Hugo Majer

1. N-GKDPZ

Bratislava, 1. 1. 2022

Úkol z predmetu Geoinformatika

Problém obchodného cestujúceho

Univerzita Karlova

Přírodovědecká fakulta

1. Zadanie

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 u 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.

2. Popis problému

Táto kapitola zo začiatku popisuje problém obchodného cestujúceho, následne túto úlohu definuje v teórií grafov a jej záver je venovaný motivácií – aplikácií a využitiu problému obchodného cestujúceho.

2.1. Problém obchodného cestujúceho

Problém obchodného cestujúceho (angl. *Traveling salesman problem*) je úloha, ktorá za svoju existenciu zožala množstvo pozornosti matematikov a programátorov a to z jednoduchého dôvodu – je ľahké ju popísať, ale veľmi ťažké vyriešiť, nakoľko sa jedná o úlohu, ktorá vo svojej všeobecnosti predstavuje ťažký kombinatorický a optimalizačný problém (Hoffman et al. 2013, Gutin and Punnen 2006).

Jedná sa o problém radený do triedy NP-úplných úloh a zároveň do NP-ťažkých úloh, čo znamená, že je ho možné vyriešiť v nedeterministickom polynomiálnom čase a zároveň ľubovoľný iný problém z NP triedy je naň polynomiálne redukovateľný. To znamená, že ak niekto nájde deterministický polynomiálny algoritmus pre problém obchodného cestujúceho, tak by bolo možné nájsť efektívny algoritmus pre všetky problémy z triedy NP-úplných úloh (Hoffman et al. 2013).

Problém obchodného cestujúceho sa dá popísať nasledovne. Cieľom je nájsť cestu obchodného cestujúceho, ktorý vychádza z počiatočnej lokácie (napr. domovského mesta), navštívi každú vopred zadanú lokalitu (každé z n miest) a vráti sa do počiatočnej lokácie (domovského mesta). Táto cesta ale musí spĺňať určité podmienky a to síce, že každá lokalita je navštívená práve jedenkrát a že celková vzdialenosť cesty je minimálna (Gutin and Punnen 2006).

Začiatky opisovanej úlohy sú nejasné. Za začiatok môže byť považovaná matematická definícia úlohy v 19. storočí írskym matematikom W. R. Hamiltonom, ktorý vytvoril hru založenú na nájdení Hamiltonovskej kružnice. Avšak úlohu vo svojej všeobecnosti študovali matematici až v prvej polovici 20. storočia, kde za zmienku stoja viedenskí matematici na čele s Karlom Mengerom, ktorý definovali *problém poštára* (angl. *Messenger problem*), ktorý predstavuje akúsi variáciu problému obchodného cestujúceho. Neskôr sa problémom zaoberali štatistici, napr. P. C. Mahalanobis v súvislosti s jeho poľnohospodárskou aplikáciou. Dnešný názov tejto úlohy pravdepodobne vznikol koncom 40. rokoch minulého storočia v USA.

Je nutné dodať, že systematické štúdium problému obchodného cestujúceho ako kombinatorického a optimalizačného problému ale začalo až v 60. rokoch 20. storočia.

2.2. Definícia problému v teórií grafov

Problém obchodného cestujúceho môžeme riešiť v rámci teórie grafov. Ak vezmeme do úvahy načrtnutý príklad problému obchodného cestujúceho v predošlej kapitole, tak si môžeme predstaviť, že *uzly* grafu sú mestá, *hrany* grafu sú prepojenia medzi mestami, pričom každá hrana má priradené *ohodnotenie* (cena, vzdialenosť) (pozri Obr. 1).

Problém obchodného cestujúceho môžeme v terminológií teórie grafov formálne definovať takto:

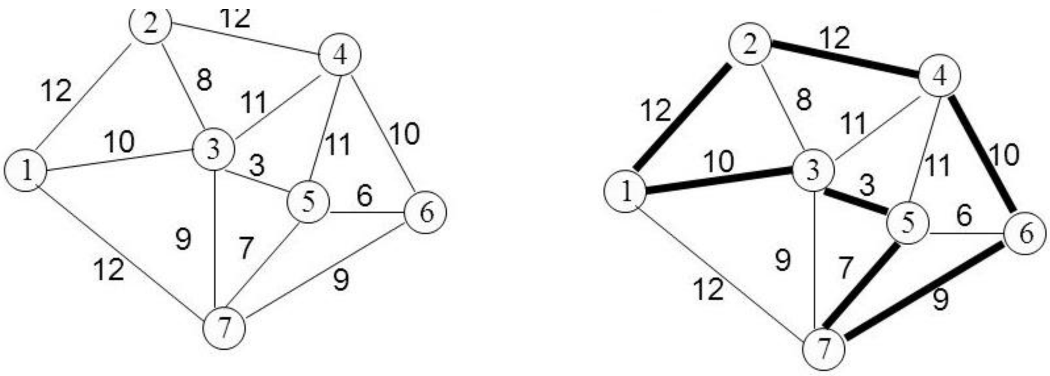
Def. 1. (*Problém obchodného cestujúceho*)

