



Duale Hochschule Baden-Württemberg Mannheim

Projektarbeit I

Variation, Analyse und Verbesserung eines Algorithmus zur heuristischen Lösung des Travelling Salesman Problems

Studiengang Wirtschaftsinformatik

Studienrichtung Software Engineering

Verfasser/in: Benno Grimm

Matrikelnummer: 5331201 Firma: SAP SE

Abteilung: Transportation Management

Kurs: WWI18SEA

Studiengangsleiter: Prof. Dr. Julian Reichwald Wissenschaftlicher Betreuer: Prof. Dr. Julian Reichwald

julian.reichwald@dhbw-mannheim.de

+49 (0)621 4105 - 1395

Firmenbetreuer: Peter Wadewitz

peter.wadewitz@sap.com

+49 6227 7-63730

Bearbeitungszeitraum: 13.05.2019 – 01.09.2019

Kurzfassung

Titel Variation, Analyse und Verbesserung eines Algorithmus zur heu-

ristischen Lösung des Travelling Salesman Problems

Verfasser/in: Benno Grimm Kurs: WWI18SEA Ausbildungsstätte: SAP SE

Hier können Sie die Kurzfassung der Arbeit schreiben.

Inhaltsverzeichnis

Abbildungsverzeichnis								
Ta	Tabellenverzeichnis Quelltextverzeichnis							
Qı								
ΑI	Algorithmenverzeichnis							
Αŀ	okürz	ungsverzeichnis	viii					
1	Einle	eitung	1					
2	Das 2.1 2.2	Travelling Salesman Problem Exakte Löungsverfahren und die Optimale Route						
3	Verv 3.1	vendete Lösungsverfahren Insert-First-Verfahren						
		3.1.1Funktionsweise	5					
	3.2	Insert-Furthest-Verfahren	9					
	3.3	3.2.2Zeitkomplexität	11					
		3.3.1Funktionsweise	13 13					
	3.4	Zusammenfassung der Schwächen und Verbesserungsvorschläge						
4		Entfernen von Überschneidungen	19					
	4.2 4.3	Nachbesserung eines Pfads	22					

DHBW Mannheim ii

5	Fallbeispiel Transportation Management SAP 5.1 Implementierung der Algorithmen	
6	Zusammenfassung und Ausblick	24
Lit	teraturverzeichnis	25
Α	Anhang A.1 Bilder	26 26 30

DHBW Mannheim iii

Abbildungsverzeichnis

Abbildung 3.1	Insert-First führt zu schlechtem Ergebnis	6
Abbildung 3.2	Insert-First führt zu guten Ergebnis	8
Abbildung 3.3	Insert-First führt zu schlechtem Ergebnis	11
Abbildung 3.4	Insert-First führt zu schlechtem Ergebnis	12
Abbildung 3.5	Der Insert-Closest Algorithmus kommt zu einem schlechten Er-	
	gebnis	15
Abbildung 4.1	Graph mit und ohne Überkreuzung (Das rote Rechteck in Abbildung a) dient späteren Illustrationszwecken)	18
Abbildung A3	Viele Bilder	28
	Pfad aus 40 Knoten mit und ohne Crossover	
Abbildung A3	Beispiel für eine Nachbesserung, Teilpfad	30

DHBW Mannheim iv

Tabellenverzeichnis

Quelltextverzeichnis

DHBW Mannheim vi

Algorithmenverzeichnis

1	Insert-First-Algorithmus	3
2	Insert-Furthest-Algorithmus	10
3	Insert-Closest-Algorithmus	14
	Erkennen von Überkreuzungen	
6	Einfügen eines neuen Knoten in einen Pfad	30
7	Tauschen von Knoten auf einem Graph zwischen zwei eingegebenen Knoten .	31
8	Berechnung der Distanz zwischen zwei Knoten	31
9	Berechnung der Gesamtdistanz eines Pfads	31

DHBW Mannheim vii

Abkürzungsverzeichnis

DHBW Duale Hochschule Baden-Württemberg

TSP Travelling Salesman Problem

LE Längeneinheiten

Alg. Algorithmus

DHBW Mannheim viii

1 Einleitung

 Das Travelling Salesman Problem (TSP) ist... Duale Hochschule Baden-Württemberg (DHBW) DHBW

2 Das Travelling Salesman Problem

- 2.1 Exakte Löungsverfahren und die Optimale Route
- 2.2 Heuristische Lösungsverfahren

3 Verwendete Lösungsverfahren

3.1 Insert-First-Verfahren

3.1.1 Funktionsweise

Das Insert-First-Verfahren ist ein heuristischer Lösungsansatz des TSPs, bei dem das Betrachten der Knoten zum Aufbau eines Graphen in zufälliger Reihenfolge, bzw. in der Reihenfolge ihrer Erzeugung geschieht. Dabei wird zu einem Zeitpunkt genau ein Knoten betrachtet und an der für ihn bestmöglichen Stelle in den bereits bestehenden Graphen eingefügt.

```
Algorithmus 1 Insert-First-Algorithmus
```

```
Require: Graph G, Pfad P
Require: G = K_1, K_2, ..., K_n, n > 2
 1: p_1 \leftarrow K_1
                                                             ⊳ Setzen der ersten beiden Knoten
 p_2 \leftarrow K_2
 3: for a \leftarrow 3, a \le n, a \leftarrow a + 1 do
        j_S \leftarrow -1
                                                                 ▶ Index der geringsten Distanz
 4:
        d_S \leftarrow -1
                                                                               ▶ Geringste Distanz
 5:
        for b \leftarrow 1, b \leq m, b \leftarrow b + 1 do
 6:
            d_C \leftarrow \text{MERGEAT}(P, b, K_a) \text{ DISTANCE}
                                                                 \triangleright Gesamtdistanz, wenn K_a am
    Index b in P eingefügt werden würde (siehe Alg. 6 auf Seite 30)
            if j_S = -1 or d_C < d_S then
 8:
                 d_S \leftarrow d_C
 9:
                                                        ▶ Kürzeste Distanz wird übernommen
                 j_S \leftarrow b
                                                                                   ▶ Und ihr Index
10:
            end if
11:
        end for
12:
        P \leftarrow \text{MERGEAT}(P, j_S, K_a)
                                                        \triangleright K_a wird am Index j_S in P eingefügt
13:
14: end for
15: return new Graph(P)
                                                     ▷ Graph mit Pfad P wird zurückgegeben
```

Als Eingabe erhält der Algorithmus einen Graphen mit einer ungeordneten Liste von n Knoten K_1, K_2, \ldots, K_n mit n > 2. Jeder Algorithmus erzeugt einen Pfad P, in dem die Knoten des Graphs G eingefügt und angeordnet werden. Dabei kann jede gefüllte Position p_k im Pfad P mit einem Knoten des Graphs G gleichgesetzt werden. Es gilt

also $\forall p \in G$. Wenn beispielsweise:

$$G = K_1, K_2, K_3, K_4$$

und

$$P = p_1, p_2, p_3$$

mit

$$p_1 = K_1, p_2 = K_4, p_3 = K_2$$

dann

$$P = K_1, K_4, K_2$$

Nun wird K_1 , der erste Knoten aus der übergebenen Liste, in den Pfad des Graphs an erster Stelle, p_1 , eingefügt. Dies geschieht so oder ähnlich bei allen Verfahren, um einen statischen Ausgangspunkt zu gewährleisten und somit vergleichbare Ergebnisse zu erzielen. Anschließend wird noch der zweite Knoten, K_2 angehängt.

