Univerzita Karlova

Přírodovědecká Fakulta

Geoinformatika

Traveling salesman problem

Adam Kulich

Geoinformatika, kartografie a dálkový průzkum Země, 1. ročník NMgr.

Praha, 2023

Zadání

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 (obě metody nechte proběhnout alespoň 10x), a zhodnoťte je.
Pro implementaci obou konstrukčních heuristik použijte programovací jazyk Python, vizualizaci výstupů proveďte ve vhodné knihovně, např. `matplotlib`.

Teoretická část

Problém obchodního cestujícího

Problém obchodního cestujícího (traveling salesman problém, TSP) je jeden z nejprominentnějších kombinatoricko-optimalizačních problémů. TSP je definovaný jako situace, kde hypotetický obchodní cestující musí projít zadaná místa, každé přesně jednou, a vrátit se zpět do počátečního. Úvaha je převedena na úlohu z oblasti teorie grafů jako hledání nejméně nejkratší (nejméně nákladné) Hamiltonovské kružnice. Kvůli rychlému přibývání počtu možných řešení se zvyšujícím se počtem uzlů v grafu se při jejím řešení nehledá optimální cesta, ale optimalizačními heuristikami se tvoří její odhad (Laporte, 1992, Muthukumaravel & Sathya, 2015). V praxi se TSP objevuje například v některých geografických problémech nebo při sekvenování DNA.

Jeděmi z prvních, kteří formulovali řešení tohoto problému byli Dantzig, Fulkerson a Johnson v roce 1954. Jejich algoritmus zkoumá všechny Hamiltonovské kružnice a vyhledá jejich minimum. V následujících letech byly implementovány i další exaktní řešení (Laporte, 1992, 2006).

Pro řešení složitějších verzí problému, obsahujících více uzlů, se implementují heuristické metody, které odhadují co nejlepší řešení. Mezi tyto patří například jednodušší heuristiky použité v tomto cvičení, Nearest neighbor a Best insertion. Mezi další, komplexnější heuristiky patří například genetické algoritmy, metoda „Tabu search“ nebo metoda mravenčích kolonií (Laporte, 1992).

Použité heuristiky

Nearest neighbor (nejbližší soused) je jednoduchá metoda konstrukční heuristiky, která je deterministická při použití stejných parametrů. Vstupem je kromě grafu také počáteční uzel. Algoritmus v každém uzlu postupně hledá nejbližší uzel, který dosud není součástí Hamiltonovské kružnice, ten k ní připojí a znovu hledá nejbližší uzel k tomuto novému uzlu (Laporte, 1992, Bayer, 2022).

Best insertion (nejlepší proložení) je složitější metoda, produkující lepší výsledky. Na počátku je vytvořena Hamiltonovská kružince definovaná třemi náhodnými uzly, která se v každém kroku rozšiřuje o náhodný uzel, přidaný na místo v kružnici, kde nejméně zvětší její délku. Výsledek závisí na náhodně zvolených třech počátečních uzlech i na pořadí v jakém jsou uzly přidávány do kružnice (Laporte, 1992, Bayer, 2022).

Metodika

Tvorba datasetu

V softwaru ArcGIS Pro byly vytvořeny dva datasety s cca 100 body. Každý z nich obsahuje seznam souřadnic x a y v S-JTSK, které byly použity v algoritmech. První dataset tvoří seznam souřadnic všech vrcholů v Česku s nejvyšším bodem výše než 1000 m n.m. (103 řádků) a druhý je seznamem souřadnic všech obcí s rozšířenou působností v Čechách a kraji Vysočina (137 řádků).

Použité heuristiky a jejich implementace

Obě heuristiky byly naprogramovány v programovacím jazyce Python a postup je uveden i v komentářích v obou přiložených souborech (viz nn.py a bi.py).

Nearest neighbor

Metoda přiřazuje ke kružnici vždy uzel nejbližší poslednímu uzlu. Délka vzniklé Hamiltonovské kružnice závisí na zadaném počátečním bodu, a proto byl do řešení úlohy zařazen cyklus, který, pokud je spuštěn, spočítá délku Hamiltonovské kružince pro každý počáteční uzel.