Problém obchodného cestujúceho definujeme ako hľadanie Hamiltonovskej kružnice s minimálnou dĺžkou v súvislom, ohodnotenom, orientovanom alebo neorientovanom grafe (Gutin and Punnen 2006).

Pozn. 1: Pre úplnosť dodajme, že graf nemusí byť súvislí, ale v nesúvislom grafe môžeme chýbajúce hrany nahradiť hranami s extrémne vysokou cenou (Gutin and Punnen 2006).

Pozn. 2: Trasou obchodného cestujúceho je každá Hamiltonovská kružnica, teda ňou môže byť aj najdlhšia Hamiltonovská kružnica a nie nutne tá najkratšia.

Pozn. 3: Pre účely tejto práce uvažujeme len neorientovaný graf.



Obr. 1: Ukážka riešenia problému obchodného cestujúceho.
Vľavo: vstupný graf; vpravo: riešenie - nájdená Hamiltonovská kružnica (tučne).
Zdroj: *neznámy*.

Pre správne a lepšie pochopenie vyššie uvedenej definície je nutné si pripomenúť základné pojmy teórie grafov, ktoré Def. 1 obsahuje. Sú definované takto:

Def. 2. (*Neorientovaný graf*)

Jedná sa o dátovú štruktúru tvorenú usporiadanou trojicou G = ⟨U, H, ρ⟩, pričom prvky množiny *U* predstavujú uzly, prvky množiny *H* hrany a zobrazenie *ρ* je incidencia grafu G, ktorá priraďuje hrane *h* dvojicu uzlov $\{u, v\}$; $h \in H(G)$, $u, v \in U(G)$. Hrany neorientovaného grafu sú teda dvojprvkové množiny, kdežto hrany orientovaného grafu sú usporiadané dvojice a teda v prípade neorientovaného grafu platí, že výraz $\{u, v\}$ a $\{v, u\}$ značí tú istú hranu (Kolář 2004).

Def. 3. (*Ohodnotený graf*)

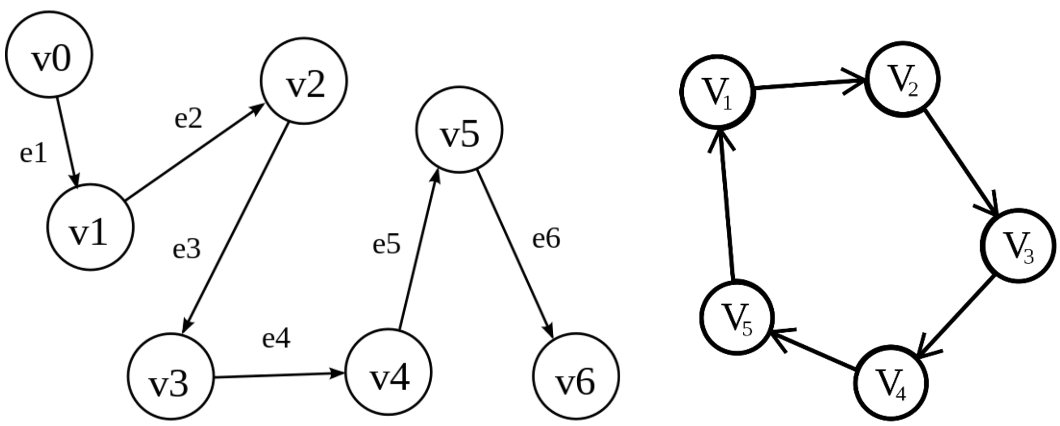
Graf G = (U, H) nazývame ohodnoteným grafom, ak každá jeho hrana *h*, $h \in H(G)$ má pridelené ohodnotenie $w(h)$ vo formáte reálneho čísla (Kovář 2012).

