Heuristiken für kombinierte Standort- und Gebietsplanung mit vorgegebenen und zusätzlichen, frei wählbaren Standorten

Diplomarbeit
von

Tim Ulrich

An der Fakultät für Informatik
Institut für Theoretische Informatik

Erstgutachter:	Prof. Dr. Stefan Nickel
Zweitgutachter:	Prof. Dr. Dorothea Wagner
Betreuende Mitarbeiter:	Alexander Butsch
	Dr. Martin Nöllenburg

Bearbeitungszeit: 15. November 2013 – 14. August 2014

Eidesstattliche Erklärung

Ich versichere hiermit wahrheitsgemäß, die Arbeit selbständig angefertigt, alle benutzten Hilfsmittel vollständig und genau angegeben und alles kenntlich gemacht zu haben, was aus Arbeiten anderer unverändert oder mit Abänderung entnommen wurde.

Datum *Name*

Inhaltsverzeichnis

1 Einleitung

Die Wahl des richtigen Standortes ist für alle betrieblichen und staatlichen Einrichtungen von zentraler Bedeutung. Aufgrund der hohen Kosten, die in der Regel mit einem Standortwechsel verbunden sind, ist die Entscheidung für einen Standort nur langfristig zu ändern. Gerade deshalb bietet die Wahl eines Standortes erhebliches Optimierungspotenzial.
Standortplanung ist der Teil des Operations Research, der sich mit der Wahl optimaler Standorte auseinandersetzt. Die Entscheidung, wo eine Anlage platziert werden soll gehört wegen ihrer langfristigen Auswirkungen zu den strategischen Entscheidungen, die ein Unternehmen zu treffen hat. Gerade wegen der Bedeutung der Standortwahl für ein Unternehmen ist der Bedarf nach Lösungsverfahren groß.
Im öffentlichen Bereich ist die Bedeutung der Standortwahl noch größer als im privaten Bereich, da die Standortentscheidungen der privaten Akteure stark davon abhängen, wo sich staatliche Einrichtungen befinden oder sie geplant werden. So ist der Zugang zu Transportwegen und anderer Infrastruktur wie Telekommunikation oder Energie für Wirtschaftsbetriebe von überragender Bedeutung. Öffentliche Standortentscheidungen beeinflussen darüber hinaus auch die Wohnortwahl der Bürger, welche maßgeblich durch Bedürfnisse wie Bildung, Gesundheit, Sicherheit und Kultur bestimmt wird.
Die Gleichheit der Wahl ist wesentlich für ein demokratisches System. Während sich Wahlgleichheit formal nur auf den Wert einer einzelnen Stimme bezieht, bedarf es für tatsächliche Wahlgleichheit des gleichen *Einflusses* jeder Stimme. Hieraus leitet sich die Forderung nach einer gerechten Einteilung der Wahlkreise ab, sodass die Bevölkerung der einzelnen Wahlkreise nicht zu stark voneinander abweicht. Mit diesem Problem verwandt ist die Einteilung von Zuständigkeitsbezirken für öffentliche Einrichtungen wie Schulen, Krankenhäuser, Feuer- und Polizeiwachen. Wegen dieser Anwendungen ist die gesellschaftliche Bedeutung der Gebietsplanung, der Einteilung eines Planungsgebiets in kleinere Teilgebiete nach formalen Kriterien, nicht zu unterschätzen.
Auch im betrieblichen Kontext spielt Gebietsplanung eine wichtige Rolle. Die Einteilung und Zuweisung von Verkaufs- und Servicegebieten ist nicht nur wesentlich für den Unternehmenserfolg. Stark voneinander abweichende Arbeitsbelastungen werden als ungerecht empfunden, wirken sich darüber hinaus auch auf die Entlohnung durch erfolgsabhängige Zahlungen aus. Beides erhöht die Unzufriedenheit der Mitarbeiter und senkt damit ihre Motivation.
Bedeutung und Vielzahl der Anwendungen von Standort- und Gebietsplanung haben eine ausgedehnte wissenschaftliche Literatur in diesen Bereichen hervorgebracht. Dabei geht im Falle der Standortplanung die wissenschaftliche Betrachtung der Anwendung deutlich voraus. Vorläufer der Standortplanung wurden bereits im 17. Jahrhundert behandelt. Als mit der Industrialisierung die Mobilität von Menschen und Gütern größer wurde, nahm auch die Bedeutung der Standortwahl für eine Produktionsstätte zu. Nach Alfred