Funkce nearest neighbor – v nn.py

```
Inicializuj seznam který označí uzly jako nenavštívené a prázdnou Hamiltonovskou kružnici
Inicializuj počáteční bod kružnice a sumu vah
Označ počáteční bod kružnice jako navštívený

Dokud existují nezpracované uzly
    nastav minimální vzdálenost jako ∞
    pro každý nezpracovaný uzel
        spočítej vzdálenost od posledního uzlu v Hamiltonovské kružnici
        pokud je vzdálenost menší než minimální
            přepiš minimální vzdálenost jako vzdálenost
            označ uzel jako navštívený
            připoj uzel k Hamiltonovské kružnici
            přičti vzdálenost k sumě vah
spočítej vzdálenost k počátečnímu bodu kružnice
přičti vzdálenost k počátečnímu bodu kružnice k sumě vah
připoj počáteční bod kružnice
vrať Hamiltonovskou kružnici a sumu vah
```

Best insertion

Tato metoda utvoří mezi uzly náhodný trojúhelník, do kterého přidává v každém kroku náhodný uzel tak, aby byla vzniklá kružnice co nejkratší. Metoda není deterministická, takže pro účel úlohy byl program spuštěn 10krát a výsledky byly srovnány mezi sebou.

Funkce best insertion – v bi.py

```
Inicializuj seznam, který označí uzly jako nenavštívené a prázdnou Hamiltonovskou kružnici
Inicializuj počáteční bod kružnice a sumu vah
Vyber tři náhodné uzly
Přidej tyto tři uzly do Hamiltonovské kružnice
Spočítej sumu vah těchto tří uzlů
Inicializuj seznam uzlů, označených jako nenavštívené
pro každý uzel v Hamiltonovské kružnici
    označ tento uzel jako navštívený
    odeber tento uzel ze seznamu nenavštívených uzlů
```

Dokud existují nezpracované uzly
vyber náhodný uzel u
nastav minimální vzdálenost jako ∞
pro každý uzel v Hamiltonovské kružnici
označ uzel jako v1
pokud uzel neleží na posledním místě v kružnici
označ další uzel v pořadí jako v2
jinak označ první uzel v kružnici jako v2
spočítej rozdíl délky po přidání uzlu mezi v1 a v2
pokud je rozdíl délky menší, než minimální vzdálenost
přepiš minimální vzdálenost jako vzdálenost
označ u jako navštívený
připoj u k Hamiltonovské kružnici mezi v1 a v2
přičti minimální vzdálenost k sumě vah
připoj počáteční bod kružnice na konec Hamiltonovské kružnice
vrať Hamiltonovskou kružnici a sumu vah

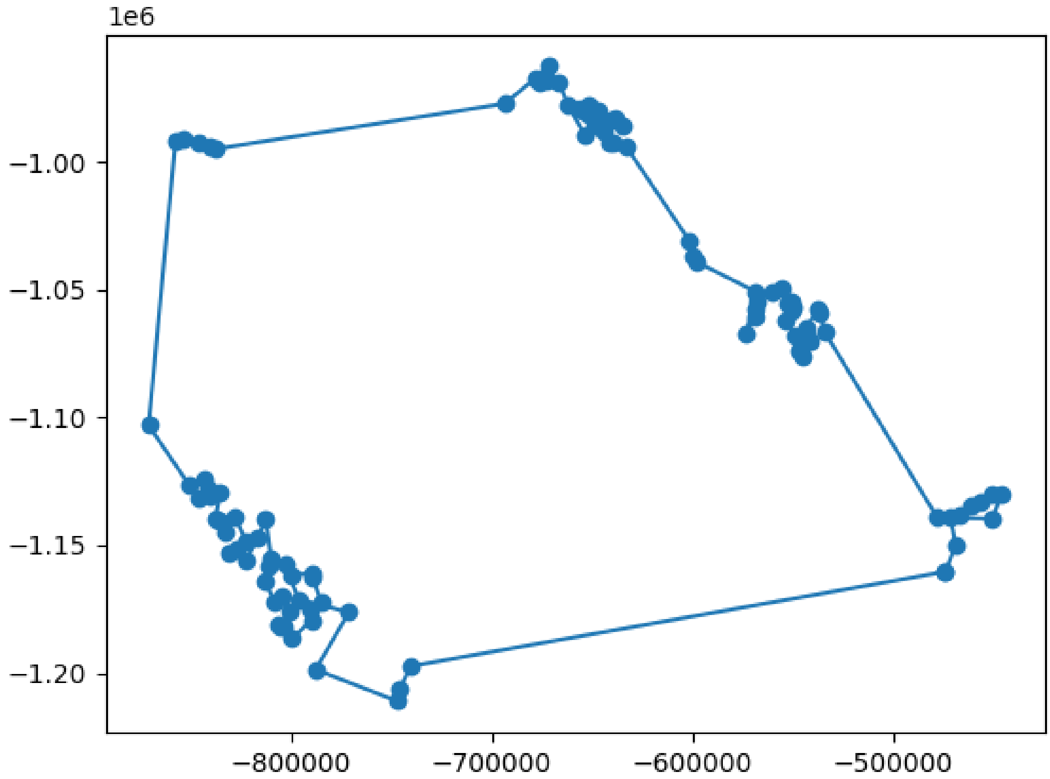
Nakonec byl na oba datasety implementován nástroj Network Analyst v ArcGIS Pro. Výstupy z všech tří heuristik byly srovnány.

