

**Untersuchung der Korrelation von  
Maßnahmen zur Verbesserung logistischer  
Netzwerke mittels Evolutionärer Algorithmen**

Masterthesis zur Erlangung des akademischen Grades:

**Master of Science (M. Sc)**

im Fach Maschinenbau

Fakultät Maschinenbau  
Fachgebiet IT in Produktion und Logistik  
Technische Universität Dortmund

**Autor:** René Zimmermann,  
rene.zimmermann@tu-dortmund.de,  
Matrikelnummer: 156995

**Ausgegeben am:** 11.07.2018

**Eingereicht am:** 27.12.2018

**Betreuer:** Univ.-Prof. Dr.-Ing. Markus Rabe  
M.Sc. Majsja Ammouriova

# Inhaltsverzeichnis

Abkürzungsverzeichnis .....	I
Abbildungsverzeichnis .....	II
Tabellenverzeichnis.....	III
Formelverzeichnis .....	IV
1. Einleitung.....	1
2. Logistische Einordnung.....	4
2.1 Supply Chains.....	4
2.2 Der Netzwerkbegriff .....	6
2.3 Elemente des logistischen Netzwerkes.....	11
2.4 Beziehungen innerhalb Logistischer Netzwerke .....	15
3. Evolutionäre Algorithmen .....	18
3.1 Terminologie.....	19
3.2 Aufbau eines Evolutionären Algorithmus.....	20
3.3 Programmiersprache .....	32
4. Initialisierung des Algorithmus .....	33
4.1 Problemstellung.....	36
4.2 Kodierung und Initialisierung.....	39
4.2.1 Binäre Kodierung.....	39
4.2.2 Anlehnung an das “Traveling Salesperson Problem“ .....	40
4.2.3 Anlehnung an das “TSP with Forbidden Neighborhoods” .....	48
5. Fortpflanzung der Eltern.....	52
5.1 Bestimmung der Fitnesswerte .....	52
5.2 Rekombination der Elternchromosome .....	54
5.3 Mutationen.....	57
6. Programmdurchlauf .....	59
6.1 Implementierung Distanzmatrix.....	60
6.2 Implementierung Crossing Over .....	60
6.3 Implementierung Mutation .....	61
6.4 Implementierung Abbruchsbedingung .....	61
7. Auswertung .....	62
8. Zusammenfassung .....	71
Literaturverzeichnis.....	73
Anhang.....	78

## Abkürzungsverzeichnis

EA	Evolutionärer Algorithmus
PN	Projektnetze
RN	Regionale Netzwerke
SC	Supply Chain
SN	Strategische Netzwerke
TSP	Traveling Salesperson Problem
VU	Virtuelle Unternehmungen

## Abbildungsverzeichnis

2.1	Supply Chain--Kette [nach COR08].....	5
2.2	Zusammenhang Netzwerke und Prozesse [BRE10].....	6
2.3	Netzwerkarten [ZIR07].....	9
2.4	Darstellung eines Logistischen Netzwerkes nach [WAG05].....	11
2.5	Zentrallagerkonzept.....	13
2.6	Dezentrallagersystem.....	13
3.1	Darstellung Chromosom nach [KOC14].....	20
3.2	Ablauf von Evolutionären Algorithmen [RUD15].....	21
3.3	Fitnesswertbestimmung [KOC14].....	23
3.4	Darstellung der Roulette-Verteilung.....	26
3.5	Sammlung von Mutationen [POH99].....	30
4.1	Abhängigkeitsverhältnisse der Variablen.....	36
4.2	Ablaufschema TSP.....	46
4.3	Abstände von einem Punkt [FIS17].....	49
7.1	Verlauf der Kosten über mehrere Generationen.....	66
7.2	Verlauf der Kosten bei Mutationsrate von 0,1.....	68
7.3	Verlauf der Kosten bei Mutationsrate von 0,2.....	68
7.4	Kostenverlauf bei Mutationsrate von 0,8 und Crossover-Wahrscheinlichkeit von 0,2.....	70

## Tabellenverzeichnis

2.1	Merkmale zur Netzwerktypologisierung [WAG05].....	7
2.2	Eignung der Lagerkonzepte.....	16
2.3	Abhängigkeiten des logistischen Netzwerkes.....	17
3.1	Selektionswahrscheinlichkeiten.....	25
3.2	n-Punkt-Crossover [KOC14].....	27
3.3	Uniform Crossover [NIS97].....	28
3.4	Eigene Darstellung der Inversion.....	29
4.1	Korrelationen der beeinflussbaren Aktionen.....	35
4.2	Ideale Aktionsliste.....	37
4.3	Zufällige Aktionsliste.....	38
4.4	Reale Aktionsliste.....	38
4.5	Kodierung für zweidimensionales System.....	42
4.6	Zusammenhang von Dimensionalität und Kodierung.....	43
4.7	Distanzzuordnung I.....	44
4.8	Dekodierung des Ablaufschemas.....	46
4.9	Merkmalsposition auf Chromosom.....	47
4.10	Distanzzuordnung II.....	50
5.1	Exemplarische Bestimmung von Fitnesswerten.....	53
5.2	Rekombination zweier Chromosome.....	54
5.3	Irreguläre Rekombination.....	55
5.4	Order Crossover OX-Verfahren.....	56
5.5	Vorgang der Punktmutation.....	58
7.1	Auswertung von Versuchen mit Anfangsgröße 10.....	62
7.2	Auswertung von Versuchen mit Anfangsgröße 5.....	63
7.3	Versuche mit variierenden Mutations- und Crossover-Wahrscheinlichkeiten.....	64
7.4	Einfluss von veränderten Mutationsverfahren.....	67
7.5	Durchschnittliche Kosten bei veränderten Mutations-Wahrscheinlichkeiten.....	67
7.6	Kostenmatrix bei verschiedenen Mutations- und Crossover-Wahrscheinlichkeiten.....	69

## Formelverzeichnis

1.	Verhältnis von Sicherheitsbestände und Lageranzahl.....	17
2.	Formel für Optimierungsproblem.....	18
3.	Fitnesswertintervall.....	22
4.	Fitnesswert als Optimierungsproblem.....	23
5.	Kehrwert Fitnessfunktion.....	23
6.	Zielfunktion.....	24
7.	Fitnesswert.....	24
8.	Selektionswahrscheinlichkeit.....	25
9.	Kumulierte Selektionswahrscheinlichkeit.....	25
10.	Intervall kumulierter Selektionswahrscheinlichkeit.....	26
11.	Modifizierter Fitnesswert.....	27
12.	Komma-Strategie.....	30
13.	Plus-Strategie.....	31
14.	Elemente einer Rundreise.....	41
15.	Fitnessfunktion.....	41
16.	Traveling Salesperson Problem with Forbidden Neighborhoods-Abstand.....	48
17.	Gesamtstrecke Chromosom.....	52
18.	Inverse Fitnessfunktion.....	52
19.	Inverse Gesamtstrecke.....	52

## 1. Einleitung

Mit Beginn der Globalisierung streben alle Bereiche der Wertschöpfungskette nach zunehmender Produktivität. Wie Produktionskonzepte und die zugehörige Technik kontinuierlich an steigende Anforderungen angepasst werden, so unterliegt auch das zugehörige Transportnetzwerk der Güter einem kontinuierlichen Wandel. Die Zeiten, in denen Güter nach ihrer Produktion zu den Distributionszentren geliefert werden, um dort auf Vorrat gelagert zu werden, sind lange vorbei. Bei der Planung aktueller logistischer Netzwerke müssen eine Vielzahl an Parametern berücksichtigt werden, beispielsweise Transportwege oder Interaktionen zwischen Lagern. Durch genauere Planung sollen Transportzeiten und Liegezeiten minimiert werden, um so Kosten senken zu können.

Der renommierte MIT-Professor Yossi Sheffi beschreibt die Diskrepanz zwischen Produktion und Logistik eingängig [FRI07, s.S.152]:

*„Making Stuff-that’s easy. Supply Chain, now that is really hard”*

In seinem überspitzt formulierten Zitat trennt er die Bereiche der Produktion und der Supply Chain deutlich. Herstellung wird von ihm als simpler Prozess dargestellt, der mit einfachen Mitteln per Reverse Engineering nachvollzogen oder gar kopiert werden kann. Durch Ausgliederung des Zu- bzw. Abtransportes von Waren oder Betriebsmitteln ist für einen Produktionsprozess lediglich das produzierende Unternehmen verantwortlich. Aufwendiger ist die Organisation von Wertschöpfungsketten. Diese Prozesse sind nicht nur abhängig von dem Handeln eines einzelnen Unternehmens. Der Erfolg beruht auf abgestimmter Kooperation zwischen allen Teilnehmern eines logistischen Netzwerkes, dies betrifft Unternehmen der Rohstoffgewinnung, der Produktion bis zu dem Transport zum Kunden. Mit steigender Anzahl an Agitatoren steigt auch die Komplexität der interlogistischen Verknüpfungen und der Planungsaufwand des Netzwerkes an.

Nach Sheffi ist das wichtigste Ziel „global optimization“ zu erreichen [CAR11]. In diesem Zustand wäre das Verhältnis aus Zuverlässigkeit des logistischen Netzwerkes und der Herstellungskosten ausgeglichen. Produktionskosten sind irrelevant, wenn der zeitliche Rahmen überschritten wurde oder es zu Lieferkomplikationen gekommen ist. Es reicht nicht aus, einen einzelnen Zulieferer auf Grund geringer Transportkosten auszuwählen, wenn dafür die Kosten der anderen Komponenten zu hoch sind. Es gilt, das günstigste Gesamtkonzept zu ermitteln.

Um den Zustand der „global optimization“ zu realisieren, wird eine große Menge an Informationen benötigt [BRE10]. Eine Auswertung bzw. die Bearbeitung dieser Big Data wird mit zunehmender Datenmenge aufwendiger. Dem zunehmenden Aufwand geschuldet werden manuelle Planungen durch steigende Komplexität und einen höheren Zeitbedarf schnell unwirtschaftlich. Unterstützend sollen Informationstechnologien eingesetzt werden [GER13]. Durch computergestützte Vorsortierung bzw. -bearbeitung der Daten und der Möglichkeit, Systeme und Prozesse simulativ nachzubilden, soll der manuelle Aufwand verringert werden.

Eine Möglichkeit der informationstechnischen Unterstützung sind Evolutionäre Algorithmen. In Anlehnung an die Evolutionstheorie werden aus dem Datenpool Individuen generiert, deren (Erb)Informationen sich aus den Daten zusammensetzen. Diese Individuen pflanzen sich fort, dabei

vererben beide Elternteile Informationen an ihre Nachkommen. Diese Individuen setzen sich aus den kombinierten Daten der Eltern zusammen. Durch die Vereinigung der Daten kann ein anfangs definiertes Problem möglicherweise gelöst werden. Andernfalls vermischen sich die jeweiligen Individuen der Nachfolgenerationen bis eine Lösung gefunden wird oder sich ein festgelegtes Abbruchskriterium erfüllt.

Das Ziel dieser Arbeit ist der Vergleich eines Evolutionären Algorithmus, welchem bekannte Zusammenhänge hinzugefügt worden sind und einer unbeeinflussten Version auf Effektivität und Rechenzeit. Um den Einfluss bereitgestellter Korrelationen auf den Algorithmus feststellen zu können, werden zunächst beeinflussbare Aktionen innerhalb logistischer Netzwerke herausgearbeitet. Diese Aktionen sollen aus planungstechnischer Sicht ausgewählt werden. So gilt beispielsweise die Auswahl eines Lagersystems als zu untersuchender Faktor, der Absatz einer Ware kann jedoch nicht direkt kontrolliert werden und gilt so nicht als Faktor. Aus dieser Vorarbeit sollen Korrelationen der Elemente logistischer Netzwerke erfasst werden. Anschließend wird ein Evolutionärer Algorithmus durch die erarbeiteten Korrelationen geplant. Der Algorithmus mit bekannten Korrelationen wird anschließend mit einem weiteren Algorithmus, ohne entsprechende Verknüpfungen, verglichen. Die beiden Evolutionären Algorithmen werden auf Funktionalität bzw. Effizienz untersucht. Die Ergebnisse sollen die verbesserte Funktionalität Evolutionärer Algorithmen zeigen, wenn diesen zusätzliche Informationen zugeführt werden.

Die erste Hälfte dieser Masterthesis widmet sich der terminologischen Klärung der titelgebenden Begriffe der „logistischen Netzwerke“ sowie „Evolutionärer Algorithmen“. Um die fachliche Abgrenzung der beiden titelgebenden Themenkomplexe hervorzuheben, werden die thematischen Grundlagen in einem logistischen und in einem informationstechnischen Kapitel abgegrenzt. Als Quellen werden hauptsächlich Fachbücher und Konferenzmitschriften verwendet.

Das einleitende Kapitel behandelt die logistischen Grundlagen dieser Thesis. Für das Verständnis logistischer Netzwerke wird zunächst auf den logistischen Prozess eingegangen. Neben allgemeinen logistischen Konzepten logistische Netzwerke betreffend, werden die Bestandteile eines solchen Systems sowie mögliche Lager- oder Verbindungsarten erklärt. Ergänzend wird die kooperative Zusammenarbeit zwischen den Teilnehmern eines solchen Netzwerkes betrachtet. Aus den Erklärungen dieser Grundlagen werden Zusammenhänge zwischen den einzelnen Elementen eines logistischen Netzwerkes herausgearbeitet. Im Fokus stehen dabei die Einflüsse von Aktionen aufeinander.

Das zweite Grundlagenkapitel thematisiert die Begrifflichkeiten des informationstechnischen Anteils dieser Masterthesis. Im Fokus stehen dabei Evolutionäre Algorithmen. A priori werden für die Erklärung Evolutionärer Algorithmen zunächst übergeordnete Begriffe definiert. Die Erläuterungen der Evolutionären Algorithmen umfassen zusätzlich zu Erklärungen der Zielsetzung auch das schematische Ablaufkonzept. Neben den thematischen Grundlagen wird auch auf deren mathematische Konzepte eingegangen. Das abschließende Kapitel des technischen Stands der Technik beschreibt Python, die Programmiersprache, in der der Algorithmus generiert wird.

In dem Hauptteil dieser Arbeit werden zuerst die Korrelationen logistischer Netzwerke spezifiziert. Zurückgreifend auf die allgemeinen Abhängigkeiten aus dem logistischen Stand der Technik soll beschrieben werden, wie sich Faktoren verändern, wenn andere Faktoren erhöht bzw. verringert werden.

Es gilt, Synergien und negative Abhängigkeiten zwischen den einzelnen Konzepten und Elementen zu erfassen. Anschließend erfolgt die Generierung eines Evolutionären Algorithmus. Die Planung der einzelnen Phasen des Evolutionären Algorithmus erfolgt unter Berücksichtigung der Vorarbeit. In den erzeugten Individuen des Algorithmus sollen die Korrelationen der logistischen Elemente einfließen. Die Phasen sollen Verbindungen aus korrelierenden Merkmalen innerhalb der Individuen fördern.

In dem letzten Kapitel wird der geplante Algorithmus mit einem weiteren Algorithmus verglichen. Der zweite Algorithmus arbeitet, ohne bekannte Korrelationen zwischen seinen Variablen. Aus diesem Vergleich soll eine Aussage über den Nutzen eingespeister Zusammenhänge getroffen werden.

## 2. Logistische Einordnung

Kern dieser Arbeit ist die Weiterentwicklung eines Evolutionären Algorithmus (EA) und der Vergleich auf Effizienz mit seiner ursprünglichen Form. Dieser Basisalgorithmus soll durch zuvor herausgearbeitete Korrelationen innerhalb eines logistischen Netzwerkes angepasst werden. In diesem Kapitel werden die logistischen Terminologien erläutert, die für das Verständnis der folgenden Kapitel nötig sind.

### 2.1 Supply Chains

Eine Supply Chain (SC) ist die logistische Verknüpfung eines Industrieunternehmens. Sie beschreibt den internen Güterfluss eines Unternehmens und die interlogistische Lieferkette von Lieferanten zu einem Unternehmen. Der Begriff Supply Chain impliziert, dass der Warenfluss eine Lieferkette sei. Vielmehr handelt es sich bei Supply Chains um Netzwerke aus verschiedenen Handelspartnern. Der Unterschied einer Logistikkette und einer Supply Chain liegt in der Betrachtungsweise der einzelnen Teilnehmer [WER17]. Der Supply Chain unterliegt einem stärkeren Abstimmungsgedanken. Ein Agitator muss mit allen verknüpften Partnern kooperieren. Bei der Logistikkette handeln die Teilnehmer isolierter. Abstimmung erfolgt nur mit vor- und nachgelagerten Unternehmen.

Gemäß einer in der Literatur häufig verwendeten Definition sind Supply Chains logistische Netzwerke von Rohstoffgewinnung bis hin zur Auslieferung an den Endkunden [ALB14]. Sie beinhalten Unternehmen für die Produktion von Waren und Zwischenprodukten oder der Bereitstellung von Dienstleistungen. Diese Unternehmen sind durch Waren- und Informationsströme miteinander verknüpft.

In diesem Unterkapitel werden zuerst im Zusammenhang der Supply Chains grundlegende logistische Sachverhalte erklärt. Das aufbauende Kapitel spezifiziert den Begriff der Netzwerke weiter.

Abbildung 2.1 zeigt den exemplarischen Ablauf einer Supply Chain [PAU10]. Die dünnen Pfeile symbolisieren den Warenfluss von der Rohstoffgewinnung bis zu dem zugehörigen Endkunden, die dickeren gestrichelten Pfeile stehen für den Informationsfluss zwischen den Handelspartnern.

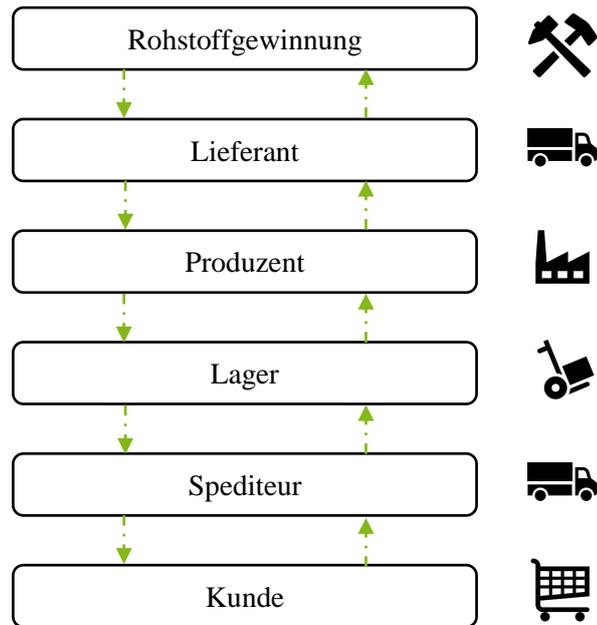


Abbildung 2.1: Supply Chain-Kette [nach PAU10]

**Rohstoffgewinnung:** Gewinnung von Ressourcen, aus denen durch Weiterverarbeitung Güter hergestellt werden können.

**Lieferant:** Wirtschaftlicher Akteur dessen Aufgabe es ist, Waren oder Dienstleistungen zu einem der Handelspartner zu liefern.

**Produzent:** Produzieren aus den gelieferten Rohstoffen die von dem Kunden bestellten Waren. Die Produzenten nehmen durch ihre Nähe zum Endkunden eine dominante Rolle im Supply Chain Management ein.

**Lager:** Gebäude, in dem produzierte Güter zeitlich begrenzt eingelagert werden und für deren spätere Verwendung verfügbar gehalten werden. Lager dienen dem zeitlichen und quantitativen Ausgleich von Nachfrageschwankungen oder Lieferproblemen.

**Spediteur:** Dienstleistungsunternehmen, welches die Waren von einem Handelspartner abholt und den Transport in Richtung eines Kunden abwickelt.

**Kunde:** Endabnehmer der produzierten Waren oder Dienstleistungen. Akteur, von dessen Auftrag der Transport initiiert wird. Produzenten können auch als Kunden anderer Produzenten auftreten und Waren für Drittparteien produzieren.

## 2.2 Der Netzwerkbegriff

Nach Wagner erfolgte noch keine allgemeingültige Definition für *Netzwerke* im logistischen Kontext [WAG04]. Dies ist für ihn auf die häufige Verwendung des Begriffes zurückzuführen, welche zu ungenauen Definitionen führt.

Als Abhilfe wird versucht, Netzwerke anhand folgender Merkmale zu charakterisieren:

- Zahlreiche und fluktuierende Akteure
- Hohe Vernetzungsdichte
- Offenheit gegenüber der Umwelt
- Wechselwirkungen
- Verteilte Aktivitäten räumlicher und zeitlicher Art
- Komplexität und Dynamik
- Exponentielles Ansteigen des Koordinationsaufwandes bei Expansion
- Unsichere Prognose des Systemverhaltens bei Parameterveränderungen
- Hybrider Charakter bzgl. Vertrauen und Kontrolle sowie Abhängigkeit und Autonomie in Netzwerken

Sydow und Manning (vgl. 2006, S.1) beschreiben den Begriff des „Netzwerkes“ im ökonomischen Kontext als die „[...] Kooperation in und/oder zwischen relativ autonomen, gleichwohl in ein Netz von Beziehungen eingebundenen Organisationen bzw. Unternehmungen (oder Organisationseinheiten).“ Der Begriff wird beschrieben als kooperative Beziehung zweier oder mehrerer autonomer Unternehmen zur unternehmensübergreifenden Leistungserstellung. Schematisch ist dies in Abbildung 2.2. dargestellt:

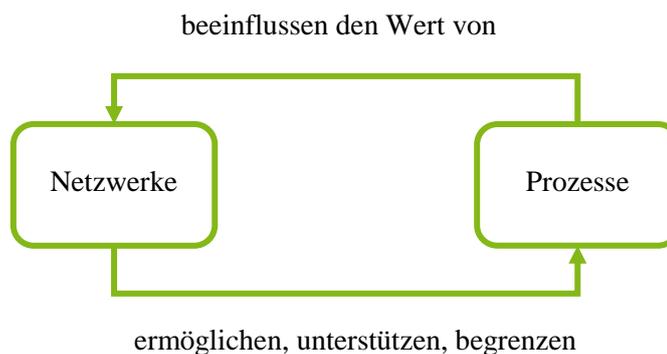


Abbildung 2.2.: Zusammenhang Netzwerke und Prozesse [BRE10]

Netzwerke sind Systeme, in denen einzelne Unternehmen hinzugefügt werden, um Synergien nutzen zu können [BRE10]. Diese Nutzeffekte steigen mit zunehmenden Verbindungen. Jeder Teilnehmer eines Netzwerkes bringt individuelle Eigenschaften in ein System ein, welche auf die anderen Handelspartner wirken. Zusätzlich können einzelne Teilnehmer in mehr als nur einem einzigen Netzwerk eingebunden sein.

## 2. Logistische Einordnung

---

Ein Netzwerk spezifiziert den Oberbegriff der Kooperation [PFO04]. Kooperationen werden im Hinblick auf gewünschte Win-Win-Situationen realisiert. Im Austausch für die partielle Aufgabe unternehmenseigener Autonomien für Interdependenz, locken gemeinsame Ziele und kollektiver Zusammenhalt. Übergeordnet steht als Primärziel aller Teilnehmer die Verbesserung bzw. Aufrechterhaltung der Wettbewerbsfähigkeit. Dieses Ziel lässt sich auf mehrere Synergieeffekte aufteilen, deren Erfüllung der allgemeinen Wettbewerbsfähigkeit dient [PFO04]:

**Kostensenkung:** Gerade im Bereich der Kostensenkung gibt es oftmals Optimierungspotenzial. Durch verbesserte Logistikplanung sollen sich die gegenwärtigen Bestandskosten verringern. Zugeschnittene Planung soll den Anteil der produzierten Waren, welche auf Vorrat gelagert werden müssen, senken. Zusätzlich können Transportkosten gesenkt werden, wenn die Teilnehmer eines Netzwerks besser aufeinander abgestimmt sind.

**Zeitreduktion:** Im Zuge der besseren Abstimmung innerhalb der SC können auch Durchlaufzeit und Lieferzeit verbessert werden.

**Qualitätsverbesserung:** Alle Netzwerk-Teilnehmer eint das Ziel, das Endprodukt in seiner bestmöglichen Qualität zu seinem Endabnehmer zu liefern. Je enger die Verflechtung der Kooperationspartner, desto wahrscheinlicher erfolgen Verbesserungen an Qualität. Einzelhändler sind näher an den Kunden und deren Wünschen bzw. Vorlieben als die Produzenten. Durch die Kommunikation über SC-Ebenen hinweg können Probleme oder Wünsche aufgenommen und vorgelagerten Ebenen weitergegeben werden.

**Flexibilitätssteigerung:** Durch Steigerung der Flexibilität ist es möglich, auf unplanmäßige Lieferungen eingehen zu können. So können auch kurzfristige Bestellungen oder Retouren bearbeitet werden.

Neben den genannten Vorteilen gibt es auch einige wenige Nachteile. Gerade kleinere Unternehmen laufen Gefahr, ungewollt Knowhow weiterzugeben oder sich von anderen Handelspartnern abhängig zu machen.

Netzwerke treten in verschiedenen Formen auf [SYD10] [WAG05]. Aufgrund der Vielzahl an möglichen Varianten ist es nicht möglich, alle Formen zu benennen. In der folgenden Tabelle 2.1 werden einige Merkmale zur Netzwerktopologisierung dargestellt:

Tabelle 2.1: Merkmale zur Netzwerktopologisierung [WAG05]

Kriterium	Ausprägung		
Kooperationsrichtung	horizontal	vertikal	diagonal
Stärke und Dauer der Wirkung	strategisch		operativ
Koordinationsform/ Struktur	heterarchisch		hierarchisch
Dauer	dauerhaft		temporär
Steuerungsform	polyzentrisch (dezentral)		fokal (zentral)

Produkte durchlaufen einen Wertschöpfungsprozess (s. Abb. 2.1). An den einzelnen Stufen eines Wertschöpfungsprozesses sind unterschiedliche Akteure beteiligt. Die **Kooperationsrichtung** gibt an, auf welcher Wertschöpfungsstufe die jeweiligen Kooperationspartner zueinanderstehen [FIS06]. Das Verhältnis wird in einer von drei Dimensionen beschrieben. *Horizontale Kooperation* beschreibt ein Kooperationsverhältnis von Akteuren, die auf derselben Wertschöpfungsstufe stehen. Da beide Partner in derselben Branche und in derselben Wertschöpfungsstufe agieren, konkurrieren diese miteinander um dieselben Kunden und Ressourcen. Im Gegensatz dazu beschreibt die *vertikale Kooperation* Koordination sukzessiver Wertschöpfungsstufen. Hintereinander gelagerte Akteure konzentrieren sich auf ihre Kernkompetenzen und lagern Schwachstellen an Partnern aus. Es handelt sich um eine klassische Zulieferer-Abnehmer-Beziehung. Als letzte Form der Kooperationsrichtung wird die *diagonale Kooperation* beschrieben. Sie bezeichnet eine branchenübergreifende Partnerschaft. Partner verschiedener Wirtschaftszweige kooperieren für ein gemeinsames Projekt.

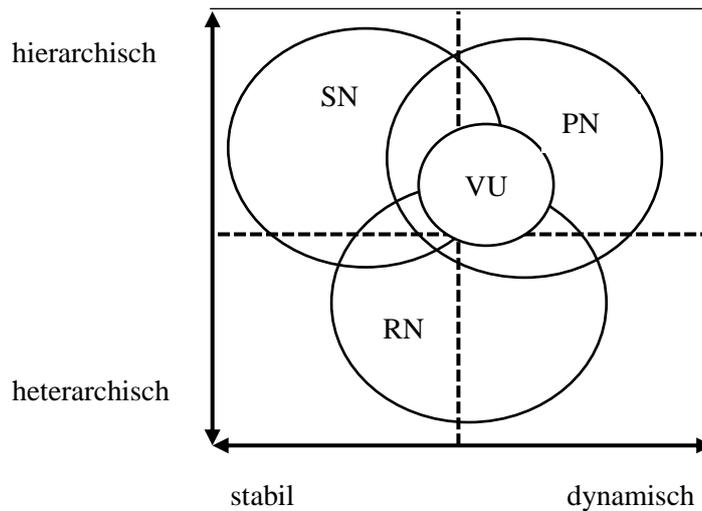
Ein weiteres Kriterium der Netzwerktypologisierung ist die Betrachtung von **Stärke und Dauer der Wirkung**. Die Wirkung wird unterschieden in operativ und strategisch. Operative Netzwerke sind zeitlich begrenzt und werden häufig nur für ein Jahr aufrechterhalten. Sie dienen der Abwicklung einzelner Projekte. Strategische Netzwerke werden langfristig angelegt, über die Dauer eines einzelnen Projektes hinaus.

Bezüglich der **Koordinationsform** gibt es zwei Kategorien, die unterscheiden, wie viele Parteien das Netzwerk steuern. *Heterarchische* Koordination verteilt die Koordinationssteuerung auf mehrere Akteure. Die Machtverteilung erfolgt symmetrisch, alle Akteure sind gleichberechtigte Partner. Als *hierarchisch* werden Systeme bezeichnet, die eine hierarchische Struktur aufweisen. Von allen Teilnehmern gibt es einen überstehenden Koordinator, der die Verantwortung übernimmt [RIE08].

Die **Bildungsdauer** von Netzwerken variiert. Es wird zwischen temporären und dauerhaften Verbindungen unterschieden. Temporäre Verbindungen werden eingegangen, um besondere Aufgabenstellungen abzuschließen, die von dem Regelbetrieb abweichen. Diese Verbindungen sind zeitlich begrenzt. Dauerhafte Netzwerke bestehen ohne zeitliche Begrenzung, sie werden für den Normalbetrieb eingegangen.

Als letztes Merkmal wird die **Steuerungsform** aufgeführt. Sie ergänzt die Ausprägungen der Koordinationsform. Heterarchische Netzwerke sind polyzentrisch. Es gibt mehrere Teilnehmer mit Entscheidungskompetenz, die Machtstruktur ist gleichmäßig verteilt. Hierarchische Netzwerke sind fokal. Die Steuerung geht alleine von dem überstehenden Teilnehmer aus [ALB03].

Innerhalb der möglichen Kombinationen gibt es keine universell optimale Zusammenstellung. Dennoch haben sich einige verbreitete Konstellationen herauskristallisiert [ZIR07]. Das Venn-Diagramm (s. Abb. 2.3) stellt die verschiedenen Netzwerkart in Bezug zu ihren Koordinationsformen und der Stabilität ihrer Verbindungen dar:



- Legende:
- SN = strategische Netzwerke
  - RN = regionale Netzwerke
  - PN = Projektnetzwerke
  - VU= virtuelle Unternehmungen

Abbildung 2.3: Netzwerkarten [ZIR07]

In dem Diagramm beschreiben die drei großen Bereiche die physischen Kooperationsformen. Diese setzen sich aus strategischen Netzwerken, regionalen Netzwerken und Projektnetzwerken zusammen. Die Abszissenachse beschreibt die zeitliche Komponente dieser Netzwerktypen. Die Stabilität dieser Kooperationen verläuft zwischen stabilen und dynamischen Beziehungen. Die andere Achse vergleicht die Koordinationsform, ob es einen zentralen Koordinator gibt. Zwischen den genannten Bereichen gibt es Überlagerungen. Der letzte Bereich beschreibt die Menge der virtuellen Unternehmungen. Wie abgebildet ist diese eine Teilmenge der Projektnetzwerke mit Überschneidungen aus den anderen beiden Mengen.