Def. 4. (*Cesta grafom*)

Cestu grafom G = (U, H) rozumieme postupnosť $P = (u_0, h_1, u_1, h_2, u_2, ..., h_n, u_n)$, pričom v tejto postupnosti sa pre každú hranu h_i žiadny uzol u_i neopakuje (Kovář 2012). Inými slovami sa v postupnosti *P* žiadny uzol nenachádza viackrát ako raz.

Def. 5 (*Kružnica grafu*)

Kružnica grafu alebo cyklus grafu C_n je spojenie *n* uzlov *n* hranami a to tak, že každý uzol je spojený s nasledujúcim uzlom a naviac posledný uzol je spojený s prvým uzlom (Kovář 2012). Alternatívnym označením kružnice grafu môže byť *K*.



Obr. 2: Porovnanie cesty grafom (vľavo) a kružnice grafu (vpravo).
Zdroj: Wikipedia 2021a, Wikipedia 2021b.

Def. 6. (*Súvislí graf*)

Súvislí graf je definovaný jednoducho a to tak, že pre jeho ľubovoľné dva uzly *u*, *v* je uzol *v* dosiahnuteľný z uzlu *u*, tj. existuje medzi nimi cesta (Kovář 2012).

Def. 7 (*Hamiltonovská kružnica*)

Hamiltonovská kružnica (cyklus) K_h je kružnica grafu (cyklus), ktorá prechádza všetkými uzlami daného grafu. Jej dĺžka *W* je rovná sume ohodnotenia hrán, ktoré ju tvoria. Graf, v ktorom existuje Hamiltonovská kružnica sa nazýva Hamiltonovský graf (Kovář 2012).

2.3. Aplikácia a využitie problému

(Matai et al. 2010) ako jednu z aplikácií problému obchodného cestujúceho uvádza situáciu, kedy kuriér má za úlohu vyzdvihnúť tovar tvorený viacerými položkami, ktoré sú skladované vo viacerých skladoch. Kuriér teda musí vyzdvihnúť všetky položky a doviesť ich ku zákazníkovi. Sklady tvoria uzly grafu, hrany sú spojnice jednotlivých skladov a sú ohodnotené časom presunu kuriéra.

Spomenutá aj v rámci viacerých publikácií (napr. Hoffman et al. 2013) je aj zaujímavá aplikácia v rámci vŕtania otvorov do dosiek plošných spojov. Tieto otvory môžu byť rôznej veľkosti a tak na vyvŕtanie rôznych otvorov sa vŕtačka musí presunúť na určité miesto, kde je možná výmena vŕtacieho zariadenia, čo zaberá čas. Je teda nutné stanoviť nejaký priemer otvoru, vyvŕtať všetky otvory tohto priemeru, vymeniť vrták a znova vyvŕtať všetky otvory, ktoré majú byť daného priemeru. Každý jeden otvor určitého priemeru je teda uzol, ohodnotenie hrán je čas nutný na presun vŕtačky z jednej polohy do druhej. Celkovým cieľom je minimalizovať čas jazdy vŕtačky (Matai et al. 2010).

Celkovo je možné povedať, že problém obchodného cestujúceho sa využíva v doprave, logistike a v plánovaní.

3. Riešenie problému

Nech riešením problému obchodného cestujúceho je najkratšia cesta medzi mestami, resp. medzi uzlami grafu, tak hľadáme Hamiltonovskú kružnicu K_h s minimálnou dĺžkou W, čo môžeme zapísať nasledovne:

$$W(K_h) = \sum_{i=1}^{n-1} w(u_i, u_{i+1}) = \text{min.}$$

Ako bolo spomenuté v prvej kapitole, problém obchodného cestujúceho je NP-ťažká úloha. Vedci teda hľadajú polynomiálne aproximačné algoritmy – heuristiky, ktoré nenájdu zaručene optimálne riešenie, ale v rozumnom čase riešenie „približné“, ale prípustné, čím sa vyhnú exponenciálnemu nárastu času trvania hľadania riešenia problému. Dá sa povedať, že u jednotlivých heuristík nás zaujíma ich rýchlosť a ako blízko sú k optimálnemu riešeniu (Nilsson 2003).

Ako príklady heuristík môžeme uviesť *konštrukčného heuristiky*, konkrétne Najbližší sused (angl. *Nearest Neighbor*), Greedy metóda, Metóda vkladania (angl. *Insertion method*), Christofidesova metóda (Nilsson 2003), ďalej *heuristiky založené na nájdení najlacnejšej kostry, heuristiky geometrického prístupu* a *vylepšovacie heuristiky* (Valent 2002).