Přehled výsledků

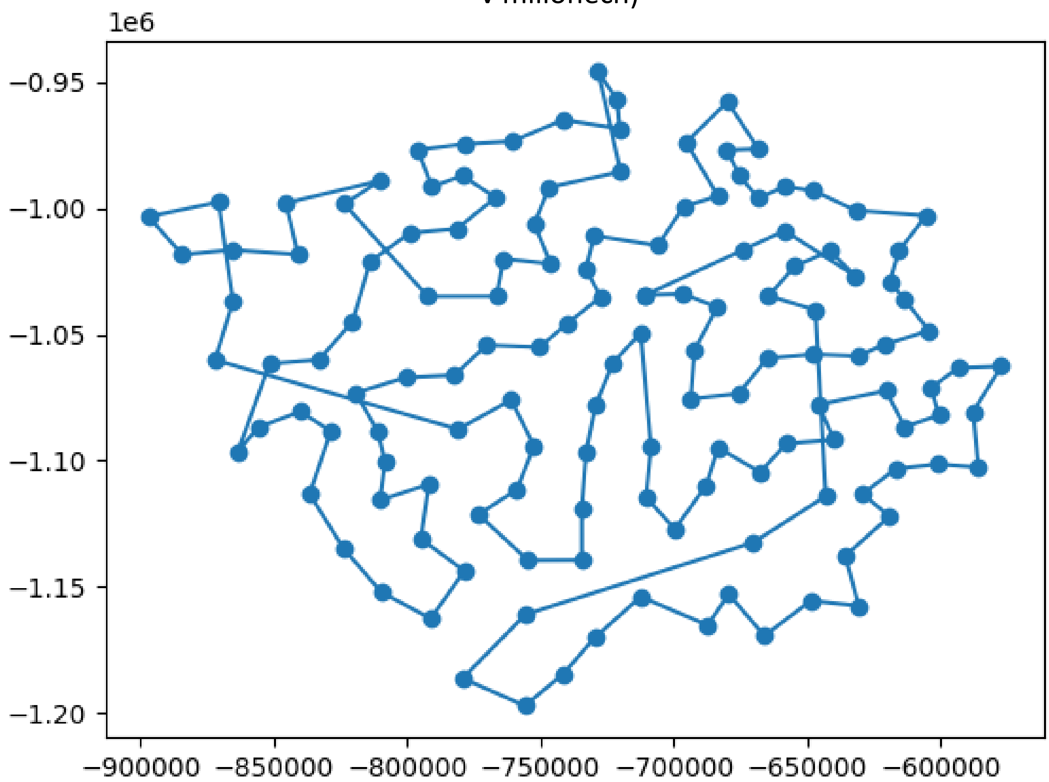
Pro metodu nejbližšího souseda bylo vybráno 10 nejkratších kružnic pro každý dataset a metoda Best insertion byla provedena na každém z nich 10krát. Výsledné délky Hamiltonovských kružnic pro každý dataset byly zaznamenány do tabulky 1 a nejlepší cesty byly vizualizovány na obr. 1-6. Plné výsledky včetně pořadí bodů jsou zaznamenány v souboru results.csv.