Das Vorgehen für das Einfügen der restlichen Knoten lässt sich wie folgt beschreiben: Sei G ein Graph mit einer ungeordneten Menge von n Knoten K_1, K_2, \ldots, K_n und bereits teilweise befülltem Pfad K_1, \ldots, K_m mit $m \geq 2$ und m < n. Die Knoten, die noch eingefügt werden müssen, werden in der Reihenfolge ihres Auftretens in der übergebenen Liste in den Graphen eingefügt, womit der als nächstes einzufügende Knoten immer K_i mit i = m + 1 ist.

Um die beste Stelle zu ermitteln, in die K_i eingefügt werden soll, wird für jeden möglichen Index, also jede mögliche Stelle, die Gesamtdistanz des entstehenden Graphen berechnet.

Das niedrigste Ergebnis dieser Möglichkeiten wird zusammen mit dem dazugehörigen Index j vermerkt. Nachdem die niedrigste Distanz für K_i errechnet wurde kann anhand des Index' der Knoten an der bestmöglichen Stelle in den Graphen eingefügt werden. Einfügen bedeutet hier, dass alle Knoten, deren Index gleich oder höher j ist um einen Platz nach hinten verschoben werden. Nachdem alle Knoten auf diese Weise verschoben wurden, kann K_i an der Stelle j eingefügt werden, ohne, dass andere Knoten verloren gehen. Beispielhaft sähe das mit den vorher festgelegten Bezeichnungen wie folgt aus:

$$P = K_1, K_2, K_4, K_3 \text{ und } K_{i-5}$$

Durch das ermitteln der Gesamtdistanzen in Abhängigkeit zu den möglichen Einfügestellen wird bekannt, dass $K_{i=5}$ mit dem Index j=4, also zwischen K_4 und K_3 bestmöglich eingefügt werden kann. Durch das Einfügen nach Algorithmus (Alg.) 6 auf Seite 30 entsteht der Pfad

$$P = K_1, K_2, K_4, K_5, K_3$$

für den Graphen.

3.1.2 Zeitkomplexität

Analysiert man den Insert-First-Algorithmus nach seiner Zeitkomplexität lassen sich, in Verbindung mit seinen Ergebnissen, die im Abschnitt 3.1.3 diskutiert werden, Aussagen über seine Effizienz treffen.

Die äußere Iteration beginnend in Zeile 3 des Algorithmus sei hier I_a und die innere beginnend in Zeile 6 I_b . I_a wird, bedingt durch die Abbruchbedingung $a \leq n$ und das Inkrementieren von a um 1 nach jeder Iteration n-2 mal durchlaufen. Die Laufzeit von I_b ist abhängig von m, der Menge der Elemente des Pfads P. Da in jedem Iterations-durchlauf ein Element in P eingefügt wird berechnet sich m in einem Iterationsschritt mit m=a-1. Für eine Funktion f(n), welche die Komplexität des Algorithmus in der Form $f: \mathbb{N} \to \mathbb{N}$ abbildet, lässt sich also sagen

$$f(n) = \sum_{a=3}^{n} m_a$$
$$f(n) = \sum_{a=3}^{n} a - 1$$
$$f(n) = \frac{(n-1) \cdot (n-2)}{2}$$

Womit sich ergibt, dass der Algorithmus eine quadratische Laufzeit der Form

$$O(n) = n^2$$

hat. Daraus lässt sich ablesen, dass die Laufzeit des Algorithmus zwar nicht exponentiell zur Eingabemenge, aber wenigstens noch quadratisch steigt. Besser wäre hier ein logarithmisches Verhalten, da Algorithmen dieser Art besser für große Eingabemengen Skalieren.¹

3.1.3 Ergebnisse und Schwächen

Bevor einige durch den Algorithmus generierte Beispiele betrachtet werden, wird hier das Szenario dieser und aller folgender Beispiele, es sei denn ist anderes angegeben, beschrieben. Alle gezeigten Knoten befinden sich auf einem zweidimensionalem Fläche mit den Maßen zehn mal zehn Längeneinheiten (LE). Folglich kann jedem Knoten eine X- und Y-Koordinate zwischen jeweils null und zehn zugeordnet werden. Dementsprechend bewegen sich auch die Gesamtdistanzen er gezeigten Graphen in dieser Größenordnung. Weiterhin werden aus Gründen der Übersichtlichkeit für den Großteil der folgenden Beispiele nur Graphen mit fünf Knoten betrachtet.

¹Gurski et al. 2010, S. 9ff.

Bei dem Einsatz des oben beschriebenen Algorithmus kommt es zu Ergebnissen, die in ihrer Qualität nah an die optimale Lösung herankommen, teilweise aber auch weit von ihr abweichen können.

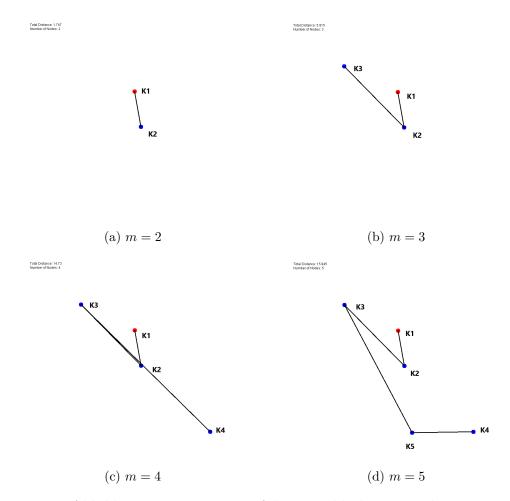


Abbildung 3.1: Insert-First führt zu schlechtem Ergebnis

Auf 3.1c lässt sich erkennen, dass das Einfügen des vierten Knoten K_4 nicht optimal geschieht. Besser für m=4 wäre hier der Pfad