### **Strategische Netzwerke:**

Strategische Netzwerke werden von einem oder mehreren Unternehmen gesteuert, jedoch koordiniert ein einziges zentrales Unternehmen das Netzwerk. Das Ziel eines solchen Netzwerkes sind stabile Verbindungen unter den einzelnen Teilnehmern, diese Verbindungen werden für einen längerfristigen Zeitraum eingegangen. Durch diese Kontinuität ist diese Art von Netzwerken vor allem für kleinere Unternehmen interessant. Sie können sich auf ihr Produkt oder ihre Dienstleistung spezialisieren, ohne Gefahr zu laufen, durch ein konkurrierendes Unternehmen ausgetauscht zu werden. Klassische strategische Netzwerke finden sich im Automotive-Bereich. Dieser Netzwerktyp wird aber auch im Dienstleistungssektor verwendet.

### **Regionale Netzwerke:**

Regionale Netzwerke heben sich durch die räumliche Nähe der einzelnen Unternehmen hervor. Meist gehen klein- oder mittelständige Unternehmen regionale Netzwerke ein. Die Unternehmen ähneln sich oftmals in ihrer jeweiligen Größe. Es gibt kein fokales Unternehmen, dieser Kooperationstyp ist vermehrt polyzentrisch ausgelegt. Regionale Netzwerke sollen längerfristig bestehen bleiben, trotzdem kann es vorkommen, dass einzelne Unternehmen aus dem Netzwerk ausgetauscht werden, bspw. im Falle von auftretenden Nachfrageänderungen oder Unstimmigkeiten unter den Handelspartnern.

Als Beispiele für regionale Netzwerke haben sich physische Cluster gebildet. Diese Cluster haben diverse Schwerpunkte, wie wissenschaftliche oder produktive Orientierung. Zu den bekanntesten regionalen Netzwerken gehören das Silicon Valley in Nordamerika oder der Shenzhen-Bereich in China.

### **Projektnetzwerke:**

Projektnetzwerke sind wie strategische Unternehmen fokal aufgebaut. Es existieren ein oder mehrere Unternehmen, welche das Beziehungsgeflecht dominieren. Heterarische Beziehungen sind auch möglich. Diese Netzwerkkategorie unterscheidet sich von den anderen beiden durch ihre zeitliche Limitierung. Die Beziehungen zwischen den handelnden Unternehmen werden nur für die Dauer eines Projektes eingegangen. Nach dessen Abschluss werden die Verbindungen nur noch latent aufrechterhalten, um im Falle einer erneuten Zusammenarbeit, bei einem weiteren Projekt, diese wieder intensivieren zu können. Projektnetzwerke werden hauptsächlich in Branchen eingegangen, die von der Kurzlebigkeit ihrer Projekte leben. Dies sind bspw. Unternehmen im Mediensektor oder Unternehmen, die in der Baubranche tätig sind.

### **Virtuelle Unternehmungen:**

Virtuelle Unternehmungen gehören zu den instabilen Netzwerkformen. Branchenführende Unternehmen unterschiedlicher Bereiche gehen eine zeitlich limitierte Verbindung ein. Diese Verbindung tritt als einheitliche Firma auf, besitzt aber keine festgelegten physischen Strukturen. Zusätzlich ist die Kooperationsform polyzentrisch aufgebaut, es gibt keine hierarchische Ordnung.

Symptomatisch für virtuelle Unternehmungen sind hohe Anforderungen an Informations- und Kommunikationssysteme sowie an die Mitarbeiter der teilnehmenden Unternehmen. Der fehlende zwischenmenschliche Kontakt und die zeitliche Begrenzung müssen durch eine erhöhte Projektidentifikation ausgeglichen werden.

Abgesehen von den vier idealisierten Netzwerkartentypen gibt es mehrere Mischformen, wie kompetenzzentrierte Unternehmensnetzwerke. Diese sollen Kunden das Unternehmen vermitteln, welches für dessen derzeitige Aufgabenstellung am besten passt. Die Unternehmen sind als gleichwertige Teilnehmer in einem Cluster angeordnet. Die Vermittlung übernimmt eine Softwareplattform, entsprechend ist kein übergeordnetes Unternehmen nötig. Der Hybridtyp gliedert sich in ein- und mehrdimensionale kompetenzzentrierte Unternehmensnetzwerke. Bei der ersten

Variante werden Unternehmen ähnlicher Bereiche gebündelt, bei der zweiten Variante werden mehrere Kompetenznetzwerke miteinander verbunden.

### 2.3 Elemente des logistischen Netzwerkes

In dem folgenden Kapitel werden die relevanten logistischen Schemata und Theorien erläutert. Auf diesen Erläuterungen baut das anschließende Kapitel auf, in dem Korrelationen zwischen den beschriebenen Elementen erklärt werden.

Diese Masterarbeit konzentriert sich auf Supply Chain-Ebenen der ersten Stufe, der Tier-1 [WAG05], von einer theoretisch beliebig erweiterbaren Zuliefereinteilung wird abgesehen. Die Produktion dient als Ausgangspunkt. Die Rohstoffgewinnung und deren logistische Verknüpfungen gehören nicht mehr zu den betrachteten Supply Chain-Stufen. So liegt der Fokus auf dem Transport zwischen Produzent und Kunden. Schematisch ist dies in der untenstehenden Abbildung 2.5 dargestellt. Die unterschiedliche Anzahl aus- und eingehender Verbindungen soll die mögliche Komplexität eines logistischen Systems darstellen. Detaillierte Darstellungen folgen:

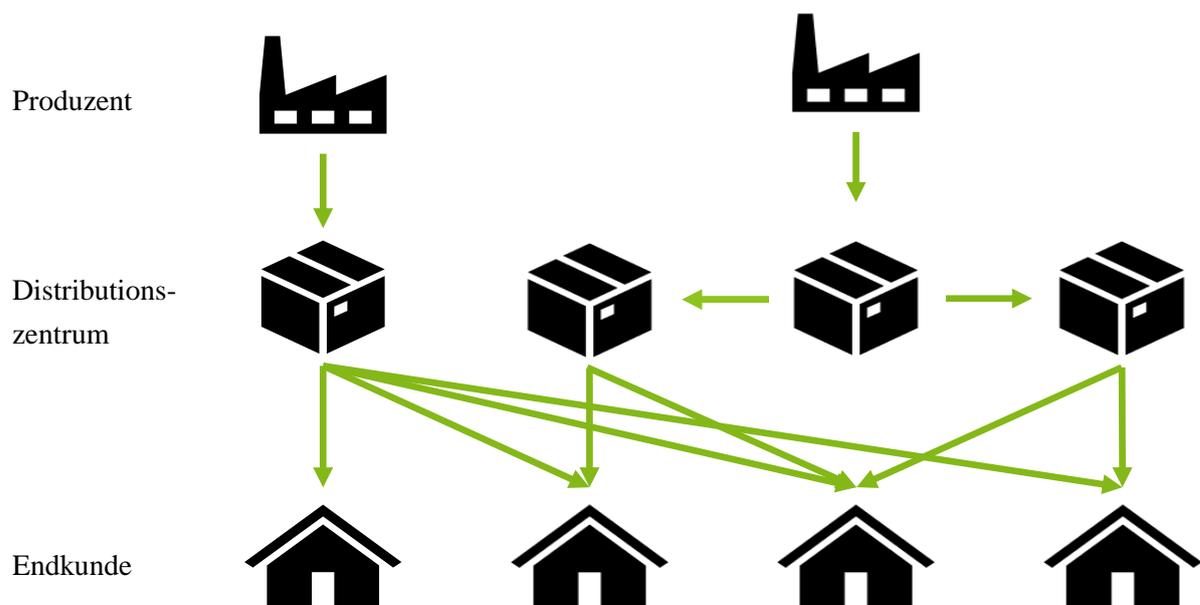


Abbildung 2.4: Darstellung eines Logistischen Netzwerkes nach [WAG05]

In der stark vereinfachten Abbildung eines logistischen Netzwerkes führen alle Pfeile zu den Lagern hin oder gehen von diesen aus. Sie werden von den Produzenten mit Waren beliefert. Ein Produzent kann ein einzelnes Lager beliefern, von welchem mehrere Kunden beliefert werden, oder ein Produzent beliefert jedes Lager. Diese Lager sind nur mit den räumlich nah gelegenen Kunden verknüpft.

Die im Folgenden gelisteten Elemente des logistischen Netzwerkes werden aufgrund ihrer Verknüpfungen untereinander weder chronologisch, noch priorisiert aufgeführt. Ihre Reihenfolge erfolgt alphabetisch.

### **Beschaffenheit der Waren**

Der elementarste Bestandteil einer Supply Chain ist die zu handhabende Ware. Anhand des jeweiligen Produkts werden zugehörige Parameter bzgl. des Transports oder auch der Lagerhaltung bestimmt. Da keine näheren Beschreibungen über das Produkt gegeben sind, wird eine grobe Klassifizierung vorgenommen, welche die Wahl des Lagertyps beeinflusst. Bei der gängigen ABC-Analyse werden Objekte in absteigender Reihenfolge priorisiert, meist geschieht dies nach Anteil am Gesamtumsatz [MAT18]. Eine überschneidende Alternative ist die Unterscheidung zwischen Schnell- und Langsamdreher. Schnelldreher sind Produkte mit einer hohen Umschlagshäufigkeit, sie entsprechen in der ABC-Analyse den A-Waren [KOE14]. Ihnen stehen die Langsamdreher gegenüber, Waren die nur in geringen Mengen umgeschlagen werden. Langsamdreher entsprechen der C-Kategorie. Waren der B-Kategorie können nicht eindeutig zugeordnet werden und können beiden Varianten entsprechen. Im Folgenden wird nur die Unterscheidung Schnell- und Langsamdreher fortgeführt, da sie sich durch ihre gröbere Klassifizierung besser für unklassifizierte Produkte eignet.

### **Endkunde**

Der Endkunde hat lediglich einen marginalen Einfluss auf ein logistisches Netzwerk. Obwohl der Endkunde den Transport durch seine Warenbestellung erst initiiert, ist sein Einfluss auf diesen begrenzt. Dabei ist der Einfluss auf eine reale Bestellung zumindest noch gegeben, durch verschiedene Lieferoptionen lässt sich besonders der Aspekt der Transportfrequenz individuell anpassen. Im Zuge dieser Thesis dient der Endkunde als Senke, zu der die Lieferung gebracht wird, um generelle Aussagen über den Transportprozess treffen zu können.

### **Lagerart**

Für eine optimale Transportkette muss jedes Element dieser Kette, egal ob Produzent, Lieferant oder Lagerung, aufeinander abgestimmt sein. Bei einer optimalen Abstimmung entsteht ein kontinuierlicher Materialfluss. Dieser Materialfluss stockt durch die zwischenzeitliche Einlagerung von Waren. Ein Lager stört nicht nur die idealisierte Transportkette, es bindet auch Kapital in Form von Waren und verursacht Unterhaltskosten. In der Realität kommen die wenigsten logistischen Netzwerke ohne ein Lager aus. Komplikationen und Nachfrageschwankungen verhindern oder erschweren die Realisierung eines lagerfreien Systems, entsprechend analysiert diese Arbeit auch nur Transportvorgänge mit Bevorratung.

Die Auswahl des Lagersystems erfolgt im Einklang der übrigen Komponenten in dem logistischen Netzwerk. Dabei wird auf die transportierten Güter und die Lagerhilfsmittel geachtet [FOR07]. Es wird zwischen zentralen Lagersystemen und dezentralen Systemen unterschieden. Zuerst wird auf das Zentrallagerkonzept eingegangen. Abbildung 2.5 verdeutlicht das Konzept genauer als das Gesamtschema in Abbildung 2.4:

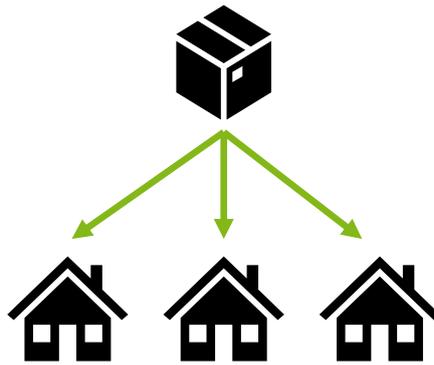


Abbildung 2.5: Zentrallagerkonzept

Unter einem Zentrallagersystem versteht man ein Konzept, bei dem jede Lieferung in Richtung eines Kunden von einem zentralen Lager ausgeht. In der Transportkette ist dem namensgebenden Zentrallager ein Werkslager vorgeschaltet. Dieses nimmt an der Produktionsstätte die hergestellten Waren auf und dient dem Sammeln von Waren bis zum Erreichen der benötigten Losgröße. Die Distributionsstruktur der zentralen Lagerung weist einige Vorteile gegenüber der dezentralen Lagerung auf.

Einige der offensichtlichsten Vorteile beziehen sich auf finanzielle Aspekte. Organisatorische Elemente wie Lagerleitung und Warenein- und Ausgang werden nur einmalig benötigt. Bei der zentralen Lagerung werden die Räumlichkeiten effizienter genutzt. Zusätzlich sind die Mindestbestände eines Zentrallagers geringer als die Summe der Bestandsreserve dezentralisierter Lager. Zusätzlich unterliegen die zentralgelagerten Artikel einer höheren Umschlagshäufigkeit als dezentralgelagerten Produkten. Dadurch ist auch der mittlere Bestand geringer. Der mittlere Bestand gibt den Durchschnittswert von Anfangs- und Endbestand an. Durch die geringeren Bestände ist auch die Kapitalbindung der gelagerten Artikel geringer.

Die Alternative zu einem zentralisierten Lager ist ein Dezentrallagerkonzept. Dieses wird in Abbildung 2.6 dargestellt.

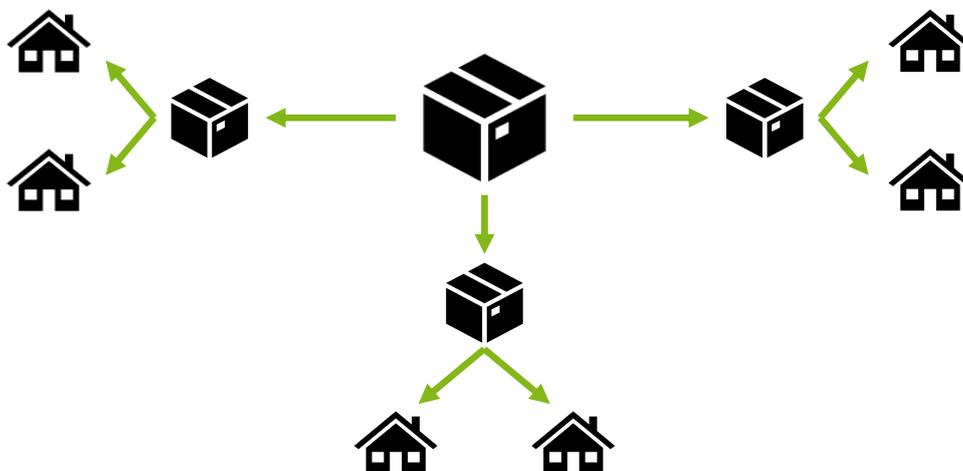


Abbildung 2.6: Dezentrallagersystem

Bei einem Dezentrallagerkonzept spannen mehrere Lager das logistische Netzwerk auf. Im Gegensatz zu einem zentralen Lagersystem existiert zwar ein Hauptlager, von dort werden die Kunden aber nicht direkt beliefert. Vor Lieferung zu einem Kunden werden die Waren zu einem weiteren Lager geliefert und temporär gelagert bzw. kommissioniert.

Das Lagerkonzept eignet sich für ein System, dessen Kunden weit verstreut liegen und nur geringe Produktmengen abnehmen. Dabei wird die Ware in Kundennähe gelagert [WEB15]. Aufgrund der vergleichsweise nahen Entfernung zwischen Kunde und Lager ergeben sich kürzere Transportzeiten. Gegen ein dezentrales Lagerkonzept sprechen die höheren Investitions- bzw. Fixkosten. Es werden mehrere Lager benötigt, die aufgebaut und ausgestattet werden müssen.

Bei der Realisierung dezentraler Systeme muss entschieden werden, in wie viele Lagerstufen sich das System aufgliedern soll. Die einzelnen Lager können sich theoretisch immer weiter aufteilen, daher muss in der Realität ein Konsens aus Nutzen und Aufwand bzw. Kosten gefunden werden. Diese Thematik thematisiert lediglich die simpelste Form der Dezentralisierung. Nach dem Hauptlager werden die Güter nur in eine einzige sukzessive Stufe übermittelt.

### **Lagerstrategie**

Lagerstrategien bestimmen den Ablauf von Ein- und Auslagerung von Waren innerhalb eines Lagers. Unterschieden werden diese Strategien nach der Reihenfolge der Entnahme der eingelagerten Güter. Die beiden verbreitetsten Lagerstrategien sind First In – First Out (FIFO) und Last In – First Out (LIFO). Bei dem FIFO - Verfahren wird jene Ware zuerst ausgelagert, die als erstes eingelagert worden ist. Die Ware, welche am längsten gelagert wurde, wird entnommen. Das gegenteilige LIFO-Verfahren lagert neu eingelagerte Waren als erstes wieder aus. Länger eingelagerte Waren werden erst entnommen, wenn die später eingelagerten Waren ausgelagert worden sind.

### **Transport**

Unter der Prämisse, dass es das Ziel eines jeden Transportvorgangs ist, seine Waren möglichst schnell zu seinem Zielort zu liefern, sind die meisten Aspekte des Transportvorgangs für diese Thematik trivial. So ist es produkt- und regionenabhängig, welches Transportmittel gewählt wird. Die Auswahl des Transportmittels korreliert nicht mit den Stellschrauben. Sowohl Lieferungen per Luftfracht als auch per Seetransport können in Lagern desselben Distributionskonzeptes eingelagert werden.

Der relevante Faktor ist die Transportfrequenz, sie beschreibt die Häufigkeit, in dem Waren zu demselben Ziel transportiert werden. Wichtig ist, dass die Transportfrequenz auf das Transportvolumen im Netzwerk angepasst ist. Eine zu hohe Transportfrequenz bei verhältnismäßig niedrigem Transportvolumen verursacht hohe Kosten.

### **Zulieferer**

In einem realen System werden Zulieferer nach diversen Faktoren ausgewählt, z.B. nach der Distanz zum Produktionsort, Kosten oder vergangene Handelsbeziehungen. Diese Faktoren können bisher für theoretische Systeme nur marginal berücksichtigt werden.

Bei der Auswahl der Zulieferer gibt es unterschiedliche Beschaffungskonzepte. Eines dieser Konzepte unterscheidet nach der Anzahl der eingebundenen Zulieferer, dabei wird zwischen Single-Sourcing und Multiple-Sourcing unterschieden [HUT15]. Bei Single-Sourcing wird das geförderte Produkt von einem einzigen Lieferanten geordert. Es ist ein Zeichen einer starken Verbindung zwischen den beiden Kooperationspartnern und diese Lieferantenbeziehungen werden im Hinblick einer langfristigen Zusammenarbeit angelegt. Das alternative Konzept wird Multiple-Sourcing genannt. Im Gegensatz zu der Ein-Zulieferer-Ein-Produkt-Beziehung existieren mehrere Zulieferer, die dasselbe Gut anbieten und liefern. Durch Multiple-Sourcing wird eine Abhängigkeit von einem einzigen Zulieferer verhindert. So stören unternehmensübergreifende Streitigkeiten oder Probleme bei dem zuständigen Zulieferer die Transportkette weniger.

Ein anderes Beschaffungskonzept unterscheidet die Zulieferer nach Entfernung. Die erste Stufe ist das Local-Sourcing. Sie beschreibt eine Lieferantenbeziehung in der Nähe der Produktionsstätte. Die nächst höhere Stufe ist das Regional-Sourcing. Der Lieferant kann hierbei aus dem nächstgelegenen Wirtschaftsraum stammen. Ein solcher Wirtschaftsraum ist bspw. die Europäische Union. Noch entferntere Distanzen werden als Global-Sourcing beschrieben. Güter können hierbei weltweit geordert werden.

Ein weiteres Unterscheidungsmerkmal unterteilt die Beschaffung zum Zeitpunkt ihrer Bestellung [TOP13]. Es wird zwischen dem Bestellpunkt- und dem Bestellrythmusverfahren unterschieden. Beim Bestellpunktverfahren wird das benötigte Gut bestellt, sobald eine eigens definierte Mindestmenge erreicht oder unterschritten worden ist. Der Zeitpunkt der Bestellung variiert, die Menge des Restbestandes in diesem Moment nicht. Das Bestellpunktverfahren eignet sich für Güter, deren Verbrauch unregelmäßig erfolgt, wie bspw. saisonale Güter. Der Kontrollaufwand ist bei dieser Variante jedoch deutlich größer. Bei dem gegenteiligen Bestellrythmusverfahren werden neue Güter in regelmäßigen Abständen bestellt. Die Zeitspanne zwischen zwei Bestellungen ist jedes Mal dieselbe, die Restmenge variiert. Dieses Verfahren eignet sich hauptsächlich für Güter mit regelmäßigem Verbrauch [TOP13].

## **2.4 Beziehungen innerhalb Logistischer Netzwerke**

Um Prozesse optimieren zu können, ist es notwendig, Abhängigkeiten innerhalb des Systems zu kennen. Im Folgenden werden die erklärten Konzepte und Elemente logistischer Netzwerke in Verbindung gesetzt. Dafür werden zunächst die Anwendungspotentiale der Lagertypologien gegenübergestellt.

Für Artikel mit einer geringen Umschlagshäufigkeit, die sogenannten Langsamdreher, empfiehlt sich die Bevorratung in einem Zentrallager [BRA18]. Diese Güter werden unregelmäßig nachgefragt, dafür aber auch in entferntere Regionen geliefert. Schnelldreher eignen sich nicht für die Lagerung in einem

zentralen Lager. Ihre Nachfrage ist regionenunabhängig groß und prognostizierbar, so dass etwaige Nachfrageschwankungen zu vernachlässigen sind.

Ein weiteres Kriterium für die Bevorratung von Gütern in einem Zentrallager ist deren Güterwert [KRO66]. Kostspielige Güter sollten zentral gelagert werden, zum einen ist die Nachfrage teurer Produkte meistens unregelmäßiger als Produkte der täglichen Notwendigkeiten. Zum anderen binden teure Produkte mehr Kapital. Deren Lagerung in diversen verstreuten Lagern ist unwirtschaftlich. Günstige Produkte, z.B. des täglichen Bedarfs, werden in regelmäßigen Abständen gekauft. Diese sollten in dezentralen Lagern gelagert werden.

Durch die Unterteilung in viele Lager werden bei einem Dezentrallagersystem viele kleinere Kunden beliefert. Von einem Zentrallager werden häufig Großkunden beliefert, die zwar größere Warenmengen abnehmen, aber eine kleinere Kundenbasis bilden. Diese Großkunden sind aber verstreuter als regionale Kunden, die Transportwege sind daher weiter.

Ausgehend von dem Lagersystem werden die vorläufigen Zusammenhänge in Tabelle 2.2 aufgeführt. Die jeweils optimalere Lagerart wird grün hervorgehoben.

Tabelle 2.2: Eignung der Lagerkonzepte

Eigenschaften		Zentrallager	Dezentrallager
<b>Ware</b>			
Warenachfrage	Langsamdreher		
	Schnelldreher		
Güterwert	teurer		
	günstiger		
<b>Kunde</b>			
Kundenanzahl	viele		
	wenige		
Entfernung Kunden	weit		
	nah		

Aus der Auswahl des Lagersystems ergeben sich weitere Handlungsparameter. Diese Parameter spezifizieren das jeweilige Transportnetz weiter.

Eine dieser Handlungen ist die Realisierung bzw. das Aufrechterhalten des Sicherheitsbestandes eines Lagers [DRA01]. Sicherheitsbestände bieten eine Sicherheit gegen fehlende Warenmengen, hervorgerufen unter anderem durch fehlerhafte bzw. verzögerte Lieferungen oder durch Nachfrageschwankungen. Die Berechnung von Mindestbeständen ist sehr aufwendig. Verschiedene Erfahrungswerte, Zeiten sowie stochastische Verfahren müssen berücksichtigt werden und erschweren die individuelle Berechnung. Abhilfe können Näherungsverfahren liefern.

Der Sicherheitsbestand eines Zentrallagers ist geringer als der addierte Sicherheitsbestand aller Lager eines dezentralen Lagersystems [SKJ07]. Dieser Zusammenhang kann mit der Quadratwurzelregel der

Lagerbestandszentralisierung engl. Square-Root-Law beschrieben werden [DUC17]. Das Verhältnis der kumulierten Sicherheitsbestände  $S_x$  zweier Netzwerksysteme ist dasselbe, wie das Verhältnis der Wurzeln der jeweiligen Lageranzahl  $n_x$ .

$$\frac{S_1}{S_2} = \frac{\sqrt{n_1}}{\sqrt{n_2}} \quad (1)$$

Beispielsweise führt das Einsetzen der Werte von einem Zentrallagersystem ( $S_1, n_1$ ) und einem Dezentralssystem ( $S_2, n_2$ ) mit fünf Lagern zu einer Reduzierung des Sicherheitsbestandes von 55 %. Darüber hinaus beschreibt das Quadratwurzelgesetz auch das Verhältnis allgemeiner Warenbestände zweier Logistiksysteme. Entsprechend begünstigt ein Zentrallagersystem die Reduzierung der Bestände. Der reduzierte Bestand wirkt sich positiv auf die Lagerkosten aus, kumulativ wird weniger Fläche benötigt und weniger Kapital gebunden.

Veränderungen der Transportparameter, Transportvolumen und -frequenz sowie der Distanz beeinflussen das logistische Netzwerk. Zusätzlich korrelieren diese auch untereinander. Transportwege mit unausgelastetem Transportvolumen müssen in einer erhöhten Frequenz wiederholt werden. Ein hohes Transportvolumen senkt die Kosten durch Reduzierung der nötigen Transportvorgänge oder deren Annullierung. Die von dem Transportvolumen beeinflusste Lieferfrequenz gibt den zeitlichen Abstand an in welchem der Warentransport getaktet ist. Je niedriger die Frequenz festgelegt ist, desto weniger Transportvorgänge verursachen Lieferkosten. Als letzten direkten Transportparameter wird der Einfluss der Lieferentfernung betrachtet. Die Überbrückung der Distanz weit entfernter Strecken bindet Transportmittel und verursacht Kosten für dessen Verwendung.

Die Abhängigkeiten innerhalb logistischer Netzwerke werden in Tabelle 2.3 zusammengefasst. Ihre Synergieeffekte, wie sich eine Aktion auf andere Aktionen auswirkt, werden in Kapitel 4.1 detaillierter beschrieben:

Tabelle 2.3: Abhängigkeiten des logistischen Netzwerkes

Eigenschaften/ Aktion		Beeinflusst
Zentralisierung	reduziert	Sicherheitsbestände
Anlegen von Sicherheitsbeständen	garantieren	Absatz
Zentralisierung	reduziert	Bestände
Anlegen von Beständen	bedingen	Absatz
Absatz	limitiert	Bestände
Anlegen von Beständen	bestimmen	Servicelevel
Servicelevel	beeinflusst	Absatz
Dezentralisierung	erhöhen	Kundenverbindungen
Transportvolumen	beeinflusst	Transportfrequenz
Transportfrequenz	beeinflusst	Transportvolumen
Kundenverbindungen	erhöhen	Transportfrequenz

### 3. Evolutionäre Algorithmen

Ein Algorithmus beschreibt die Vorgehensweise, wie ein Problem zu lösen ist [MEN97]. In einem Vorgehensplan werden sukzessive Einzelschritte beschrieben, welche ein Problem lösen sollen. Algorithmen sind von elementarer Bedeutung in der Mathematik und in der Programmierung. Sie sind nicht an eine Programmiersprache oder eine Formulierung gebunden. Sie können auch analog vermittelt werden. Diese Thesis arbeitet mit einer Unterkategorie der Algorithmen, sogenannten Evolutionären Algorithmen (EA).

Evolutionäre Algorithmen (EA) sind heuristische Verfahren für die Lösung eines Optimierungsproblems  $(\Omega, f, >)$  [KRU10]. Diese sind gekennzeichnet durch ein Problem, dessen Lösungswert möglichst groß werden soll. Mögliche Lösungen entstammen einem festgelegten Suchraum  $\Omega$ , diese Lösungskandidaten werden mittels einer Bewertungsfunktion  $f: \Omega \rightarrow \mathbb{R}$  bewertet. Die gewünschte Optimierung kann sich auf ein Minimierungs- oder Maximierungsproblem beziehen. Eine Vergleichsrelation  $> \in \{<, >\}$  ermöglicht beiderlei Problemstellungen. Die Menge der globalen Optima  $H \subseteq \Omega$  ist definiert als:

$$H = \{x \in \Omega_1 \mid \forall x' \in \Omega: f(x) > f(x')\} \quad (2)$$

Gesucht wird ein Element  $x$  aus dem Suchraum  $\Omega$ , für welches die Funktion  $f$  einen optimalen Wert annimmt.

Die Funktionsweise von EAs ist an evolutionsbiologische Prozesse angelehnt. Der Ablauf EAs beruht auf den Theorien der Wissenschaftler Darwin und Mendel. In Charles Darwins „*Über die Entstehung der Arten*“ wird dessen Evolutionstheorie als „*Survival of the Fittest*“ beschrieben. Evolution wird als kontinuierlicher Prozess angenommen, bei dem die stärksten Individuen einer Art die höchsten Überlebenschancen besitzen. Durch bessere Anpassung an ihre Umgebung können diese Individuen ihre Erbinformationen an ihre Nachfahren vererben. Schlecht angepasste Individuen pflanzen sich nur vermindert oder gar nicht fort. Am stärksten wird der Genpool einer Generation durch die am besten angepassten Individuen der Vorgängergeneration geprägt.

Zeitgleich mit Darwin lebte Gregor Johann Mendel, der dessen Konzepte mit Fokus auf die Vererbung weiterführte. Er führte eine Vielzahl an Versuchsreihen mit Kreuzungen von Erbsenpflanzen durch. Deren Samen untersuchte er auf Form und Farbe, so konnte er einige Gesetzmäßigkeiten erkennen. Aus diesen konnte er die Mendelschen Gesetze ableiten [HES72]: Uniformitätsgesetz, Spaltungsgesetz, Rekombinationsgesetz.

- **Uniformitätsgesetz:** Bestimmte Merkmale reiner Rassen einer Art setzen sich generationsübergreifend durch.
- **Spaltungsgesetz:** Bei der Kreuzung von Mischlingen spalten sich die Merkmale auf, wobei auch die Merkmale der Eltern wieder auftreten.
- **Rekombinationsgesetz:** Erbanlagen unterschiedlicher Rassen werden unabhängig vererbt.

Seit den frühen 60er Jahren werden diese Gesetzmäßigkeiten für Such- und Optimierungsaufgaben verwendet. Aus den EAs haben sich einige Bereiche herausgebildet [STR02]: Genetische Algorithmen (GA), Genetische Programmierung (GP), Evolutionsstrategien (ES), Evolutionäre Programmierung (EP). Diese Teilgebiete ähneln sich sehr stark, ihre Unterscheidung ist historisch bedingt. Trotzdem gibt es Unterschiede zwischen ihnen [KOC14].