	Dataset 1 (nejvyšší body)		Dataset 2 (ORP)	
	Počáteční bod	Délka kružnice (m)	Počáteční bod	Délka kružnice (m)
Nearest neighbor	**23**	**1449816**	**13**	**2897604**
	11	1458439	6	2898550
	22	1460180	22	2901751
	7	1470186	49	2911724
	19	1470186	1	2915618
	32	1474270	56	2917604
	31	1474300	51	2920486
	5	1476468	63	2925593
	85	1483392	65	2925593
	48	1490145	38	2928373
	Číslo pokusu	Délka kružnice (m)	Číslo pokusu	Délka kružnice (m)
Best insertion	**5**	**1405010**	**8**	**2769824**
	10	1419928	5	2783321
	7	1420397	10	2784410
	4	1425763	7	2824529
	2	1429915	4	2826023
	1	1437987	6	2827034
	9	1449344	3	2834109
	6	1453167	2	2834810
	3	1462926	9	2860260
	8	1482140	1	2888574
Network Analyst	**1**	**1380111**	**1**	**2601950**

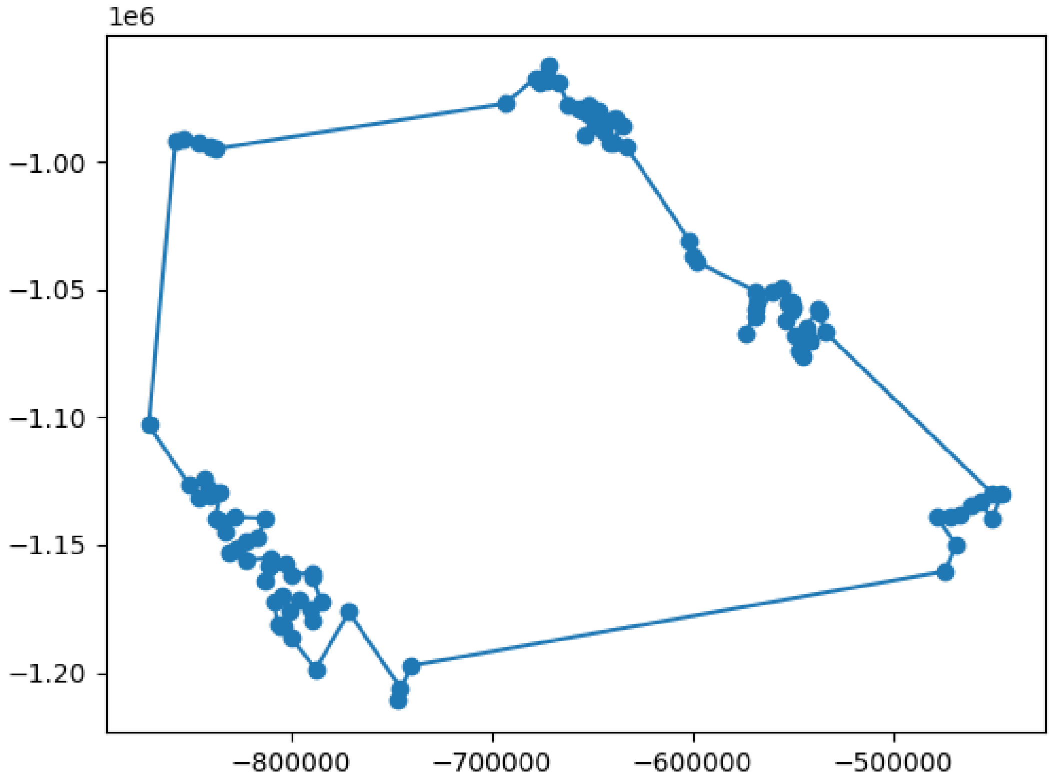
Tabulka 1: Délky Hamiltonovských kružnic po aplikaci jednotlivých algoritmů seřazené od nejkratší po nejdelší