Weber, der zu Beginn des 20. Jahrhunderts die erste Anwendung der Standortplanung beschrieb, wurde das Problem, den gewichteten Median einer Punktmenge zu finden, Weber-Problem genannt. Der gewichtete Median ist auch heute noch eine wichtige Zielfunktion der Standortplanung. Die zweite wichtige Zielfunktion der Standortplanung ist das (gewichtete) Zentrum, welches von Seifollah Hakimi eingeführt wurde. Diese beiden Ziele dominieren auch heute noch die Literatur zur Standortplanung.
Es existiert eine umfangreiche Literatur zur Standortplanung. Betrachtet werden verschiedene Zielfunktionen und Bedingungen, unter denen die Standortwahl erfolgt. Da es keinen effizienten Algorithmus gibt, der optimale Lösungen für Standortprobleme berechnet, wird auf Heuristiken zurückgegriffen. Üblich ist auch die Modellierung als ganzzahliges lineares Programm, welches dann mit den bekannten Verfahren gelöst werden kann. Diese Lösungsansätze gehen alle davon aus, dass keine vorhandenen Zentren existieren, welche in die Planung mit einbezogen werden müssen.

1.1 Gegenstand dieser Arbeit

In den letzten Jahren ist zumindest die Modellierung dieses Spezialfalls Gegenstand der Forschung geworden. Jedoch existieren neben der Möglichkeit das ganzzahlige Programm zu lösen, keine Verfahren zur Lösung eines Standortproblems unter Berücksichtigung bereits vorhandener Zentren. Im Rahmen dieser Arbeit wird ein solches heuristisches Verfahren entwickelt. Dieses Verfahren baut auf dem von [Kal06] entwickelten Recursive-Partitioning-Algorithmus (Rec-Part) zur Gebietsplanung auf. Dieser Algorithmus zerlegt ein Gebietsplanungsproblem solange in kleinere Teilprobleme bis eine elementare Größe erreicht wurde. Aus diesen elementaren Teilproblemen wird die Lösung zusammengesetzt. Werden für die elementaren Teilprobleme Zentren berechnet, kann der Algorithmus auch zur Standortplanung eingesetzt werden. Das in dieser Arbeit beschriebene Verfahren unterteilt das ursprüngliche Problem so in Teilprobleme, dass genau so viele elementare Teilprobleme keine der vorhandenen Zentren enthalten wie neue Standorte ausgewählt werden sollen. Für diese Teilprobleme werden dann Zentren bestimmt, welche zusammen die Lösung bilden.

1.2 Aufbau dieser Arbeit

Zunächst werden im nächsten Abschnitt grundlegende Begriffe der Standort- und Gebietsplanung eingeführt. Anschließend werden die beiden Themenfelder abgegrenzt und formalisiert. Insbesondere werden die verwendeten Ziele und Zielfunktionen sowie Eigenschaften eines Standortproblems beschrieben.
Abschnitt 3 enthält eine Beschreibung des Recursive-Partitioning-Verfahren. Nach einem kurzen Überblick werden die Aufteilungen der Teilprobleme, die Partitionen, definiert. Es folgt die Beschreibung des Ablaufs des Algorithmus und des Umgangs mit nicht lösbaren Teilproblemen. Anschließend wird gezeigt, wie der Algorithmus zur Standortplanung

unter Einbeziehung von Facilities verwendet werden kann. Der Recursive-Partitioning-Algorithmus verwendet sogenannte Maße, um die verschiedenen Aufteilungen eines Teilproblems zu bewerten. Von diesen Maßen wurden einige im Rahmen dieser Arbeit entwickelt. Eine Beschreibung und Kategorisierung der verwendeten Maße findet sich in Abschnitt 4.
Da der Recursive-Partitioning-Algorithmus modular aufgebaut ist, kann er einfach erweitert werden. Für diese Arbeit wurden einige Erweiterungen entwickelt, die in Abschnitt 5 vorgestellt und in Abschnitt 6 getestet werden. Der sechste Abschnitt enthält auch die Tests der verschiedenen Maße. Abschließend werden die Ergebnisse zusammengefasst.

2 Grundlagen

Bezeichnungen können in verschiedenen Kontexten unterschiedliche Bedeutungen haben. Ebenso können kleine Unterschiede in der Definition von Begriffen zu Missverständnissen führen. Um solche Missverständnisse zu vermeiden, werden zunächst die in dieser Arbeit verwendeten Begriffe eingeführt.