$$P = K_1, K_3, K_2, K_4$$

Dieser wird allerdings nicht durch das Insert-First-Verfahren gebildet, da dies eine Änderung des bereits erzeugten Graphen in 3.1b erfordern würde. Dies ist jedoch nicht möglich, da K_4 nur zwischen bereits im Pfad des Graphen vorhandenen Knoten eingefügt werden kann, sodass der schlussendlich generierte Graph eine Gesamtlänge

von 15,945 LE hat. Hier lässt sich auch das grundlegende Problem des Algorithmus erkennen: Das Erstellen einer Route ohne vorherige Betrachtung der Gesamtheit der Knoten. Einzelne Teilschritte des Graphen können gut erzeugt werden, wie beispielsweise im Schritt von 3.1a auf der vorherigen Seite zu 3.1b auf der vorherigen Seite. Andere hingegen, wie vorher erwähnt, nicht. Grund hierfür ist die alleinige Betrachtung des Knotens K_i . Spezifischer bedeutet das, dass das frühe Einfügen von Knoten in den Pfad eines Graphen später zu Komplikationen führen kann, da es objektiv besser gewesen wäre einen anderen Knoten früher einzufügen. Am konkreten Beispiel führt die generierte Reihenfolge von

$$P = K_1, K_2, K_3$$

in 3.1b auf der vorherigen Seite dazu, dass K_4 nur unter einen vergleichsweise großen Gesamtdistanzzuwachs in den Graphen eingefügt werden kann.

Konträr zu diesem schlechten Beispiel ist der Insert-First-Algorithmus auch in der Lage gute bis optimale Ergebnisse zu generieren.

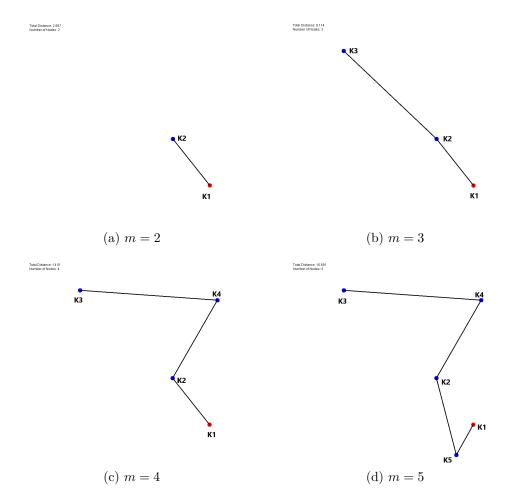


Abbildung 3.2: Insert-First führt zu guten Ergebnis

Am Beispiel in 3.2 lässt sich erkennen, wie der Insert-First-Algorithmus einen optimalen Pfad mit den gegebenen Knoten generiert. Gerade im Schritt von 3.2c zu 3.2d ist ein funktionierendes und korrektes Einfügen des Knotens in den Graphen zu sehen, bei dem der Anstieg der Gesamtdistanz der Route sehr gering gehalten wird. Hier wird der aktuelle Knoten mit geringem Zuwachs der schlussendlichen Gesamtdistanz in den Graphen eingefügt, sodass die Gesamtdistanz zum Ende bei 16,691 LE liegt. Die scheint zwar höher als das vorherige schlechte beispiel, das liegt aber an den Positionen der einzelnen Knoten.

Für die Bewertung des Algorithmus müssen also beide Seiten betrachtet werden. Zwar ist Insert-First in der Lage eine gute oder auch optimale Route zu erstellen, allerdings beeinflusst die Reihenfolge der Betrachtung der Knoten stark die Qualität des Endergebnisses.

3.2 Insert-Furthest-Verfahren

3.2.1 Funktionsweise

Aufbauend auf den Erkenntnissen des Insert-First-Verfahrens können experimentell einige Verbesserungsideen abgeleitet und ihre Auswirkungen auf das Erzeugen eines Graphen betrachtet werden. Beim Insert-First-Verfahren wurde festgestellt, dass eine große Schwäche des Algorithmus die Reihenfolge der Betrachtung der Knoten sein kann. Ein möglicher Ansatz, dieser in 3.1.3 auf Seite 5 beschriebenen Schwäche entgegenzuwirken, ist die Einführung eines Kriteriums zur Betrachtung der Knoten. Eine mögliche Umsetzung eines solchen Kriteriums ist das Insert-Furthest-Verfahren. Hier wird der als nächstes einzufügende Knoten (K_i) durch seine Distanz zum Vorgänger (K_{i-1}) bestimmt.

Ähnlich dem Insert-First-Verfahren wird auch hier ein Graph mit einer Liste von Knoten und einem zu Beginn leerem Pfad erzeugt. Auch hier wird wieder der erste Knoten der Liste K_1 als initialer Knoten p_1 des Pfades P gesetzt. Der nächste zu betrachtende Knoten ist nun aber nicht K_2 , sondern wird durch die Distanz zu K_1 bestimmt. Ausgewählt wird der Knoten, der am weitesten von K_1 , bzw. allgemein am weistesten von K_{i-1} , entfernt ist und nicht bereits Teil des Pfads ist. Dieser Knoten wird nun auf die gleiche Weise wie die Knoten beim Insert-First-Verfahren in den Pfad des Graphen eingefügt; die Stelle mit der geringsten Distanzerhöhung für den Graphen wird gesucht und K_i an dieser Stelle nach dem im Alg. 6 auf Seite 30 beschriebenen Verfahren eingefügt.

Der Gedanke hinter der dieser Veränderung ist der Versuch Knoten mit größerer Vorraussicht als im Insert-First-Verfahren in den Pfad des Graphen einzufügen. Ziel ist es mit den ersten paar Knoten einen Pfad zu generieren, der einen großen Teil der Fläche überspannt, auf der sich Knoten befinden. Das kann insofern zu einem besserem Ergebnis führen, dass die ersten Knoten zwar unter einer, relativ zur schlussendlichen Gesamtlänge des Graphen, hohen Distanzerhöhung eingefügt werden, die nachfolgenden Knoten aber durch geringe Umwege des bestehenden Pfads in den Graphen eingebunden werden können. Auf diese Weise sollen Komplikationen beim Einfügen der letzten Knoten verhindert werden und so suboptimale Graphen wie in 3.1 auf Seite 6 umgangen werden.