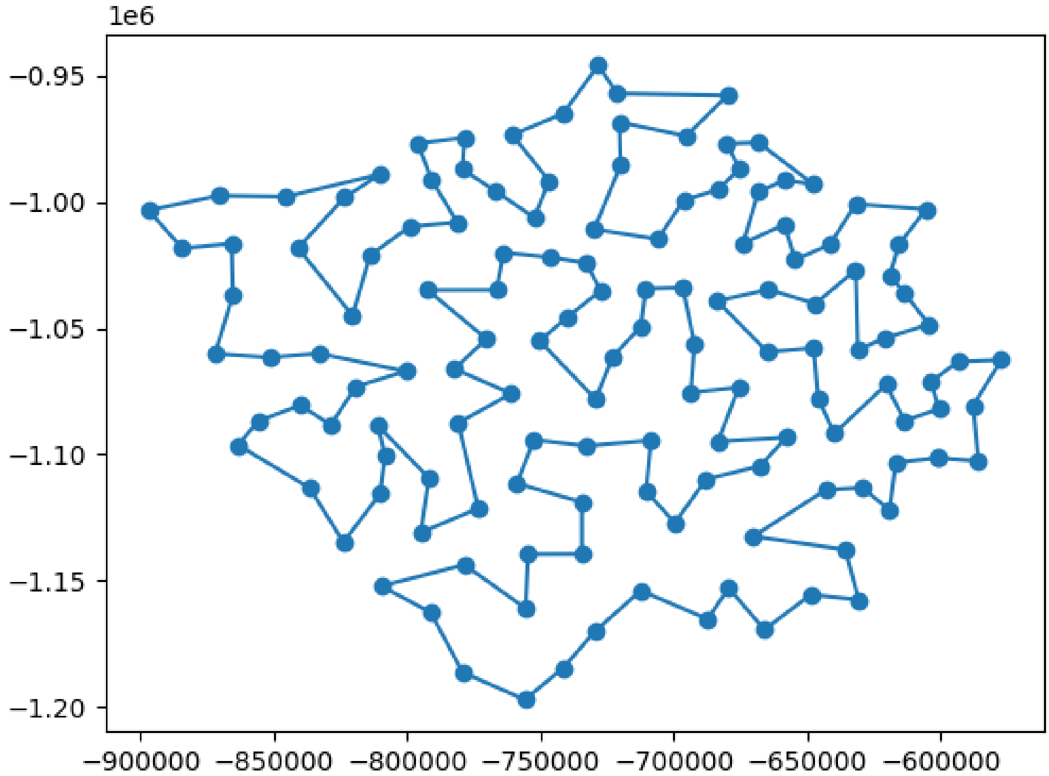
Obr. 1: Nejkratší Hamiltonovská kružnice vytvořená heuristikou Nearest neighbor na datasetu 1 (pozn.: osa y zobrazuje hodnoty v milionech)