V rámci tejto práce sa zameriame na dve konštrukčné heuristiky – Najbližší sused a *Best Insertion*.

3.1. Metóda najbližšieho suseda

Metóda najbližšieho suseda je považovaná za najjednoduchšiu a za najpriamočiarejšiu. Táto metóda konštruuje Hamiltonovskú kružnicu pripájaním vždy najbližšieho uzlu (Nilsson 2003). Pseudokódom by sa táto metóda dala vyjadriť nasledovne:

1) Inicializácia všetkých uzlov ako „nespracované“ *N*, inicializácia dĺžky Hamiltonovskej kružnice *W* hodnotou 0.
2) Stanovenie štartovného uzlu cesty, v prípade nášho algoritmu je ním prvý uzol, ten sa stáva aktuálnym uzlom a označujeme ho „spracovaný“ *C*.
3) Pokým existuje „nespracovaný“ uzol, vykonaj nasledujúce:
   a. Zo všetkých „nespracovaných“ uzlov vyber ten, ktorý minimalizuje prírastok ohodnotenia cesty, tj. taký, ktorý je najbližšie aktuálnemu uzlu tvoriacemu cestu.

b. Daný uzol zaraď do cesty, označ ho ako „spracovaný“, aktualizuj dĺžku cesty o daný prírastok ohodnotenia.
c. Daný uzol považuj za aktuálny uzol.
4) Dotvorenie kružnice spojením posledného uzlu so štartovným a aktualizácia dĺžky cesty, resp. kružnice.

Táto metóda je častokrát využívaná vďaka svojej jednoduchej implementácií a pre úlohy s malým počtom miest, resp. uzlov, dosahuje uspokojivé výsledky, no pre objemnejšie vstupné množiny bodov si počína výrazne horšie v porovnaní s inými algoritmami. Metóda sa spočiatku vyznačuje pomerne kvalitným zapájaním uzlov, kedy vyberá skutočne najlepšie uzly avšak v priebehu algoritmu zapája uzly na značne horšie pozície ako by sa zdala byť optimálna pozícia. Táto vlastnosť vrcholí na konci algoritmu, kedy sa dotvára Hamiltonovská kružnica daného grafu – spojenie posledného uzlu so štartovným býva veľmi nákladné (Valent 2002).

3.2. Best Insertion

Metódy vkladania sú taktiež považované za relatívne priamočiare a existujú rôzne alternatívy a modifikácie. Základným princípom týchto metód je začiatok z kružnice grafu tvorenej určitou podmnožinou uzlov (napr. troma uzlami) a následne vkladanie zvyšných uzlov na základe určitej heuristiky (Nilsson 2003).

Pseudokódom by sa táto metóda dala vyjadriť nasledovne:

1) Inicializácia dĺžky Hamiltonovskej kružnice W hodnotou 0.
2) Náhodný výber 3 uzlov, ktoré inicializujú hľadanú Hamiltonovskú kružnicu zo zoznamu uzlov a ich odstránenie zo zoznamu uzlov.
3) Určenie dĺžky inicializovanej kružnice.
4) Pokým v zozname uzlov zostáva nejaký uzol, tak vykonaj:
    a. Náhodná voľba uzlu zo zoznamu uzlov. Označme ho u.
    b. Z dosiaľ vytvorenej kružnice grafu nájdi takú hranu medzi uzlami u_i a u_{i+1}, aby vzniknutá cesta cez uzol u predĺžila tvorenú kružnicu minimálne a tým pádom minimalizovala prírastok ohodnotenia Δw
    $$\Delta w = \min w(u_i, u) + w(u, u_{i+1}) - w(u_i, u_{i+1})$$
    c. Uzol u pridaj do tvorenej kružnice medzi uzly u_i a u_{i+1}, aktualizuj dĺžku kružnice a odstráň uzol u zo zoznamu uzlov.