Algorithmus 2 Insert-Furthest-Algorithmus

```
Require: Graph G, Pfad P
Require: G = K_1, K_2, \dots, K_n, n > 2
                                                                                 ⊳ Setzen des ersten Knoten
 1: p_1 \leftarrow K_1
 2: i \leftarrow -1
 3: d_i \leftarrow -1
 4: for a \leftarrow 1, a \leq n, a \leftarrow a + 1 do
         j_F \leftarrow -1
         d_F \leftarrow -1
 6:
         for b \leftarrow 2, b \leq n, b \leftarrow b+1 do \triangleright Finde K_b mit der höchsten Distanz zu p_m
 7:
              d_C \leftarrow \text{DISTANCE}(K_b, p_m) \quad \triangleright \text{ Distanz zwischen } K_b \text{ und letztem Knoten } p_m
 8:
              if K_b \notin P and (j_F = -1 \text{ (or) } d_C > d_F) then
 9:
                   d_F \leftarrow d_C
10:
                   j_F \leftarrow b
11:
12:
              end if
         end for
13:
         i_S \leftarrow -1
14:
         d_S \leftarrow -1
15:
          for b \leftarrow 2, b < m, b \leftarrow b + 1 do
16:
              d_C \leftarrow \text{MERGEAT}(P, b, K_{i_F}) \text{ DISTANCE}
17:
              if i_S = -1 or d_C < d_S then
18:
19:
                   i_S \leftarrow b
                   d_S \leftarrow d_C
20:
              end if
21:
22:
          end for
          P \leftarrow \text{MERGEAT}(P, K_{i_F}, i_S)
                                                                                  ⊳ Siehe Alg. 6 auf Seite 30
23:
24: end for
25: return new Graph(P)
```

3.2.2 Zeitkomplexität

Der Insert-Furthest-Algorithmus fügt verglichen mit dem Insert-First-Algorithmus ein Auswahlkriterium hinzu.

Dieses Kriterium drückt sich im Pseudocode im Alg. 2 durch eine zusätzliche Schleife in Zeile vier aus. Um nun die Zeitkomplexität zu ermitteln, reicht es die Laufzeit dieser Schleife, n^2 , auf die in 3.1.2 auf Seite 5 berechnete zu addieren, wodurch sich

$$f(n) = \frac{n^2 + n}{2} + n^2$$

und dadurch auch hier eine Komplexität von

$$O(n) = n^2$$

ergibt. Damit skaliert dieser Algorithmus bei sich verändernder Eingabe in etwa genauso wie Insert-First.

3.2.3 Ergebnis und Schwächen

Testet man des Insert-Furthest-Verfahren anhand der Knoten des Beispiels 3.1 auf Seite 6 wird sichtbar, dass der Algorithmus tatsächlich in der Lage ist einen besseren Pfad zu generieren als das Insert-First-Verfahren.

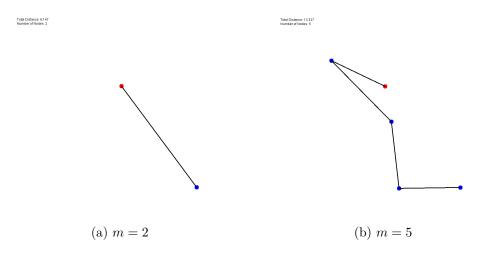


Abbildung 3.3: Insert-First führt zu schlechtem Ergebnis

In 3.4a auf der nächsten Seite ist zu erkennen, dass, anstatt wie in 3.1a auf Seite 6 K_2 als erster Knoten, K_4 eingefügt wird. Dieses Verhalten ist nach der in 3.2.1 auf Seite 9 definierten Funktionsweise zu erwarten, da K_4 der Knoten mit der größten Distanz zu K_1 ist und daher als erstes in den Pfad des Graphen eingefügt wird. Nach der Ausführung aller Schritte des Algorithmus erhält man den in 3.4b auf der nächsten Seite zu sehenden Graphen. Dieser hat eine Gesamtdistanz von 13,327 LE und ist somit im Vergleich mit dem in 3.1 auf Seite 6 durch das Insert-First-Verfahren erzeugten Graph 2,618 LE oder 16,41% kürzer.

Ebenso wie das Insert-First-Verfahren kann das Insert-Furthest-Verfahren auch Graphen generieren die in ihrer Gesamtdistanz vom Optimum abweichen. Am folgenden

Beispiel wird deutlich, dass auch dieser Algorithmus von ähnlichen Schwächen betroffen ist.

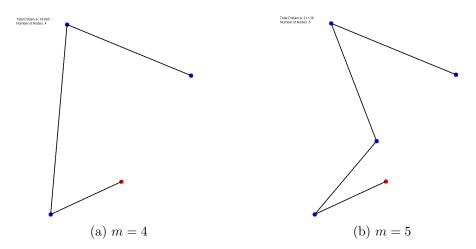


Abbildung 3.4: Insert-First führt zu schlechtem Ergebnis

Die der Abbildung 3.4 vorhergehenden Schritte m=1 bis m=3 sind im Anhang zu finden.

Auch hier lässt sich das Problem der Reihenfolge beobachten, welches in 3.1.3 auf Seite 5 beschrieben wurde. Durch das Einfügen des letzten Knotens K_4 in den Pfad entsteht ein suboptimaler Graph. In diesem konkreten Beispiel ist das mit dem menschlichen Auge jedoch nicht sofort ersichtlich. Vergleicht man aber die Distanzen zwischen den Knoten K_4 und K_1 (6,129 LE) mit den zwischen K_1 und K_2 (5,019 LE) wird schnell ersichtlich, dass eine Verminderung der Distanz durch das Umlegen der Knoten erreicht werden kann. Verursacht wird diese Abweichung vom Optimum dadurch, dass K_4 als letztes in den Graphen eingefügt wird, da der Knoten sich, relativ zu den restlichen Knoten, in der Mitte der Fläche befindet und somit aufgrund seiner geringeren Entfernung vom Algorithmus als betrachtet wird. Die optimale Route

$$P = K_1, K_3, K_4, K_2, K_1$$

würde es jedoch erfordern, dass K_4 früher betrachtet und in den Pfad eingefügt wird. Der durch den Algorithmus erzeugte Graph hat eine Gesamtlänge von 21,138 LE, während durch das Umlegen zum optimalen Graph eine Länge von 20,028 LE, also Reduktion der Distanz um 1,11 LE oder 5,251% erreicht werden kann.

Das Insert-Furthest-Verfahren verhält sich in einigen Fällen, wie in 3.3 auf der vorherigen Seite gezeigt, besser als das Insert-First-Verfahren, weist aber immer noch

eindeutige Schwächen, gerade im Bezug auf die Reihenfolge der Betrachtung der Knoten. Das Beispiel in 3.4 auf der vorherigen Seite zeigt deutlich, wie auch hier die Positionierung der Knoten Einfluss auf das Endergebnis hat.