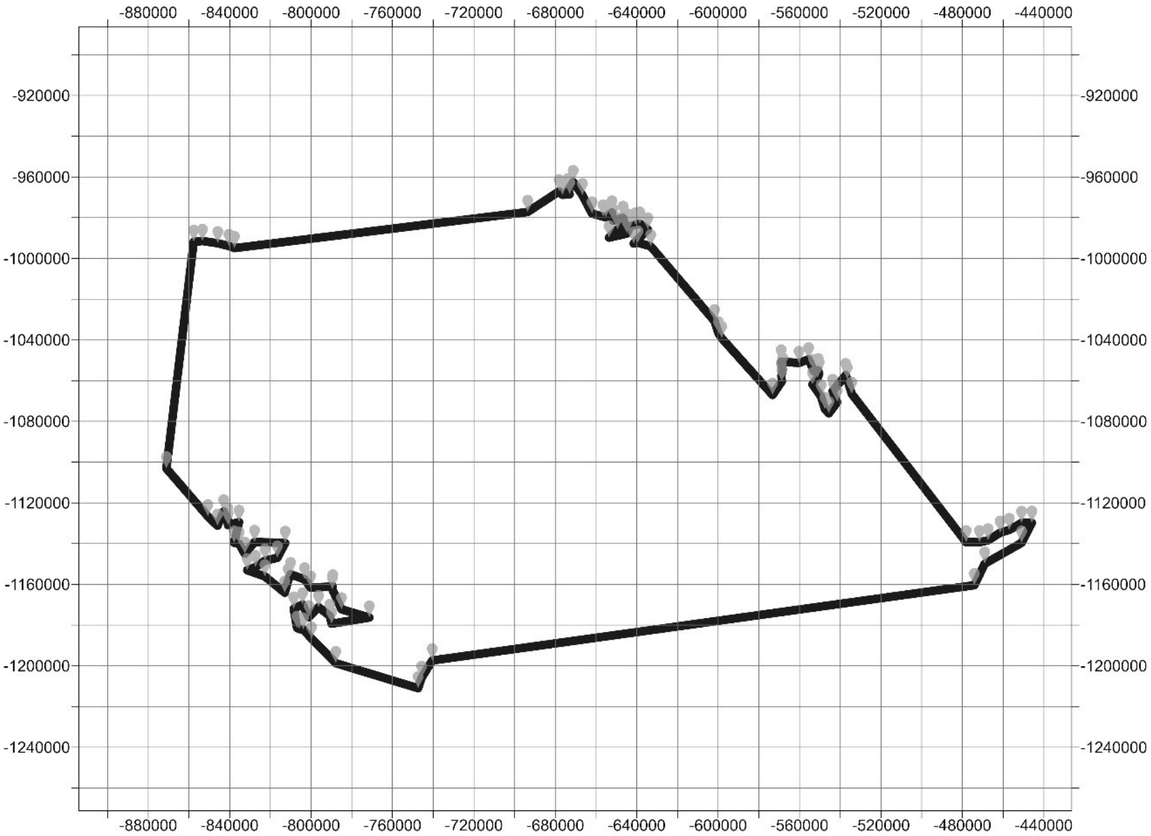
Obr. 2: Nejkratší Hamiltonovská kružnice vytvořená heuristikou Nearest neighbor na datasetu 2 (pozn.: osa y zobrazuje hodnoty v milionech)



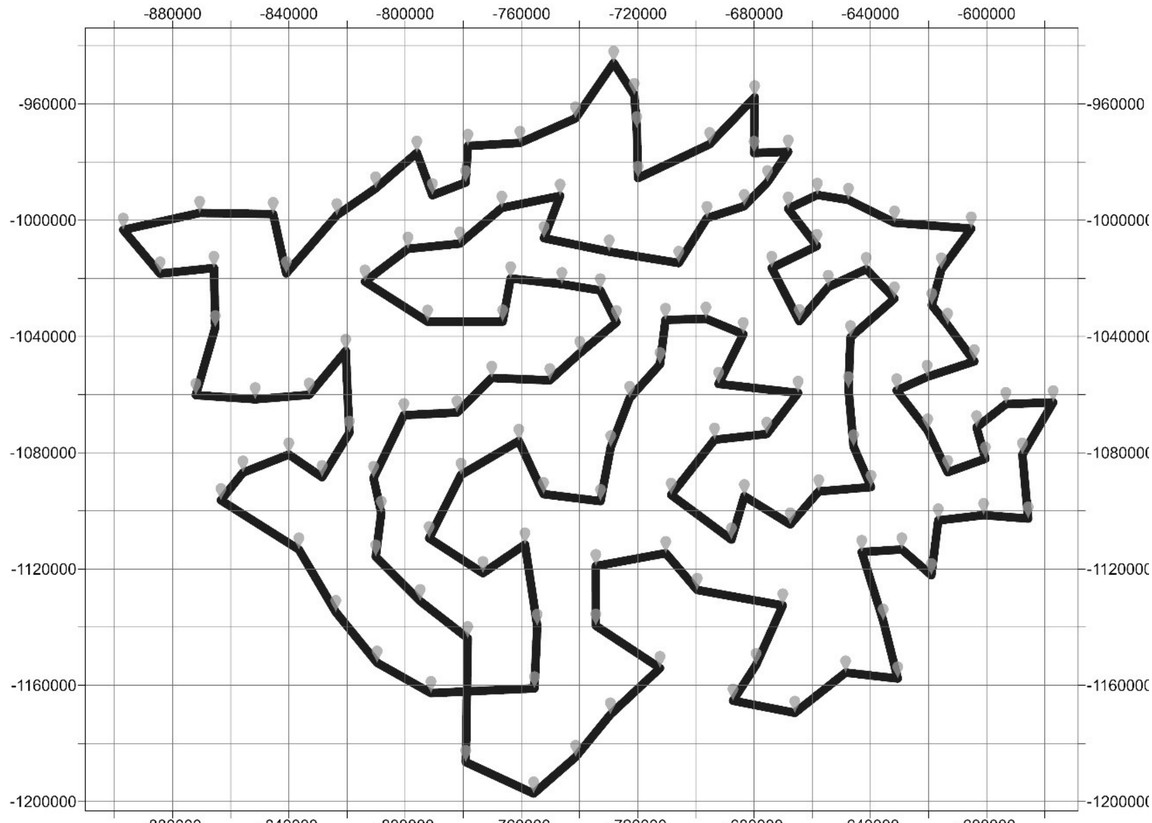
Obr. 3: Nejkratší Hamiltonovská kružnice vytvořená heuristikou Best insertion na datasetu 1 (pozn.: osa y zobrazuje hodnoty v milionech)