Grundbegriffe der Standortplanung

Kunde Kunden sind die Objekte der Standortplanung, welche die Dienstleistung einer Einrichtung in Anspruch nehmen. Ihnen ist ein Wert zugeordnet, der die Nachfrage nach der Dienstleistung beschreibt. Dieser Wert wird Bedarf genannt. Die Bezeichnung Kunde kann auch dann gewählt werden, wenn es sich um eine öffentliche Dienstleistung handelt, wie bei Rettungsdiensten oder Schulen. Kunden werden durch einen eindeutigen Index voneinander unterschieden. Im dieser Arbeit wird für die Menge der Kunden die Bezeichung V gewählt.

Standort, Einrichtung (Facility) Ein Standort ist der Ort, an dem nach erfolgter Standortplanung eine besondere Einrichtung (engl. Facility) errichtet werden soll. Aufgrund der zahlreichen Anwendungen der Standortplanung gibt es verschiedene Arten von Einrichtungen wie Transportlager, Supermärkte, Krankenhäuser oder Mülldeponien. Die Menge der möglichen Standorte wird mit X, die Anzahl der auszuwählenden Standorte mit p bezeichnet. Einer Facility kann eine Kapazität zugeordnet, die angibt wie viel Bedarf diese Facility maximal bedienen kann.

Planungsregion Die Menge aller Kunden sowie der vorhandenen und möglichen Standorte wird als Planungsregion bezeichnet. Da die meisten Anwendungen der Standort- und Gebietsplanung einen geographischen Hintergrund haben, entspricht die Planungsregion in der Regel einem geographischen Gebiet.

Grundbegriffe der Gebietsplanung

Basisgebiet Basisgebiete sind in der Gebietsplanung die Einheiten, die zu übergeordneten Einheiten, den Distrikten, zusammengefasst werden. Basisgebiete können in der Praxis Kunden des Außendienstes, Städte, die zu Wahlkreisen zusammengefasst werden sollen oder ähnliches sein. Einem Basisgebiet wird die sogenannte Aktivität zugeordnet, eine Zahl, welche die für die Problemstellung relevante Größe angibt, z.B. Bedarf eines Kunden oder Einwohner einer Stadt. Auch in der Standortplanung können die Objekte, an denen die zu planenden Standorte ausgerichtet werden, als Basisgebiete bezeichnet werden. Die Begriffe Kunde und Basisgebiet werden in dieser Arbeit synonym verwendet.

Aktivität, Bedarf Aktivität ist eine für die Gebietsplanung relevante Eigenschaft eines Basisgebiets. Eine Eigenschaft wird dann als Aktivität bezeichnet, wenn sie dem Balancekriterium genügen soll, also gleichmäßig auf die Distrikte aufgeteilt werden soll. In der Standortplanung wird die entsprechende Eigenschaft Bedarf genannt. Eine weitere Bezeichnung ist Gewicht (des Kunden/Basisgebiets). Die Aktivität des Basisgebiets mit Index i wird mit $w(i)$ bezeichnet.

Distrikt Ein Distrikt ist eine Menge an Basisgebieten. Einem Distrikt können Zentren zugeordnet, welche die entsprechende Aufgabe für dieses Gebiet übernehmen. Diese Zentren entsprechen den Standorten bzw. Facilities der Standortplanung. Alternative Bezeichnungen sind Territorium oder Gebiet.

Gebietsplan Das Ergebnis einer Gebietsplanung ist der sogenannte Gebietsplan. Dieser ist die Menge aller Distrikte, die während der Planung in die Lösung aufgenommen wurden.

2.1 Standortplanung

Standortplanung im engeren Sinne bezeichnet die Auswahl einer bestimmten Anzahl an Standorten aus einer vorgegebenen Menge möglicher Standorte nach gewissen Entscheidungskriterien. Die Anzahl der Standorte kann fest vorgegeben sein oder im erst Verlauf der Planung festgelegt werden. Die Auswahl von Standorten wird als Location-Problem bezeichnet. Im weiteren Sinne bezeichnet Standortplanung zusätzlich die Zuordnung des Bedarfs (Kunden oder Basisgebiete) zu den Standorten. Dies wird Allocation-Problem genannt. Viele Algorithmen führen neben der Auswahl der Standorte auch eine Zuordnung der Kunden durch. In diesem Fall wird von der Location- und der Allocation-Phase gesprochen (siehe Abbildung 1).
Für die Lösung eines Standortproblems müssen unter anderem die folgenden Fragen beantwortet werden[Das95]:

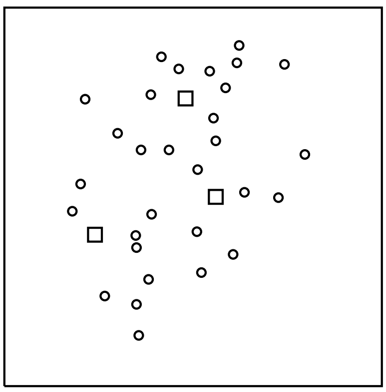
- Wie viele Standorte sollen ausgewählt werden?