3.3 Insert-Closest-Verfahren

3.3.1 Funktionsweise

Aufbauend auf den durch das Insert-First- und Insert-Furthest-Verfahren gewonnen Erkenntnissen ist es mögliche weitere Variationen der Heuristik zu entwickeln und deren Ergebnisse zu betrachten. Da sowohl in 3.1 auf Seite 3 als auch in 3.2 auf Seite 9 festgestellt wurde, dass ein Grund für suboptimal erstellte Routen die Reihenfolge der Betrachtung der Knoten ist, wird für das Insert-Closest-Verfahren eine weitere anderes Kriterium für eben diese Reihenfolge festgelegt. Wie der Name des Verfahrens schon suggeriert, geschieht hier die Auswahl der Knoten wieder nach ihrer Distanz.

Das Insert-Closest-Verfahren bezeichnet im Grundprinzip die Umkehrung des Insert-Furthest-Prinzips. Anstatt des am weitesten entfernten Knotens wird hier der dem aktuellen Knoten nächste betrachtet.

Zu Beginn beschreibt sich der Algorithmus identisch zum Insert-Furthest-Verfahren. Auch hier wird ein neuer Graph mit einem leerem Pfad path und einer Liste von Knoten K_1, \ldots, K_n der Länge n erzeugt. Auch hier wird der erste Knoten K_1 als erster Knoten des Pfades festgelegt. Der als nächstes einzufügende Knoten wird wie beim Insert-Furthest-Verfahren durch seine Distanz zum vorherigen bestimmt. Für das Insert-Closest-Verfahren wird der dem Vorherigen Knoten nächste Knoten, unter der Bedingung, dass dieser nicht bereits Teil des Graph ist, in den Pfad eingefügt. Um die beste Stelle zum Einfügen des Knotens zu ermitteln wird das gleiche Verfahren wie bei den beiden vorherigen Algorithmen angewandt; die Stelle, die den geringsten Anstieg für die Gesamtdistanz des Pfads wird ausgewählt. Auch das Vorgehen beim Einfügen orientiert sich hier an den vorherigen Algorithmen. Der aktuelle Knoten K_i wird nach dem in 6 auf Seite 30 dargestelltem Prinzip an dem vorher festgelegten Index j in den Graph eingefügt.

Der beschriebene Algorithmus befindet sich als Pseudocode in etwas ausführlicher Version unter Alg. 2 auf Seite 10.

3.3.2 Zeitkomplexität

Da es sich beim Insert-Closest-Algorithmus im eine Abwandlung des Insert-Furthest handelt, ist die Struktur beider Algorithmen gleich – nur eine Bedingung ändert sich.

Algorithmus 3 Insert-Closest-Algorithmus

```
Require: Graph G, Pfad P
Require: G = K_1, K_2, ..., K_n, n > 2
                                                                                 ⊳ Setzen des ersten Knoten
 1: p_1 \leftarrow K_1
 2: i \leftarrow -1
 3: d_i \leftarrow -1
 4: for a \leftarrow 1, a \le n, a \leftarrow a + 1 do
         j_S \leftarrow -1
 5:
          d_S \leftarrow -1
 6:
          for b \leftarrow 2, b \leq n, b \leftarrow b+1 do \triangleright Finde K_b mit der geringsten Distanz zu p_m
 7:
               d_C \leftarrow \text{DISTANCE}(K_b, p_m) \quad \triangleright \text{ Distanz zwischen } K_b \text{ und letztem Knoten } p_m
 8:
              if K_b \notin P and (j_S = -1 \text{ (or) } d_C > d_S) then
 9:
10:
                   d_S \leftarrow d_C
                   j_S \leftarrow b
11:
               end if
12:
          end for
13:
          i_S \leftarrow -1
14:
          d_S \leftarrow -1
15:
          for b \leftarrow 2, b < n, b \leftarrow b + 1 do
16:
              d_C \leftarrow \text{MERGEAT}(P, b, K_{j_F}) \text{ DISTANCE}
17:
               if i_S = -1 or d_C < d_S then
18:
                   i_S \leftarrow b
19:
                   d_S \leftarrow d_C
20:
               end if
21:
22:
          end for
          P \leftarrow \text{MERGEAT}(P, K_{j_S}, i_S)
                                                                                  ⊳ Siehe Alg. 6 auf Seite 30
23:
24: end for
25: return new Graph(P)
```

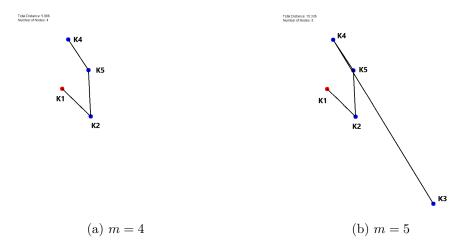


Abbildung 3.5: Der Insert-Closest Algorithmus kommt zu einem schlechten Ergebnis

Daher kann auch für diesen Algorithmus eine Zeitkomplexität von

$$O(n) = n^2$$

angenommen werden, womit auch dieser Algorithmus wie Insert-First skaliert.

3.3.3 Ergebnis und Schwächen

Auch die Ergebnisse dieses Algorithmus gestalten sich sehr divers. Wie bei den beiden anderen Verfahren entstehen hier Graphen, die nahe an die Optimale Route heranreichen oder ihr teilweise auch entsprechen, aber auch solche, die weit von der optimalen Lösung abweichen. Übergibt man dem Algorithmus beispielsweise die gleichen Knoten wie in Abbildung 3.1 auf Seite 6, einem Szenario, bei dem der Insert-First Algorithmus einen suboptimalen Graph generiert, kommt der Insert-Closest Algorithmus, so wie der Insert-Furthest Algorithmus auch, auf die optimale Route.

So wie die anderen Algorithmen stößt aber auch der Insert-Closest Algorithmus bei bestimmten Konstellationen von Knoten an seine Grenzen.

Wie in Abbildung 3.5 zu sehen ist, plant der Algorithmus einen Pfad, der deutlich erkennbar nicht optimal ist. Die einzelnen Schritte, die zur Generierung des Graphen führen finden sich im Anhang in der Abbildung A.1 auf Seite 29. Der in 3.5b gezeigt Pfad hat eine Gesamtdistanz von 15,336 LE. Durch das Ändern der Reihenfolge der Knoten zu

$$P = K_1, K_4, K_5, K_2, K_3$$

könnte eine Verringerung der Distanz im Vergleich zum vorherigen Resultat um 3,208 LE, bzw. 20,918% erreicht werden. Auch hier wird der suboptimal geplante Pfad durch die Reihenfolge der Betrachtung der Knoten verursacht. Im konkreten Beispiel werden die ersten vier Knoten K_1, K_2, K_4 und K_5 zu einem für sich optimalen Pfad zusammengefügt. K_3 wird aufgrund seiner hohen Distanz zu den übrigen Knoten als letztes in den Pfad eingefügt. Im gezeigten Beispiel kommt es also genau zu der Umkehrung des Problems, welches beim Insert-Furthest Algorithmus besteht; dort werden Knoten aufgrund ihrer zu niedrigen Distanz teilweise zu spät eingefügt. Diese Problem hat hier zur Folge, dass K_3 als letzter Knoten des Pfads nach dem von ihm am weitesten entfernten Knoten eingefügt wird, was zu einem hohen Zuwachs der Gesamtdistanz führt.