Obr. 4: Nejkratší Hamiltonovská kružnice vytvořená heuristikou Best insertion na datasetu 2 (pozn.: osa y zobrazuje hodnoty v milionech)



Obr. 5: Hamiltonovská kružnice vytvořená v softwaru ArcGIS Pro na datasetu 1



Obr. 6: Hamiltonovská kružnice vytvořená v softwaru ArcGIS Pro na datasetu 2

Diskuse

Z výsledků vyplývá, že všechny použité metody poskytují různě přesné řešení problému. V tomto případě není jasné, jaká je ideální cesta, a proto není možné exaktně měřit přesnost použitých metod ve srovnání s ideálním řešením. Lze však porovnávat heuristiky mezi sebou. Možnosti metody Nearest neighbor byly vyčerpány tím, že byly vyzkoušeny všechny možnosti počátečních uzlů. Lepších výsledků už tedy nemůže dosáhnout. Metoda Best insertion zjevně dosahuje výrazně lepších výsledků a obecně je lepším řešením problému. Teoreticky by mohla dojít i k lepším řešením, než jakých bylo dosaženo po 10 pokusech. Ani jedna z metod však nedosahuje exaktního řešení, jen se snaží o jeho aproximaci. Ještě výrazně lépe si vedla metoda, kterou používá nástroj Network Analyst v ArcGISu. Zřejmě se jedná o složitější metodu, která v tomto případě vytvořila nejkratší Hamiltonovskou kružnici na obou z použitých datasetů. Pro lepší analýzu je tedy určitě vhodné použít komplexnější heuristiku než Nearest neighbor a Best insertion, kterých existuje celá řada.

Závěr

Úloha problému obchodního cestujícího byla úspěšně řešena třemi různými heuristikami na dvou datasetech. Všechna řešení dosáhla uspokojivého výsledku. Metoda Nearest neighbour vyprodukovala nejkratší cesty na prvním a druhém datasetu o délce 1449,8 a 2897,6 km. Metoda Best insertion po 10 pokusech vytvořila nejkratší cesty ve stejném pořadí o délce 1405,0 a 2769,8 km. Algoritmus z nástroje Network Analyst softwaru ArcGIS Pro vytvořil ze všech nejkratší Hamiltonovské kružnice o délkách 1380,1 a 2601,9 km.

Zdroje

BAYER, T. (2021): Úloha: Řešení problému obchodního cestující, Výukový materiál k předmětu Geoinformatika https://web.natur.cuni.cz/bayertom/images/courses/Geoinf/tsp_zadani.pdf.

BAYER, T. (2022): Problém obchodního cestujícího, konstrukční heuristiky: stručný návod na cvičení, Výukový materiál k předmětu Geoinformatika https://web.natur.cuni.cz/bayertom/images/courses/Geoinf/tsp_uloha.pdf.

LAPORTE, G. (2006): History of the Traveling salesman problem, Centre for Research on Transportation (CRT) and GERAD HEC Montréal

LAPORTE, G. (1992): The traveling salesman problem: An overview of exact and approximate algorithms, European Journal of Operational Research, 59, 2, 231-247.

SATHYA, N. and MUTHUKUMARAVE, A. (2015): A Review of the Optimization Algorithms on Traveling Salesman Problem, Indian Journal of Science and Technology, Vol 8(29)