Ihre Differenzierung kann anhand folgender Merkmale erfolgen:

- Selektionsoperatoren
- Repräsentation bzw. Kodierung der Lösung
- Reproduktionsoperatoren
- Auswahlmechanismus für Reproduktion

Die Differenzierungsmerkmale zeigen, dass Selektion und Reproduktion auch bei EA von elementarer Bedeutung sind. Wie diese realisiert werden, wird in den nachfolgenden Kapiteln erklärt.

## 3.1 Terminologie

Die verwendeten Begriffe EA sind ebenfalls an die Evolutionstheorie angelehnt. Im Folgenden werden die wichtigsten Begriffe erklärt, die in den aufbauenden Kapiteln verwendet werden [GER13]:

**Individuum/Chromosom:** In der Biologie setzt sich ein Individuum aus einer Vielzahl an Chromosomen zusammen. In diesen ist der individuelle Bauplan eines Lebewesens hinterlegt. Bei EA beinhalten Chromosome die Lösungen eines Optimierungsproblems [KOC14]. Der größte Unterschied bei Chromosomen im biologischen und im mathematischen Kontext liegt in der Anzahl der Chromosomen pro Individuum. Bei EAs besteht ein Individuum normalerweise nur aus einem einzelnen Chromosom, während Lebewesen meistens aus einer Vielzahl an Chromosomen bestehen. Daher werden die Begriffe kontextbezogen synonym verwendet. Individuen werden meistens als Bitstränge dargestellt, seltener als reellwertige Vektoren oder Integerlisten [RÄU14].

**Gen:** Ein Gen ist eine Sequenz oder eine einzelne Stelle eines Individuums/Chromosoms. Der Kontext bestimmt den genauen Sachverhalt, welchen ein Gen abbildet.

**Allel:** Die Merkmalsausprägung eines Gens wird Allel genannt. In der Biologie kann dies z.B. die Blütenfarbe beschreiben. Das Allel an bestimmter Stelle im Chromosom bestimmt, ob die Blüte weiß oder rot blühen würde. Im Kontext der EAs entspricht ein Allel einem Zeichen- bzw. Zahlenwert eines Gens.

In der Abbildung 3.1 werden diese Begriffe bildlich dargestellt:

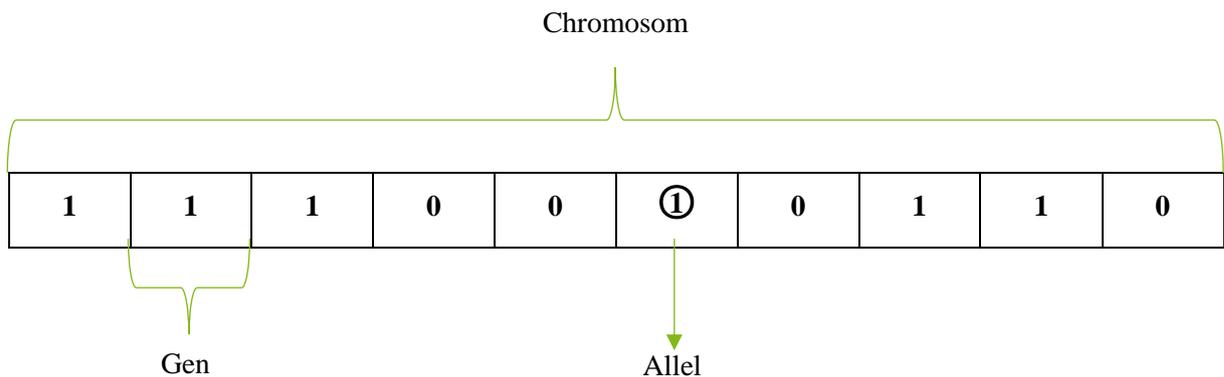


Abbildung 3.1: Darstellung Chromosom nach [KOC14]

**Länge eines Chromosoms:** Bei Evolutionären Algorithmen bestehen Individuen nur aus einem Chromosom, daher ist die Anzahl der Gene sowohl gleich der Länge des Chromosoms als auch der Gesamtchromosomlänge.

**Population:** Eine festgelegte Anzahl von gleichartigen Individuen wird als Population bezeichnet. Jedes Individuum beschreibt einen Punkt im Suchraum und enthält Parameter einer potentiellen Lösung. So besteht eine Population aus einer festgelegten Anzahl an potentiellen Lösungen.

**Generation:** Die Betrachtung von Populationen über mehrere Zeitpunkte wird als Generation bezeichnet.

**Genotyp** [BLU09]: Gesamtheit der Allele eines Organismus.

**Phänotyp:** Gesamtheit der Merkmalsausprägungen eines Individuums. Erscheinungsbild eines Individuums.

## 3.2 Aufbau eines Evolutionären Algorithmus

Die Generierung Evolutionärer Algorithmen erfolgt meistens einer gleichbleibenden Abfolge. Im Folgenden wird der Ablauf zuerst in seinem Pseudocode und dann in einer grafischen Darstellung vorgestellt. Der Pseudocode beschreibt den Ablauf des Algorithmus unabhängig der verknüpften Technologien [LI14].

Pseudocode

```
begin  
  Generiere InitialPopulation;  
  while (bis Stopp-Kriterium)  
    for (Jedes Chromosom)
```

```

    Berechne Fitnesswert;
    Selektion (Auswahl der stärksten Individuen);
    Crossover (hier werden neue generationen erzeugt);
    if (Es existieren Chromosome)
        Mutation (Austausch von Genen);
    end
    end
    Generiere neue Population;
end
end

```

Der Pseudocode ist ein grober Entwurf, der den Ablauf des Algorithmus aufführen soll. Er dient als informelle Anleitung der Programmierung. Aus dem Pseudocode wird nicht ersichtlich, welchen Nutzen die einzelnen Zeilen haben. Deswegen wird der Ablauf des Algorithmus in Abbildung 3.2 zusätzlich grafisch abgebildet. [RUD15]. Die einzelnen Bereiche sind chronologisch nummeriert und werden unter der Abbildung erläutert.

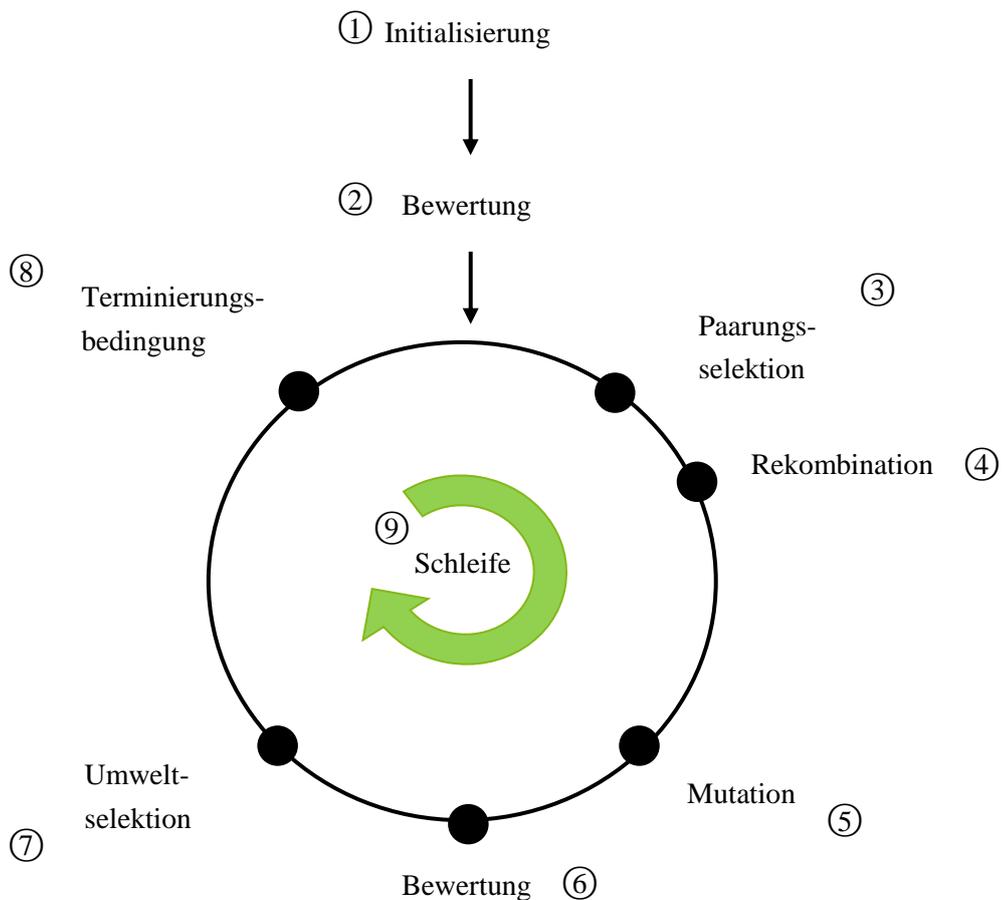


Abbildung 3.2: Ablauf von Evolutionären Algorithmen [RUD15]

Die einzelnen Elemente des Ablaufes eines EA werden im Folgenden kompakt erklärt, eine ausführlichere Beschreibung erfolgt anschließend [LIN03]:

1. Population wird initialisiert.
2. Bewertung der Population durch zu optimierende Funktion.
3. Jeweils zwei Elternteile werden mittels gewählter Selektionsvariante selektiert.
4. Aus Elterninformationen werden mittels Kreuzungsvariante Nachkommen erzeugt.
5. Die Allele der Nachkommen können mutieren.
6. Anhand der zu optimierenden Funktion wird die gegenwärtige Population bewertet.
7. Die Population wächst um die neugenerierten Nachkommen, nach einem Ersetzungsschema wird ein Teil der Ursprungspopulation durch neue Individuen ersetzt.
8. Die Schritte 3.-7. werden wiederholt, bis das Abbruchkriterium erfüllt wird.
9. Anhand der Terminierungsbedingung wird geprüft, ob der Algorithmus stoppt oder weitere Populationen generiert werden.

**Initialisierung:** Die Initialisierung beschreibt den Anfang eines EA. Es wird eine Startpopulation aus einer variablen Individuenzahl generiert. Dabei ist die Wahl der Populationsgröße wichtig. Eine zu kleine Population kann zu schnell in Richtung Optimum konvergieren. Eine zu große Population kann Rechenleistung blockieren [ZBI98]. Bei der Initialisierung wird zwischen zufälliger und nicht zufälliger Initialisierung unterschieden [POH99].

- Zufällige Initialisierung: Zum Startzeitpunkt sind keine Erkenntnisse über den gegebenen Suchraum bekannt.
- Nichtzufällige Initialisierung: Es sind bereits Teillösungen bekannt. Kenntnisse über bessere Individuen werden in den Optimierungsprozess miteinbezogen. Alternativ wird von einer erweiterten Initialisierung gesprochen.

**Bewertung:** Anhand einer Bewertungsfunktion wird die Eignung eines Individuums zur Lösung des Problems bemessen [SEL03]. Die Eignung zur Fortpflanzung wird mittels einer Fitnessfunktion ermittelt. Eine Fitnessfunktion wird im Kontext der Optimierungsprobleme auch Zielfunktion genannt. Der Begriff der Zielfunktion macht ihren Zweck deutlicher, sie deklarieren, welche Lösung als gültig erachtet wird.

Bei einer Fitnessfunktion handelt es sich um eine Funktion, welche jedem Element aus dem zugeordneten Suchraum eine meist positive Zahl zuordnet [KOC14]. Je höher der zugeordnete Fitnesswert, desto höher die Überlebenswahrscheinlichkeit des Individuums:

$$fit: S \rightarrow [0, \infty) \quad (3)$$

Der Ablauf der Fitnesswertbestimmung wird in Abbildung 3.3 schematisch dargestellt:

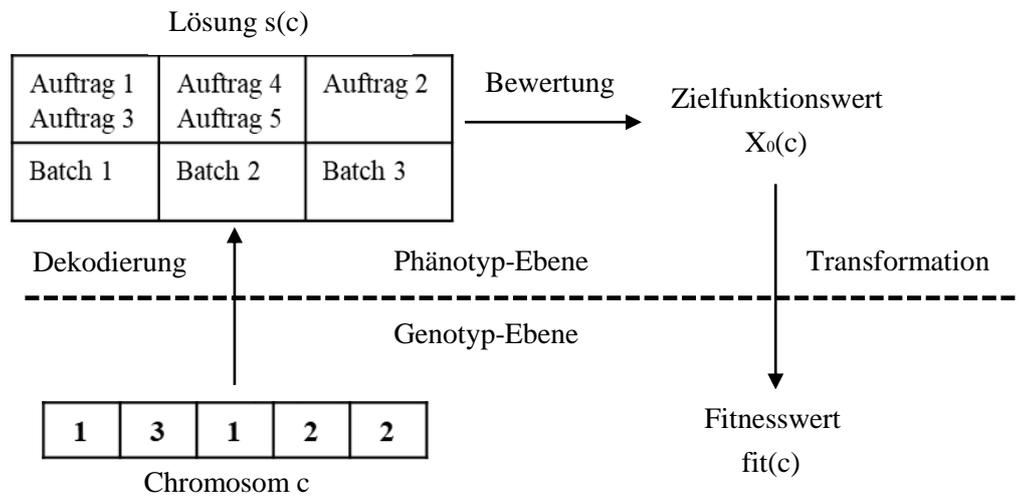


Abbildung 3.3: Fitnesswertbestimmung [KOC14]

Auffallend ist der Wechsel zwischen der Phänotyp-Ebene und der Genotyp-Ebene. Für die Bewertung der Fitnesswerte wird zunächst das Chromosom  $c$  dekodiert. Durch die Dekodierung erfolgt ein Wechsel zwischen der Genotyp- und der Phänotyp-Ebene. Diese Lösung wird anschließend bewertet und bekommt einen Zielfunktionswert zugeordnet. Falls das Optimierungsproblem eine maximale Lösung finden soll, entspricht der Zielfunktionswert  $x_0(c)$  dem Fitnesswert  $fit(c)$  und kann übernommen werden:

$$fit(c_k) = x_0(c_k) \quad k=1,2,\dots,n \quad (4)$$

Falls die Lösung des Optimierungsproblems möglichst klein sein soll, oder die Zielfunktionswerte negativ sind, müssen die Zielfunktionswerte in die Genotyp-Ebene transformiert werden, um verwendet werden zu können. Dies ist für eine der folgenden Schritte der EAs notwendig. Die Bestimmung der Selektionswahrscheinlichkeiten setzt voraus, dass gute Lösungen mit niedrigem Zielfunktionswert trotzdem einen hohen Fitnesswert zugeordnet bekommen. Vice versa müssen den schlechten Lösungen niedrige Fitnesswerte trotz hohem Zielfunktionswert zugeordnet werden. Die einfachste Transformation erfolgt über den Kehrwert:

$$fit(c_k) = \frac{1}{x_0(c_k)} \quad k=1,2,\dots,n \quad (5)$$

Eine Alternative zur Fitnesswertbestimmung eines möglichst geringen Lösungswertes ist der Umweg über eine selbstgewählte Konstante. Die Konstante muss mindestens so groß sein, wie der größte

Funktionswert. So wird vermieden, dass die Fitnesswerte negativen Zahlen zugewiesen werden. Meistens wird für den Wert der Konstante der Zielfunktionswert  $x_0(c_{\text{worst}})$  des schlechtesten Chromosoms  $c_{\text{worst}}$  der Gesamtpopulation gewählt. Als das schlechteste Chromosom wird das Chromosom bezeichnet, welches den größten Zielfunktionswert bezeichnet:

$$x_0(c_{\text{worst}}) \geq x_0(c_k) \quad k=1,2,\dots,n \quad (6)$$

Der Fitnesswert wird durch folgende Formel berechnet:

$$fit(c_k) = x_0(c_{\text{worst}}) - x_0(c_k) + \varepsilon \quad k=1,2,\dots,n \quad (7)$$

Der Berechnung des Fitnesswerts wird ein weiterer Wert  $\varepsilon$  hinzuaddiert. Dieser wird aufsummiert, um sicherzustellen, dass jedes Individuum, unabhängig von seiner Güte, einem positiven Fitnesswert zugeordnet wird. Durch einen Funktionswert größer als Null, nehmen nicht nur die elitären Individuen an der Reproduktion teil, sondern alle Individuen.

**Selektion und Rekombination:** Selektion und Rekombination stellen den evolutionären Anteil EA dar [KOC14]. Die Selektion bezieht sich auf die Auswahl der Chromosomen einer Generation, welche miteinander kombiniert werden sollen. Die Selektion baut auf der vorgelagerten Bestimmung des Fitnesswertes auf. Anhand der Fitnesswertbestimmung wird den einzelnen Individuen ein Rang zugeordnet. Je höher der Fitnessrang, desto höher der Rang. So bekommt das beste Individuum den ersten Rang zugeordnet.

Es gibt mehrere Unterscheidungsmerkmale bei Selektionsverfahren. Ein Merkmal ist die Unterscheidung zwischen statischen und dynamischen Verfahren. Bei statischen Selektionsverfahren werden die Chromosomen anhand ihres zugeteilten Ranges ausgewählt. Jedem Rang wird eine Selektionswahrscheinlichkeit zugeordnet. Je höher der Rang, desto höher die Wahrscheinlichkeit. Mit abnehmendem Rang sinkt auch die Selektionswahrscheinlichkeit. Die Zuordnung von Wahrscheinlichkeit und Rang bleibt bei statischen Selektionsverfahren über alle Generationen gleich. Das jeweilige Individuum des ersten Ranges besitzt generationsübergreifend dieselbe Wahrscheinlichkeit. Die Selektionswahrscheinlichkeit bei dynamischen Verfahren ist individueller verteilt. Sie wird von der Fitness des jeweiligen Individuums bestimmt. Da die Wahrscheinlichkeiten nicht länger konstant sind, sondern über die Fitness bestimmt werden, müssen die Selektionswahrscheinlichkeiten pro Generationswechsel berechnet werden.

Selektion kann entweder diskriminierend oder nichtdiskriminierend sein. Diskriminierende Selektion grenzt aus. Die Individuen bestimmter Eigenschaften nehmen nicht an der Selektion teil, ihre Selektionswahrscheinlichkeit wird auf Null gesetzt. Diskriminierende Selektion ist weiter unterteilt in rechts- und linksseitige Selektion. Rechtsseitige Selektion senkt die Selektionswahrscheinlichkeit der Individuen, die aufgrund eines schlechten Fitnesswerts am evolutionären Fortschritt nur einen geringen Anteil haben werden. Die Alternative ist eine linksseitige Selektion, bei der die besten Individuen

ausgeschlossen werden. Eine nicht diskriminierende Selektion wird angewendet, wenn die Selektionswahrscheinlichkeiten unverändert bleiben sollen.

Ein weiteres Unterscheidungsmerkmal ist die Unterscheidung zwischen reiner Selektion und Elitenselektion. Die reine Selektion basiert auf dem Zufall. Die Auswahl der Chromosomen erfolgt zufällig. Bei Elitenselektion wird eine festgelegte Menge der elitären Chromosomen mit zufälligen Chromosomen aufgefüllt. Elitäre Chromosomen beschreiben die geeignetsten Individuen und besitzen die höchsten Fitnesswerte.

Das verbreitetste Selektionsverfahren ist das *Roulette Wheel Selection*. Für dessen Initiierung wird zunächst die Selektionswahrscheinlichkeit jedes Chromosoms bestimmt. Aus dem Quotienten Fitnesswert des Chromosoms und der Summe der Gesamtfitnesswerte berechnet sich die Selektionswahrscheinlichkeit  $p_{sel}(c)$ :

$$p_{sel}(c_k) = \frac{fit(c_k)}{\sum_{u=1}^f fit(c_u)} \quad f=1,2,\dots,n \quad (8)$$

Nach Berechnen der Selektionswahrscheinlichkeit, muss die kumulierte Selektionswahrscheinlichkeit bestimmt werden. Sie errechnet sich aus der Summe der vorherigen Wahrscheinlichkeit mit dem betrachteten Chromosom:

$$p_{kum}(c_k) = \sum_{u=1}^k p_{sel}(c_u) \quad k=1,2,\dots,n \quad (9)$$

Beispielhaft wurden in der folgenden Tabelle 3.1 Berechnungen für die Selektionswahrscheinlichkeit und der kumulierten Wahrscheinlichkeit durchgeführt.

Tabelle 3.1: Selektionswahrscheinlichkeiten

Individuum	C <sub>1</sub>	C <sub>2</sub>	C <sub>3</sub>	C <sub>4</sub>	C <sub>5</sub>
Fitnesswert $fit(c_k)$	27	12	33	15	21
Selektionswahrscheinlichkeit $p_{sel}(C_k)$	$\frac{27}{108}$ =0.25	0.1111	0.3056	0.1389	0.1945
Kumulierte Selektionswahrscheinlichkeit $p_{kum}(c_k)$	0,25	0,3611	0,6667	0,8056	1,0000

Aus diesen Chromosomen wird zufällig ausgewählt, welches als Elternteil fungiert. Dies geschieht über die Generierung einer Zahl innerhalb [0,1]. Anschließend erfolgt die Auswahl über die zugehörige kumulierte Wahrscheinlichkeit. Bei einer Zufallszahl von 0,3 würde Individuum C<sub>2</sub> ausgewählt werden. Die kumulierte Wahrscheinlichkeit spannt ein Intervall zwischen dem Wert des vorherigen Individuums und dem Wert des betrachteten Individuums auf. Dies wird in der zugehörigen mathematischen Beziehung deutlich:

$$p_{kum}(c_{k-1}) < z \leq p_{kum}(c_k) \quad (10)$$

Damit das Individuum  $C_1$  ebenfalls entsprechend seiner kumulierten Wahrscheinlichkeit ausgewählt werden kann, wird ein imaginäres Chromosom  $c_0$  mit einer kumulierten Wahrscheinlichkeit von 0 erschaffen.

Der Name der beschriebenen Methode *Roulette Wheel Selection* ist an ein Roulette-Rad angelehnt. Den einzelnen Bereichen des Roulette-Rads werden die verschiedenen Chromosomen zugeordnet. Die Felder sind abhängig von dem Fitnesswert des zugehörigen Individuums. Das Generieren der Zufallszahl bzw. die Auswahl des Elternchromosoms entspricht einer Drehung des Roulette-Rads. Pro Drehung wird eine Zahl bzw. Chromosom per Zeiger ausgewählt. Die Anzahl der benötigten Eltern entspricht der Anzahl der Ziehungen der Zufallszahlen.

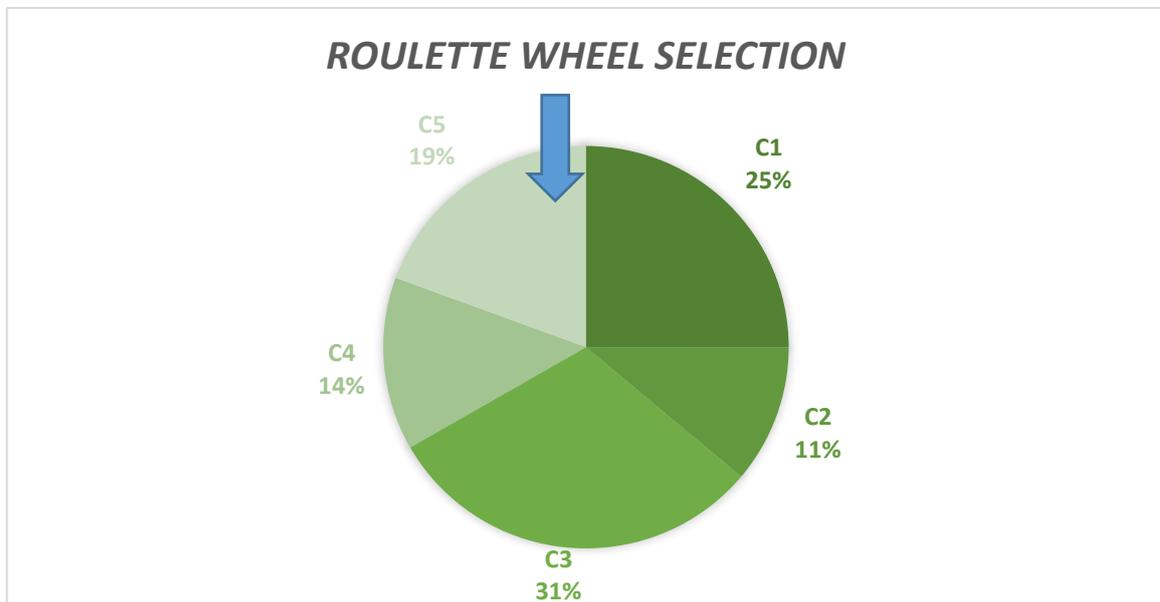


Abbildung 3.4: Darstellung der Roulette-Verteilung

Eine Alternative zur *Roulette Wheel Selection* ist das *Stochastic Universal Sampling*. Bei dieser Methode werden mehrere Eltern gleichzeitig ausgewählt. Die Initiierung verläuft wie bei der *Roulette Wheel Selection*. Für jedes Chromosom wird ebenfalls dessen Selektionswahrscheinlichkeit zugeordnet. Bei der *Roulette Wheel Selection* besitzt das symbolische Roulette Rad mehrere Zeiger, welche in gleichen Abständen befestigt sind. Das Roulette muss nur einmal gedreht werden, um die benötigten Chromosomen zu bestimmen.

Eine weitere Alternative ist das *Remainder Stochastic Sampling*. Erneut wird zunächst die Selektionswahrscheinlichkeit  $p_{sel}(c_k)$  bestimmt. Anschließend erfolgt eine Auswahl an Kopien jedes Chromosoms. Die Anzahl berechnet sich durch  $p_{sel} \times (c_k) \times n$ . Die restlichen Chromosomen werden aus der Gesamtzahl  $n$  über den modifizierten Fitnesswert  $fit_{mod}(c_k)$  bestimmt:

$$fit_{mod}(c_k) = p_{sel}(c_k)r[p_{sel}(c_k)n] \quad k=1,2,\dots,n \quad (11)$$

Von dem *Remainder Stochastic Sampling* gibt es zwei verschiedene Ausprägungen. Zum einen eine Variante mit Zurücklegen und eine Variante ohne Zurücklegen. Bei der ersten Variante werden neue Selektionswahrscheinlichkeiten berechnet, dies erfolgt diesmal über die modifizierten Fitnesswerte. Dann greift die Roulette Wheel Selection, mit der die restlichen Chromosomen bestimmt werden. In der Variante ohne Zurücklegen entscheidet der jeweiligen Fitnesswerte eines Chromosoms darüber, ob dieses ausgewählt wird. Diese Selektion läuft bis zum Erreichen der gewünschten Anzahl an Chromosomen.

Für die rangbasierte Selektion werden nur die relativen Unterschiede zwischen den Fitnesswerten berücksichtigt. Dem Verfahrensnamen entsprechend wird zunächst eine Rangfolge erstellt. Die Rangfolge erfolgt anhand ihrer Fitnesswerte. Die Ränge bestimmen die Selektionswahrscheinlichkeit. Dabei ist die Differenz der Selektionswahrscheinlichkeiten zwischen den Plätzen identisch.

Ein abweichendes Verfahren ist die *Wettkampfselektion*. Zuerst wird eine bestimmte Anzahl an Chromosomen aus der Gesamtpopulation gezogen. Diese Chromosomen werden anhand ihrer Fitnesswerte verglichen. Das Individuum mit dem höchsten Fitnesswert wird anschließend als Elternteil angenommen. Die restlichen Chromosomen werden nach Auswahl wieder zurückgelegt und stehen für eine erneute Auswahl zur Verfügung. Das Selektionsverfahren wird durchgeführt, bis die gewünschte Anzahl an Chromosomen ausgewählt worden ist.

Unter Rekombination versteht man einen weiteren an die Evolution angelehnten Schritt. Aus den Informationen beider Elternteile werden zwei Kinder erzeugt. Die Informationen werden dabei durchmischt. Dies wird auch als Crossing Over bezeichnet. Wie bei der Selektion existieren auch bei der Rekombination mehrere Verfahren [KOC14].

Ein weit verbreitetes Verfahren ist das *n-Punkt-Crossover*. Bei dem Verfahren werden die Informationen beider Eltern an n Stellen zerschnitten werden. Anschließend werden die Bereiche gerader Nummerierung für die Nachkommen ausgetauscht. Meist werden bei dem n-Punkt-Crossover -Verfahren verwendet, bei denen n gerade ist. In der folgenden Tabelle 3.2 wird das Verfahren mit n=2 exemplarisch dargestellt:

Tabelle 3.2: n-Punkt-Crossover [KOC14]

Bereich	1	2	3
<b>Elter<sub>1</sub></b>	1111	111	111
<b>Elter<sub>2</sub></b>	0000	000	000
<b>Kind<sub>1</sub></b>	1111	000	111
<b>Kind<sub>2</sub></b>	0000	111	000

Mit steigendem  $n$  nimmt das Position Bias ab [BOD06]. Dies ist eines der Merkmale zum Bewerten von Rekombinationsansätzen, dabei soll sein Wert möglichst gering ausfallen. Position Bias beschreibt die Austauschwahrscheinlichkeit eines Gens anhand seiner Lage in dem Chromosom. So nimmt bspw. die Wahrscheinlichkeit des 1-Punkt-Verfahrens zu, dass ein Gen ausgetauscht wird, je weiter es auf dem Chromosom liegt. Dieser Effekt ist unerwünscht, da er verhindert, dass weit entfernte Gene zusammen vererbt werden.

Dafür nimmt das Distributional Bias für ein anderes Merkmal mit steigendem  $n$  zu. Das sogenannte Distributional Bias beschreibt die fehlende Gleichverteilung ausgetauschter Gene.

Ein weiteres Rekombinationsverfahren ist das Uniform Crossover [NIS97]. Es werden keine Schnittpunkte festgelegt, die die auszutauschenden Segmente festlegen. Für jedes Elterngen entscheidet eine Zufallsvariable über dessen Austausch. Gängige Wahrscheinlichkeiten für einen Austausch liegen zwischen 0,5 und 0,8. Das Uniform Crossover wird in der nachfolgenden Tabelle 3.3 exemplarisch durchgeführt:

Tabelle 3.3: Uniform Crossover [NIS97]

<b>Elter<sub>1</sub></b>	<b>1</b>									
<b>Elter<sub>2</sub></b>	<b>0</b>									
<b>Austausch (ja/nein)</b>	<b>n</b>	<b>n</b>	<b>j</b>	<b>n</b>	<b>j</b>	<b>j</b>	<b>n</b>	<b>j</b>	<b>n</b>	<b>n</b>
<b>Kind<sub>1</sub></b>	<b>1</b>	<b>1</b>		<b>1</b>			<b>1</b>		<b>1</b>	<b>1</b>
<b>Kind<sub>2</sub></b>			<b>1</b>		<b>1</b>	<b>1</b>		<b>1</b>		

Bei dem Uniform Crossover-Verfahren wird jedes Gen mit gleicher Wahrscheinlichkeit ausgetauscht, daher besitzt es keinen Positional Bias. Das andere Merkmal, das Distributional Bias, ist deutlich stärker ausgeprägt. Die ausgetauschten Bits sind binomial verteilt.