3.4 Zusammenfassung der Schwächen und Verbesserungsvorschläge

Die Ergebnisse, welche die Algorithmen erzeugen, können zum jetzigen Zeitpunkt nur begrenzt miteinander verglichen werden, da noch zu wenige Beispiele vorliegen. Allerdings lassen sich durch die wenigen Beispiele die vorliegen bereits einige Beobachtungen treffen, anhand deren die erzeugten Graphen nachträglich noch verbessert werden können. Betrachtet man nun diese Graphen, die durch alle drei Algorithmen erzeugt werden, so fallen bei allen schnell Schwächen auf, die ihr Ergebnis beeinträchtigen. Anhand der vorherigen Beispiele wurde deutlich, dass die Algorithmen zwar in der Lage sind gute Ergebnisse zu erzeugen, gleichzeitig aber auch auf sich allein gestellt nicht sehr zuverlässig sind. Dies trifft sowohl auf das Insert-First Verfahren, als auch auf die anderen vorgestellten Algorithmen zu.

Bezüglich des Verhältnis' von Komplexität und Nutzen lässt sich keine genaue Aussage treffen, welcher Algorithmus effizienter arbeitet. Alle drei Algorithmen weisen eine Komplexität von $O(n) = n^2$ und erzeugen gleichzeitig Ergebnisse, die in ihrer Qualität von optimal bis erkennbar verbesserbar reichen.

Weiterhin ist in Abbildung 3.5 auf der vorherigen Seite ist ein Phänomen zu beobachten, welches gerade bei Anwendung der beschriebenen Algorithmen mit mehr Knoten häufig auftritt. Das Entstehen von Überkreuzungen von Kanten zwischen zwei Knoten (siehe hierzu Abbildung A.1 auf Seite 30). Gerade bei solchen Überkreuzungen besteht immer die Möglichkeit den Graph so umzulegen, dass die Gesamtdistanz sinkt. Daher ist es sinnvoll einen Algorithmus zu entwickeln, der versucht Überkreuzungen aufzulösen.

Weiterhin fallen bei manchen Graphen Konstellationen von Knoten und bestehenden Kanten auf, bei denen durch ein einfaches Umlegen der Kanten, bzw. der Knotenreihenfolge, signifikante Verringerungen in der Distanz erzielt werden können. Beispiels-

weise erkennt man in der Veränderung von Abbildung A.-2h auf Seite 30 zu A.-3i auf Seite 30 ein solches Umlegen und die einhergehende Distanzverminderung. Auch hier kann ein Algorithmus, der gezielt nach solchen Konstellationen sucht und diese aufhebt, Abhilfe schaffen.

Nachdem diese Algorithmen im nächsten Kapitel implementiert werden folgt eine Analyse der Ergebnisse verschiedener Kombinationen der vorgestellten Algorithmen, basierend auf mehreren Testläufen mit zufällig erzeugten Knoten. Dadurch soll es möglich sein die Algorithmen untereinander und in Verbindung miteinander zu vergleichen.

4 Verbesserung eines bestehenden Pfads

4.1 Entfernen von Überschneidungen

4.1.1 Beschreibung des Problems

Wie in Abschnitt 3.4 auf Seite 16 angedeutet kommt es kommt es bei generierten Graphen zu Überkreuzungen von Teilrouten zwischen jeweils zwei Knoten. Betrachtet man solche Überkreuzungen im Detail fallen einige Gemeinsamkeiten zwischen ihnen auf. So ist es beispielsweise immer möglich eine Überkreuzung durch das Verändern der Reihenfolge der Knoten im Pfad aufzulösen und so eine Verringerung in der Gesamtdistanz zu erreichen.

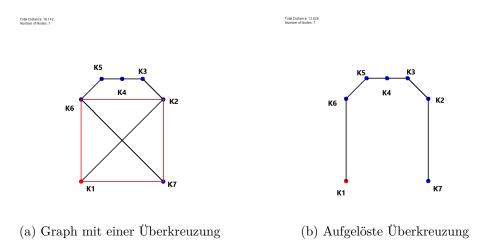


Abbildung 4.1: Graph mit und ohne Überkreuzung (Das rote Rechteck in Abbildung a) dient späteren Illustrationszwecken)

Anhand dieses Beispiels wird nun das Entstehen, Erkennen und Auflösen von Überkreuzungen erläutert.

Eine Überkreuzung repräsentiert das Auftreten eines Schnittpunkts von zwei Kanten eines Graphen in einem für den Graphen relevanten Bereich. Ein Schnittpunkt von zwei Kanten bedeutet hier, dass keine der vier Knoten gleich sein dürfen. Besteht eine Überkreuzung also aus den Knotenpaaren A und B mit $A = K_{A_1}, K_{A_2}$ und

 $B=K_{B_1},K_{B_2}$, dann muss gelten $K_{A_1} \neq K_{A_2} \neq K_{B_1} \neq K_{B_2}$. Ist diese Bedingung nicht erfüllt, kann es keine Überkreuzung geben. Sind nur drei der vier benötigten Knoten einzigartig kann es nicht zu einer Überkreuzung kommen, da dies einen zusammenhängenden Streckenabschnitt der Form K_1, K_2, K_3 darstellen würde.