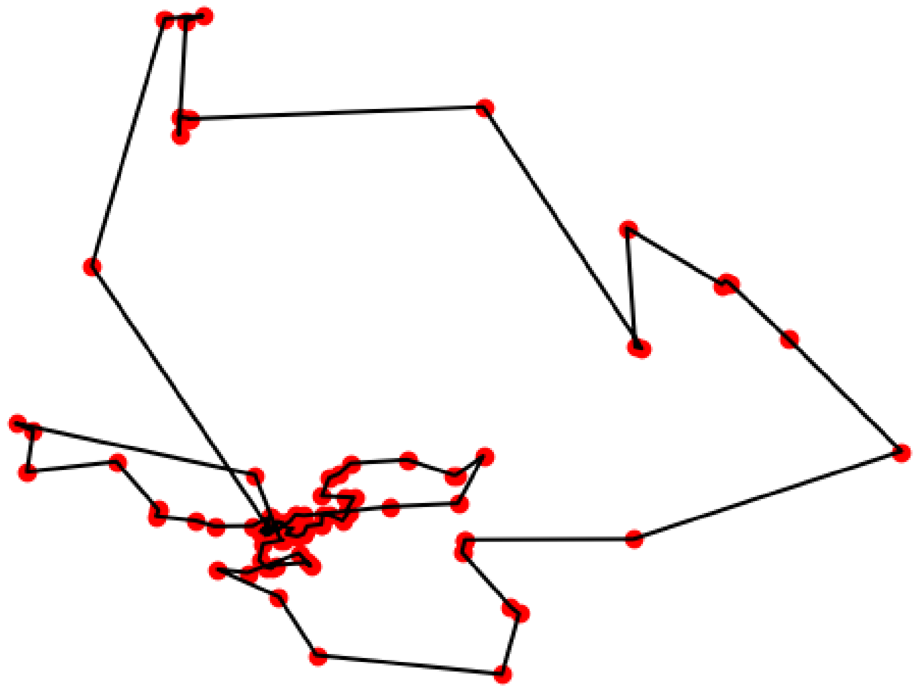
4. Výsledky a záver

Ako vstupné datasety pre otestovanie zdrojových kódov poslúžili geografické, voľne dostupné dáta OpenStreetMap prevedené do .txt formátu. Sú súčasťou odovzdania. Prvý dataset *Dataset 1* je tvorený 134 bodmi, ktoré reprezentujú bary v Bratislavskom kraji, druhý dataset *Dataset 2* tvorí 121 bodov, ktoré predstavujú už záležitosť hodnejšiu prírodovedy - vrcholy Malých Karpát. Ohodnotenie hrán predstavuje vzdušná vzdialenosť v metroch.

Aplikáciu jednotlivých metód na spomenuté datasety hodnotíme dĺžkou nájdenej Hamiltonovskej kružnice *W*. Výsledok dosiahnutý metódou najbližšieho suseda je uvedený v Tab. 1. Vizualizácia výsledku je na Obr. 3 pre *Dataset 1*, pre *Dataset 2* na Obr. 4.

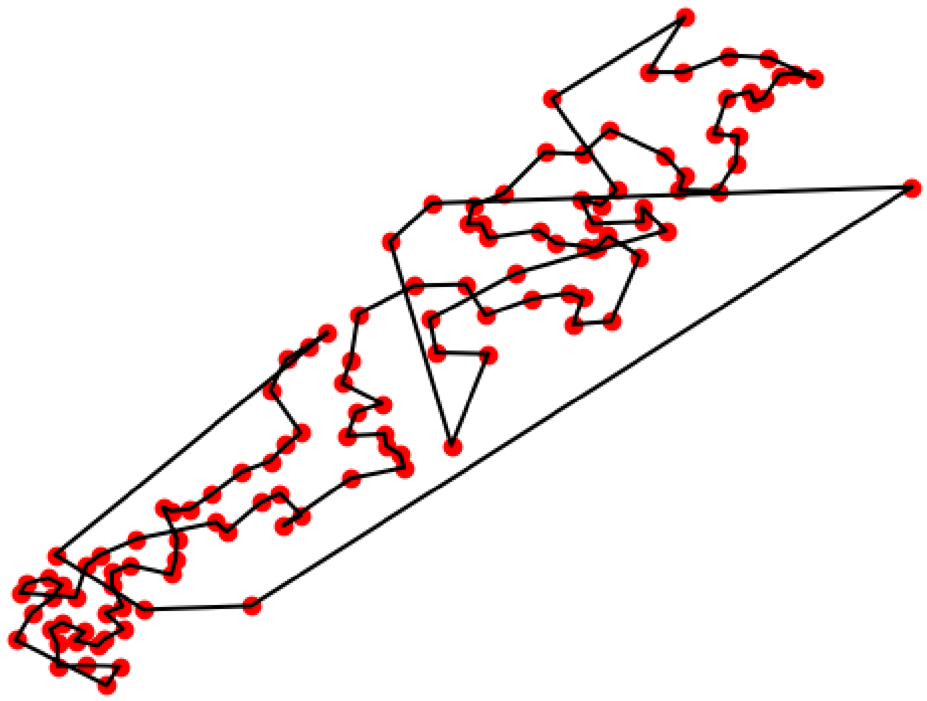
Najbližší sused, W (m)	
Dataset 1 (bary)	Dataset 2 (vrcholy)
237 244,52	221 166,21

Tab. 1: Výsledky dosiahnuté metódou najbližšieho suseda pre obidva datasety.