Ein gänzlich anderes Verfahren ist die Inversion. Die Rekombination geschieht nicht zwischen zwei Eltern. Es wird nur ein Chromosom eines Elternteils vererbt. Wie bei dem  $n$ -Punkt-Verfahren werden Schnittpunkte festgelegt, die den Schnittbereich von dem restlichen Chromosom trennt. Die eingeschlossenen Bits werden invertiert, dies bedeutet, ihre jeweilige Position innerhalb des Bereiches wird umgekehrt. Dies ist in der Tabelle 3.4 dargestellt:

Tabelle 3.4: Eigene Darstellung der Inversion

<b>Elter</b>	1	0	1	0	1	1	1	1	0	1
<b>Kind</b>	1	0	1	1	0	1	1	1	0	1

Die beschriebenen Rekombinationsverfahren sind die am häufigsten verwendeten Operatoren. Es gibt weitere Verfahren, die auf spezielle Problemstellungen zugeschnitten sind. Aufgrund ihrer verminderten Relevanz wurde nur auf die wichtigsten Operatoren eingegangen.

**Mutation:** Im Zuge der Generierung der neuen Generation werden diese zufällig einer möglichen Mutation unterworfen. In diesem Fall werden den Variablen Fehler zugeordnet. Es existieren mehrere Arten von Mutationen, die nach ihrem Definitionsraum oder ihrer Funktionsweise unterschieden werden [POH99]:

- Mutation binärer Variablen
  - Veränderung Variablenwert: Die Mutation von Bitstrings erfolgt durch Bit Flips an einer oder mehreren Stellen.



- Mutation der Reihenfolge
  - Austausch von Variablen: Zwei Positionen jeweils an beliebiger Stelle der Individuen werden zufällig ausgewählt. Anschließend tauschen beide Stellen ihre jeweilige Position.



- Einfügen von Variablen: Zwei Positionen an beliebiger Stelle werden ausgewählt. Danach wird eine der beiden Variablen vor bzw. hinter die andere ausgewählte Variable verschoben.



### 3. Evolutionäre Algorithmen

- Reverse Mutation: Zwei Positionen an beliebiger Stelle werden ausgewählt. Im Unterschied zu den vorher genannten Mutationsarten werden nicht nur eine oder beide ausgewählten Positionen verändert. Stattdessen wird der gesamte Bereich, den die Positionen umschließen, invertiert.



- Scramble Mutation: Die beliebig ausgewählten Positionen werden zufällig neu angeordnet. Die restlichen Variablen behalten ihre Position.



Abbildung 3.5: Sammlung von Mutationen [POH99]

**Umweltselektion:** Nachdem die Nachkommen mutiert und mittels Zielfunktion bewertet wurden, tritt die Umweltselektion ein. Die Umweltselektion hat die Aufgabe, die Populationen der Eltern und der Nachfahren zu einer gemeinsamen Population zu vereinen, um die nächste Generation zu bilden [KOC14]. Dabei muss entschieden werden, welche Nachkommen zu der bestehenden Population hinzugefügt werden sollen sowie deren Anzahl. Im Gegenzug muss die Elternpopulation durch die hinzugefügten Individuen reduziert werden.

Der Parameter, der das Einsetzen steuert, ist die Wiedereinfügerate. Sie gibt an, wie viele Individuen der Elterngeneration maximal durch ihre Nachkommen auszutauschen sind. Bei einer Wiedereinfügerate von 1,0 werden alle Individuen ersetzt. Die Population von Eltern und Nachkommen wäre gleich groß. Der Parameter *generation gap* beschreibt den ergänzenden Teil dieses Austausches. Er beschreibt, wie viele Nachkommen produziert werden. Wiedereinfügerate und *generation gap* beschreiben zusammen, wie viele Individuen in die Grundpopulation eingefügt werden. Bei einer Wiedereinfügerate größer gleich des *generation gaps* werden alle Nachfahren der Population hinzugefügt. Übersteigt der Parameter *generation gap* die Wiedereinfügerate, so wird nur eine limitierte Anzahl an Individuen eingeführt.

Es haben sich mehrere Einsetzungsvarianten herausentwickelt, die sich in ihrer Gewichtung von Wiedereinfügerate und *generation gap* unterscheiden [DAN07]:

- Komma-Strategie: Austauschstrategie, bei der die neue Population nur aus Nachkommen gebildet wird.
- 

$$Population_{n+1} \subseteq Kinder \quad (12)$$

Die Wiedereinsetzungsrate und der *generation gap* liegen beide bei 1,0. Es werden so viele Nachkommen erzeugt, wie es Individuen in der Elternpopulation gibt. Die Lebensdauer eines

Individuums beträgt nur eine Generation, da jede Generation ersetzt wird, ist nicht gesichert, dass die Güte nicht schlechter wird.

- Plus-Strategien: Kombinationsstrategien, bei der die neue Generation aus Individuen der Elternpopulation und der Kinderpopulation gebildet wird.

$$Population_{n+1} \subseteq (Kinder \cup Population_n) \quad (13)$$

Es gibt zwei Untervarianten, die eine neue Population aus Eltern und Kindern zusammensetzen. Dies geschieht per zufälligem bzw. elitärem Wiedereinfügen oder per Auswahl der Nachkommen. Erstere Variante tritt ein, wenn entweder Wiedereinsetzungsrate  $< 1,0$  oder generation gap  $< 1,0$ . Der neuen Population werden mehr Individuen der Elternpopulation als Nachfahren hinzugefügt. Entweder werden die älteren durch neuere Individuen zufällig ausgetauscht oder die Auswahl betrifft die schlechtesten der Elternpopulation.

Nur die besten Individuen der Nachkommen (Wiedereinfügerrate/generation gap) werden in die Population eingesetzt.

**Terminierungsbedingung:** Der Optimierungsprozess benötigt eine Terminierungsbedingung. Dies ist ein künstliches Kriterium, bei dessen Erfüllung der Algorithmus beendet wird. Die zum Abbruchzeitpunkt besten Chromosomen gelten als Ergebnis des EAs. Die Individuen werden auf Erfüllung des Problems überprüft, dies können bspw. qualitäts- oder zeitabhängige Bedingungen sein [KOK02]. Es gibt mehrere gebräuchliche Abbruchkriterien verschiedener Art, die in unterschiedlicher Häufigkeit verwendet werden [KOC14]:

- Abbruch nach Zykluszahl
- Abbruch nach Rechenzeit
- Abbruch bei fehlender Verbesserung
- Abbruch bei fehlender Verbesserung des besten Zielfunktionswertes innerhalb definierter Zykluszahl
- Abbruch, wenn Verbesserung generationsübergreifend nicht hoch genug ist

Die am häufigsten verwendete Terminierungsbedingung ist ein Abbruch nach einer festgelegten Zykluszahl. Sie gilt als vergleichsweise einfach anzuwenden und findet daher oft Anwendung. Ein anderes Abbruchkriterium ist der Abbruch nach einer festgelegten Rechenzeit. Dieses Kriterium wird verwendet, um zwei oder mehrere Algorithmen miteinander vergleichen zu können. Der Unterschied zwischen Algorithmen liegt in der Anzahl an Schritten, die sie in derselben Rechenzeit vollziehen können. Der Vergleich über die Zykluszahl wäre daher irreführend. Die Rechenzeit betrifft alle Algorithmen gleichermaßen.

Die unteren drei Abbruchkriterien unterscheiden sich von den beiden oberen. Der Abbruch nach Zyklen und der Abbruch nach Rechenzeit wird vor Beginn des EA definiert. Die drei anderen Terminierungsbedingungen sind variabel gehalten, sie brechen den EA ab, wenn sie erfüllt bzw. nicht erfüllt werden. Durch ihre Variabilität kann die Rechenzeit angepasst werden

## 3.3 Programmiersprache

Der EA dieser Thesis wird in Python realisiert. Python ist eine der geläufigsten Programmiersprachen [LUT00]. Durch den einfachen Aufbau und die Möglichkeit einer schnellen Programmbeschreibung eignet sich die Programmiersprache vor allem für Naturwissenschaften. Gegenüber vergleichbaren Programmiersprachen ähnlicher Verbreitung, wie C++ oder Java, reduziert sich der Programmcode. Python ist eine Open Source-Programmiersprache.

Neben seiner kostenlosen Bereitstellung spricht für Python eine breite Anzahl an Bibliotheken. Neben eigens auf Python zugeschnittene Programmbibliotheken kann auch auf Bibliotheken von C und C++ zurückgegriffen werden. Die Programmierung kann objektorientiert ablaufen. Die Programmierung des Gesamtprogramms wird in einzelne Objekte unterteilt, welche eigenen Objektattributen unterworfen sind, aber miteinander interagieren können. Durch die Modularisierung in Teilelementen ist das Programm leicht zu warten und zu erweitern.

Der Syntax ähnelt den der anderen geläufigen Programmiersprachen. Zu sehen ist dies an dem üblichen „Hello World“-Programm:

```
print ('Hello World')
```

Die simple print-Funktion stellt einen variablen Text dar. Bspw. funktioniert die Textwiedergabe bei Java-ähnlichen System.out.println("Hello, World").

Ergänzend zu Python existiert auch ein Python-Kommandointerpreter namens Python-Shell, welcher Anweisungen unmittelbar ausführt. Python-Shell ist auf ein Betriebssystem angepasst. In der Shell steht der Interaktivmodus zur Verfügung. Wie eine Schale schützt es das System vor dem Benutzer, zusätzlich ermöglicht es das schnelle Testen von Anweisungen, ohne ein ganzes Programm zu schreiben.

### 4. Initialisierung des Algorithmus

Vor Einordnung der Auswirkungen der Korrelation einzelner Parameter eines logistischen Netzwerks müssen zuerst einige grundlegende Kriterien festgelegt werden, anhand derer die Effekte bewertet werden können. Die Beziehungen aus Kapitel 2.4 werden in diesem Kapitel genauer betrachtet und mit ihren Auswirkungen beschrieben. Mehrere Determinanten geben Aufschluss auf die Tauglichkeit logistischer Netzwerke [HAG12] [HER11]. Aufbauend auf diesen Beziehungen werden diverse Ansätze evolutionärer Algorithmen erörtert. Einer dieser Ansätze wird nach Eignung in Bezug zu der Problemstellung und der gegebenen Variablen ausgewählt, bevor dieser in den anschließenden Kapiteln angepasst wird.

#### **Logistikkosten:**

- Lagerkosten: allgemeine Kostenträger, wie Personal-, Energie- oder Raumkosten, z. B. Miete
- Bestandskosten: Kosten, die durch intralogistische Bewegung von Produkten im Lager entstehen. Hierzu zählen auch Zinsen auf die gelagerten Produkte sowie Kosten für Obsoletbestände, welche nicht verwendet werden können.
- Transportkosten: Alle Transportkosten, die durch den vorhandenen Lagertyp bedingt werden. Hierzu gehören sowohl Kosten zur Überbrückung intralogistischer Transportwege als auch aus- bzw. eingehende Transporte.
- Handlingskosten: Erweiterte Kosten zur Vor- bzw. Nachbereitung von Transportprozessen. Betrachtung Kosten für Kommissionierung und Verpackung.

#### **Zeit:**

- Der zeitliche Aspekt ist eine weitere Zielgröße. Transport- und Wartezeit sowie zeitliche Flexibilität sind Indikatoren für die Leistungsfähigkeit.

#### **Quantität:**

- Kapazitiver Faktor der jeweiligen Lager

#### **Qualität:**

- Qualitative Unterschiede entstehen vor allem durch das Warenhandling bzw. deren Manipulation.

Zentrale Unterscheidung ist die Wahl der Lagerart. Die Entscheidung zwischen Zentralisierung oder Dezentralisierung bestimmt das ganze logistische Netzwerk. Nur eine von beiden Methodiken kann realisiert werden. Die Zentralisierung von Waren reduziert die benötigten Sicherheitsbestände, verglichen zu den kumulierten Sicherheitsbeständen eines dezentralen Systems. Durch den geringeren Bedarf an bevorrateten Gütern sinkt der Anteil an gebundenem Kapital. Zentralisierung wirkt sich auf

Sicherheitsbestände positiv aus. Zusätzlich muss beachtet werden, dass Sicherheitsbestände angelegt werden sollten. Sie ermöglichen einen Absatz auch über die prognostizierten Nachfrage hinaus.

Selbiges gilt auch für die Bevorratung der Waren. Ausgenommen einer Just-in-Time Lieferkette werden fast in allen Transportbeziehungen Bestände benötigt. Bestände ermöglichen es, die Nachfrage von Waren zu decken und unterstützen so deren Absatz. Entsprechend existiert keine negative Korrelation bezüglich der Bevorratung von Beständen. Die Bestandsmenge muss zu dem jeweiligen Produkt bzw. dem zugehörigen System passen. Der Absatz bzw. die Nachfrage reguliert den Bestand. Die benötigten Bestände eines Zentrallagers sind geringer, als die zusammenaddierten Bestände aller Lager des komparativen Dezentallagersystems. Durch die reduzierte Bindung des Kapitals bzw. der bevorrateten Güter spart ein zentrales System an Kosten. Konträr verursacht eine Dezentralisierung höhere Kosten, da mehr gebundenes Kapital in dem Gesamtsystem vorliegt.

Das Service-Level beschreibt die Wahrscheinlichkeit, keinen Fehlbestand vorzufinden. Mit der Vermeidung von Fehlbeständen wird auch der zugehörige Umsatz beeinflusst. Das gewünschte Gut ist zu der Wahrscheinlichkeit im Lager vorrätig. Das Service-Level ist keine unmittelbar einstellbare Größe oder beeinflussbare Aktion, es wird bestimmt durch das Verhältnis von Beständen und Nachfrage. Bei gleich prognostizierter Nachfrage würden höhere Bestände die Nachfrage zu einer höheren Wahrscheinlichkeit decken können, als eine geringere Menge an bevorrateten Waren. Die Menge an gelagerten Gütern bestimmt aber gleichzeitig die Bestandskosten. Es muss ein Kompromiss zwischen einer höheren Verfügbarkeit und höheren Kosten gefunden werden.

Durch ein höheres Service-Level steigt die Transportfrequenz [BdKEP16]. Mit zunehmender Lieferverlässlichkeit steigt die Zahl an nötigen Lieferfahrten, um diese garantieren zu können. Ein geringes Service-Level erfordert nur eine geringe Transportfrequenz. Es werden nur wenige Transportfahrten gebraucht, um die benötigten Güter zu beschaffen, welche die Nachfrage decken sollen.

Ein weiterer Faktor thematisiert die Anzahl der Kundenverbindungen. Dezentralisierung eignet sich für viele Kleinkunden. Durch die größere Anzahl an Lagern sind diese näher an den Kunden. Die kürzere Distanz führt zu geringeren Transportzeiten und -kosten. Gegenteiliges gilt für zentralisierte Lager. Die Kundenstruktur ist häufig von Großkunden geprägt, welche weit verstreut liegen. Dadurch entstehen höhere Transportzeiten und -kosten.

Das Transportvolumen bestimmt die benötigte Transportfrequenz. Zusammen bestimmen sie den Transportaufwand einer Lieferung. Ein niedrig ausgelastetes Transportvolumen eines Fördermittels benötigt eine erhöhte Transportfrequenz, um dieselbe Gütermenge bei ausgelastetem Transportvolumen transportieren zu können. Bei einem hohen Transportvolumen kann die Transportfrequenz minimiert werden. Da sich jeder Transportvorgang finanziell auswirkt, gilt es, die Anzahl an benötigten Fahrten gering zu halten. Der Zusammenhang zwischen Transportvolumen und -frequenz gilt auch vice versa. Das Volumen könnte auch der Frequenz angepasst werden. Bei einer hohen Frequenz verringert sich das Transportvolumen.

Die herausgearbeiteten Korrelationen werden in der Tabelle 4.1 tabellarisch dargestellt. In der ersten Spalte werden die beeinflussbaren Aktionen als Variablen aufgeführt. Die zweite Spalte beschreibt das

#### 4. Initialisierung des Algorithmus

Verhältnis zwischen den Variablen. Die korrelierten Variablen werden in der dritten Spalte zugeordnet. Die dabei auftretenden Synergien oder negativen Effekte werden in der letzten Spalte dargestellt.

Tabelle 4.1: Korrelationen der beeinflussbaren Aktionen

Aktion 1	Verhältnis	Aktion 2	Effekte
Anlegen von Sicherheitsbeständen	garantieren	Absatz	-
Zentralisierung	erhöht	Sicherheitsbestände pro Lager	verringert Sicherheitsbestände insgesamt
Dezentralisierung	verringert	Sicherheitsbestände pro Lager	erhöht Sicherheitsbestände insgesamt
Anlegen von Beständen	bedingen	Absatz	-
Zentralisierung	erhöht	Bestände pro Lager	verringert Bestände insgesamt
Dezentralisierung	reduziert	Bestände pro Lager	erhöht Bestände insgesamt
Anlegen von Beständen	bestimmen	Verfügbarkeit	-
hohe Bestände	benötigen	hohe Transportvolumen	h. Quantität
		hohe Transportfrequenz	h. Bestandskosten
niedrige Bestände	benötigen	niedrige Transportvolumen	n. Bestandskosten
		niedrige Transportfrequenz	n. Quantität
Transportfrequenz	korreliert	Transportvolumen	-
niedriges Transportvolumen	benötigen	h. Transportfrequenz	h. Kosten
hohes Transportvolumen	benötigen	n. Transportfrequenz	n. Kosten
niedrige Transportfrequenz	benötigt	hohe Transportvolumen	n. Kosten
hohe Transportfrequenz	benötigt	niedrige Transportvolumen	h. Kosten

Einige der tabellarisch dargestellten Korrelationen sind reziprok. Ihre Auswahl begünstigt sich wechselseitig. Abbildung 4.1 hebt diese Beziehungen deutlicher hervor. Ausgehend von der Auswahl des Lagertyps werden die anderen Variablen mit Richtungspfeilen verbunden.

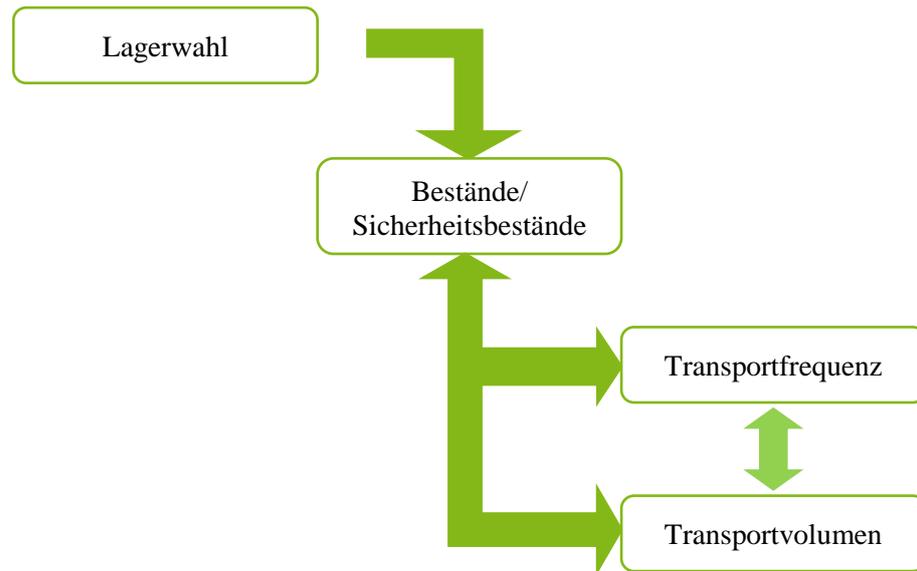


Abbildung 4.1: Abhängigkeitsverhältnisse der Variablen

Die Auswahl des Lagertyps korreliert einseitig mit der Höhe der Bestände und Sicherheitsbestände. Die Wahl, ob eine Ware zentralisiert oder dezentralisiert wird, beeinflusst die Bevorratung in einem Lager. Umgekehrt wirkt sich eine Verringerung oder Erhöhung der Bestände nicht auf eine eventuelle Zentralisierung aus. Abgesehen von der Lagerauswahl beeinflussen sich die anderen Parameter wechselseitig. Bestände benötigen einen angepassten Abtransport, auf den die Parameter Transportfrequenz und -volumen zugeschnitten sind. Umgekehrt erfüllt eine Transporttaktung nur ihren Zweck, wenn die benötigten Waren in einer ausreichenden Menge vorrätig sind.

### 4.1 Problemstellung

Tabelle 4.1 stellt die Korrelationen aller möglichen Aktionen in logistischen Netzwerken dar. Eine ideale Aktionsliste setzt sich aus einer Lagerart und vier synergetischen Variablenausprägungen zusammen.

In Tabelle 4.2 wird eine solche Aktionskette dargestellt. Die genauen Waren und Distributionszentren (DC) sind unspezifisch gehalten. Durch die unspezifische Darstellung wird eine allgemeingültige Aktionskette idealisiert.

#### 4. Initialisierung des Algorithmus

Tabelle 4.2: Ideale Aktionsliste

Aktionsliste				
1	2	3	4	5
Zentralisierung von Ware <sub>x</sub> in DZ <sub>x</sub>	h.Bestände Ware <sub>x</sub> in DZ <sub>x</sub>	h.Sicherheitsbestände von Ware <sub>x</sub> in DZ <sub>x</sub>	h.Transportvolumen Ware <sub>x</sub> in DZ <sub>x</sub>	n.Transportvolumen DZ <sub>x</sub>

Diese Darstellung der idealisierten Logistikkette suggeriert, dass das logistische Netzwerk ein unabhängiges System bildet. Die einzelnen Aktionen der Transportkette beeinflussen von ihnen abhängige Aktionen. Durch die abhängigen Verknüpfungen zwischen unspezifischen Aktionen erscheint die vorgestellte Aktionskette trivial. Real existierende Logistiknetze sind aber selten derart vereinfacht, dass sie nur ein eindimensionales System umfassen. Ein logistisches Netzwerk umfasst selten nur einen einzigen Warentyp. Ebenso existieren meistens mehrere Distributionszentren in einem logistischen Netzwerk. Da Waren allen möglichen Distributionszentren zugeordnet werden können, spannt sich ein mehrdimensionales System auf. Die Aufteilung des Gütertransportes in mehrere unterschiedliche Waren und Distributionszentren erschwert die Gesamtübersicht und erhöht so den Aufwand etwaiger Analysen des logistischen Netzwerkes.

Durch die Initialisierung des Evolutionären Algorithmus wird eine Elterngeneration generiert. Diese besteht aus Individuen, deren einzelne Gene meistens zufällig ausgewählt werden. Tabelle 4.3 zeigt ein mögliches Chromosom bzw. ein Individuum, welches durch die Initialisierung erzeugt werden kann. Es wird eine Aktionsliste aus fünf Elementen exemplarisch dargestellt. Die Anzahl an aufgeführten Aktionen ist an die Anzahl der optimalen Aktionskette bzw. der Anzahl der korrelierenden Variablen aus Tabelle 4.2 angelehnt und wird im Verlauf dieser Thesis beibehalten.

Die Aktionsmerkmale auf den Genen des Individuums wurden so ausgewählt, dass die fünf Aktionselemente nicht miteinander korrelieren. Es soll exemplarisch gezeigt werden, wie variabel die Merkmalskombinationen sein können. Im Hinblick dessen wird auf korrelierende Faktoren verzichtet, um zu zeigen, dass auch Individuen mit gänzlich unabhängigen Merkmalskombinationen generiert werden können.

Das erste ausgewählte Element ist dasselbe, wie das in Tabelle 4.2. In der optimalen Aktionskette würde der Zentralisierung der Ware<sub>1</sub> in dem Distributionszentrum<sub>1</sub> eines der korrelierenden Elemente aus Erhöhung der Sicherheitsbestände oder der Erhöhung der allgemeinen Bestände folgen. Die weiteren Aktionen weichen von der optimalen Form ab. In dem zweiten Segment bezieht sich die aufgeführte Aktion auf einen anderen Warentyp Ware<sub>2</sub>. Beide Waren eint nur ihr Distributionszentrum, ansonsten korrelieren die beiden Güter nicht miteinander und stehen auch in keiner anderen Verbindung zueinander. Die restlichen Elemente der Aktionsliste spannen das logistische Netzwerk noch weiter auf. Segment Drei führt sowohl eine weitere Ware als auch ein weiteres Distributionszentrum ein. Dies geschieht auch bei der folgenden vierten Aktion. Die Bestände der Ware<sub>2</sub> sollen in Distributionszenter<sub>5</sub> verringert werden. Hier wird wieder Bezug zum zweiten Element aufgenommen. Durch diese Aktion wurden die Bestände der Ware<sub>2</sub> in einem Lager erhöht, während durch Aktion Vier die Bestände in einem Lager reduziert werden. Nicht nur die Bestandszahl von Waren kann durch verschiedene

#### 4. Initialisierung des Algorithmus

Aktionen beeinflusst werden, auch Distributionszentren können verschiedenen Waren zugeordnet werden. Das letzte Element beschreibt die Erhöhung der Bestände von der Ware<sub>5</sub> in Distributionszentrum<sub>2</sub>, sowie Aktion Zwei ebenfalls die Bestände der Ware<sub>1</sub> in diesem Distributionszentrum erhöht.

Tabelle 4.3: Zufällige Aktionsliste

Aktionsliste				
1	2	3	4	5
Zentralisierung von Ware <sub>1</sub> in DZ <sub>1</sub>	h.Bestände Ware <sub>2</sub> in DZ <sub>1</sub>	Zentralisierung von Ware <sub>3</sub> in DZ <sub>2</sub>	n.Bestände Ware <sub>2</sub> in DZ <sub>5</sub>	h.Bestände Ware <sub>5</sub> in DZ <sub>2</sub>

Die Generierung von Individuen mit zufälligen Zusammenstellungen von Allelen ist verhältnismäßig einfach. Die Initialisierung der Elternpopulation unter Berücksichtigung bekannter Abhängigkeiten der Merkmale ist deutlich komplexer und aufwendiger.

Gesucht wird eine Initialisierung, die zufallsbasierte Individuenerzeugung und die Auswahl nach bekannten Abhängigkeiten kombiniert. Merkmale sollen wie in Tabelle 4.3 zufällig ausgewählt und dem Chromosom eines Individuums hinzugefügt werden. Zusätzlich soll ein bereits ausgewähltes Merkmal die Auswahl des nächsten Merkmals beeinflussen. Gemäß Tabelle 4.2 soll eine Aktion die Auswahl weiterer Aktionen begünstigen, mit denen es einen Strang der optimalen Aktionskette bildet.

Ein mögliches Resultat einer solchen kombinierten Initialisierung aus Zufall und Abhängigkeit wird in Tabelle 4.4 dargestellt. Die drei mittleren Elemente sind der optimalen Aktionsliste entnommen. Sie sind auch in der optimalen Reihenfolge angeordnet. Die beiden äußeren Elemente wurden zufällig ausgewählt. Sie korrelieren nicht mit den drei mittleren Elementen.

Tabelle 4.4: Reale Aktionsliste

Aktionsliste				
1	2	3	4	5
Zentralisierung von Ware <sub>3</sub> in DZ <sub>2</sub>	Zentralisierung von Ware <sub>1</sub> in DZ <sub>1</sub>	h.Bestände Ware <sub>1</sub> in DZ <sub>1</sub>	h.Sicherheitsbestände von Ware <sub>1</sub> in DZ <sub>1</sub>	h.Bestände Ware <sub>2</sub> in DZ <sub>1</sub>

Die einzelnen Aktionen einer kombinierten Methodik sollen weiterhin zufällig ausgewählt werden können. Ein ausgewähltes Element soll die Wahrscheinlichkeit begünstigen, dass ein korrelierendes Element ausgewählt wird. Wenn das im Beispiel angegebene Segment Zwei ausgewählt werden würde, Ware<sub>1</sub> in Distributionszentrum<sub>1</sub> zu zentralisieren, so soll die Wahrscheinlichkeit der Auswahl der Aktion der Bestandserhöhung der Ware<sub>1</sub> in Distributionszentrum<sub>1</sub> gesteigert werden. Eine Abweichung von der idealen Aktionskette führt zu einer neuen Ausrichtung an dem abweichenden Element. An dieser Aktion wird der Ablauf der idealen Aktionskette weitergeführt. Wenn bspw. dem zuvor beschriebenen Element

der Bestandserhöhung der Ware<sub>2</sub>, eine Aktion der Zentralisierung der Ware<sub>1</sub> folgt, soll das anschließende Element wahrscheinlich die Bestandserhöhung der Ware<sub>1</sub> beschreiben.

### 4.2 Kodierung und Initialisierung

Kodierung und Initialisierung sind die initialen Phasen eines evolutionären Algorithmus. Die Planung dieser Phasen bestimmt die Güte eines Evolutionären Algorithmus. Je besser diese Phasen auf ein Problem zugeschnitten werden, desto genauer die Ergebnisse des EA. Durch die Phasen der Kodierung und der Initialisierung wird eine Population mit ersten Individuen erzeugt. Mit den Individuen kann der Algorithmus arbeiten.

Der für diese Arbeit relevante Algorithmus soll durch bekannte Zusammenhänge seiner Variablen angepasst werden. Diese Erkenntnisse fließen bereits in die Planung der Kodierung und Initialisierung mit ein. Regulär werden die Individuen der Elternpopulation meist zufällig in der Initialisierungsphase generiert. Die zufällige Auswahl der Allele soll beeinflusst werden. Die Auswahl einzelner Allele soll zufällig erfolgen. Die Auswahl weiterer Allele eines Individuums soll aber von den zuvor ausgewählten Merkmalen mitbeeinflusst werden. Für die angepasste Individuenauswahl bzw. der ihnen zugeordneten Gene reicht eine Anpassung der Initialisierungsphase nicht aus. Die Kodierung muss zu der Initialisierung passen und vice versa. Da sich beide initialen Phasen gegenseitig beeinflussen, werden beide Phasen zusammen in einem Kapitel geplant.

#### 4.2.1 Binäre Kodierung

Intuitiver Ansatz ist eine binäre Kodierung [YU10]. Da binäre Kodierungen zu den gebräuchlichsten Kodierungen gehören, wird zuerst eine mögliche Anwendung erörtert. Jede der Variablen besteht aus genau zwei Ausprägungen, diesen wird jeweils eine „1“ oder eine „0“ zugeordnet. Auf den ersten Blick überzeugt die Binärkodierung durch ihre Einfachheit. Den einzelnen Variablen würden feste Gene auf dem Chromosom zugeordnet werden. Bei jedem Individuum stünde an einer festen Genposition eine veränderliche Merkmalsausprägung einer Variablen. Durch ihre eindeutige Lage und festgelegten Ausprägungen wäre durch eine binäre Kodierung auch eine schnelle Dekodierungen möglich. So könnte die Codeform problemlos in ihre Bestandteile zerlegt werden.

Idee war eine Kodierung in kumulierten Einser- und Nullerketten. Die Bestandteile einer optimalen Aktionskette sollten einheitlich kodiert werden. Der Zentralisierung und ihren begünstigten Merkmalen sollten so wahlweise durch eine Reihe von Einsen oder Nullen beschrieben werden. Bei der Dezentralisierung sollte dies durch die übrige Kodierung geschehen. Dann würde gelten: Je mehr gleichkodierte Bits, desto ähnlicher einem der beiden optimalen Pfade. Ansatz war die Bildung möglichst langer gleichkodierter Ketten.

Nach ersten Überlegungen wurde der Ansatz der binären Kodierung schnell wieder verworfen. Ein evolutionärer Algorithmus dieser Kodierung ist zwar übersichtlich und einfacher zu gestalten, für das bestehende Problem jedoch nicht geeignet. Eine Variable belegt durchgehend das gleiche Gen. So würden in jeder Generation dieselben Variablen an denselben Positionen eines Codes liegen. Zusätzlich funktioniert eine binäre Kodierung nur in einem eindimensionalen System. Eine Variable besitzt in der binären Kodierung nur zwei Ausprägungen. In dem vorliegenden Fall wären dies jeweils die

gegenteiligen Aktionen einer der Variablen, die jeweils eine Ware und ein Distributionszentrum betreffen. In dem System sollen aber verschiedene Waren verschiedenen Distributionszentren zugeordnet werden können. Dies ist bei einer binären Kodierung nicht möglich. Aus demselben Grund ist eine realwertige Kodierung ebenfalls nicht möglich. Bei der realwertigen Kodierung sind die Merkmale auf festgelegten Positionen des Chromosoms angeordnet. Die Lokalisierung der Merkmale variiert nicht.