- Welche Standorte sollen ausgewählt werden?
- Wie groß sollen die einzelnen Facilities an den Standorten sein?
- Wie soll der Bedarf auf die Facilities aufgeteilt werden?

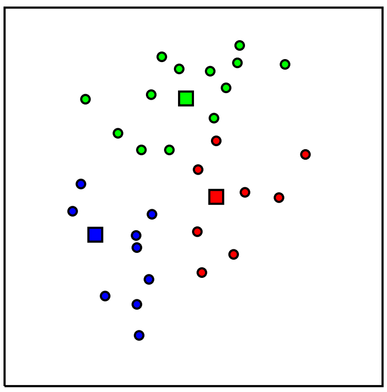
Die ersten beiden Fragen sind der Standortplanung im engeren Sinne zuzurechnen und werden im Rahmen jedes Standortproblems gelöst, während die letzten beiden Fragen im Rahmen der Standortplanung im weiteren Sinne behandelt werden. Wird im Rahmen eines Standortplanung der Bedarf nicht den einzelnen Standorten zugeordnet, werden auch die beiden letzten Fragen nicht beantwortet. In diesem Fall wird nur das Location-Problem, nicht aber das Allocation-Problem gelöst.
Die verschiedenen Aspekte der Standortplanung sind miteinander verflochten. So hängt etwa die Wahl der Standorte von der Anzahl der auszuwählenden Standorte ab, da die n-Mediane im Allgemeinen keine Teilmenge der (n+1)-Mediane ist. Weiterhin sinkt die Größe einer Facility mit steigender Anzahl, da der Bedarf auf mehr Standorte verteilt wird. Nicht immer wird den Standorten ein gleich großer Bedarf zugeteilt. Daher spielt bei der Wahl ihrer Größe die Zuteilung des Bedarfs eine wichtige Rolle.

2.1.1 Entwicklung der Standortplanung

Pierre de Fermat formulierte bereits im 17. Jahrhundert das Problem, innerhalb eines Dreiecks ABC einen Punkt D zu finden, so dass die Summe der Abstände zwischen diesem Punkt und den Eckpunkten des Dreiecks minimal wird [DKSW95]. Nach dieser Formulierung wird das Problem auch Fermat-Problem genannt, auch wenn nicht geklärt ist, ob es frühere Formulierungen gibt. Während das Problem mehr als 200 Jahre theoretisch blieb, fand Alfred Weber 1909 eine Anwendung der verallgemeinerten Version des Problems mit gewichteten Entfernungen für die Standortplanung. In seiner Formulierung,



(a) Auswahl der Standorte



(b) Zuordnung der Kunden

Abbildung 1: Standortplanung mit Location-Phase (a) und Allocation-Phase (b)