Obr. 3: Vizualizácia výslednej Hamiltonovskej kružnice vytvorenou metódou najbližšieho suseda pre *Dataset 1*.

Zdroj: autor.



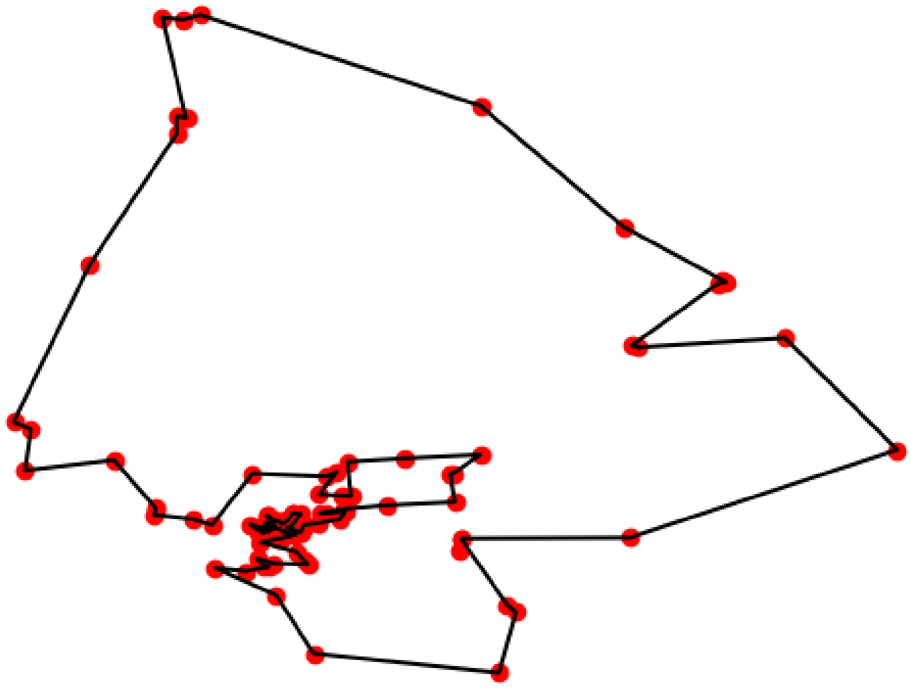
Obr. 4: Vizualizácia výslednej Hamiltonovskej kružnice vytvorenou metódou najbližšieho suseda pre *Dataset 2*.

Zdroj: autor.

V súlade so zadaním úlohy bola metóda *Best Insertion* otestovaná 10x pre každý dataset. Výsledky sú uvedené v Tab. 2. Na Obr. 5 a Obr. 6 je zobrazený najlepší výsledok pre *Dataset 1*, resp. *Dataset 2*.

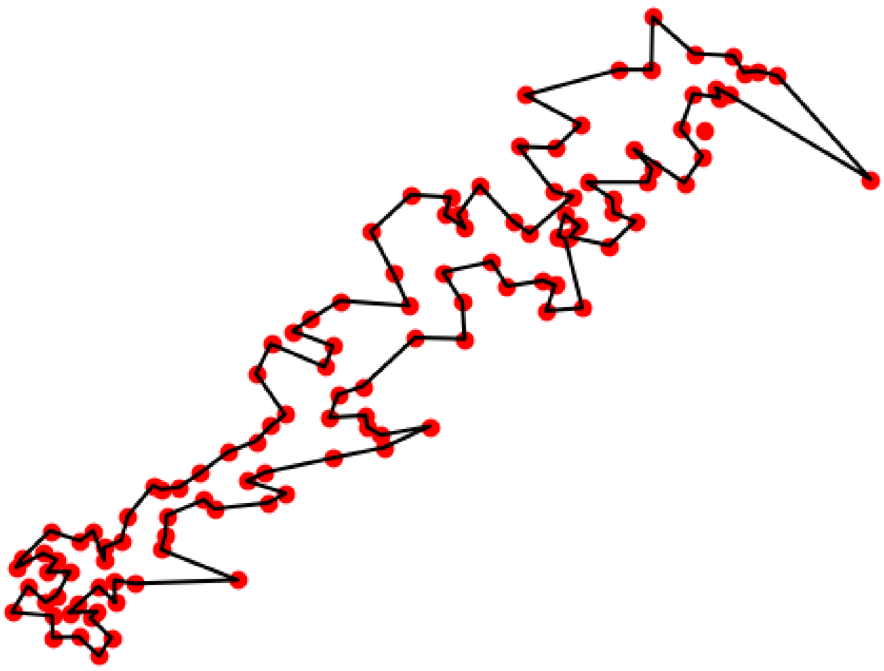
	Best Insertion, W (m)	
	Dataset 1 (bary)	Dataset 2 (vrcholy)
1	206 835,05	165 232,22
2	210 840,20	164 774,39
3	217 774,71	163 987,66
4	210 640,18	165 864,94
5	***197 074,04***	165 675,98
6	200 650,44	166 777,76
7	206 792,83	***160 621,83***
8	203 009,82	168 632,48
9	205 695,82	163 273,96
10	204 520,03	162 015,78