Aus den Überlegungen mit der unpassenden Binärkodierung ließen sich zwei grundlegende Erkenntnisse gewinnen, die zukünftige Ansätze erfüllen müssen.

- Es muss ein Ansatz gewählt werden, der ohne festgelegte Gene funktioniert. Aktionen müssen an jeder Stelle eines Chromosoms auftreten können.
- Die Kodierung darf nicht eingrenzend sein. Eine binäre Kodierung reicht bei weitem nicht aus, alle möglichen Kombinationen aus Merkmalsausprägung und den zugehörigen Waren und Distributionszentren abdecken zu können.

### 4.2.2 Anlehnung an das “Traveling Salesperson Problem“

Eine Anwendungsmöglichkeit Evolutionärer Algorithmen ist das bekannte Problem des Handlungsreisenden. Dies ist kein exklusives Anwendungsfeld Evolutionärer Algorithmen, doch wird das Problem häufig mittels EA gelöst. In der Theorie muss ein Händler durch eine Reihe von Städten reisen. Diese Städte sollen alle bereist werden, ohne dass eine Stadt mehrmals besucht wird. Durch die verstreute Anordnung der Städte ist der kürzeste Weg durch alle Städte nicht erkennbar. Evolutionäre Algorithmen sollen die Wegstrecke minimieren.

Auf den ersten Blick haben der Ansatz für das Problem des Handelsreisenden und die Problemstellung wenig gemein. Sie eint nur ihre quasi-randomisierte Variablenauswahl. Bei dem Problem des Handelsreisenden wählt der Algorithmus seinen Streckenstartpunkt zufällig aus und verbindet alle weiteren Punkte sukzessiv. Näher gelegene Punkte werden mit einer höheren Wahrscheinlichkeit verknüpft. Verbindungen zwischen weit entfernten Punkten sind zwar möglich, aber Verbindungen mit geringerer Entfernung halten die kumulierte Gesamtstrecke geringer. Durch die wahrscheinlich kürzere Gesamtstrecke werden kurze Paarungen eher eingegangen.

Das bestehende logistische Problem und das Distanzproblem des Handlungsreisenden unterscheiden sich hauptsächlich in ihrer Variablenauswahl. Das Problem des Handelsreisenden arbeitet mit Daten, die bspw. als eine Distanztabelle einzelner Städte oder Orte vorliegen. Diese Daten beschreiben Punkte, für deren Auswahl nur ihre Distanz zueinander relevant ist. Die Punkte sind alle gleichwertig und einzigartig. Es erfolgt weder eine Priorisierung noch existieren verschiedene Ausprägungen von ihnen.

Im Falle von Problemen Handelsreisender wird die Gesamtstrecke durch die Orte als Chromosom bezeichnet [DEV13]. Die Rückreise zu dem initialen Startpunkt ist optional und kann problemspezifisch berücksichtigt werden. Jede Rundreise ist eine mögliche Lösung  $x$  des Problems. Eine Rundreise besteht als Vektor aus  $n$  Elementen. Jedes Element ist ein Integer zwischen 1 und  $n$ , der Anzahl der zu besuchenden Städte. Eine Rundreise bzw. die Lösung  $x$

$$x = (x_1, \dots, x_n) \in P \quad (14)$$

wird auch als Chromosom bezeichnet. Wenn der zulässige Bereich  $P$  fünf Orte umfasst, ist  $[2\ 1\ 5\ 4\ 3]$  eine mögliche Lösung und ein Chromosom [SCH18] [STA14]. Die Reihenfolge der Orte auf dem Chromosom bestimmt dessen Eignung zur Fortpflanzung. Für die Lösung des Problems des Handlungsreisenden wird die Ziel- bzw. Fitnessfunktion  $f(x)$  über die Abstände der Gene eines Chromosoms berechnet:

$$f(x) = \|p_{x_1} - p_{x_2}\| + \|p_{x_2} - p_{x_3}\| + \dots + \|p_{x_{n-1}} - p_{x_n}\| \quad (15)$$

Das logistische Problem verwendet Variablen, die in zweifacher Ausprägung und in verschiedenen Waren und Distributionszentren auftreten können. Die Kodierung nach Variablen in all ihren Ausprägungen wird in Systemen höherer Ordnung schnell kompliziert. Die Beziehungen zwischen allen Waren und Distributionszentren kann durch klassische Kodierung nicht berücksichtigt werden. Eine Alternative ist die separate Kodierung aller Merkmalsausprägungen.

Die einzelnen Merkmalsausprägungen werden in Punkte überführt. Dabei wird jede Merkmalsausprägung genau einem Punkt zugeordnet. Diese Punkte werden mit einer eindeutigen Nummerierung kodiert. Diese Punkte beschreiben nicht die Variable, wie es bei einer binären Kodierung der Fall wäre, sondern jeweils nur eine der Merkmalsmöglichkeiten. So ist die Auswahl aller Ausprägungen der Variablen möglich. In Anlehnung an den Handlungsreisenden, der diverse Orte besucht, beschreibt eine Reihe von Punkten die Auswahl verschiedener Merkmalsausprägungen. Merkmalsausprägungen der gleichen Variable, aber unterschiedlicher Waren oder Distributionszentren können miteinander verkettet sein. Binäre Kodierung unterscheidet nur zwischen den Variablen so dass eine Kette der Merkmale „Zentralisierung der Ware<sub>1</sub> in Distributionszentrum<sub>1</sub>“ und „Zentralisierung der Ware<sub>2</sub> in Distributionszentrum<sub>2</sub>“ nicht möglich ist. Eine Kette mit binärer Kodierung beinhaltet ausschließlich verschiedene Variablen, deren Position auf dem Chromosom festgelegt ist. Für die Möglichkeit gleiche Variablen zu verketten, werden die Merkmale einzeln kodiert. Die Merkmale werden als Punkte angesehen die einzeln besucht werden können. Diese Punkte unterliegen ebenfalls keiner Priorisierung und beinhalten keine verschiedenen Ausprägungen. Durch die Möglichkeit, Variablen doppelt auszuwählen, werden den Variablen keine festen Gene zugeordnet. Entsprechend der Repräsentation der Rundreise als Chromosom können die Variablen an jeder Position eines Chromosoms stehen.

Tabelle 4.5 zeigt eine Kodierung für ein zweidimensionales System. Die Darstellung der Kodierung höher dimensionierter Systeme würde zu viel Platz beanspruchen. In der aufgeführten Tabelle stehen verschiedene Variablen in separaten Spalten. Ihre Ausprägungen stehen untereinander. Die zugehörigen Kodierungen stehen in der nächsten Spalte. Es werden zwei verschiedene Waren betrachtet, die in zwei verschiedenen Distributionszentren gelagert werden können.

#### 4. Initialisierung des Algorithmus

Tabelle 4.5: Kodierung für zweidimensionales System

	Code		Code		Code		Code		Code
Zentralisierung $W_1$ in $DZ_1$	01	Erhöhen $B_1$ in $DZ_1$	11	Erhöhen $SB_1$ in $DZ_1$	21	Erhöhen TF: $SB_1$ in $DZ_1$	31	Erhöhen TV: $SB_1$ in $DZ_1$	41
Zentralisierung $W_1$ in $DZ_2$	02	Erhöhen $B_1$ in $DZ_2$	12	Erhöhen $SB_1$ in $DZ_2$	22	Erhöhen TF: $SB_1$ in $DZ_2$	32	Erhöhen TV: $SB_1$ in $DZ_2$	42
Zentralisierung $W_2$ in $DZ_1$	03	Erhöhen $B_2$ in $DZ_2$	13	Erhöhen $SB_2$ in $DZ_2$	23	Erhöhen TF: $SB_2$ in $DZ_2$	33	Erhöhen TV: $SB_2$ in $DZ_2$	43
Zentralisierung $W_2$ in $DZ_2$	04	Erhöhen $B_2$ in $DZ_2$	14	Erhöhen $SB_2$ in $DZ_2$	24	Erhöhen TF: $SB_2$ in $DZ_2$	34	Erhöhen TV: $SB_2$ in $DZ_2$	44
Dezentralisierung $W_1$ in $DZ_1$	05	Verringern $B_1$ in $DZ_1$	15	Verringern $SB_1$ in $DZ_1$	25	Verringern TF: $SB_1$ in $DZ_1$	35	Verringern TV: $SB_1$ in $DZ_1$	45
Dezentralisierung $W_1$ in $DZ_2$	06	Verringern $B_1$ in $DZ_2$	16	Verringern $SB_1$ in $DZ_2$	26	Verringern TF: $SB_1$ in $DZ_2$	36	Verringern TV: $SB_1$ in $DZ_2$	46
Dezentralisierung $W_2$ in $DZ_1$	07	Verringern $B_2$ in $DZ_2$	17	Verringern $SB_2$ in $DZ_2$	27	Verringern TF: $SB_2$ in $DZ_2$	37	Verringern TV: $SB_2$ in $DZ_2$	47
Dezentralisierung $W_2$ in $DZ_2$	08	Verringern $B_2$ in $DZ_2$	18	Verringern $SB_2$ in $DZ_2$	28	Verringern TF: $SB_2$ in $DZ_2$	38	Verringern TV: $SB_2$ in $DZ_2$	48

Die Codierung der Merkmalsausprägungen wurden so gewählt, dass alle Kodierungen einer Variable dieselbe Zehnerziffer aufweisen. Dies erhöht die Übersichtlichkeit und erleichtert die anschließende Dekodierung. Die Zuordnung von kodierter Form und dekodierter Beschreibung funktioniert schneller, wenn bekannt ist, welche der Variablen mit welcher Ziffer anfängt.

Die Art der Kodierung kommt mit zunehmender Dimensionalität des Systems an seine Grenzen. Mit erhöhter Waren- und Distributionszentrenanzahl funktioniert die Aufteilung nicht mehr. Schon die Erhöhung zu einem dreidimensionalen System verhindert die obige Kodierung. Da  $3^3 = 27$  reicht eine Dekodierung in Zehnerpotenzen nicht aus. Da jeder Variablen zwei Ausprägungen zugeordnet sind, muss eine Kodierung mindestens 54 Elemente pro Variable abdecken. In einem dreidimensionalen System werden die Variablen bis 1000 kodiert. Anhand der Hundertziffer eines Merkmals kann die

#### 4. Initialisierung des Algorithmus

---

Variable schnell abgelesen werden. Mit jeder zusätzlichen Dimension bedarf die Kodierung mindestens einer zusätzlichen Zehnerpotenz. Der Zusammenhang zwischen Dimension und benötigten Codes wird in Tabelle 4.6 dargestellt. Die benötigte Zehnerpotenz ist jeweils um eine Stelle höher als die potenzierte Dimensionalität. So kann eine Kodierung direkt ihrer Variable zugeordnet werden, unabhängig der Dimensionalität. Die aufgeführten Dimensionen werden bis zur fünften Stufe aufgeführt, deckungsgleich mit der betrachteten Variablenzahl des Problems. So kann jeder der fünf Variablen eine andere Waren/Distributionszentrum-Beziehung beschreiben.

Tabelle 4.6: Zusammenhang von Dimensionalität und Kodierung

Dimensionalität				
1	2	3	4	5
1	4	27	256	3125
Zehnerpotenz zur Kodierung				
$10^0=1$	$10^1=10$	$10^2=100$	$10^3=1000$	$10^4=10000$

Ausschlaggebend für eine hinreichend gute Lösung eines Problems des Handlungsreisenden sind möglichst kurze Entfernungen der Punkte zueinander. In realen Systemen oder gegebenen Distanztabelle bestimmen Maßangaben die Position der Punkte zueinander. Punkte kürzerer Distanz werden wahrscheinlicher nacheinander besucht, um keine Umwege zu entfernten Punkten zu verursachen. Innerhalb der logistischen Variablen sollen auch bestimmte Merkmalsausprägungen vermehrt gemeinsam ausgewählt werden. Bei Auswahl eines Merkmals soll die Wahrscheinlichkeit steigen, ein korrelierendes Merkmal ebenfalls auszuwählen. Es soll aber immer noch möglich sein, unabhängige Merkmale auswählen zu können, wenn auch mit reduzierter Wahrscheinlichkeit. In Anlehnung an das Distanzproblem des Handlungsreisenden werden künstliche Abstände zwischen den Punkten erstellt. Da sich den logistischen Variablen trivialerweise keine genauen Abstände zuordnen lassen, sollen diese lediglich ein Indiz für die Beziehung der Merkmale untereinander sein. Korrelierenden Merkmalsausprägungen wird eine kürzere Distanz zugewiesen, so dass die Wahrscheinlichkeit steigt, diese nacheinander auszuwählen. Für die anderen Merkmalsausprägungen, welche nicht korrelieren, wird ein standardisierter Abstand festgelegt. Dieser einheitliche Abstand soll eine zufällige Auswahl der unkorrelierenden Faktoren ermöglichen. Sollte der nah gelegene korrelierende Faktor nicht ausgewählt werden, wird einer der unkorrelierenden Faktoren zufällig bestimmt. Da sie alle dieselbe Distanz zu dem letzten Punkt aufweisen, besitzen alle nichtkorrelierenden Merkmale dieselbe Wahrscheinlichkeit ausgewählt zu werden.

In Tabelle 4.7 ist ein Ausschnitt einer möglichen Distanztabelle für ein zweidimensionales System aufgeführt. Die Distanzen sind als euklidische Abstände eines kartesischen Koordinatensystems angegeben [ECK02]. Die Merkmale sind in ihrer kodierten Form angegeben, so kommt es zu einem Bruch der Zahlenfolge in der ersten Zeile. Die Tabelle gibt nur einen Ausschnitt der benötigten Distanzen wieder. Die Darstellung ist ideal für die hier gewollte Präsentation. Eine endgültige Distanztabelle ist im Anhang zu finden.

#### 4. Initialisierung des Algorithmus

In dem Tabellenausschnitt werden in der Vertikalen die Abstände der Punkte bzgl. der Lagerart aufgeführt. In der Horizontalen werden sowohl die Abstände der Punkte Merkmale bezüglich der Lagerung als auch die der Bevorratung angegeben.

Tabelle 4.7: Distanzzuordnung I

Nach Von	1	2	3	4	5	6	7	8	11	12	13	14	15	16	17	18
1	0	1	1	1	2	1	1	1	0,5	1	1	1	2	1	1	1
2	1	0	1	1	1	2	1	1	1	0,5	1	1	1	2	1	1
3	1	1	0	1	1	1	2	1	1	1	0,5	1	1	1	2	1
4	1	1	1	0	1	1	1	2	1	1	1	0,5	1	1	1	2
5	2	1	1	1	0	1	1	1	2	1	1	1	0,5	1	1	1
6	1	2	1	1	1	0	1	1	1	2	1	1	1	0,5	1	1
7	1	1	2	1	1	1	0	1	1	1	2	1	1	1	0,5	1
8	1	1	1	2	1	1	1	0	1	1	1	2	1	1	1	0,5

Da in der Tabelle nur wenige Merkmale aufgeführt werden können, wird auch nur eine Auswahl an Korrelationen dargestellt. In der Tabelle werden Distanzen, welche von einem Punkt ausgehen und auch wieder zu diesem führen, mit einer „0“ angegeben. Diese Zeilenwerte werden so egalisiert. Die Berücksichtigung der Distanz von einem Punkt zu demselben Punkt ist im Falle der Auswahl der Lagerart Unsinn. Die Verbindung der anderen Variablen mit sich selber ist jedoch möglich. Durch doppelte Nennung können diese Variablen entsprechend skaliert werden. Ihnen wird die selbe Distanz zugeordnet wie unkorrelierende Merkmalen. Merkmalskombinationen einer Variablen, die sich ausschließen wird eine erhöhte Distanz zugeordnet. Bspw. soll kein Chromosom entstehen mit dem Gen der Zentralisierung einer Ware in einem Distributionszentrum und einem anderen Gen der Dezentralisierung mit den gleichen Parametern. Die Auswahl gegensätzlicher Kombinationen ist noch unerwünschter als unkorrelierende Kombinationen. Als Ansatz wird die Distanz im Vergleich zu unkorrelierten Merkmalen vorläufig verdoppelt. Ob die Distanz ausreicht, wird erst in den nächsten Schritten sichtbar. Gegebenenfalls müssen die Entfernungen nachträglich angepasst werden. Durch die größere Distanz soll die Wahrscheinlichkeit erhöht werden, nähere Verbindungen auszuwählen. Die gegensätzlichen Beziehungen sind in der Tabelle 4.7 rot hervorgehoben. Grün hervorgehoben sind positive Korrelationen der Merkmalsausprägungen. Die Zentralisierung einer beliebigen Ware in ein Distributionszentrum korreliert mit der Erhöhung der Bestände der selben Ware im selben Distributionszentrum. Damit diese Verbindungen öfter eingegangen werden, wird ihre Distanz im Vergleich zu den unkorrelierten Abständen halbiert. Ihre Auswahl fördert das Erreichen einer möglichst distanzarmen Verkettung der einzelnen Punkte. Die Distanzmatrix ist nicht symmetrisch. Zwei Punkte können unterschiedliche Distanzen aufweisen, je nachdem von welchem Punkt die Verbindung betrachtet wird. So können reziproke und einseitige Korrelationen berücksichtigt werden.

Das Problem des Handlungsreisenden gilt dann als näherungsweise gelöst, wenn die kumulierte Streckendistanz durch alle Punkte möglichst gering ausfällt. Dafür werden alle existierenden Punkte

miteinander verbunden. Das vorliegende logistische Problem sucht jedoch nur eine begrenzte Anzahl an Merkmalen, die ausgewählt werden und die Variablenpositionen belegen. Diese Verbindungen sind auch verkettet, jedoch nur unter den ausgewählten Merkmalsausprägungen. Die nicht ausgewählten Merkmale werden für diesen Durchlauf nicht berücksichtigt. Entsprechend muss der Algorithmus die Strecke nach Auswahl des fünften Punktes abbrechen. So werden nur die gewünschte Anzahl an Merkmalsausprägungen als Punkte einbezogen.

Eine exakte Darstellung der Punkte mit ihren Entfernungen ist nicht möglich. Durch die standardisierte Entfernung nicht korrelierender Merkmale ist es nicht möglich, die Punkte mit den zugeordneten Distanzen darzustellen. Da die meisten Merkmale nicht miteinander korrelieren, weisen sie größtenteils eine einheitliche Distanz zueinander auf. In realen Systemen können nur wenige Punkte ein einheitliches System aus gleichen Abständen bilden.

Ein System aus drei Faktoren ist das größtmögliche System, dessen Punkte sich in einer Ebene so anordnen lassen, dass ihre Entfernung zueinander identisch ist. Das Ergänzen eines vierten Punktes macht die Anordnung in der Ebene unmöglich. Vier Punkte müssen bereits in einem dreidimensionalen Raum beschrieben werden. Systeme höherer Ordnung existieren nur theoretisch. Selbst ein dreidimensionaler Raum reicht nicht mehr aus, um fünf Punkte gleicher Distanz zu verbinden. Die Kodierung aller Merkmalsausprägungen übersteigt die Grenze von fünf Faktoren deutlich. Bereits im eindimensionalen Raum besitzen alle fünf Variablen zwei Ausprägungen, die untereinander verbunden werden müssen. Die Verbindung der Merkmale ist weder darstellbar noch vorstellbar. Das System existiert nur theoretisch.

Im Fall eines zweidimensionalen Systems müssen von jedem Punkt 39 Verbindungen ausgehen. Die 39 Verbindungen setzen sich aus den restlichen sieben Merkmalsausprägungen der eigenen Variable und den übrigen vier Variablen in beiden ihrer Ausprägungen zusammen. Von diesen 39 Verbindungen besitzen 36 Punkte die gleiche Entfernung zu dem betrachteten Punkt. Ein System mit derart vielen gleichlangen Distanzen kann real nicht dargestellt werden. Von den restlichen Punkten weisen zwei Punkte eine doppelt so lange Distanz auf, ihre Merkmale widersprechen sich. Der letzte Punkt liegt näher an dem betrachteten Punkt. Ihre Korrelation spiegelt sich in ihrer Distanz wieder. Nahe Verbindungen, welche aus korrelierenden Merkmalen bestehen sollen mit einer erhöhten Wahrscheinlichkeit eingegangen werden.

Abbildung 4.1 stellt einen Ansatz in Anlehnung an ein Problem des Handlungsreisenden, vereinfacht dar. Die Distanzen sind nicht relationsgetreu modelliert. Zusätzlich wird die Anzahl an Merkmalen zu Gunsten der Übersichtlichkeit reduziert. Neben einer möglichen Gesamtstrecke aus fünf Punkten werden einige ergänzende Punkte aufgeführt.

Der Startpunkt des Weges wird bei Faktor „Zwei“ initiiert. Ausgehend von „Zwei“ verläuft der Pfad zu Punkt „11“ und anschließend zu Punkt „21“. Der erste Weg ist länger als die zweite Verbindung. Die ersten beiden Variablen korrelieren nicht miteinander, daher ist ihre theoretische Entfernung voneinander größer. Die Punkte „11“ und „21“ liegen näher beieinander, ihre dekodierten Merkmale korrelieren miteinander. Der nächste Weg verbindet die korrelierenden Faktoren „21“ und „31“. Aufgrund ihrer Abhängigkeit voneinander liegen sie ebenfalls nah beieinander, ihre gemeinsame Auswahl ist dadurch wahrscheinlicher. Abschließend endet der Pfad bei Punkt „26“. Die letzte

#### 4. Initialisierung des Algorithmus

Verbindung wird erneut über eine größere Distanz eingegangen, ihre dekodierten Merkmale korrelieren nicht miteinander.

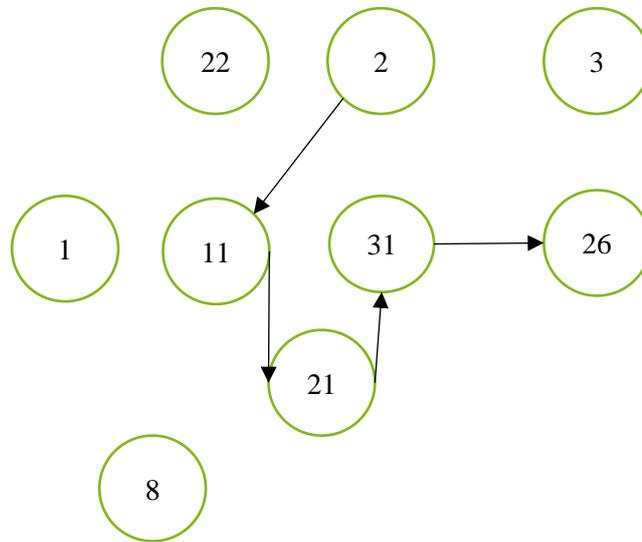


Abbildung 4.2: Ablaufschema TSP

Tabelle 4.8 stellt die kodierte Abfolge von Abbildung 4.1 dekodiert dar. In der dekodierten Schreibweise wird deutlich, dass die drei mittleren Punkte miteinander korrelieren. Sie beschreiben alle Aktionen, die dieselbe Ware und dasselbe Distributionszentrum beinhalten. Zusätzlich ist ihre Reihenfolge dem optimalen Ablauf entsprechend.

Tabelle 4.8: Dekodierung des Ablaufschemas

Aktionsliste				
1	2	3	4	5
Kodierung				
2	11	21	31	26
Dekodierung				
Zentralisierung W1 in DZ2	Erhöhen B1 in DZ1	Erhöhen SB1 in DZ1	Erhöhen TF: SB1 in DZ1	Verringern SB1 in DZ2

Durch die Aufteilung der einzelnen Ausprägungsmerkmale in Punkten wird das Problem der festen Positionen der Merkmale auf dem Chromosom gelöst. Durch den randomisierten Initialpunkt können diverse Kombinationen generiert werden. Die Position auf dem Chromosom ist nicht mehr von ihrem Bezug zur Lagerart abhängig. Beispielsweise könnte auf dem ersten Gen eines Chromosoms entweder eine Erhöhung der Transportfrequenz beschrieben werden oder auch eine Verringerung einer Bestandsmenge. Auch im Hinblick auf spätere Mutationen müssen Merkmale nicht länger ausschließlich durch andere Varianten ihrer Ausprägungen getauscht werden. Bei einer binären

#### 4. Initialisierung des Algorithmus

Kodierung wäre die Inversion eine der wenigen Mutationsarten, die zulässig wären, um die Gene eines Chromosoms zu ändern. Um die Position des Merkmals nicht zu ändern, könnte es nur zu seiner gegensätzlichen Ausprägung mutiert werden. Die variable Abfolge der Merkmale ermöglicht das Wirken weiterer Mutationsarten. Merkmale könnten ihre Positionen austauschen oder durch beliebige andere Variablenausprägungen ersetzt werden.

Der Unterschied eines Chromosoms mit festgelegten Positionen der Variablen und einer Anordnung über die Nächsten-Nachbar-Beziehung wird in Tabelle 4.9 dargestellt. Die erste Zeile beschreibt die ideale Zusammenstellung der Variablen. Diese Anordnung kann sowohl durch eine binäre Kodierung als auch durch eine Nächsten-Nachbar-Beziehung auftreten. Bei dem Binärcode können die einzelnen Variablen auch in ihrer alternativen Form auftreten, stehen diese stehen jedoch immer an ihrer jeweiligen Position. Die Kette ist der Lagerart nach ausgerichtet. Bei der Nächsten-Nachbar-Beziehung kann die erste Zeile auftreten, wenn die Strecke an dem Punkt „Zentralisierung: W1 in DZ1“ initiiert wird. Die nachfolgenden Punkte sind die jeweils sukzessiv nächsten Nachfolger und werden aufgenommen.

Die Punkte werden gemäß der zweiten Zeile angeordnet, wenn die Strecke mit dem Punkt „Erhöhung: SB2 in DZ1“ initiiert wird. Der optimale Weg bricht nach zwei Verbindungen ab. Der weitere Weg wird zufällig ausgewählt. Damit der Zufall wirkt, müssen die unkorrelierenden Variablen die gleiche Distanz aufweisen. Ab diesem Punkt gelten die Nächsten-Nachbar-Beziehungen und die von diesem Punkt ausgehend kürzeste Verbindung wird ausgewählt.

Tabelle 4.9: Merkmalsposition auf Chromosom

Merkmalspositionen auf Chromosom				
1	2	3	4	5
Zentralisierung: W1 in DZ1	Erhöhung: W1 in DZ1	Erhöhung: SB1 in DZ1	Erhöhung: T.Frequenz W1 von DZ1	Erhöhung: T.Volumen: W1 von DZ1
Erhöhung: SB2 in DZ1	Erhöhung T.Frequenz W2 von DZ1	Erhöhung T.Volumen: W2 von DZ1	Zentralisierung: W1 in DZ1	Erhöhung: W1 in DZ1

Ein an das Problem des Handlungsreisenden angelehntes Vorgehen führt zu einem anderen Hindernis. Bei dem ursprünglichen Problem soll die kumulativ am besten geeignete Strecke gefunden werden. Dies kann über die Fahrtkosten oder die Streckendistanzen geschehen. Daher werden nähergelegene Verbindungen mit einer höheren Wahrscheinlichkeit eingegangen. Bei dem Problem-des-Handlungsreisenden ist dies gewünscht, da alle Punkte besucht werden müssen. In der beschreibenden Abwandlung werden pro Strecke nur eine begrenzte Anzahl an Punkten besucht. Da nicht die kürzeste Gesamtstrecke gesucht wird, können kürzere Verbindungen nicht durch verbliebene entferntere Strecken ausgeglichen werden. Durch die reduzierte Anzahl an Punkten würden immer die kürzesten Verbindungen gesucht werden. Die kürzeren Verbindungen beschreiben zwar eine stärkere Beziehung der Punkte, zufällige Verbindungen werden aber nicht eingegangen.

Aus dem Ansatz, der sich an das Problem des Handlungsreisenden anlehnt, lassen sich weitere Erkenntnisse gewinnen. Die Kodierung aller Merkmalsausprägungen in einzelne Punkte ist deutlich zielführender als die zuvor überlegte binäre Kodierung. Merkmale können in dem letzten Ansatz an allen Positionen des Chromosoms stehen und alle möglichen Waren und Distributionszentren betreffen. Trotz der Annäherung an die Problemlösung reichen die herausgearbeiteten Überlegungen noch nicht aus, das Problem gänzlich zu lösen. Folgende Schwierigkeiten müssen noch überwunden und berücksichtigt werden:

- Das Vorkommen gegensätzlicher Merkmalsausprägungen in einem Chromosom muss verhindert werden.
- Der Zufall muss einen stärkeren Einfluss auf die Merkmale haben. Gegenwärtig werden näher gelegene Punkte zu stark priorisiert.

Gegenteilige Merkmalsausprägungen können nach wie vor gemeinsam innerhalb eines Chromosoms auftreten. In Tabelle 4.7 wurden diesen widersprüchlichen Merkmalsverbindungen höhere Distanzen zugeordnet. Dadurch sollten andere Verbindungen zwischen Punkten, die sich nicht ausschließen, wahrscheinlicher ausgewählt werden. Chromosomen mit Merkmalen, die sich ausschließen, wird ein schlechterer Fitnesswert zugeordnet, so dass diese mit einer geringeren Wahrscheinlichkeit zur Fortpflanzung beitragen. Die Möglichkeit der Auswahl widersprüchlicher Merkmale innerhalb eines Chromosoms besteht aber weiterhin, wenn diese auch mit verminderter Wahrscheinlichkeit erzeugt werden bzw. sich seltener vermehren. Im Folgenden wird eine alternative Distanzfestlegung für gegensätzliche Merkmale erörtert. Bei diesem Verfahren werden gegensätzliche Merkmale in einem Chromosom gänzlich unterbunden. In der Planung der anschließenden Schritte des EAs wird die besser geeignete Variante herausgearbeitet.

### 4.2.3 Anlehnung an das “TSP with Forbidden Neighborhoods”

Es existieren Sonderformen des Problems des Handlungsreisenden, die mit unerwünschten Strecken arbeiten können. Eines dieser Verfahren ist das Traveling Salesperson Problem with Forbidden Neighborhoods (TSPFN) [FIS17]. Bei dieser Abwandlung des klassischen Distanzproblems wird die Verbindung eines Punktes zu seinen nächstgelegenen Nachbarn verhindert. TSPFN werden in dreidimensionalen und ebenen Anwendungsfällen verwendet. In dem zweidimensionalen Fall liegen die Punkte in einem  $m \times n$ -Gitter.

Der Abstand kann dabei variabel gewählt werden, und für die zugehörige Problemstellung definiert werden. Für eine Rundreise durch alle Punkte  $(p_0, \dots, p_{mn-1})$  muss  $i \in \{0, \dots, mn-1\}$  erfüllt sein:

$$\|p_i - p_{i+1}\| > r \quad (16)$$

Der Betrag des euklidischen Abstands zwischen einem Punkt und einem weiteren Punkt muss größer als der festgelegte Abstand  $r$  sein [FIS17].