4.1.2 Erkennnen von Überkreuzungen

Um Überkreuzungen erkennen zu können ist die Bedingung "in einem für den Graphen relevanten Bereich" wichtig. Betrachtet man zwei zufällig ausgewählte Kanten als unendliche Linien, stellt man sie also in der zweidimensionalen Ebene als lineare Funktion, mit einer Steigung und einem Schnittpunkt mit der Ordinate, dar, dann schneiden sich alle diese Funktionen an irgendeinem Punkt, es sei denn sie sind parallel zueinander. Um zu überprüfen, ob eine Überkreuzung im für die Generierung eines Graphen relevanten Bereich ist, wird zuerst der Schnittpunkt der beiden Kanten berechnet. Dazu wird aus zwei Knoten einer Kante eine lineare Funktion der Form f(x) = mx + n simuliert, wobei $m = \frac{\Delta y}{\Delta x}$ und n = y - mx. Sind die Knoten der beiden Kanten nun A_1 , A_2 und B_1 , B_2 , dann ergibt sich für die Berechnung der Schnittstelle:

$$x_S = \frac{(y_{B_1} - \frac{y_{B_2} - y_{B_1}}{x_{B_2} - x_{B_1}} \cdot x_{B_1}) - (y_{A_1} - \frac{y_{A_2} - y_{A_1}}{x_{A_2} - x_{A_1}} \cdot x_{A_1})}{(\frac{y_{A_2} - y_{A_1}}{x_{A_2} - x_{A_1}}) - (\frac{y_{B_2} - y_{B_1}}{x_{B_2} - x_{B_1}})}$$
(4.1)

$$y_S = \frac{y_{A_2} - y_{A_1}}{x_{A_2} - x_{A_1}} \cdot x_S + \left(y_{A_1} - \frac{y_{A_2} - y_{A_1}}{x_{A_2} - x_{A_1}} \cdot x_{A_1}\right) \tag{4.2}$$

Mit x_S und y_S lässt sich der Punkt $S(x_S|y_S)$ konstruieren.

Mit Hilfe des Punkts S gilt es nun zu überprüfen, ob sich dieser im relevanten Bereich befindet. Um dies zu bestimmen wird um die Knoten beider Kanten jeweils ein Recht simuliert, wie es beispielhaft in Abbildung 4.1a auf der vorherigen Seite eingezeichnet ist. Eine Überkreuzung ist genau dann für den Algorithmus relevant, wenn sie in den Rechtecken beider Kanten liegt. Um dies zu überprüfen wird folgender Algorithmus angewandt:

Algorithmus 4 Erkennen von Überkreuzungen

```
Require: Knoten A_1, A_2, Punkt P
 1: score \leftarrow 0
 2: if x_{A_1} > x_{A_2} then
 3:
        if x_P > x_{A_2} and x_P < x_{A_1} then
 4:
            score \leftarrow score +1
        end if
 5:
 6: else
        if x_P < x_{A_2} and x_P > x_{A_1} then
            score \leftarrow score +1
 8:
 9:
        end if
10: end if
11: if y_{A_1} > y_{A_2} then
        if y_P < y_{A_1} and y_P > y_{A_2} then
12:
            score \leftarrow score +1
13:
        end if
14:
15: else
        if y_P > y_{A_1} and y_P < y_{A_2} then
16:
17:
            score \leftarrow score +1
18:
        end if
19: end if
20: return (score == 2) \triangleright score == 2 gibt wahr zurück und signalisiert, dass P im
    Rechteck von A_1 und A_2 ist
```

An dieser Stelle sei angemerkt, dass das Problem der Überkreuzungserkennung auch mit Hilfe von Vektorenskalierung lösbar ist. Dieses Verfahren wird jedoch in dieser Arbeit nicht diskutiert.

4.1.3 Auflösen von Überkreuzungen

Um einen Algorithmus zur Auflösung von Überkreuzungen entwickeln zu können, ist es wichtig die Knoten um eine Überkreuzung herum vor und nach deren Auflösung zu betrachten. Dazu kann als Beispiel wieder Abbildung 4.1 auf Seite 18 dienen. Reihenfolge der Knoten in Abbildung 4.1a auf Seite 18 ist

$$P_{alt} = K_1, K_2, K_3, K_4, K_5, K_6, K_7$$

Die Reihenfolge der Knoten nach dem Auflösen in Abbildung 4.1b auf Seite 18 ist

$$P_{neu} = K_1, K_6, K_5, K_4, K_3, K_2, K_7$$

In diesem Beispiel seien die betroffenen Kanten A und B mit den Knoten $A_1 = K_1$, $A_2 = K_2$ und $B_1 = K_6$, $B_2 = K_7$. Die hier interessanten Knoten sind K_2 bis K_6 , da sich deren Reihenfolge umkehrt. Daraus kann gefolgert werden, dass zum Auflösen einer Überkreuzung das Umkehren der betroffenen Knoten in der Mitte reicht. Diese betroffenen Knoten bestimmen sich durch die Kanten A und B – der erste umzukehrende Knoten ist immer A_2 , während der letzte B_1 ist. Dies ist allerdings nur unter der Bedingung wahr, dass A im Graph vor B ist. Ist dies nicht der Fall kehren sich die Rollen der Knoten um und B_1 ist der erste, während A_2 der letzte umzukehrende Knoten ist. Ein Algorithmus, der auf einem Pfad mit Überkreuzung und bekannten A und B eben dieses Tauschen ausführt, findet sich im Anhang unter Algorithmus 7 auf Seite 31.

Ein vollständiger Algorithmus, der die aufgeführten Methodiken anwenden soll, um Überkreuzungen zu erkennen und aufzulösen, müssen also alle Kanten gegeneinander geprüft werden. Eine simple Umsetzung davon, die alle beschriebenen Methodiken mit einschließt kann sein

Algorithmus 5 Erkennen und Auflösen von Überkreuzungen auf einem Pfad

```
Require: Pfad P
Require: P = p_1, p_2 \cdots, p_n, n \ge 4, \forall p \in G
 1: for a \leftarrow 2, a \le n, a \leftarrow a + 1 do
          for b \leftarrow b, b \le n, b \leftarrow b+1 do
 2:
              if not (p_a \neq p_b \text{ and } p_a \neq p_{b-1} \text{ and } p_{a-1} \neq p_b \text{ and } p_{a-1} \neq p_{b-1} \text{ then}
 3:
 4:
              end if
 5:
              x_S \leftarrow \text{nach } 4.1 \text{ auf Seite } 19
 6:
                                                                  \triangleright p_a und p_{a-1} entsprechen A_1 und A_2
              y_S \leftarrow \text{nach } 4.2 \text{ auf Seite } 19
                                                                  \triangleright p_b und p_{b-1} entsprechen B_1 und B_2
 7:
              if CHECK(p_a, p_{a-1}, S(x_S|y_S)) and CHECK(p_b, p_{b-1}, S(x_S|y_S)) then
     CHECK repräsentiert dabei Alg. 4 auf der vorherigen Seite
                   P \leftarrow \text{RESOLVE}(P, p_a, p_{b-1})
                                                              ▶ Auflösen der Überkreuzung mit Alg. 7
 9:
     auf Seite 31
              end if
10:
          end for
11:
12: end for
```

Wird das Entfernen von Überkreuzungen nach diesem Prinzip implementiert, ergeben sich einige Randbedingungen, die es Wert sind erwähnt zu werden. Aufgrund der geschachtelten Iterationen über die Kanten des Graphs lässt sich eine Zeitkomplexität von

$$O(n) = n^2$$

ermitteln, womit der Algorithmus im Rahmen der polynomialen Zeitkomplexitätsklasse liegt. Dies bedeutet, dass die Laufzeit des Algorithmus proportional zum Quadrat

seiner Eingabemenge wächst. 2 Als Eingabemenge können hier Knoten bzw. Kanten eines Graphen behandelt werden, wobei n die Menge der Knoten repräsentiert. Ein Algorithmus, der Überkreuzungen aus einem Graph entfernt, arbeitet also mit einer ähnlichen Laufzeit wie die Heuristiken, die den Graph vorher erzeugen.