Tab. 2: Výsledky dosiahnuté metódou *Best Insertion* pre obidva datasety. Zvýraznená hodnota značí najlepšiu dosiahnutú hodnotu.



Obr. 5: Vizualizácia najlepšej výslednej Hamiltonovskej kružnice vytvorenou metódou *Best Insertion* pre *Dataset 1*.

Zdroj: autor.



Obr. 6: Vizualizácia najlepšej výslednej Hamiltonovskej kružnice vytvorenou metódou *Best Insertion* pre *Dataset 2*.

Zdroj: autor.

Následne boli výsledky porovnané s výstupom nástroja Network Analyst v ArcMap (Obr. 7 pre *Dataset 1*, Obr. 8 pre *Dataset 2*) a to formou pomeru k, definovaného ako:

$$k = \frac{W}{W_O},$$

pričom W_O značí „optimálne“ riešenie, za ktoré je v tomto prípade považovaný výstup Network Analyst. Hodnoty W_O sú uvedené v Tab. 3. Hodnoty pomeru *k* pre metódu najbližšieho suseda sú v Tab. 3, pre metódu *Best Insertion* v Tab. 4 (porovnanie len s najlepšou dosiahnutou hodnotou).

ArcMap Network Analyst, W_o (m)	
Dataset 1 (bary)	Dataset 2 (vrcholy)
194 250,94	153 098,80

Tab. 3: Dĺžka Hamiltonovskej kružnice W_0 vygenerovanej nástrojom Network Analyst v ArcMap.

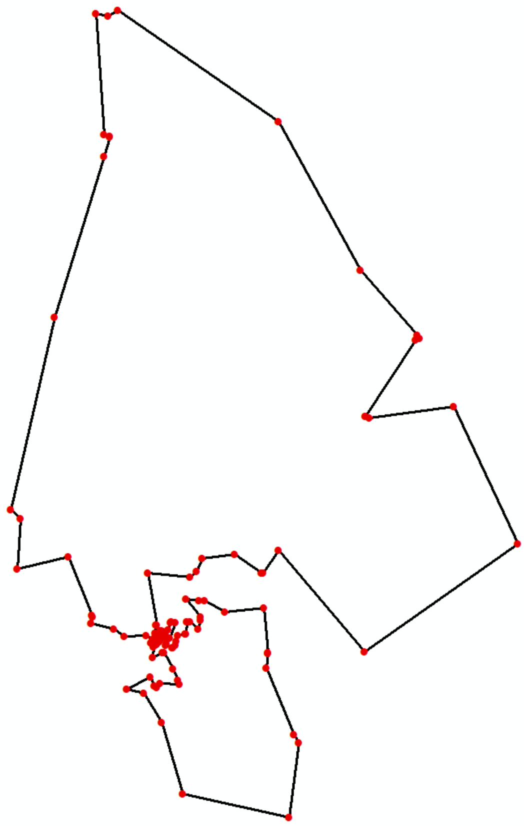
Najbližší sused, k (%)	
Dataset 1 (bary)	Dataset 2 (vrcholy)
122,13	144,46

Tab. 4: Pomer *k* pre výsledné hodnoty metódy najbližšieho suseda.