#### 4. Initialisierung des Algorithmus

In Abbildung 4.2 wird der Einfluss verschiedener Abstände auf ihre Verbindungsauswahl grafisch dargestellt. Die vier Quadrate stellen ein Raster möglicher Verbindungen dar. Dargestellt wird jeweils eine Anordnung in der Ebene. In der Mitte der vier Quadrate steht der jeweilige Punkt, zu dem die Abstände betrachtet werden. Die grün markierten Felder stehen für die weiter entfernte Nachbarschaft, deren Punkte über dem festgelegten Abstand  $r$  stehen. Die rotgefärbten Felder stehen für die nächstgelegenen Punkte. Diese Felder werden von der jeweiligen Distanz beeinflusst. Je höher die festgelegte Distanz, umso größer die mindest notwendige Entfernung zu einem möglichen Punkt.

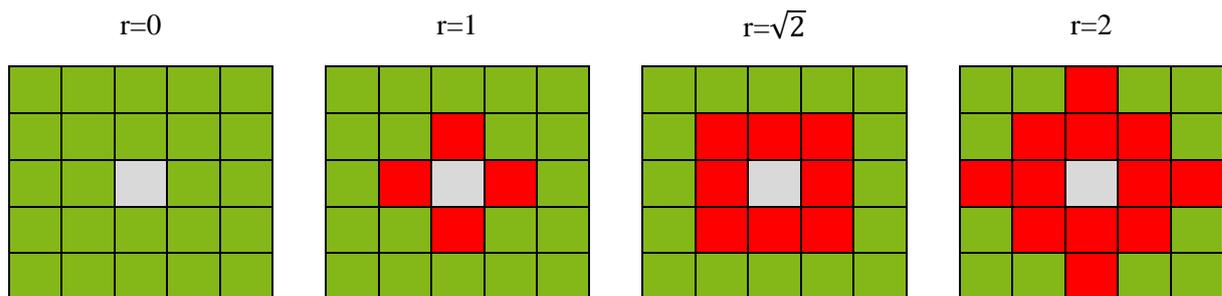


Abbildung 4.3: Abstände von einem Punkt [FIS17]

Bei Betrachtung der Visualisierung des Einflusses verschiedener Abstände wird deutlich, dass sich Komplikationen unerwünschter Verbindungen am leichtesten umgehen lassen, wenn störende Verbindungen konsequent verhindert werden. Die Abstände der gegensätzlichen Kombinationen werden auf Null limitiert. Im Gegensatz zu dem vorherigen Ansatz, sich ausschließende Kombinationen zu verhindern, in dem ihre Distanz zueinander erhöht wird, wird deren Distanz bei TSPFN reduziert. Die unerwünschten Verbindungen erhalten die gleiche Distanz wie ein Punkt zu sich selbst. Da beide Verbindungen irregulär sind, werden diese gar nicht erst eingegangen. So wird ihre Distanz auf Null festgelegt. Durch einen Abstand von Null wird vermieden, dass ihre Verbindung doch in das Chromosom mit aufgenommen werden.

Tabelle 4.10 stellt die aktualisierten Distanzen der Entfernungsmatrix dar. Die Entfernungen unterscheiden sich minimal von den ursprünglichen Distanzen aus Tabelle 4.7. Die Änderungen beziehen sich auf die rot markierten Diagonalen, die gegensätzlichen Beziehungen beschreiben, wie gleichzeitige Zentralisierung und Dezentralisierung. An dieser Stelle wird erneut nur ein Ausschnitt der Matrix angegeben. Die vollständige Distanzmatrix eines zweidimensionalen Systems ist im Anhang I hinterlegt. Die Distanzen gegensätzlicher Merkmalskombinationen werden offengehalten, um auch für die alternative Distanzmatrix ihre Gültigkeit zu bewahren. Die entsprechenden Felder sind rot hervorgehoben. Bei einem höher dimensionalisierten System wird jede Zeile mit weiteren unkorrelierenden Verbindungen ergänzt. Die Anzahl korrelierender Faktoren in einer Zeile bleibt identisch, daher ist die Matrix als Beispiel auch für komplexere Systeme verwendbar.

#### 4. Initialisierung des Algorithmus

Tabelle 4.10: Distanzordnung II

Von \ Nach	1	2	3	4	5	6	7	8	11	12	13	14	15	16	17	18
1	0	1	1	1	0	1	1	1	0,5	1	1	1	0	1	1	1
2	1	0	1	1	1	0	1	1	1	0,5	1	1	1	0	1	1
3	1	1	0	1	1	1	0	1	1	1	0,5	1	1	1	0	1
4	1	1	1	0	1	1	1	0	1	1	1	0,5	1	1	1	0
5	0	1	1	1	0	1	1	1	0	1	1	1	0,5	1	1	1
6	1	0	1	1	1	0	1	1	1	0	1	1	1	0,5	1	1
7	1	1	0	1	1	1	0	1	1	1	0	1	1	1	0,5	1
8	1	1	1	0	1	1	1	0	1	1	1	0	1	1	1	0,5

Die Überlegung, alle Distanzen zu verdoppeln, wird verworfen. Die Distanzen können auch Kommazahlen beinhalten. Durch die Einschränkung, Verbindungen der Länge „0“ zu verhindern, werden auch reguläre Distanzen kleiner als 1 nicht beeinflusst. Verbindungen unkorrelierender Merkmale der Distanz 1 und Distanzen von korrelierenden Merkmalen der Distanz 0 können weiterhin eingegangen werden.

Trotz Anpassung der Distanzmatrix bleibt ein Problem bestehen:

- Der Zufall muss einen stärkeren Einfluss auf die Merkmale haben. Gegenwärtig werden näher gelegene Punkte zu stark priorisiert. Von der idealen Transportkette abweichende Strecken müssen ermöglicht werden.

Durch die Ausrichtung an das Traveling Salesperson Problem with Forbidden Neighborhoods können zum einen alle Merkmalsausprägungen an jeder Stelle in einem Chromosom stehen, zum anderen werden gegensätzliche Merkmale in einem Chromosom verhindert. Welche Distanzen den gegensätzlichen Merkmalen zugeordnet werden, wird in den Planungen der nachfolgenden Schritte diskutiert.

Problematisch bleibt die Möglichkeit der zufälligen Auswahl weiterer Merkmale nach Auswahl des ersten Punktes. Gegenwärtig wird zufällig ein Punkt gewählt, von diesem werden die weiteren Punkte besucht. Da die Bedingung, dass alle Punkte einmal besucht werden müssen, nicht existiert, fehlt ein randomisierender Faktor. In Kombination zu der Bedingung, dass nur eine begrenzte Anzahl an Punkten besucht wird, werden nach Auswahl eines zufälligen Merkmals hauptsächlich korrelierende Faktoren ausgewählt, da ihre kodierten Punkte näher beieinander liegen. Die Verbindung der korrelierenden Faktoren wird zwangsläufig die geringste Gesamtstrecke ergeben. Größere Distanzen werden nur eingegangen, wenn eine Kette aus korrelierenden Merkmalen beendet ist und noch nicht alle Gene des Chromosoms gefüllt sind.

Unter der Prämisse, dass die Orientierung an einem Problem des Handlungsreisenden den bisher aussichtsreichsten Ansatz zur Problemlösung darstellt, wird eine weitere Veränderung das Distanzproblem weiter anpassen. Überlegungen, das Verfahren zu ändern, in bspw. einen

schwarmbasierten Ameisenalgorithmus, wird verworfen, da deren Auswahl der Merkmale ebenfalls nicht zufällig erfolgt.

Die zufällige Auswahl von Merkmalen ist eine Abweichung der kombinierten Merkmale bzw. ein Fehler der Gesamtstrecke des Handlungsreisenden. Da diese Abweichung der Sollzusammensetzung die besuchten Punkte betrifft, muss bereits deren Auswahl angepasst werden. Bezüglich der Auswahl eines Punktes muss dem Code eine Fehlerhäufigkeit hinzugefügt werden. Zu einer bestimmten Wahrscheinlichkeit soll nicht der optimale Punkt ausgewählt werden, sondern ein zufälliger Punkt der Gesamtmenge an Merkmalen. Von dem zufällig ausgewählten Punkt sollen dann wieder die korrelierenden Merkmale ausgewählt werden.

## 5. Fortpflanzung der Eltern

### 5.1 Bestimmung der Fitnesswerte

Nach der Generierung der Elternchromosome muss festgelegt werden, wie aus den Chromosomen der Eltern die Generierung der Individuen der Kindgeneration erfolgt. Dazu muss zuerst festgelegt werden, welche der Elternchromosome sich fortpflanzen. Zuerst wird eine Elternpopulation generiert. Diese besteht aus mehreren verschiedenen Elternindividuen. Die Auswahl der miteinander zu kombinierenden Eltern erfolgt über die Zuordnung der Fitnesswerte. Eltern deren Chromosome Merkmale kürzerer Distanz aufweisen, besitzen einen höheren Fitnesswert als Individuen deren Merkmale nicht korrelieren und so weiter voneinander entfernt liegen. Aufgrund ihres höheren Fitnesswertes werden sie als mögliches Elternteil priorisiert.

Der Fitnesswert eines Individuums hängt alleine von der Gesamtstrecke der einzelnen Punkte eines Chromosoms ab [SAB12].

$$l = distance(c_n, c_1) + \sum_{i=1}^{n-1} distance(c_i, c_{i+1}) \quad (17)$$

Der erste Summand gilt für die Rückführung von dem letzten Punkt der Rundreise zurück zu dem Initialpunkt. Im gegebenen Fall wird auf diesen Operator verzichtet, da eine Rückführung nicht erwünscht ist. Der zweite Summand summiert die restlichen Distanzen aller gegebenen Punkte. In dem vorliegenden Problem verläuft eine Tour durch eine begrenzte Zahl an Punkten einer größeren Menge an möglichen Punkten. So gilt  $n-1$  gleich die Anzahl an zu besuchenden Punkten. In dem gegebenen Fall muss ein Chromosom mit fünf Genen gefüllt werden. So gilt  $n-1=5$ . Der Fitnesswert soll proportional zur Eignung eines Individuums sein. Umgekehrt soll sich der Fitnesswert bei Minimierungsproblemen antiproportional zu den entstehenden Kosten verhalten. Die Kosten sollen möglichst gering gehalten werden. Im Falle des Problems des Handlungsreisenden drücken sich die Kosten als Entfernung zwischen den Punkten aus. Um ein Chromosom bewerten zu können, muss die Inverse einer Gesamtstrecke berechnet werden [AHM10].

$$F(x) = \frac{1}{f(x)} \quad (18)$$

$$F(x) = \frac{1}{l} = \frac{1}{\sum_{i=1}^5 distance(c_i, c_{i+1})} \quad (19)$$

In der Tabelle 5.1 werden exemplarisch die Fitnesswerte von sieben verschiedenen Chromosomen berechnet. Die sieben Chromosome wurden so gewählt, dass sie möglichst verschiedene Touren repräsentieren. Diese Touren unterscheiden sich in ihrer Auswahl an Merkmalen und der Anzahl an Korrelationen zwischen diesen. Die Auswahl der verschiedenen Punkte führt zu unterschiedlich hohen Fitnesswerten. Je mehr korrelierende Faktoren, desto kürzer die Gesamtstrecke und umso höher der

Fitnesswert. Sechs der Touren sind allgemeingültige Touren und können unabhängig von dem Umgang mit unerwünschten Kombinationen auftreten. Die siebte Tour ist nur möglich, wenn unerwünschte Verbindungen durch höhere Distanzen vermieden werden und nicht durch nicht existierende Verbindungen.

Tabelle 5.1: Exemplarische Bestimmung von Fitnesswerten

Chromosom	Merkmalsposition auf Chromosom					Fitnesswert
	1	2	3	4	5	
1	1	11	21	31	41	0,5
2	2	22	32	3	13	0,4
3	12	22	25	35	45	0,4
4	2	12	22	31	41	0,4
5	4	21	3	13	23	0,333
6	1	12	23	34	35	0,25
Chromosom mit gegensätzlichen Merkmalen						
7	1	11	21	31	15	0,222

In dem aufgeführten Beispiel sind die Chromosomen bereits nach Höhe ihrer Fitnesswerte angeordnet. Chromosom 1 besteht aus einer optimalen Transportkette aus korrelierenden Merkmalen. Alle Gene des Chromosoms werden von Merkmalen belegt, die eine Entfernung von 0,5 aufweisen. Da dies der kleinstmögliche Abstand zweier Punkte ist, nimmt der Fitnesswert den größten Wert an, den ein Chromosom in dem behandelnden Fall besitzen kann. Abweichungen von einer Kette aus korrelierenden Merkmalen führt durch die höhere Distanz zwischen unkorrelierenden Merkmalen zu erhöhten Fitnesswerten. Die Chromosomen 2, 3 und 4 weisen alle einen identischen Fitnesswert von 0,4 auf. In jedem dieser Chromosomen weicht ein Punkt von einer optimalen Tour ab. Bei jedem der drei Chromosomen wird einmalig ein Umweg zu einem unkorrelierenden Punkt gemacht. Die erhöhte Entfernung wirkt sich negativ auf den Fitnesswert aus. In Chromosom 5 umfasst eine Tour zwei Verbindungen zwischen Punkten, die nicht korrelieren. Dadurch liegt der zugehörige Fitnesswert noch unter dem Wert der Chromosomen mit einer einzigen Abweichung der optimalen Lieferkette. Das vorletzte der exemplarisch betrachteten Chromosomen besteht nur aus unkorrelierenden Merkmalsausprägungen. Daher ist die Entfernung zwischen den Punkten jeweils maximal. Eine Tour in Reihenfolge der aufgeführten Punkte besteht aus der höchsten Gesamtstrecke.

Da noch keine Empfehlung über den Umgang mit Distanzen gegensätzlicher Merkmalskombinationen getroffen werden kann, wird Chromosom 7 exemplarisch für eine unerwünschte Genzusammenstellung aufgeführt. Solche Chromosome können auftreten, wenn versucht wird, gegensätzliche Merkmale durch erhöhte Distanzen zu vermeiden. Die kodierten Merkmale 31 und 15 schließen sich aus. Abgesehen von dieser unerwünschten Beziehung stellt Chromosom Sieben eine optimale Merkmalskette dar. Der Fitnesswert eines Chromosoms mit unerwünschten Verbindungen muss schlechter sein, als der Fitnesswert eines Chromosoms, dessen Merkmale sich neutral zueinander verhalten. Die Distanz von

einer unerwünschten Verbindung wird auf 3 festgelegt, so liegt der Fitnesswert des Chromosoms 7 trotz seiner sonst optimalen Zusammenstellung unter allen anderen Fitnesswerten.

Durch die hohe Gesamtstrecke ist der Fitnesswert des Chromosoms 6 niedriger als der Fitnesswert der anderen Chromosomen.

Anhand der Fitnesswerte werden Elternteile bestimmt, aus denen Nachfolgeindividuen erzeugt werden. Die Auswahl erfolgt nach Höhe des Fitnesswertes. Von den sechs Chromosomen besitzt Chromosom 1 den höchsten Wert. Dieses Individuum würde als eines der Elternteile ausgewählt werden. Das zweite Elternteil wird aus der Gruppe der Chromosomen mit dem nächsthöheren Fitnesswert ausgewählt. Die Auswahl betrifft die Chromosomen 2, 3 und 4. Da die drei Chromosomen denselben Fitnesswert aufweisen, erfolgt die Auswahl des übrigen Elternteils unter ihnen zufällig.

## 5.2 Rekombination der Elternchromosome

Für die Erzeugung von Nachkommen vermischen die ausgewählten Elternindividuen ihre Gene. Beide Chromosomen werden gegenübergestellt, per Crossover werden Merkmalssequenzen der Elternteile ausgetauscht. Die Nachkommen bilden sich aus den Sequenzen beider Elternchromosome. Die Chromosomen der Eltern bleiben auch nach Rekombination identisch in ihren vorhandenen Merkmalen und deren Reihenfolge. Nicht alle möglichen Rekombinationsverfahren eignen sich für alle Probleme. Im Folgenden wird erörtert, welches Verfahren sich für die gewählte Kodierung am besten eignet [RAA02].

Die verschiedenen Crossover-Verfahren werden nach Anzahl der ausgetauschten Merkmalssequenzen unterschieden. Das Ein-Punkt-Crossover-Verfahren vertauscht eine Sequenz ab einer zufälligen Position zwischen 1 und  $n-1$ , wobei  $n$  für die Gesamtlänge des Chromosoms steht. In dem vorliegenden Fall gilt  $n=5$  für fünf mögliche Merkmale. Ab der bestimmten Position bis zum Ende des Chromosoms werden die Gene ausgeschnitten und zur Generierung der Nachfolger mit dem entsprechenden Gegenpart des anderen Elternchromosoms ausgetauscht.

In Tabelle 5.2 wird eine Rekombination exemplarisch durchgeführt. Als Elternchromosome werden zwei Chromosome aus Tabelle 5.1 übernommen. Das erste Elternteil ist das Chromosom 1, welches den höchsten Fitnesswert der Elternpopulation besitzt. Als zweites Chromosom wird Chromosom 2 zufällig als eines von drei Chromosomen gewählt, dessen Fitnesswert am nächst höchsten ist.

Tabelle 5.2: Rekombination zweier Chromosome

Chromosom	Merkmalsposition auf Chromosom				
	1	2	3	4	5
Elter <sub>1</sub>	1	11	21	31	41
Elter <sub>2</sub>	2	22	32	3	13
Rekombination					
Kind <sub>1</sub>	1	11	21	3	13
Elter <sub>2</sub>	2	22	32	31	41

## 5. Fortpflanzung der Eltern

Das Crossover findet an der vierten Position der Elternchromosome statt. Die ersten drei Gene eines Elternteiles und eines Kindes sind jeweils identisch. Ab dem vierten Chromosom unterscheiden sich die Kinder von ihren Vorgängern. Kind 1 weist jetzt eine Sequenz des zweiten Elternteils auf und Kind 2 an gleicher Stelle die Sequenz von Elternteil 1.

Die exemplarisch gewählten Elternchromosome können problemlos Sequenzen per Crossover austauschen, da sich ihre Merkmale nicht überschneiden. Chromosom 1 und Chromosom 2 bestehen aus gänzlich verschiedenen Merkmalen, so dass sie beliebige Sequenzen austauschen können. Crossover wird problematisch, wenn Eltern mit gleichen Merkmalen an unterschiedlicher Position rekombiniert werden sollen. Tabelle 5.3 zeigt dieses Problem:

Tabelle 5.3: Irreguläre Rekombination

Chromosom	Merkmalsposition auf Chromosom				
	1	2	3	4	5
Elter <sub>1</sub>	21	11	1	31	41
Elter <sub>2</sub>	2	22	32	1	31
Rekombination					
Kind <sub>1</sub>	21	11	1	1	31
Elter <sub>2</sub>	2	22	32	31	41

Verboten ist das Auftreten der gleichen Merkmalsausprägung der Variable der Lagerart in beiden Elternchromosomen an unterschiedlichen Positionen. In zwei Bereichen werden die Elternpaare und die resultierenden Nachfahren dargestellt. Bei den Elternchromosomen ist das reguläre Chromosom grau hinterlegt, dessen unerwünschte Kombination in dem Kindchromosom wird rot markiert.

An vierter Stelle weist das Elternchromosom 2 das Merkmal der Kodierung 1 auf. Das gleiche Merkmal steht bei dem ersten Elternteil an Position 3. Ein Crossover vor der vierten Genposition führt zu einer Kombination beider Elternchromosomen. Die Gene an vierter und fünfter Stelle werden ausgetauscht. Das erste Chromosom der Kindgeneration besitzt das Merkmal 1 an zwei verschiedenen Positionen seines Chromosoms. Das doppelte Auftreten eines Merkmals wurde bei der Elterngeneration noch verhindert, durch Crossover könnte eine doppelte Kombination aber auftreten.

Aufgrund dieser Problematik sollte ein klassisches Rekombinationsverfahren vermieden werden. Bei der Auswahl der Alternative muss berücksichtigt werden, dass die Chromosomenstränge möglichst erhalten bleiben sollen. Da die Chromosomenbildung maßgeblich durch die Nähe der Verbindungen bestimmt wird, bestehen die Chromosomen zu Teilen aus korrelierenden Merkmalen. Die Sequenzen aus korrelierenden Merkmalen gilt es beizubehalten. Crossover-Verfahren mit Austausch zufälliger Gene sind nicht zielführend, sie unterbrechen bestehende Ketten. Aus demselben Grund eignen sich Rekombinationsverfahren mit einer hohen Anzahl an ausgetauschten Segmenten nicht. Je mehr Segmente ausgetauscht werden, desto wahrscheinlicher wird eine bestehende Abfolge aus korrelierenden Faktoren getrennt. Bestehende Sequenzen sollten nach Möglichkeit gemeinsam vererbt werden.

Als Crossover-Verfahren, bei dem die relativen Positionen der Gene möglichst bestehen bleibt, wird das Order Crossover OX-Verfahren empfohlen [PUL13]. Wie bei einem Zwei-Punkt-Verfahren werden die Chromosomen an zwei Stellen zerschnitten. Die separierten Segmente werden aber bei den Kindchromosomen nicht direkt ausgetauscht. Der Bereich zwischen den Schnitten bleibt auch bei je einem Individuum der Nachfahren identisch. Die Randbereiche werden mit den Merkmalen des anderen Elternteiles gefüllt. Zusätzlich werden nur Merkmale eingesetzt, welche nicht bereits in dem übernommenen Kernsegment stehen. Tabelle 5.4 stellt Order Crossover OX-Verfahren exemplarisch vor:

Tabelle 5.4: Order Crossover OX-Verfahren

Chromosom	Merkmalsposition auf Chromosom				
	1	2	3	4	5
Elter <sub>1</sub>	1	11	21	31	41
Elter <sub>2</sub>	11	32	42	41	1
Vorläufige Nachkommen					
Kind <sub>1</sub>	X	11	21	31	X
Elter <sub>2</sub>	X	32	42	41	X
Nachkommen					
Kind <sub>1</sub>	32	11	21	31	42
Elter <sub>2</sub>	1	32	42	41	11

Aus zwei beliebigen Chromosomen werden zwei Nachfahren generiert. Die Merkmalssequenzen zwischen den Positionen 2 bis 4 werden von je einem Elternteil übernommen. Die entsprechenden Sequenzen sind hervorgehoben. Für die Generierung des ersten Kindes werden von den Merkmalen des zweiten Elternteils jene Merkmale gestrichen, die bereits von dem ersten Elternteil übernommen worden sind. Bei dem ersten Kind ist dies die 11 auf der ersten Position des zweiten Elternchromosoms. Die restlichen Merkmale des anderen Elternteils, welche nicht bereits in der übernommenen Sequenz stehen, füllen die Bereiche außerhalb der Schnitte auf. Das Einsetzen erfolgt chronologisch. Niedrigere Positionen werden zuerst mit Merkmalen gefüllt, die nicht bereits übernommen wurden. Die Auffüllung erfolgt in Leserichtung.

Das Order Crossover OX-Verfahren fördert zwar nicht den Austausch von Merkmalen zu korrelierenden Transportketten, es erhält aber zumindest Teilsequenzen der Elternchromosomen. So können Korrelationen mit übernommen werden. Problematisch ist nach wie vor die Möglichkeit, unerwünschte Merkmalskombinationen innerhalb der Kindchromosomen zu generieren. Das Crossover-Verfahren verhindert zwar ungültige Verbindungen in dem übernommenen Bereich, trotzdem können die ausgetauschten Chromosome unerwünschte Verbindungen herstellen. Die Elternchromosome verhindern dies noch, da sie jene Distanzen nicht eingehen oder ihr Fitnesswert zu schlecht ist, um seine Gene an Nachkommen weitergeben zu können. Die Möglichkeit der zu besuchenden Punkte, eine Mindestlänge zuzuweisen, funktioniert bei der Erzeugung der Nachkommen nicht. Die verschiedenen Crossingover-Verfahren sind zufallsbasiert, so dass auch ungültige Kombinationen entstehen können.

Das Crossover-Verfahren verhindert nur eine doppelte Vererbung des gleichen Merkmals innerhalb des gleichen Kindchromosoms. Andere irreguläre Kombinationen können nicht verhindert werden.

Aufgrund der Möglichkeit, Kinder mit unzulässigen Genkombinationen zu erzeugen, wird der Ansatz, einer Mindestentfernung der zu besuchenden Merkmale endgültig verworfen. Die Distanzen gegensätzlicher Merkmalskombinationen werden erhöht. Durch die Festlegung der Distanz auf 3 wird jedem Individuum mit einer irregulären Kombination ein schlechter Fitnesswert zugordnet. Diese Individuen können zwar erzeugt, aber nicht zur Fortpflanzung ausgewählt werden. So wird auch die Problematik bei der Vererbung unerwünschter Kombinationen behoben. Kinder mit derartigen Merkmalskombinationen können generiert werden, besitzen aber einen niedrigen Fitnesswert. Durch den niedrigen Fitnesswert werden die entsprechenden Kombinationen egalisiert.

Um zu vermeiden, dass Individuen mit unerwünschten Kombinationen in der Kindergeneration entstehen, muss die Selektionsart entsprechend angepasst werden. Die Nachkommen dürfen die Eltern nicht gänzlich verdrängen. Alle Individuen müssen mit ihrem jeweiligen Fitnesswert gegenseitig konkurrieren. Der kategorische Ausschluss aller Elternindividuen nach einer einmaligen Rekombination könnte theoretisch dazu führen, dass eine Kindpopulation generiert wird, die alle irregulären Kombinationen ihrer Merkmale aufweisen. Durch die erhöhte Auswahl an konkurrierenden Individuen ist die Wahrscheinlichkeit höher, Chromosomen mit gewünschten Genzusammenstellungen in der Population zu haben. Durch ihren höheren Fitnesswert werden deren zulässige Genkombinationen an die Nachfahren weitervererbt.

### 5.3 Mutationen

Durch die Wahl der Mutationsart werden die Gene der existierenden Chromosomen umsortiert und ersetzt. Für den bestehenden Anwendungsfall eignen sich nur Mutationen, welche Chromosomen austauschen. Bei den diversen positionsabhängigen Mutationsverfahren wird lediglich die Reihenfolge der Merkmale innerhalb eines Chromosoms vertauscht. Die Gesamtpopulation besteht in jeder Generation aus dem gleichen Genpool. Einzelne Gene werden weder entfernt noch hinzugefügt.

Klassische Traveling Salesperson Probleme verwenden konstante Genpools. Unabhängig der Distanzen zueinander, müssen für deren spezifische Problemstellungen alle Punkte besucht werden. Da von vornherein alle Punkte als Gene in einem Chromosom vorliegen, müssen Gene nicht mehr ergänzt oder hinzugefügt werden. Der zuvor beschriebene Ansatz lehnt aber nur an ein TSP an. Bei der bestehenden Problemstellung müssen nicht alle Punkte besucht werden, sondern nur eine begrenzte Menge. Da der Genpool viel mehr Gene aufweist als ein mögliches Chromosom, muss ein Mutationsverfahren gewählt werden, welches einzelne Gene austauscht. So gelangen neue Gene in eine bestehende Population.

Einfachstes Mutationsverfahren für das Ersetzen von Genen durch andere Gene ist eine Ein-Punkt-Mutation. Bei dieser Mutationsart wird an einer zufälligen Position ein Gen ausgetauscht und durch ein anderes Gen aus dem Genpool ersetzt. Eine entsprechende Mutation wird in Tabelle 5.5 dargestellt:

Tabelle 5.5: Vorgang der Punktmutation

Punktmutation					
Chromosom	1	2	3	4	5
Vor Mutation	1	11	21	2	41
Nach Mutation	1	11	21	31	41

Das Gen an der vierten Position des Chromosoms wurde zufällig ausgewählt. Es wird durch die Punktmutation durch ein anderes Gen ersetzt. Das ersetzte Gen steht an gleicher Stelle, wie das zuvor ausgetauschte Gen.

In dem Beispiel bildet das Chromosom durch das ausgetauschte Gen an vierter Position eine Aktionskette aus korrelierenden Merkmalen. Idealerweise setzt eine Mutation an einer Position an, welche eine korrelierende Kette unterbricht und tauscht das entsprechende Gen durch ein korrelierendes Merkmal. In der Praxis existieren solche Mutationsverfahren nicht bzw. deren Planung wäre zu kompliziert. Etwaige Mutationen müssten ansetzen, nachdem eine kürzere Distanz zwischen zwei Punkten ausgewählt wurde. Einer der Punkte müsste dann durch einen Punkt mit einer kürzeren Distanz ausgetauscht werden. Da Mutationen nicht gezielt eingesetzt werden können, müssen diese zufällig ablaufen. Es muss möglich sein, jedes Gen an jeder Position durch ein beliebiges Element des Genpools per Mutation des Genpools auszutauschen. Entsprechend der ungeplanten Mutation können irreguläre Kombinationen auftreten. Durch eine Punktmutation kann ein Gen durch ein gegensätzliches Gen einer anderen Position ersetzt werden. Während bei der Initialisierung der Elternindividuen auf irreguläre Genkombinationen Rücksicht genommen werden kann, können unerwünschte Kombinationen nicht verhindert werden. Da diese gegensätzlichen Merkmalskombinationen nicht verhindert werden können, muss verhindert werden, dass diese Individuen ihre Gene an nachfolgende Generationen weitergeben. Durch die Distanzzuweisungen der Merkmale werden auch den irregulären Verbindungen der Nachfolgenerationen erweiterte Distanzen zugeordnet. Die höheren Distanzen wirken sich negativ auf die Fitnesswerte der Individuen aus. Die Fitnesswerte bestimmen auch bei den Nachfolgenerationen die Eignung zur Fortpflanzung. Chromosome mit niedrigen Fitnesswerten werden ignoriert und die mit höheren Werten bevorzugt. Da die Chromosomen mit den niedrigeren Fitnesswerten Vergleichswerte benötigen, welche die Fitnesswerte übersteigen, müssen die Elternchromosomen weiter einbezogen werden.