Weiterhin ist es möglich, dass durch das Auflösen einer Überkreuzung eine weitere, neue entsteht. Falls dies so geschieht, dass in den restlichen Iterationen über die Kanten diese Überkreuzung nicht mehr erkannt wird, beispielsweise, wenn dies in der letzten Iteration passiert, dann wird die Überkreuzung nicht vom Algorithmus aufgelöst. Eine Möglichkeit dies zu umgehen ist durch das rekursive Aufrufen des Algorithmus, damit mehrmals auf Überkreuzungen überprüft wird. Allerdings besteht hier Bedarf eine solche Implementierung genauer zu untersuchen, was in dieser Arbeit nicht behandelt wird.

4.2 Nachbesserung eines Pfads

4.3 Komplexität und Nutzen der Verbesserungen

²Gurski et al. 2010, S. 15.

5 Fallbeispiel Transportation Management SAP

- 5.1 Implementierung der Algorithmen
- 5.2 Auswirkungen auf die Routenplanung

6 Zusammenfassung und Ausblick

Literaturverzeichnis

Gurski, Frank et al. (2010). Exakte Algorithmen für schwere Graphenprobleme. Berlin und Heidelberg: Springer. ISBN: 978-3-642-04499-1. DOI: 10.1007/978-3-642-04500-4.

A Anhang

A.1 Bilder

Total Distance: 0.0 Number of Nodes: 1



(a)
$$m = 2$$

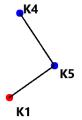
Anhang A

Total Distance: 1.565 Number of Nodes: 2



(b)
$$m = 3$$

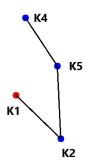
Total Distance: 3.353 Number of Nodes: 3



(c)
$$m = 4$$

Anhang A

Total Distance: 5.986 Number of Nodes: 4



(d) m = 5



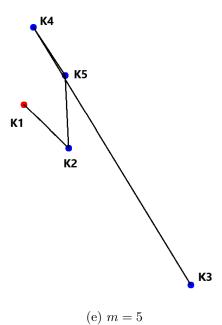
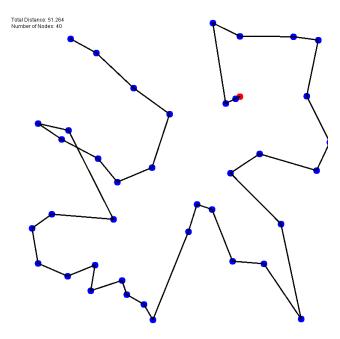
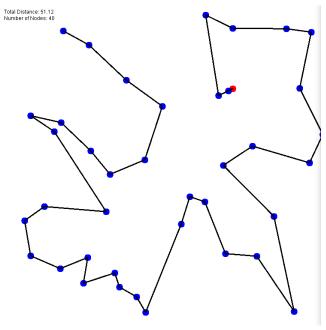


Abbildung A.-3: Viele Bilder

Anhang A



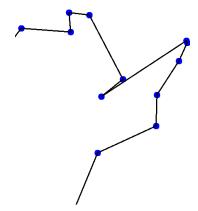
(f) Pfad mit einem Crossover



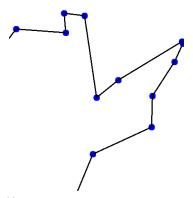
(g) Pfad mit aufgelöstem Crossover

Abbildung A.-3: Pfad aus 40 Knoten mit und ohne Crossover

Anhang A Anhang



(h) Teilpfad vor Nachbesserung



(i) Teilpfad nach Nachbesserung

Abbildung A.-3: Beispiel für eine Nachbesserung, Teilpfad

A.2 Algorithmen

Algorithmus 6 Einfügen eines neuen Knoten in einen Pfad

Require: Pfad $P = p_1, \dots, p_n$ mit $\forall p \in G$ \triangleright Jedes p ist Knoten in Graph G

Require: Knoten K^* , Index $i, i \leq n+1$ > neuer Knoten K^* , einzufügen an Index i

1: for $a \leftarrow n$, $a \ge i$, $a \leftarrow a + 1$ do

- 2: $p_{a+1} \leftarrow p_a$
- 3: end for
- $4: p_i \leftarrow K^*$
- 5: return P

Anhang A Anhang

Algorithmus 7 Tauschen von Knoten auf einem Graph zwischen zwei eingegebenen Knoten

```
Require: Graph G, Knotenpaare A_1, A_2 und B_1, B_2
```

Require: $G = K_1, K_2, \dots, K_n, n > 4$

- 1: $i_{A_2} \leftarrow \text{INDEX}(A_2)$
- 2: $i_{B_1} \leftarrow \text{INDEX}(B_1)$
- 3: **if** $i_{A_2} > i_{B_1}$ **then**
- 4: SWAP (i_{A_2}, i_{B_1}) > SWAP weißt beiden Parametern den Wert des anderen zu
- 5: end if
- 6: while $i_{A_2} < i_{B_1}$ do
- 7: SWAP $(K_{i_{A_2}}, K_{i_{B_1}})$
- 8: end while

Algorithmus 8 Berechnung der Distanz zwischen zwei Knoten

Require: Knoten A, Knoten B

- 1: $d \leftarrow \sqrt{|x_A x_B|^2 + |y_A y_B|^2}$
- 2: return d

Algorithmus 9 Berechnung der Gesamtdistanz eines Pfads

Require: Pfad P

Require: $P = p_1, \dots, p_n, n \ge 2$

- 1: sum $\leftarrow 0$
- 2: for $a \leftarrow 2$, $a \le n$, $a \leftarrow a + 1$ do
- 3: $\operatorname{sum} \leftarrow \operatorname{sum} + \operatorname{DISTANCE}(p_{a-1}, p_a)$
- 4: end for
- 5: return sum

Ehrenwörtliche Erklärung

Ich versichere hiermit, dass ich die vorliegende Arbeit mit dem Thema: Variation, Analyse und Verbesserung eines Algorithmus zur heuristischen Lösung des Travelling Salesman Problems selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe. Ich versichere zudem, dass die eingereichte elektronische Fassung mit der gedruckten Fassung übereinstimmt.

Ort, Datum Benno Grimm