedeny délky cest problému obchodního cestujícího vypočítané jednotlivými metodami. Výsledky algoritmů nearest neighbor (NN) a best insertion (BI) jsou aritmetickými průměry n hodnot (každá byla získána náhodným určením vstupních parametrů – počátečních uzlů), kde n odpovídá počtu uzlů. Za W_0 byla zvolena nejnižší hodnota, tedy vzdálenost vypočítaná v softwaru ArcGIS for Desktop.

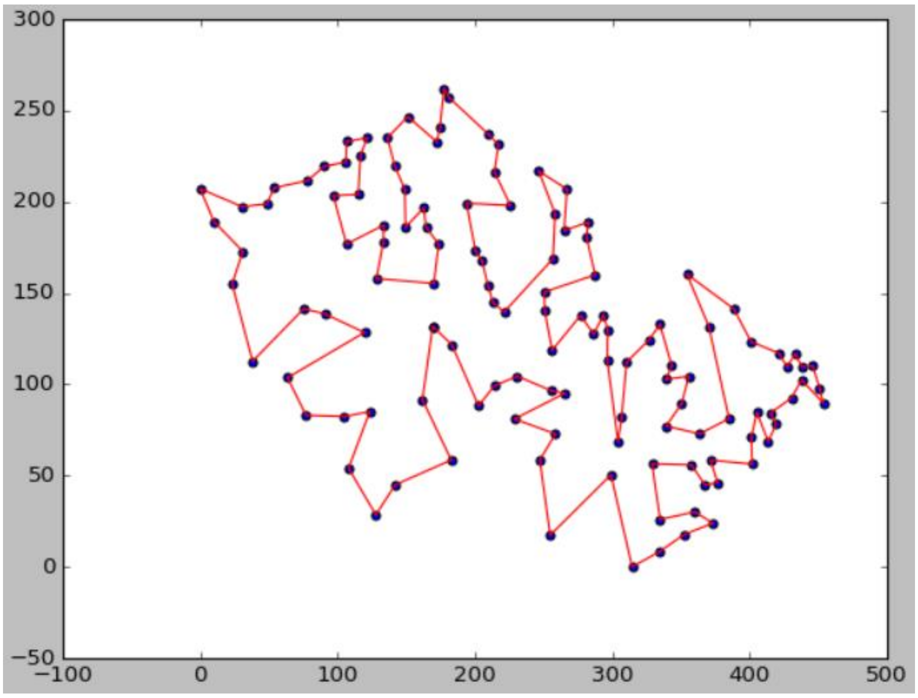
dataset	W_0 (km)	W_{NN} (km)	W_{BI} (km)	k_{NN} (%)	k_{BI} (%)
obce	349,852	442,038	373,614	1,26	1,07
ORP	2543,236	3121,723	2739,814	1,23	1,08

V následující tabulce jsou pro oba datasety uvedeny výsledky z deseti opakování pro každou metodu, minimální hodnota je poté tučně zvýrazněna.