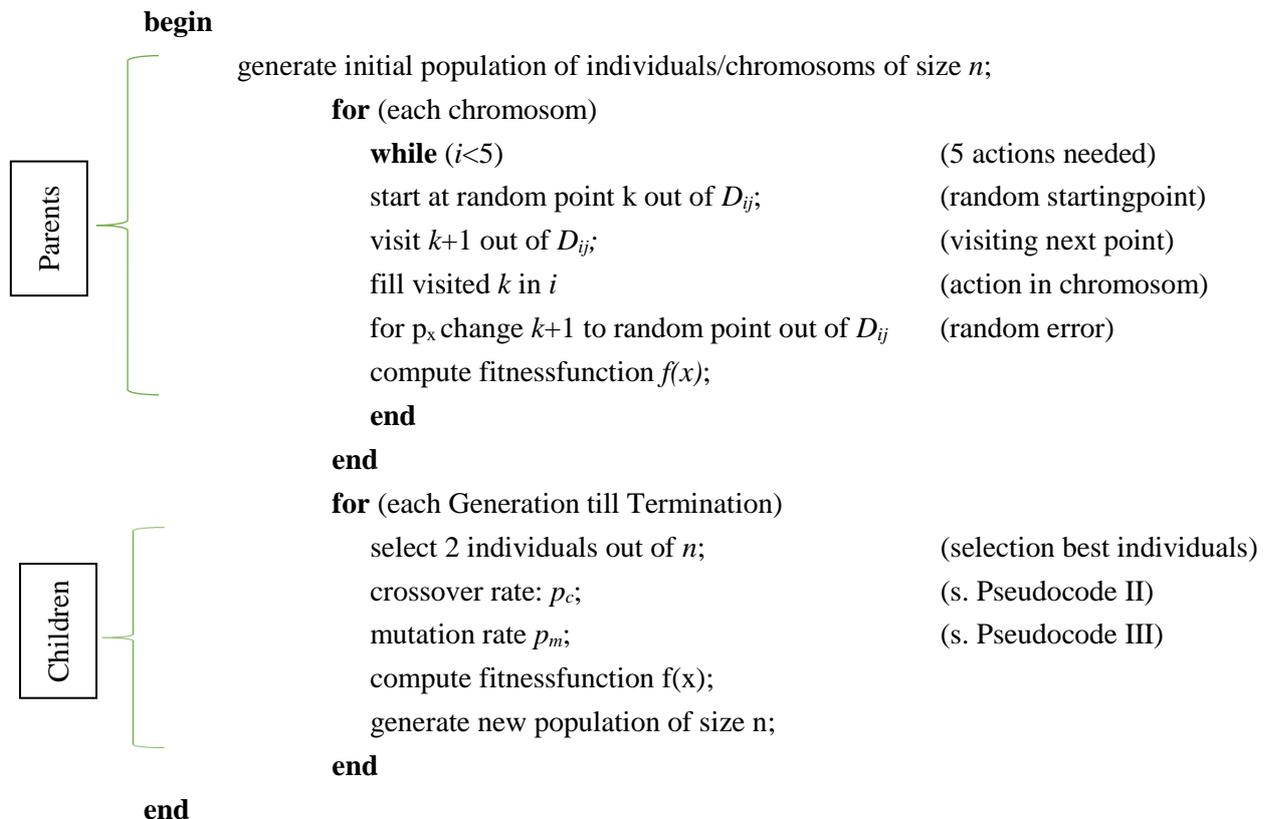
## 6. Programmdurchlauf

Dieses Kapitel behandelt die Implementierung des Algorithmus in Python und seinen Durchlauf. Der anschließende Vergleich der Ergebnisse, mit einem Algorithmus, welcher ohne eingespeiste Zusammenhänge arbeitet, liefert Aussagen über die Wirksamkeit der Implementierung der Vorarbeit. Die beiden Algorithmen werden auf Service Level und Kosten verglichen. Die genaue Berechnung dieser Faktoren ist unerheblich und kann als Black Box aufgefasst werden, welche von einem Fremdprogramm gesteuert wird. Für den Faktor Service Level, sprich das Maß für die Verfügbarkeit eines Produktes bei Nachfrage, wird ein hoher Wert angestrebt. Die Kosten sollen im Gegenzug gering ausfallen.

Auf die detaillierte Beschreibung des Gesamtcodes wird an dieser Stelle verzichtet, da die Programmierung des Codes nicht Ziel dieser Thesis war. Anhand der Pseudocodes und ergänzenden Beschreibungen einiger Codesegmente bzw. Bausteine wird das Vorgehen der Implementierung erörtert. Als Nebeneffekt werden die zuvor erarbeiteten Phasen des EA rekapituliert. Dies dient der Übersichtlichkeit und erleichtert das Nachvollziehen des Vorganges.

### Pseudocode I: Evolutionärer Algorithmus

**use** distancematrix  $D_{ij}$



$i$ = Gen/ Position eines Merkmals in Chromosom

$k$ =Punkt in Matrix

$d_{ij}$ = Matrix

### 6.1 Implementierung Distanzmatrix

Die Distanzmatrix  $D_{ij}$  kann wahlweise importiert oder in dem Code erstellt werden. In der endgültigen Variante des Algorithmus wird diese innerhalb des Codes generiert. Durch die selbstständige Generierung spannen die Variablen die Matrix selbstständig auf. Weitere Variablen können auch nachträglich ergänzt werden, welche die Matrix selbstständig erweitern. Importierte Matrizen greifen auf die gegebenen Informationen zurück. Ihre Werte sind vor Programmausführung gegeben, daher ist die Erweiterung von weiteren Variablen aufwendiger.

In der generierten Matrix wird davon ausgegangen, dass jede Variablenbeziehung die Entfernungsdistanz von 1 aufweist. Abweichende Entfernungen werden exkludierend beschrieben. Der Beziehung der Ausprägungen „Zentralisierung“ und „Dezentralisierung“ der Variable „Lagerart“ zu sich selbst wird direkt beschrieben. Die Distanz dieser Punkte zu sich selbst wird erhöht. Abweichende Beziehungen, welche die Einheitsmatrix stören, werden aufwendiger programmiert. Über Listen wird den einzelnen Variablen zugeordnet, bei welchen Variablen ihre Distanz von 1 abweicht. So werden korrelierende und gegenteilige Beziehungen beschrieben. Zur Reduzierung von Programmieraufwand werden unerwünschte Kombination, wie die Lagerart mit sich selbst oder gegensätzliche Merkmale die gleichen Distanzen zugeordnet. Diese Distanzen werden so hoch gesetzt, dass Individuen, die sie beinhalten, auf jeden Fall durch die Fitnessfunktion aus sortiert werden.

#### Pseudocode II: Distancematrix

**begin**

    build matrix with variables, all with same variable

    exclude variable storagetyp with itself

    implement list of relationships for variables which are related

**end**

### 6.2 Implementierung Crossing Over

Für das Crossing Over werden zwei Zahlen zwischen Eins und Fünf zufällig generiert. Die Zahlen entsprechen den Positionen der Liste eines Individuums. Werden bspw. die Zahlen Zwei und Vier ausgewählt, so wird die Liste mit den Merkmalen an diesen Stellen geschnitten. Kind<sub>1</sub> übernimmt die Merkmale innerhalb der Schnitte von Elter<sub>1</sub>, selbige Beziehung gilt für Kind<sub>2</sub> und Elter<sub>2</sub>. Die restlichen Positionen der Listen, der Kinder wird mit 0 aufgefüllt. Anschließend werden diese Positionen mit einer 0 mit den Merkmalen des gegenteiligen Elternteils aufgefüllt. Die Kinderlisten werden von links nach rechts aufgefüllt, bei Auftreten eines Merkmals, welches bereits nach den Schnitten kopiert worden ist, wird diese Position bei dem Elternteil übersprungen.

### Pseudocode III: Crossing Over

**begin**

choose 2 random positions out of  $i$

take the elements on the positions between them over to next generation

fill empty spots with elements of the other parent

**end**

## 6.3 Implementierung Mutation

Zu einer variabel einstellbaren Wahrscheinlichkeit, wird eine Position der Liste eines Individuums nach dem Crossing Over ausgewählt und durch ein zufälliges Element aus dem gesamten Suchraum aller Merkmale ausgetauscht.

### Pseudocode III: Mutation

**begin**

choose 1 random position out of  $i$  with  $p_m$

Fill position  $i$  with random point  $k$

**End**

## 6.4 Implementierung Abbruchsbedingung

Drei Terminierungsbedingungen bestimmen die Dauer eines Experiments. Nach einer frei festlegbaren Anzahl an Generationen, bricht das Experiment ab. Dies geschieht unabhängig davon, ob noch Veränderungen zu erwarten wären. Sollten keine Veränderungen zu erwarten sein, wenn die Änderungen stagnieren, bricht das Element ebenfalls ab. Für das dritte Kriterium wird der Versuch durch eine maximal vorgegebene Rechenzeit begrenzt. Jedes dieser Kriterien ist einstellbar.

## 7. Auswertung

Die Auswertung des EA erfolgt schrittweise. In einzelnen Versuchsreihen werden die Parameter sukzessiv angepasst, um Aussagen über das Verhalten des EA treffen zu können. Eine Versuchsreihe umfasst mehrere Versuche mit unterschiedlichem Seed-Wert, zur Generierung verschiedener Werte.

Für die ersten Versuche wird den Individuen eine größere Anzahl an Merkmalen zugeordnet, als in aufbauenden Untersuchungen. Die Versuche laufen über eine Mindestanzahl von 10 Generationen mit je 20 Individuen, die jeweils die Informationen von 10 Merkmalen besitzen. Verglichen wird der Einfluss von Korrelationen auf Kosten und Service Level. Neben den synergetischen und nachteiligen Beziehungen ist von dieser Unterscheidung auch das Crossover-Verfahren betroffen. Bei den Versuchen, mit Korrelationen wird das beschriebene Crossover-Verfahren mit zwei Schnitten verwendet. Bei den Untersuchungen, ohne Korrelationen, erfolgt die Rekombination über einen einzelnen Punkt, der die auszutauschenden Informationen markiert.

Tabelle 6.1 stellt die wichtigsten Resultate der Versuche tabellarisch dar. Die Kosten werden anhand ihrer maximalen, minimalen Werte, ihres Medians und ihrer durchschnittlichen Werten angegeben. Alle Versuche wurden bei einer Crossover-Wahrscheinlichkeit von 0,8 und einer Mutationsrate von 0,2 berechnet.

Tabelle 7.1: Auswertung von Versuchen mit Anfangsgröße 10

	Anfangsgröße der Individuen	Individuen je Generation	Min. Anzahl an Generationen	Crossover-typ	Kosten	Service Level
Ohne Korrelationen						
Max	10	20	10	einfach	88.449,16	85.25
Min	10	20	10	einfach	86.800,75	85.27
Med	10	20	10	einfach	87.770,78	85.32
Durchschnitt:					87.745,04	85.45
Mit Korrelationen						
Max	10	20	10	2-Punkt	90.180,95	85.17
Min	10	20	10	2-Punkt	88.107,15	85.45
Med	10	20	10	2-Punkt	89.205,54	85.12
Durchschnitt:					89.062,59	85.30

Die Kosten der Versuche ohne korrelierende Merkmale, sind geringer als unter Rücksichtnahme der Korrelationen. Diese Relation bestätigt sich für jede der Zeilen. Am Beispiel des Durchschnitts wird dies deutlich. Die durchschnittlichen Kosten bei Versuchen ohne Korrelation liegen mit 87745.04 KE unter dem Wert der korrelierenden Versuche von 89062.59 KE. Parallel übersteigt das durchschnittliche Service Level von 85.45, der unkorrelierenden Versuche das der korrelierenden Versuche von 85.30.

## 7. Auswertung

Beide relevanten Zielkategorien sprechen gegen die Berücksichtigung von Korrelationen. Sowohl die Kosten, als auch das Service Level favorisieren unkorrelierende Merkmale der Individuen.

In den aufbauenden Versuchen muss getestet werden, warum sich korrelierende Merkmale nachteilig auswirken. In den folgenden Untersuchungen wird die Anzahl der Merkmale pro Individuum auf Fünf reduziert. Die restlichen populationsbasierten Parameter, wie Wahrscheinlichkeiten, Mindestanzahl an Generationen und die Individuenzahl je Generation, werden beibehalten. Ergänzend werden die Abbruchsursachen jedes Versuches untersucht. Tabelle 6.2 stellt die Versuchsreihen bei reduzierter Merkmalszahl tabellarisch dar. Die angegebenen Daten werden ausgehend der entstehenden Kosteneinheiten generiert.

Tabelle 7.2: Auswertung von Versuchen mit Anfangsgröße 5

	Anfangsgröße der Individuen	Min. Generationen	Kosten	Service Level	Abbruchsart	Generation
Ohne Korrelationen						
Max	5	10	91.767,61	85.14	Max. Generationen	101
Min	5	10	89.010,31	84.96	Max. Generationen	101
Durchschnitt:			90.710,21	85.05	-	-
Mit Korrelationen						
Max	5	10	93.598,33	84.63	Stagnation	11
Min	5	10	91.045,88	84.78	Stagnation	21
Durchschnitt:			92.674,92	84.76	-	-

Die zweite Versuchsreihe bestätigt die Resultate der vorherigen Versuche. Auch bei reduzierten Merkmalsanzahl pro Individuum übersteigen die Kosten der Versuche der korrelierenden Versuche, die Kosten der unkorrelierenden Versuche. Trotz erhöhter Kosten ist der Wert des Service Levels ebenfalls geringer, als in den Versuchen ohne Korrelationen. Die tabellarische Anordnung nach dem Service Level liefert vergleichbare Rückschlüsse und wird daher vernachlässigt. Neben den beiden Zielparametern wurden die Abbruchsbedingungen untersucht. Bei den beiden aufgeführten Versuchen mit korrelierenden Merkmalen bricht der Versuch nach Erreichen der maximalen Generationenzahl von 101 ab. Die beiden ausgewählten Versuche mit korrelierenden Merkmalen brechen durch stagnierende Werte ab. Die restlichen Versuche, außer den beiden Extrema, werden ebenfalls durch Stagnation abgebrochen. Die weiteren Untersuchungen außerhalb der beiden Peaks der unkorrelierenden Merkmale ergeben, dass die Versuche zu gleichen Teilen durch Stagnation und durch Erreichen der maximalen Generationsanzahl abgebrochen werden.

Der Vorteil eines Abbruchs durch Stagnation liegt in der reduzierten Generationsanzahl. Sobald sich die Werte nur noch minimal verändern, wird der betreffende Versuch abgebrochen. Die restlichen Versuche bis zu dem Erreichen der maximalen Generationszahl müssen nicht mehr berechnet werden und können

## 7. Auswertung

übersprungen werden. Je mehr Versuche durch Stagnation abgebrochen werden, desto stärker summiert sich die zeitliche Ersparnis. Die Versuche ohne Korrelationen werden teilweise gänzlich bis Erreichen der maximalen Generationszahl durchgeführt, daher benötigen die Versuche in ihrer Gesamtheit länger als die Versuche mit den unkorrelierenden Merkmalen. Dieses Gefälle wird auch bei den angegebenen Werten ersichtlich. Die Versuche mit korrelierenden Merkmalen benötigen bei maximalen und minimalen Kosten mit 11 bzw. 21 Generationen deutlich weniger Rechenzeit als Versuche die nach 100 Generationen zwangsweise abgebrochen werden.

In weiteren Untersuchungen wird der Einfluss der Crossover-Wahrscheinlichkeit auf die Zielparameter betrachtet. In drei Versuchsreihen mit Crossover-Wahrscheinlichkeiten von 0.8, 0.65, und 0.5 werden Versuche mit Mutations-Wahrscheinlichkeiten von 0.1 und 0.2 untersucht. Tabelle 6.3 stellt die Extrema der Kosten jedes Versuches dar. Das Optimum einer Versuchsreihe ist Grün hervorgehoben.

Tabelle 7.3: Versuche mit variierenden Mutations- und Crossover-Wahrscheinlichkeiten

	Mutations-Wahrscheinlichkeit	Abbruchart	Generation	Kosten	Service Level
Crossover-Wahrscheinlichkeit: 0,8					
Max	0,2	Stagnation	11	93.598,33	84.63
Min	0,2	Stagnation	21	91.045,88	84.78
Durchschnitt				92.674,92	84.76
Max	0,1	Stagnation	15	94.109,02	84.63
Min	0,1	Stagnation	18	92.358,6	84.7
Durchschnitt:				93.322,42	84.76
Gesamt-Durchschnitt				92.998,67	84.76
Crossover-Wahrscheinlichkeit: 0,65					
Max	0,2	Stagnation	30	93.615,95	85.04
Min	0,2	Stagnation	68	91.759,74	84.81
Durchschnitt				92.664,02	84.79
Max	0,1	Stagnation	11	94.599,95	84.63
Min	0,1	Stagnation	37	92.318,91	84.7
Durchschnitt:				93.486,7	84.72
Gesamt-Durchschnitt				93.075,36	84.75
Crossover-Wahrscheinlichkeit: 0,5					
Max	0,2	Stagnation	17	93.749,47	84.63
Min	0,2	Stagnation	25	92.461,59	84.7
Durchschnitt				93.211,4	84.7
Max	0,1	Stagnation	12	94.205	84.63
Min	0,1	Stagnation	20	92.683,63	84.96
Durchschnitt:				93.472,83	84.7
Gesamt-Durchschnitt				93.342,11	84.72

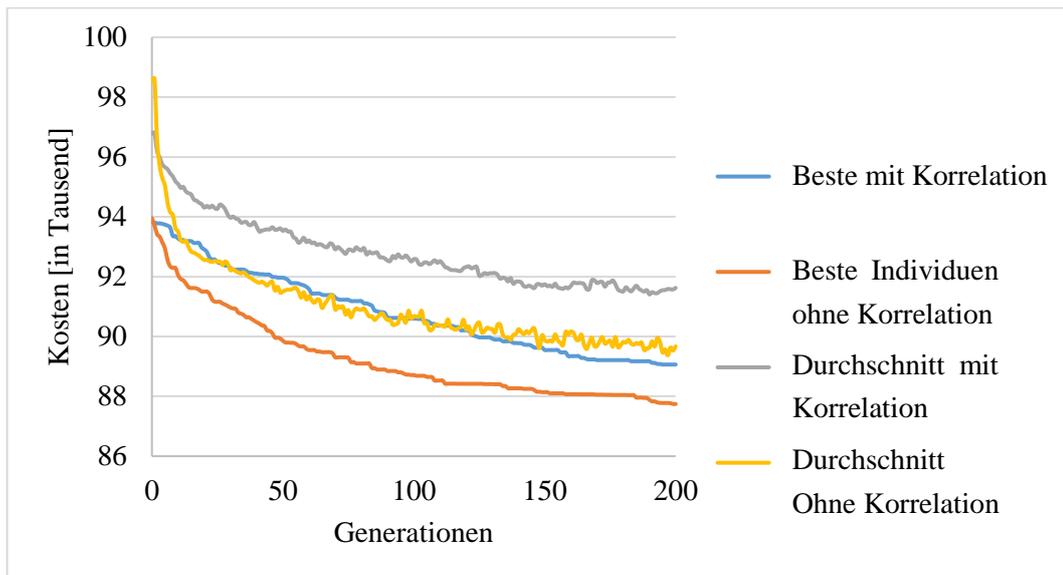
Aus der Tabelle lassen sich mehrere Ergebnisse ablesen. Mit steigender Crossover-Wahrscheinlichkeit sinken die durchschnittlichen Kosten leicht auf ein Minimum von 92.664,02 KE, bei einer Crossover-Wahrscheinlichkeit von 0,65 und einer Mutations-Wahrscheinlichkeit von 0,2. Gleichzeitig bewirkt die mittlere Crossover-Rate die höchsten Kosten von 93.486,7 KE, bei einer reduzierten Mutations-Wahrscheinlichkeit von 0,1. Bei den beiden anderen Versuchsdurchgängen wirkt sich eine erhöhte Mutationsaffinität ebenfalls positiv auf die Kosten aus.

Da die beiden Extrema der Kosten in demselben Versuchsdurchlauf generiert wurden, spannt die mittlere Crossover-Rate den größten Raum auf, in dem Kosten liegen. Die durchschnittlichen Kosten der mittleren Crossover-Wahrscheinlichkeit liegen zwischen den beiden anderen Versuchsdurchläufen. Die durchschnittlichen Kosten sinken mit zunehmenden Crossover-Raten. Die grün unterlegten Werte bestätigen dies. Für jeden Versuchsdurchlauf werden die geringsten Kosten bei der maximalen Mutationsrate von 0,2 erreicht.

Den Einfluss von Mutation und Crossover auf das Service Level festzustellen, ist durch die geringen Änderungen ungenauer. Der Zielparameter verändert sich nur minimal, es steigt von 84.72 auf maximal 84.76. Die Angaben von Maxima und Minima sind für das Service Level zu vernachlässigen, da die Spalten ausgehend der Werte der Kosten erstellt wurden. Dadurch sind die Spalten der optimalen Werte beider Zielparameter nicht immer identisch. Trotz fehlender Extrema, erlaubt die Tabelle generelle Aussagen über den Einfluss von Mutation und Crossover auf das Service Level. Das durchschnittliche Service Level steigt mit zunehmender Crossover-Wahrscheinlichkeit. Bei konstanter Crossover-Rate stagniert oder steigt das Service Level bei erhöhter Mutationsrate.

Erhöhte Mutations-Wahrscheinlichkeiten und Crossover-Raten führen zwar zu verbesserten Zielparametern bei korrelierenden Merkmalen, die durchschnittlichen Werte von Kosten und Service Level sind jedoch schlechter als bei den unkorrelierenden Versuchen. Das Diagramm in Abbildung 6.1 zeigt die Höhe der Kosten bei zunehmender Generationszahl für einen beliebigen Versuch. Durch die Darstellung verschiedener Linien soll visualisiert werden, wie sich die Kosten von Versuchen mit und ohne Korrelation entwickeln.

Tabelle 7.1: Verlauf der Kosten über mehrere Generationen



Das Diagramm bestätigt die durchschnittlich höheren Kosten bei Versuchen mit korrelierenden Merkmalen. Die Kosten übersteigen fast in jeder Generation die der unkorrelierenden Versuche. Lediglich in der initialen Generation fallen die Kosten geringer aus. Die Generierung von Individuen per Korrelationen eignet sich für die Initialisierung. Die zyklischen Mutations- und Rekombinationsvorgänge vermindern den Nutzen des EA. Der Graph der Versuche ohne Korrelation flacht viel stärker ab, als der Graph der korrelierenden Merkmale. Zu den selben Rückschlüssen kommt man bei Betrachtung der besten Individuen. Der anfängliche Kostenvorteil der besten Individuen fällt zwar geringer aus, trotzdem ist die Berücksichtigung von Korrelationen förderlich für die Generierung der Initialpopulation. Mit zunehmender Generationsanzeige flachen beide Graphen ab, der Graph der

unkorrelierenden Versuche flacht stärker ab und liefert so bessere Ergebnisse. Die Betrachtung der Graphen für das Service Level liefert ebenfalls vergleichbare Resultate. Die Graphen des Service Levels liegen näher beieinander, daher sind sie unübersichtlicher. Auf ihre Darstellung wird aus Redundanz in weiteren Tabellen und Abbildungen verzichtet. Die Initialisierung wird durch berücksichtigte Korrelationen verbessert. Mit zunehmender Generationszahl wirken sich die derzeitigen Verfahren negativer auf die Ergebnisse aus.

Da der nachlassende Nutzen der Korrelation, auf die derzeitigen Vererbungsverfahren der Mutation und des Crossingovers zurückzuführen ist, müssen diese angepasst werden. Beide Verfahren sind stark zufallsabhängig und zerstören gewollte Verbindungen. Auch wenn der eigentliche Nutzen von Mutation und Rekombination das Erzeugen von Diversität ist, müssen diese Verfahren für das spezifische Problem eingeschränkt werden.

Das derzeitige Crossover-Verfahren berücksichtigt in geringen Maßen Abhängigkeiten, daher wird zunächst das gegenwärtige Mutationsschema verändert. Aktuell wird ein Merkmal an einer zufälligen Position eines Individuums ausgewählt und mit einem Merkmal aus dem gesamten Suchraum

ausgetauscht. Das neue Schema soll auf Beziehungen der Merkmale eingehen. Dafür ist es nötig, ein Mutationsverfahren abseits von den herkömmlichen Verfahren zu erstellen.

Entsprechend dem Grundgedanken soll Mutation immer noch dem Zufall unterliegen. Daher wird weiterhin ein Merkmal auf einer Position des Individuums zufällig ausgewählt. Von diesem Merkmal aus wird das vorige Merkmal betrachtet. Anhand dieses Merkmals wird ein Element der zugehörigen Spalte aus der Distanztabelle der Initialisierung ausgewählt und für das zufällig ausgewählte Merkmal ersetzt. Dies soll die Gesamtstrecke eines Chromosoms verringern und so das Fitnesslevel erhöhen. Dies soll wiederum zu einem besseren Individuum und zu einer erhöhten Weitervermehrung führen.

Das veränderte Mutationsverfahren muss auf seine Tauglichkeit überprüft werden. In Tabelle 6.4 werden die durchschnittlichen Kosten des neuen Mutationsverfahrens mit den Kosten aus untersuchten Versuchen verglichen. Für den Vergleich werden je zwei unterschiedliche Wahrscheinlichkeiten, mit bekannten Ergebnissen aus vorherigen Ergebnissen, herangezogen.

Tabelle 7.4: Einfluss von verändertem Mutationsverfahren

Mutations-Wahrscheinlichkeit	Durchschnittliche Kosten	
	Altes Verfahren	Neues Verfahren
Crossover-Wahrscheinlichkeit: 0,8		
0,1	93.322,42	93.013,26
0,2	92.674,92	92.649,3
Crossover-Wahrscheinlichkeit: 0,5		
0,1	93.472,83	93.498,433
0,2	93.211,4	92.673,477

Bei den meisten Versuchen verbessern sich die Kosten bei Anpassung des Mutationsschemas. Exemplarisch verbessern sich die Kosten von 93.322,42 KE auf 93.013,26 KE, bei einer Mutationsrate von 0,1 und einer Crossover-Wahrscheinlichkeit von 0,8. Unstimmigkeiten gibt es bei den Kosten für die Versuche, bei einer Mutationsrate von 0,1 mit einer Crossover-Wahrscheinlichkeit von 0,5. Die Kosten des neuen Mutationsverfahrens liegen mit 93.498,433 KE leicht über den Kosten von 93.472,83 KE des vorherigen Verfahrens. Die Höhe der Abweichungen lassen sich auf zufällige Schwankungen innerhalb der Versuche zurückführen. Die Verbesserungen der restlichen Versuche sind ausschlaggebender und favorisieren das neuere Mutationsverfahren.

Tabelle 7.5: Durchschnittliche Kosten bei veränderten Mutations-Wahrscheinlichkeiten

Mutations-Wahrscheinlichkeit	Durchschnittliche Kosten	
	Neues Verfahren	unkorreliert
Crossover-Wahrscheinlichkeit: 0,8		
0,1	93.013,26	90962,93
0,2	92.649,3	90.710,21

Trotz Verbesserung sind die durchschnittlichen Kosten bei Versuchen mit unkorrelierenden Merkmalen weiterhin niedriger. Für beide aufgeführten Mutations-Wahrscheinlichkeiten übersteigen die Kosten der korrelierten Versuche die Kosten der unkorrelierten Versuche. Entweder ist das neuere Verfahren ebenfalls unzureichend, oder das Zusammenwirken mit dem angewendeten Crossover-Verfahren ist nicht zielorientiert genug.

In den folgenden Abbildungen wird für je einen Versuch ein beliebiger Seed-Wert gewählt und der Verlauf der Kosten mit zunehmenden Generationen dargestellt. Es werden die Graphen für die durchschnittlichen Kosten eines Individuums und die Kosten des besten Individuums einer Generation dargestellt. Abbildung 6.2 wurde mit den Werten für eine Mutationsrate von 0,1 und einer Crossover-Rate von 0,8 erstellt. Der Einfluss einer Mutationsrate von 0,2 wird in Abbildung 6.3 dargestellt.

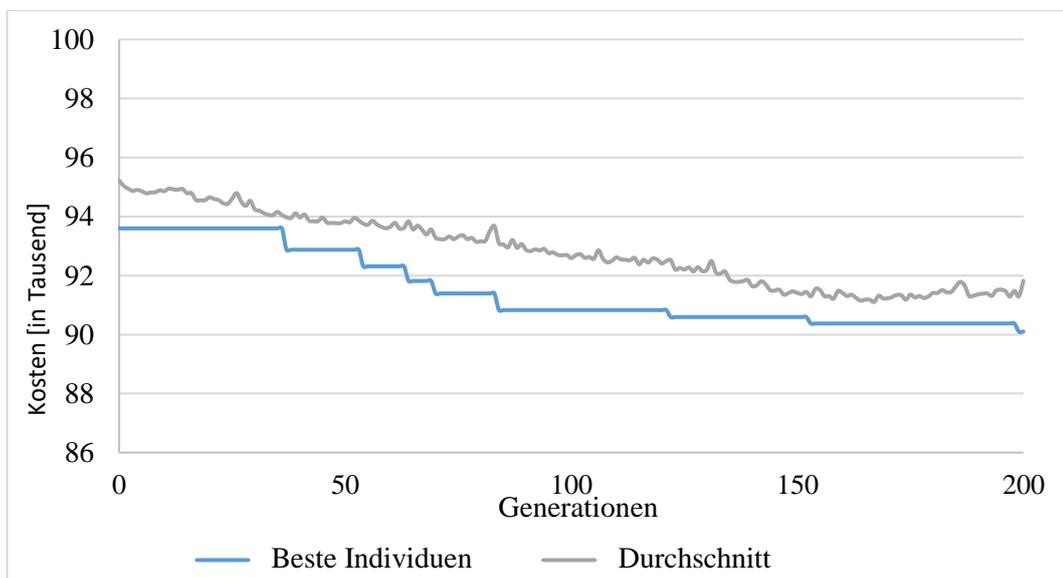


Abbildung 7.2: Verlauf der Kosten bei Mutationsrate 0,1

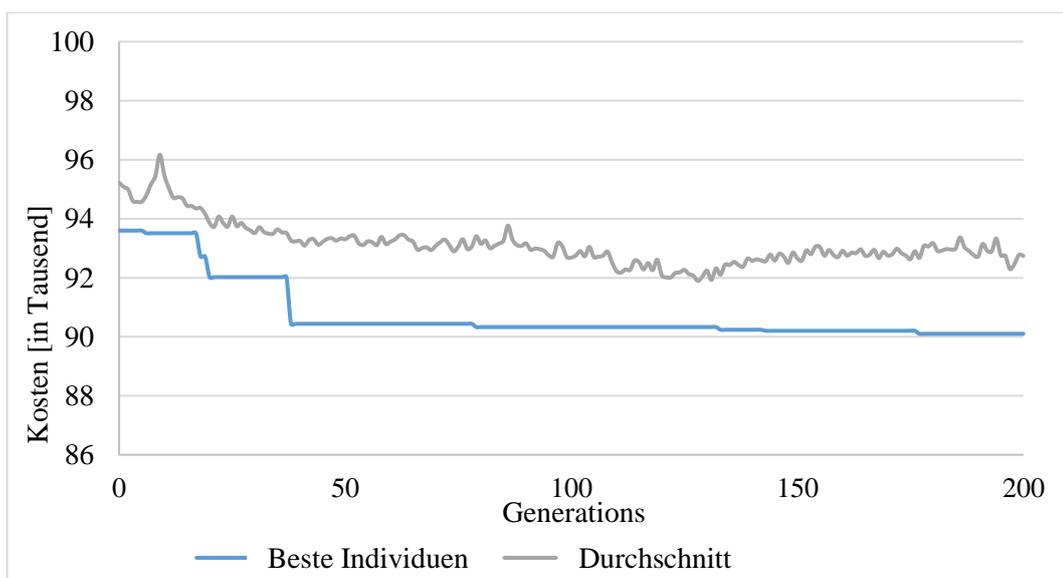


Abbildung 7.3: Verlauf der Kosten bei Mutationsrate von 0,2

Die beiden Diagramme stellen exemplarisch Verläufe für beliebige Seed-Werte dar. Aussagen unterliegen einer Unschärfe, da sich die Verläufe bei verändertem Seed-Wert ändern. Tendenzen können durch andere Graphen weiterer Seed-Werte gestützt werden.

Der Graph der besten Individuen verläuft stufenförmig. Über mehrere Generationen hinweg tritt keine Verbesserung ein. Dann flacht der Graph plötzlich ab, um den Wert wieder über mehrere Generationen beizubehalten. Mit zunehmender Generationszahl verlieren die Einbrüche an Intensität. Die Verläufe lassen auf ein konvergierendes Verhalten für spätere Generationen schließen. Die Versuche mit erhöhter Mutationsrate scheinen schneller zu konvergieren. Sowohl die Werte der besten Individuen als auch der durchschnittliche Wert flachen schneller ab, als bei dem oberen Graphen.