Best Insertion, k (%)	
Dataset 1 (bary)	Dataset 2 (vrcholy)
101,45	104,91

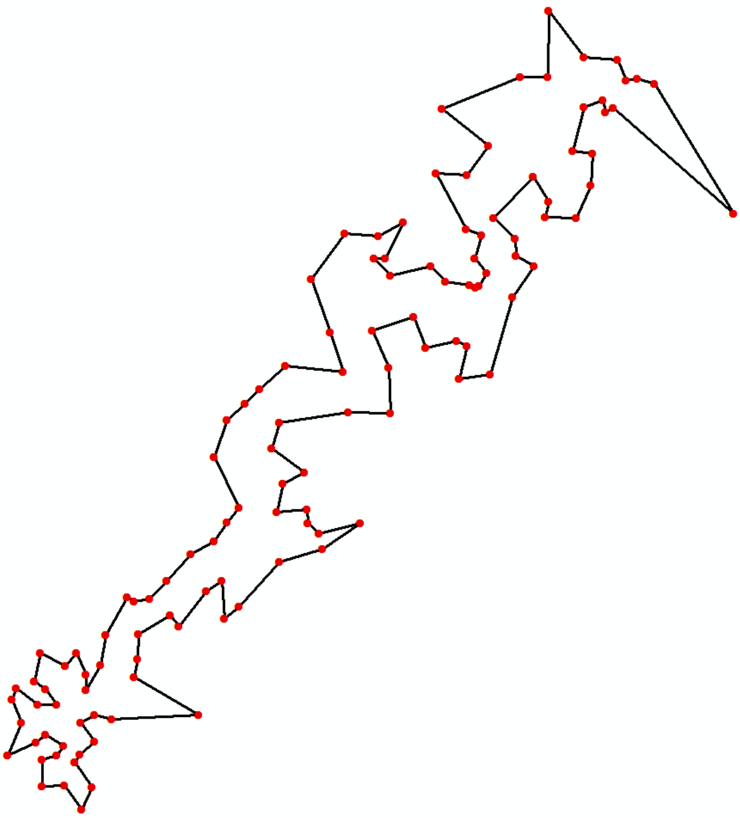
Tab. 5: Pomer *k* pre najlepšie výsledné hodnoty metódy *Best Insertion*.

Výsledné hodnoty pomeru *k* hovoria, o koľko je našimi algoritmami Hamiltonovská kružnica predĺžená oproti výstupu Network Analyst. V tomto prípade dosahuje metóda *Best Insertion* lepšie výsledky ako metóda najbližšieho suseda, čo bolo očakávané a to vzhľadom k tomu, že algoritmus vkladania je sofistikovanejší a oproti metóde najbližšieho suseda mierne zložitejší. V prípade oboch datasetov boli v rámci metódy *Best Insertion* dosiahnuté kvalitné výsledky, kedy kružnica je predĺžená oproti optimu o jednotky percent. Viacerými opakovaniami prebehnutia tejto metódy by potenciálne mohli byť dosiahnuté ešte o niečo lepšie výsledky.



Obr. 7: Výstup nástroja Network Analyst v ArcMap pre *Dataset 1*.

Zdroj: autor.



Obr. 8: Výstup nástroja Network Analyst v ArcMap pre *Dataset 2*.

Zdroj: autor.

5. Použité zdroje

Literatúra:

GUTIN, G., PUNNEN, A. P. (2006): The traveling salesman problem and its variations. Springer Science & Business Media. ISBN 9780306482137.

HOFFMAN, K. et al. (2013): Traveling salesman problem. *Encyclopedia of operations research and management science*, s. 1573-1578.

KOLÁŘ, J. (2004): Teoretická informatika. Skriptum ČVÚT Praha. ISBN 80-900853-8-5.

KOVÁŘ, P. (2012): Teorie grafů. Skriptum VŠB–TU Ostrava.

NILSSON, C. (2003): Heuristics for the Traveling Salesman Problem.

MATAI, R. et al. (2010): Traveling Salesman Problem: An Overview of Applications, Formulations, and Solution Approaches. *Traveling Salesman Problem, Theory and Applications*. DOI: 10.5772/12909.

VALENT, T. (2002): Problém obchodného cestujúceho. Diplomová práca FMPH UK Bratislava.

Obrázky:

WIKIPEDIA (2021a): Cesta (teória grafov) [online]. Dostupné z: https://sk.wikipedia.org/wiki/Cesta_(te%C3%B3ria_grafov) [cit. 1. 1. 2022].

WIKIPEDIA (2021b): Kružnica (teória grafov) [online]. Dostupné z: https://sk.wikipedia.org/wiki/Kru%C5%BEnica_(te%C3%B3ria_grafov) [cit. 1. 1. 2022].