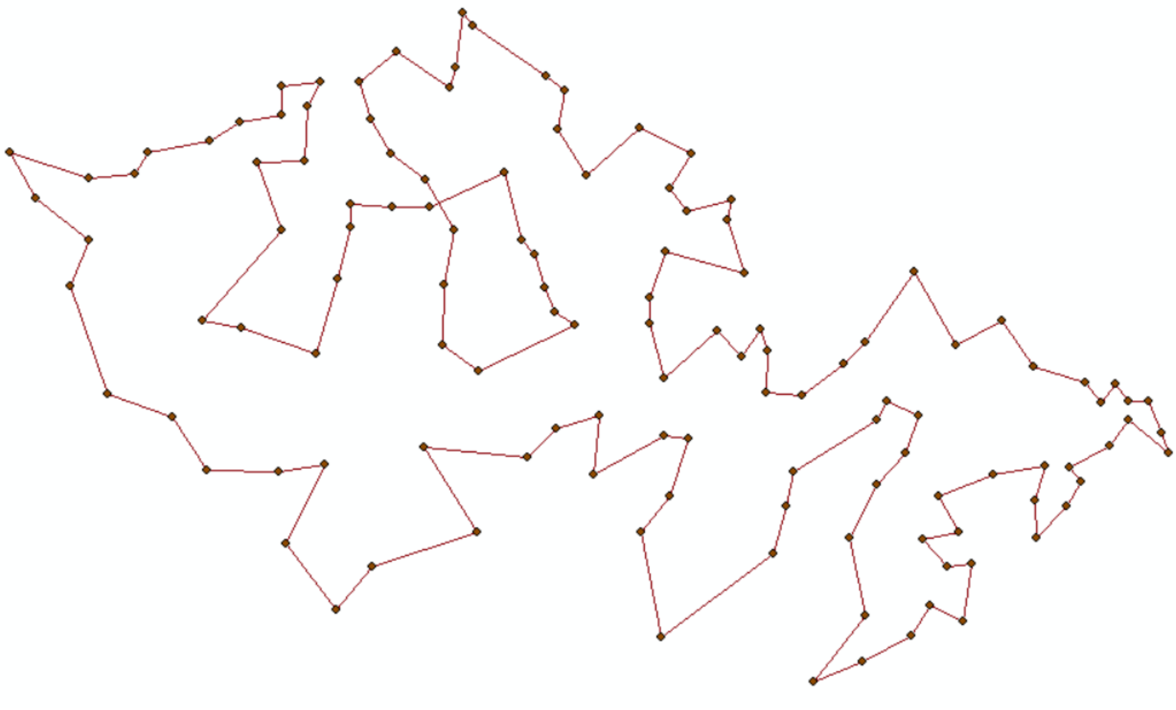
soubor obcí	NN	**418**	462	456	446	454	436	440	451	437	462
	BI	384	**362**	370	372	378	368	368	367	366	379
soubor ORP	NN	2967	3182	3173	3160	3227	**2964**	3021	3130	3246	3017
	BI	2705	2820	2705	2775	**2658**	2736	2696	2713	2766	2716



Obr. č. 4: Metoda Nearest Neighbor provedená nad ORP



Obr. č. 5: Metoda Best Insertion provedená nad ORP



Obr. č. 6: Funkce rozšíření Network Analyst softwaru ArcGIS for Desktop nad ORP

Závěr

V této práci byly prezentovány výsledky dvou nejsnadnějších heuristik pro TSP. Vzhledem ke způsobu metody Nearest Neighbor není překvapením, že nevrací uspokojivé výsledky. Na druhé straně metoda Best Insertion funguje mnohem lépe a jen o trochu zaostává za algoritmem využívaným softwarem ArcGIS for Desktop v rámci Network Analyst.

V určitém případě vzniká u obou metod nejednoznačné řešení. U algoritmu Nearest Neighbor v momentě, když existuje 2 a více nejbližších sousedů. V tomto případě metoda upřednostní prvního z těchto uzlů v seznamu. V metodě Best Insertion tato situace vzniká v případě dvou a více různých míst v Hamiltonovské kružnici se stejným nejmenším přírůstkem ohodnocení kružnice. V tomto případě je upřednostněno první takové nalezené místo.

Další vylepšení by bylo možné udělat v metodě Best Insertion. Místo výběru náhodného uzlu by bylo třeba vybrat takový uzel, který nejméně zvýší přírůstek ohodnocení Hamiltonovské kružnice v daném kroku. Dále by bylo možné určit iniciální Hamiltonovskou kružnici ve tvaru trojúhelníku za pomoci konvexní obálky.

Seznam literatury:

HOFMANN, K., PADBERG, M., RINALDI, G. (2001): Traveling salesman problem. George Mason University, Fairfax, Virginia, USA.

TURKENSTEEN, M., GHOSH, D., GOLDENGORIN, B., (2006): Iterative patching and the asymmetric traveling salesman problem. Discrete optimization, 3(1), 63–77.