Versuche mit veränderten Mutationsverfahren liefern bessere Ergebnisse als das vorherige Mutationsverfahren und konvergieren schneller als unkorrelierende Versuche. Trotz dieser hinreichend zufriedenstellenden Eigenschaften, bleiben die Zielparameter unter denen der Vergleichsversuche. Die unzureichenden Ergebnisse liegen wahrscheinlich an dem derzeitigen Zusammenwirken von Mutation und Crossingover.

*Tabelle 7.6: Kostenmatrix bei verschiedenen Mutations- und Crossover-Wahrscheinlichkeiten*

Wahrscheinlichkeiten	Durchschnittliche Kosten	
	Mutation: 0,8	Mutation: 0,2
Crossover: 0,8	92.296,78	92.649,3
Crossover: 0,2	92.667,15	-

Wie in den vorherigen Auswertungen, generieren hohe Mutations- und Crossover-Wahrscheinlichkeiten die niedrigsten Kosten. Interessanter ist der Vergleich der Auswirkungen ausgetauschter Wahrscheinlichkeiten. Eine Wahrscheinlichkeit wird bei 0,2 festgesetzt und die andere bei 0,8. Der Austausch der Wahrscheinlichkeiten wirkt sich nur minimal aus. Die durchschnittlichen Kosten sind bei einer Mutationsrate von 0,2 kaum geringer als bei einer Crossover-Wahrscheinlichkeit von 0,2. 92.667,15 KE liegt nur geringfügig über den Kosten von 92.649,3 KE. Die Gewichtung der Wahrscheinlichkeiten hat demzufolge ebenfalls kaum Einfluss auf die Ergebnisse. Diese Schlussfolgerung wird von dem Kostenverlauf bei einer Mutationsrate von 0,8 und der Crossover-Wahrscheinlichkeit von 0,2 gestützt. Abbildung 6.4 stellt diesen Verlauf über mehrere Generationen dar.

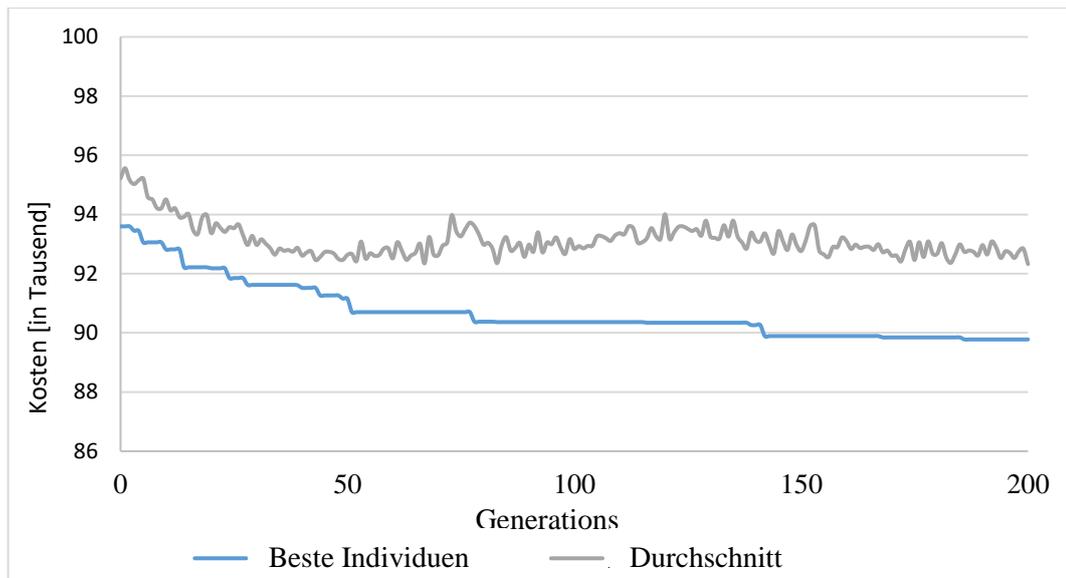


Abbildung 7.4: Kostenverlauf bei Mutationsrate von 0,8 und Crossover-Wahrscheinlichkeit von 0,2

Durch die zufällige Auswahl eines Versuchs, sind explizite Aussagen erneut nicht möglich. Tendenzen decken sich aber mit den beiden vorherigen Graphen über die Kostenverläufe. Die Kosten nehmen in den ersten Generationen stark ab. Die Steigung sinkt aber mit zunehmenden Generationen. In dem dargestellten Graphen liegt das beste Individuum unter 90.000 KE und liegt somit sogar unter den durchschnittlichen Ergebnissen mit unkorrelierenden Versuchen. Die durchschnittlichen Kosten aller Versuche dieser Mutations- und Crossover-Wahrscheinlichkeiten, schwanken um die angegebenen Kosten von 92.667,15 KE.

## 8. Zusammenfassung

Wie gezeigt werden konnte, lassen sich Evolutionäre Algorithmen durch eine entsprechende Vorarbeit beeinflussen. Die Versuchsergebnisse führten zwar nicht zu dem Nachweis der positiven Beeinflussung von korrelierenden Merkmalen auf die Zielparameter, die Zwischenergebnisse sind aber hinreichend zufriedenstellend, so dass auf diesen Versuchen aufgebaut werden kann.

Die besten Werte für die Zielparameter „Service Level“ und „Kosten“ werden bei unkorrelierenden Versuchen festgestellt. Diese Ergebnisse übersteigen ihre komparativen Versuche mit korrelierenden Merkmalen, unabhängig der jeweiligen Mutations- bzw. Crossover-Wahrscheinlichkeiten. Trotzdem sind die Ergebnisse der Versuche mit korrelierenden Merkmalen vielversprechend. Die Betrachtung der zeitlichen Verläufe der Zielparameter zeigt, dass die Versuche ohne korrelierende Merkmale sich erst nach der Initialisierung besser als ihre komparativen Versuche entwickeln. Die Individuen, die durch Berücksichtigung korrelierender Merkmale erstellt wurden, weisen in dieser Phase die besseren Werte auf. Erst durch die Veränderung der Individuen durch Mutationen und Rekombinationen, werden die Zielparameter von den unkorrelierenden Versuchen überholt. Mit zunehmenden Generationen verstärkt sich dieser Effekt.

Theoretisch wird der Evolutionäre Algorithmus positiv durch die Berücksichtigung korrelierender Merkmale beeinflusst. Dass sich der anfängliche Vorteil gegenüber den unkorrelierenden Merkmalen nicht durchsetzt, liegt an den Planungen der anschließenden Phasen. Mutation und Crossover sollen Evolutionären Algorithmen divers halten und dadurch helfen, neue Lösungen zu generieren. Durch das gegebene Verhältnis von dem großen Suchraum aller Merkmale, zu der relativ geringen Anzahl an Korrelationen eines Merkmals, ist die Wahrscheinlichkeit höher, Merkmale mit Korrelationen zu trennen, als neue Korrelationen zu erzeugen. Die getrennten Merkmale schädigen die Ergebnisse des Evolutionären Algorithmus mit korrelierenden Merkmalen. Um die bessere Ausgangslage nutzen zu können, muss der Suchraum der Mutationen und das Verfahren des Crossingovers angepasst werden.

Der Suchraum für den Austausch durch Mutationen wurde bereits im Verlauf der Auswertung angepasst. Merkmale, die von einer Mutation betroffen sind, werden durch Merkmale ausgetauscht, die die Distanztabelle vorgibt. Ausgehend des Merkmals, welches vor dem mutierenden Merkmal liegt, wird die Mutation gesteuert. Die kalkulierten Ergebnisse sind nur geringfügig besser, das Mutationsschema arbeitet trotzdem problemspezifischer. Das Crossingover erfolgt durch ein durchgängig standardisiertes Verfahren. Trotz seiner spezifischen Wirkweise, ist es nicht auf das gegebene Problem angepasst. Für verbesserte Resultate sollte auch das Crossover mehr auf das gegebene Problem angepasst werden. Eine Möglichkeit wäre es, das Crossover mit der erzeugten Distanzmatrix zu verknüpfen. Das Crossingover sollte Distanzen dieser Matrix berücksichtigen und entsprechend Merkmale austauschen. Da der Austausch derselben Positionen zweier Chromosomen sich selten mit den kürzesten Abständen deckt, müssten die ausgetauschten Positionen für beide Chromosomen variabel gehalten werden und lediglich die gleiche Anzahl an ausgetauschten Merkmalen festgelegt werden. Der Austausch selbst könnte sich, wie bei der neuen Mutationsart, an der jeweils vorigen Position richten, nach der die nächste Position der Distanzmatrix ausgewählt wird.

Die Berücksichtigung korrelierender Merkmale für einen Evolutionären Algorithmus hilft bei der Generierung guter Ausgangsindividuen. Um den Effekt nutzen zu können, müssen die anschließenden Phasen aber problemspezifischer eingesetzt werden.

## Literaturverzeichnis

- [FRI07]        **Friedman, Thomas L.:** The World Is Flat 3.0: A Brief History of the Twenty-first Century, 3. Auflage, Picador, London (2007).
- [CAR11]        **Charles, Scott:** Quicklet on Thomas Friedman's The World Is Flat (CliffNotes-like Book Summary), 3. Auflage, San Francisco, Hyperink (2011).
- [BRE10]        **Bretzke, Wolf-Rüdiger:** Logistische Netzwerke, 2. Auflage, Springer, Berlin (2010).
- [GER13]        **Gerdes, Ingrid; Klawoon, Frank; Kruse, Rudolf:** Evolutionäre Algorithmen: Genetische Algorithmen-Strategien und Optimierungsverfahren-Beispielanwendungen, Springer, Berlin (2013).
- [WER17]        **Werner, Hartmut:** Supply Chain Management: Grundlagen, Strategien, Instrumente und Controlling, 6. Auflage, Springer, Berlin (2017).
- [ALB14]        **Albrecht, Wolfgang:** Integrierte Netzwerk- und Liquiditätsplanung von Supply Chains, 5. Auflage, Springer, Berlin (2014).
- [PAU10]        **Paul, Joachim:** Praxisorientierte Einführung in die allgemeine Betriebswirtschaftslehre: Mit Beispielen und Fallstudien, 6. Auflage, Gabler, Berlin (2010).
- [WAG04]        **Wagner, Michael:** Business Networking im Internet: Interaktive Anbahnung von Kooperationen in Unternehmensnetzwerken (Markt- und Unternehmensentwicklung Markets and Organisations), Gabler, Wiesbaden (2004).
- [SYD06]        **Sydow, Jörg; Manning, Stephan:** Netzwerke beraten: über Netzwerkberatung und Beratungsnetzwerke, Gabler, Wiesbaden (2006).
- [PFO04]        **Pfohl, Hans-Christian:**  
Erfolgsfaktor Kooperation in der Logistik: Outsourcing – Beziehungsmanagement – Finanzielle Performance, In (Unternehmensführung und Logistik, Band 22), Erich Schmidt, Berlin (2004).
- [SYD10]        **Sydow, Jörg:** Management von Netzwerkorganisationen: Beiträge aus der "Managementforschung", 5. Auflage, Gabler, Wiesbaden (2010).
- [FIS06]        **Fischer, Bettina:** Vertikale Innovationsnetzwerke: Eine theoretische und empirische Analyse, Gabler, Wiesbaden (2006).

- [RIE08]       **Rief, Alexander:** Entwicklungsorientierte Steuerung strategischer Unternehmensnetzwerke, Gabler, Wiesbaden (2008).
- [ALB03]       **Albers, Sönke:** Management virtueller Unternehmen (Betriebswirtschaftliche Aspekte lose gekoppelter Systeme und Electronic Business), Deutscher Universitäts-Verlag, Wiesbaden (2003).
- [ZIR07]       **Zirpins, Christian:** Interaktionsorientierte Komposition virtueller Dienstleistungsprozesse, Habilitation, Universität Hamburg, Institut für Informatik (2007).
- [WAG05]       **Wagner, Markus:** Rahmenbedingungen zur dezentralen Koordination logistischer Netzwerke, Josef Eul, Lohmar (2005).
- [MAT18]       **Matys, Erwin:** Praxishandbuch Produktmanagement: Grundlagen und Instrumente: [Produktentwicklung, Markteinführung, Produkt-Lebenszyklus, Markt-Positionierung, Sicherung von Marktanteilen], 7. Auflage, Campus-Verlag, Frankfurt (2018).
- [KOE14]       **Koether, Reinhard:** Distributionslogistik: Effiziente Absicherung der Lieferfähigkeit, 2. Auflage, Springer, Berlin (2014).
- [FOR07]       **Fortmann, Klaus-Michael; Kallweit, Angela:** Logistik, 2. Auflage, W. Kohlhammer, Stuttgart (2007).
- [WEB15]       **Weber, Rainer:** Lageroptimierung: Bestände, Abläufe, Organisation, Datenqualität, Stellplätze, 3. Auflage, Expert, Renningen (2015)
- [HUT15]       **Hutzschenreuter, Thomas:** Allgemeine Betriebswirtschaftslehre: Grundlagen mit zahlreichen Praxisbeispielen, 6. Auflage, Springer, Berlin (2015).
- [TOP13]       **Toporowski, Waldemar:** Logistik im Handel: Optimale Lagerstruktur und Bestellpolitik einer Filialunternehmung, Springer, Berlin (2013).
- [BRA18]       **Bräbänder, Christian:** Stochastisches Bestandsmanagement: Grundmodelle für Betriebswirte, Springer-Verlag, Berlin (2014).
- [KRO66]       **Kroeber-Riel, Werner:** Beschaffung und Lagerung: Betriebswirtschaftliche Grundfragen der Materialwirtschaft, Springer, Berlin (1966).
- [DRA01]       **Draenert, Patric:** Kooperative Absatzplanung. Einführungsstrategie für den Prognosedatenaustausch, Deutscher Universitätsverlag, Wiesbaden (2001).

- [SKJ07] **Skjott-Larsen, Tage, Schary, Philip; Mikkola, Juliana:** Managing the Global Supply Chain, Copenhagen Business School Press DK, Kopenhagen (2007).
- [DUC17] **Duchowski, Andrew, Krejtz, Krzysztof; Biele, Cezary:** An inverse-linear logistic model of the main sequence, Journal of Eye Movement Research, Bern Open Publications (2017).
- [MEN97] **Menzel, Klaus:** Algorithmen: Vom Problem zum Programm (Mathematik-ABC für das Lehramt), Springer, Berlin (1997).
- [KRU10] **Kruse, Rudolf; Moewes, Christian:** Evolutionäre Algorithmen, Fachartikel, Universität Magdeburg, Fakultät für Informatik (2003), [https://www.researchgate.net/publication/224790979\\_Evolutionare\\_Algorithmen](https://www.researchgate.net/publication/224790979_Evolutionare_Algorithmen).
- [HES72] **Hess, Dieter:** Genetik: Grundlagen, Erkenntnisse, Entwicklungen d. modernen Vererbungsforschung. 4.Auflage, Herder, Freiburg (1972).
- [STR02] **Streichert, Felix:** Introduction to Evolutionary Algorithms, Frankfurt MathFinance Workshop, Frankfurt (2002).
- [KOC14] **Koch, Sören:** Genetische Algorithmen für das Order Batching-Problem in manuellen Kommissioniersystemen (Produktion und Logistik), Springer, Berlin (2014).
- [RÄU14] **Räuber, Rin:** Artificial Art: Erzeugung künstlicher Bilder mit evolutionären Algorithmen in Javascript, Dortmund. Bachelorarbeit, FH Dortmund, Institut für Informatik (2014).
- [BLU09] **Blume, Christian; Wilfried, Jakob:** GLEAM - general learning evolutionary algorithm and method: ein evolutionärer Algorithmus und seine Anwendungen, KIT Scientific Publishing, Karlsruhe (2009).
- [LI14] **Shan Li; Liying Kang; and Xing-Ming Zhao:** A Survey on Evolutionary Algorithm Based Hybrid Intelligence in Bioinformatics, Hindawi Publishing Corporation, BioMed Research International, Volume 2014, Article ID 362738 (2014)
- [RUD15] **Prof. Dr. Rudolf; Günter:** Evolutionäre Algorithmen, Fachartikel, Universität Dortmund, Fakultät für Informatik, Dortmund (2011), <https://ls11-www.cs.tu-dortmund.de/de/rudolph/ea>.
- [LIN03] **Lindemann, Djamila:** Evolutionsstrategien, Fachartikel, Universität Dortmund, Fakultät für Informatik (2003), <https://ls11-www.cs.tu-dortmund.de/lehre/SoSe03/PG431/Ausarbeitungen/Evolutionsstrategien.pdf>.

- [ZBI98]       **Zbigniew, Michalewicz:** Genetic Algorithms + Data Structures = Evolution Programs, 3. Auflage, Springer, Berlin (1998).
- [POH99]       **Pohlheim, Hartmut:** Evolutionäre Algorithmen: Verfahren, Operatoren und Hinweise für die Praxis (VDI-Buch), Springer, Berlin (1999).
- [SEL03]       **Selzam, Bianca:** Genetische Algorithmen, Universität Dortmund, Fakultät für Informatik (2003).  
[https://ls11-www.cs.tu-dortmund.de/lehre/SoSe03/PG431/Ausarbeitungen/GA\\_Selzam.pdf](https://ls11-www.cs.tu-dortmund.de/lehre/SoSe03/PG431/Ausarbeitungen/GA_Selzam.pdf).
- [BOD06]       **Bodendorf, Freimut:** Daten- und Wissensmanagement, 2. Auflage, Springer, Berlin (2006).
- [NIS97]       **Nissen, Volker:** Einführung in Evolutionäre Algorithmen: Optimierung nach dem Vorbild der Evolution (Computational Intelligence), Vieweg + Teubner, Wiesbaden (1997).
- [DAN07]       **Danylevych, Olha:** Stratifizierte Transaktionen, Dissertation, Universität Stuttgart, Institut für Architektur von Anwendungssystemen (2007).
- [KOK02]       **Kókai, Gabriella:** Erfolge und Probleme evolutionärer Algorithmen, induktiver logischer Programmierung und ihrer Kombination, Habilitation, Universität Erlangen-Nürnberg, Institut für Informatik (2002)
- [LUT00]       **Lutz, Mark; Ascher, David:** Einführung in Python, Habilitation. O'Reilly Verlag GmbH & Co. KG, Köln (2000).
- [HAG12]       **Hagedorn, Lotar-Müller; Toporowski, Waldemar, Zielke Stephan:** Der Handel: Grundlagen - Management – Strategien, 2. Auflage, W. Kohlhammer W. GmbH, Stuttgart (2012).
- [HER11]       **Hertel, Joachim; Zentes, Joachim; Schramm-Klein, Hanna:** Supply-Chain-Management und Warenwirtschaftssysteme im Handel. 3.Auflage, Springer, Berlin (2011).
- [BdKEP16]      **BdKEP, DIN e.V., Händlerbund e.V.:** Logistik auf der ersten und letzten Meile im Wandel. Beuth, Berlin (2016).
- [YU10]       **Yu, Xinjie; Gen, Mitsuo:** Introduction to Evolutionary Algorithms. Springer Science & Business Media, Berlin (2010).

- [DEV13]        **Dewdney, A.K.:** Der Turing Omnibus: Eine Reise durch die Informatik mit 66 Stationen. Springer, Berlin (2013).
- [SCH18]        **Scholz, Daniel:** Optimierung interaktiv: Grundlagen verstehen, Modelle erforschen und Verfahren anwenden. Springer, Berlin (2018).
- [STA14]        **Stadler Peter F.; Will, Sebastian:** Optimierung interaktiv: ADS: Algorithmen und Datenstrukturen 2. Skript. Institut für Informatik & Interdisziplinäres Zentrum für Bioinformatik Universität Leipzig (2014).
- [ECK02]        **Eckey, Hans Friedrich; Kosfeld, Reinhold; Rengers, Martina:** Multivariate Statistik. Springer, Berlin (2002).
- [FIS17]        **Fischer, Anja; Hungerländer, Philipp:** The traveling salesman problem on grids with forbidden neighborhoods. Journal of Combinatorial Optimization, New York (2017).
- [SAB12]        **Saban, Asher; Waltuch, Liat:** Genetic Algorithm for the Traveling Salesman Problem. Project: Introduction to Artificial Intelligence, Hebrew University of Jerusalem, Jerusalem (2012).
- [AHM10]        **Ahmed, Zakir Hussain:** Genetic Algorithm for the Traveling Salesman Problem using Sequential Constructive Crossover Operator. CiiT International Journal of Biometrics and Bioinformatics, Malaysia (2010).
- [RAA02]        **Raab, Matthias:** Anwendung: Das Travelling Salesman Problem. Proseminar Genetische und Evolutionäre Algorithmen, Universität Ulm (2002).
- [PUL13]        **Puljic, Krunoslav; Manger, Robert:** Anwendung: Comparison of eight evolutionary crossover operators for the vehicle routing problem, MATHEMATICAL COMMUNICATIONS Vo. 18, Fakultät für Mathematik, Universität Zagreb, Kroatien (2013).

## Anhang

Nach Von	1	2	3	4	5	6	7	8	11	12	13	14	15	16	17	18
1	0	1	1	1		1	1	1	0,5	1	1	1		1	1	1
2	1	0	1	1	1		1	1	1	0,5	1	1	1		1	1
3	1	1	0	1	1	1		1	1	1	0,5	1	1	1		1
4	1	1	1	0	1	1	1		1	1	1	0,5	1	1	1	
5		1	1	1	0	1	1	1		1	1	1	0,5	1	1	1
6	1		1	1	1	0	1	1	1		1	1	1	0,5	1	1
7	1	1		1	1	1	0	1	1	1		1	1	1	0,5	1
8	1	1	1		1	1	1	0	1	1	1		1	1	1	0,5
11	1	1	1	1	1	1	1	1		1	1	1		1	1	1
12	1	1	1	1	1	1	1	1	1		1	1	1		1	1
13	1	1	1	1	1	1	1	1	1	1		1	1	1		1
14	1	1	1	1	1	1	1	1	1	1	1		1	1	1	
15	1	1	1	1	1	1	1	1		1	1	1		1	1	1
16	1	1	1	1	1	1	1	1	1		1	1	1		1	1
17	1	1	1	1	1	1	1	1	1	1		1	1	1		1
18	1	1	1	1	1	1	1	1	1	1	1		1	1	1	
21	1	1	1	1	1	1	1	1	0,5	1	1	1		1	1	1
22	1	1	1	1	1	1	1	1	1	0,5	1	1	1		1	1
23	1	1	1	1	1	1	1	1	1	1	0,5	1	1	1		1
24	1	1	1	1	1	1	1	1	1	1	1	0,5	1	1	1	
25	1	1	1	1	1	1	1	1		1	1	1	0,5	1	1	1
26	1	1	1	1	1	1	1	1	1		1	1	1	0,5	1	1
27	1	1	1	1	1	1	1	1	1	1		1	1	1	0,5	1
28	1	1	1	1	1	1	1	1	1	1	1		1	1	1	0,5
31	1	1	1	1	1	1	1	1	0,5	1	1	1		1	1	1
32	1	1	1	1	1	1	1	1	1	0,5	1	1	1		1	1
33	1	1	1	1	1	1	1	1	1	1	0,5	1	1	1		1
34	1	1	1	1	1	1	1	1	1	1	1	0,5	1	1	1	
35	1	1	1	1	1	1	1	1		1	1	1	0,5	1	1	1
36	1	1	1	1	1	1	1	1	1		1	1	1	0,5	1	1
37	1	1	1	1	1	1	1	1	1	1		1	1	1	0,5	1
38	1	1	1	1	1	1	1	1	1	1	1		1	1	1	0,5
41	1	1	1	1	1	1	1	1	0,5	1	1	1		1	1	1
42	1	1	1	1	1	1	1	1	1	0,5	1	1	1		1	1
43	1	1	1	1	1	1	1	1	1	1	0,5	1	1	1		1
44	1	1	1	1	1	1	1	1	1	1	1	0,5	1	1	1	
45	1	1	1	1	1	1	1	1		1	1	1	0,5	1	1	1
46	1	1	1	1	1	1	1	1	1		1	1	1	0,5	1	1
47	1	1	1	1	1	1	1	1	1	1		1	1	1	0,5	1

Nach Von	21	22	23	24	25	26	27	28	31	32	33	34	35	36	37	38
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
8	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
11	0,5	1	1	1	1	1	1	1	0,5	1	1	1	1	1	1	1
12	1	0,5	1	1	1	1	1	1	1	0,5	1	1	1	1	1	1
13	1	1	0,5	1	1	1	1	1	1	1	0,5	1	1	1	1	1
14	1	1	1	0,5	1	1	1	1	1	1	1	0,5	1	1	1	1
15	1	1	1	1	0,5	1	1	1	1	1	1	1	0,5	1	1	1
16	1	1	1	1	1	0,5	1	1	1	1	1	1	1	0,5	1	1
17	1	1	1	1	1	1	0,5	1	1	1	1	1	1	1	0,5	1
18	1	1	1	1	1	1	1	0,5	1	1	1	1	1	1	1	0,5
21	1	1	1	1	1	1	1	1	0,5	1	1	1	1	1	1	1
22	1	1	1	1	1	1	1	1	1	0,5	1	1	1	1	1	1
23	1	1	1	1	1	1	1	1	1	1	0,5	1	1	1	1	1
24	1	1	1	1	1	1	1	1	1	1	1	0,5	1	1	1	1
25	1	1	1	1	1	1	1	1	1	1	1	1	0,5	1	1	1
26	1	1	1	1	1	1	1	1	1	1	1	1	1	0,5	1	1
27	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0,5	1
28	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0,5
31	0,5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
32	1	0,5	1	1	1	1	1	1	1	1	1	1	1	1	1	1
33	1	1	0,5	1	1	1	1	1	1	1	1	1	1	1	1	1
34	1	1	1	0,5	1	1	1	1	1	1	1	1	1	1	1	1
35	1	1	1	1	0,5	1	1	1	1	1	1	1	1	1	1	1
36	1	1	1	1	1	0,5	1	1	1	1	1	1	1	1	1	1
37	1	1	1	1	1	1	0,5	1	1	1	1	1	1	1	1	1
38	1	1	1	1	1	1	1	0,5	1	1	1	1	1	1	1	1
41	0,5	1	1	1	1	1	1	1	0,5	1	1	1	1	1	1	1
42	1	0,5	1	1	1	1	1	1	1	0,5	1	1	1	1	1	1
43	1	1	0,5	1	1	1	1	1	1	1	0,5	1	1	1	1	1
44	1	1	1	0,5	1	1	1	1	1	1	1	0,5	1	1	1	1
45	1	1	1	1	0,5	1	1	1	1	1	1	1	0,5	1	1	1
46	1	1	1	1	1	0,5	1	1	1	1	1	1	1	0,5	1	1
47	1	1	1	1	1	1	0,5	1	1	1	1	1	1	1	0,5	1
48	1	1	1	1	1	1	1	0,5	1	1	1	1	1	1	1	0,5

Von Nach	41	42	43	44	45	46	47	48
1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	1	1
3	1	1	1	1	1	1	1	1
4	1	1	1	1	1	1	1	1
5	1	1	1	1	1	1	1	1
6	1	1	1	1	1	1	1	1
7	1	1	1	1	1	1	1	1
8	1	1	1	1	1	1	1	1
11	0,5	1	1	1		1	1	1
12	1	0,5	1	1	1		1	1
13	1	1	0,5	1	1	1		1
14	1	1	1	0,5	1	1	1	
15		1	1	1	0,5	1	1	1
16	1		1	1	1	0,5	1	1
17	1	1		1	1	1	0,5	1
18	1	1	1		1	1	1	0,5
21	0,5	1	1	1		1	1	1
22	1	0,5	1	1	1		1	1
23	1	1	0,5	1	1	1		1
24	1	1	1	0,5	1	1	1	
25		1	1	1	0,5	1	1	1
26	1		1	1	1	0,5	1	1
27	1	1		1	1	1	0,5	1
28	1	1	1		1	1	1	0,5
31	0,5	1	1	1		1	1	1
32	1	0,5	1	1	1		1	1
33	1	1	0,5	1	1	1		1
34	1	1	1	0,5	1	1	1	
35		1	1	1	0,5	1	1	1
36	1		1	1	1	0,5	1	1
37	1	1		1	1	1	0,5	1
38	1	1	1		1	1	1	0,5
41	1	1	1	1		1	1	1
42	1	1	1	1	1		1	1
43	1	1	1	1	1	1		1
44	1	1	1	1	1	1	1	
45		1	1	1	1	1	1	1
46	1		1	1	1	1	1	1
47	1	1		1	1	1	1	1
48	1	1	1		1	1	1	1

# Eidesstattliche Versicherung

## (Affidavit)

Name, Vorname  
(Last name, first name)

Matrikelnr.  
(Enrollment number)

Ich versichere hiermit an Eides statt, dass ich die vorliegende Bachelorarbeit/Masterarbeit\* mit dem folgenden Titel selbstständig und ohne unzulässige fremde Hilfe erbracht habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie wörtliche und sinngemäße Zitate kenntlich gemacht. Die Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

I declare in lieu of oath that I have completed the present Bachelor's/Master's\* thesis with the following title independently and without any unauthorized assistance. I have not used any other sources or aids than the ones listed and have documented quotations and paraphrases as such. The thesis in its current or similar version has not been submitted to an auditing institution.

Titel der Bachelor-/Masterarbeit\*:  
(Title of the Bachelor's/ Master's\* thesis):

\*Nichtzutreffendes bitte streichen

(Please choose the appropriate)

Ort, Datum  
(Place, date)

Unterschrift  
(Signature)

### Belehrung:

Wer vorsätzlich gegen eine die Täuschung über Prüfungsleistungen betreffende Regelung einer Hochschulprüfungsordnung verstößt, handelt ordnungswidrig. Die Ordnungswidrigkeit kann mit einer Geldbuße von bis zu 50.000,00 € geahndet werden. Zuständige Verwaltungsbehörde für die Verfolgung und Ahndung von Ordnungswidrigkeiten ist der Kanzler/die Kanzlerin der Technischen Universität Dortmund. Im Falle eines mehrfachen oder sonstigen schwerwiegenden Täuschungsversuches kann der Prüfling zudem exmatrikuliert werden. (§ 63 Abs. 5 Hochschulgesetz - HG - ).

Die Abgabe einer falschen Versicherung an Eides statt wird mit Freiheitsstrafe bis zu 3 Jahren oder mit Geldstrafe bestraft.

Die Technische Universität Dortmund wird gfls. elektronische Vergleichswerkzeuge (wie z.B. die Software „turnitin“) zur Überprüfung von Ordnungswidrigkeiten in Prüfungsverfahren nutzen.

Die oben stehende Belehrung habe ich zur Kenntnis genommen:

### Official notification:

Any person who intentionally breaches any regulation of university examination regulations relating to deception in examination performance is acting improperly. This offense can be punished with a fine of up to €50,000.00. The competent administrative authority for the pursuit and prosecution of offenses of this type is the chancellor of TU Dortmund University. In the case of multiple or other serious attempts at deception, the examinee can also be unenrolled, section 63, subsection 5 of the North Rhine- Westphalia Higher Education Act (*Hochschulgesetz*).

The submission of a false affidavit will be punished with a prison sentence of up to three years or a fine.

As may be necessary, TU Dortmund will make use of electronic plagiarism-prevention tools (e.g. the "turnitin" service) in order to monitor violations during the examination procedures.

I have taken note of the above official notification:\*\*

Ort, Datum  
(Place, date)

Unterschrift  
(Signature)

**\*\*Please be aware that solely the German version of the affidavit ("Eidesstattliche Versicherung") for the Bachelor's/ Master's thesis is the official and legally binding version.**