Weber-Problem genannt, wird der optimale Standort einer Industrie gesucht, so dass die Transportkosten minimiert werden. Diese ergeben sich dadurch, dass die zwei Rohstoffe der Produktion an je einem festen Ort zur Verfügung stehen und das zu produzierende Gut zu einem festen Markt transportiert werden muss. Später wurde das Weber-Problem auf mehrere Basisgebiete und Standorte erweitert. Der Varignon'sche Apparat ist eine mechanische Umsetzung des Weber-Problems mit einem Standort. Der Apparat besteht aus einer Platte, in die an den Punkten Löcher gebohrt werden, die den Basisgebieten des Problems entsprechen. Durch diese Löcher werden Schnüre geführt, an denen Gewichte befestigt werden, deren Masse den Gewichten im Weber-Problem entspricht. Werden die Schnüre auf der Platte verknotet, wird dieser Knoten durch die Gewichte zu dem Punkt gezogen, der dem gewichteten geometrischen Median entspricht.[Wes93]
Das Weber-Problem erlaubt aufgrund seiner Modellierung lediglich die Berechnung des (gewichteten) geometrischen Medians. Damit ist auch die Zielfunktion festgelegt: minimiert wird die gewichtete Summe der Distanzen vom Standort (dem geometrischen Median) zu den Kunden, im Fall des klassischen Weber-Problems sind die Kunden die Orte der Rohstoffe und des Marktes. Erst Hakimi führte 1964 mit dem Zentrum (engl. center) eine andere Zielfunktion ein [Dre95a]: das (absolute) Zentrum ist der Punkt mit minimalem (gewichteten) maximalen Abstand eines Kunden zum Standort [Hak64, Dre95a]. [Hak83] untersuchen Standortplanung im Wettbewerb mit bereits bestehenden, aber konkurrierenden Standorten. „Schädliche“ (engl. (ob-)noxious) oder abstoßende Facilities wurden erstmals durch [GD75, CG78] bearbeitet. Für diese Art von Facilities werden Standorte gesucht, die im Rahmen der Restriktionen möglichst weit von den Kunden entfernt sind. Ein Beispiel für eine solche „schädliche“ Facility ist eine Mülldeponie, die möglichst weit entfernt von bewohnten Gebieten liegen soll.
Die Begriffe bedingtes Zentrum (engl. conditional center) und bedingter Median (engl. conditional median) wurden durch [Min80] eingeführt. Diese Begriffe berücksichtigen bereits vorhandene Facilities. Dabei ist das bedingte Zentrum ist der Punkt, für den die maximale bedingte Distanz minimal ist. Die bedingte Distanz ist definiert als das Minimum der Distanz zum möglichen Standort und den vorhandenen Standorten. Entsprechend ist der bedingte Median derjenige Punkt für den die Summe der bedingten Distanzen minimal ist. Diese Begriffe wurden von [Dre89] auf mehrere zu wählende bedingte Zentren und [Dre95b] auf bedingte Mediane erweitert.

2.1.2 Eigenschaften von Standortproblemen

Die zahlreichen Anwendungen der Standortplanung und die umfangreiche Forschung in diesem Bereich haben dazu geführt, dass in der Literatur viele unterschiedliche Arten an Standortproblemen beschrieben wurden. **(author?)** [Das95] geben eine Klassifizierung der verschiedenen Problemstellungen an. Die Unterschiede der einzelnen Problemformulierungen resultieren aus der Modellierung des Bedarfs und der Standorte, der verwendeten

Metrik und der Anzahl der Ziele.
Je nachdem wo der Bedarf anfällt, wird zwischen planaren, Netzwerk- und diskreten Standortproblemen unterschieden. Bei planaren Problemen wird der Bedarf durch eine stetige, zweidimensionale Funktion modelliert und fällt somit auf der gesamten Fläche an, die Gegenstand der Planung ist. Im Gegensatz dazu entsteht der Bedarf sowohl bei Netzwerk- als auch bei diskreten Problemen an bestimmten Punkten. Die Anzahl dieser Punkte ist endlich[Das95]. Im Unterschied zu diskreten Problemen entsprechen die Punkte bei Netzwerkstandortproblemen den Knoten eines Graphen. Die Wahl der Standorte kann ebenso auf die ganze untersuchte Fläche oder nur diskrete Punkte bzw. Knoten eines Netzwerks eingeschränkt sein. Hierbei ist zu beachten, dass ein diskretes Problem einfach in ein Netzwerkproblem umgewandelt werden kann, indem die (euklidischen) Distanzen zwischen einigen oder allen Punkten berechnet und diese Werte als Kantengewichte betrachtet werden. Das nächste Unterscheidungsmerkmal ist die Modellierung, wie sich der Bedarf verhält. Ist der Bedarf während des betrachteten Planungshorizontes veränderlich, wird das Modell dynamisch genannt, ansonsten statisch. Entsprechend verhält sich der Bedarf bei probabilistischen Modellen stochastich und bei deterministischen Modellen deterministisch. Weiterhin ist es in elastischen Modellen möglich, dass der Bedarf von der Entfernung zur nächsten Facility abhängt. So ist es möglich, dass er umso größer ist, je kürzer die Distanz zwischen Basisgebiet und Facility ist, da es in diesem Fall einfacher ist, diese zu nutzen. In einem inelastischen Modell ist die Entfernung zur Facility kein Bestimmungsfaktor für die Höhe des Bedarfs. In einem Ein-Produkt-Modell wird nur eine Art an Bedarf modelliert, in Mehr-Produkt-Modellen entsprechend mehrere, deren Bedarf unabhängig voneinander sein kann.
Ein weiteres Merkmal eines Standortproblems ist die Modellierung der Facilities. Eine wichtige Eigenschaft ist die Anzahl der auszuwählenden Standorte. Ist sie exogen bestimmt, heißen die Probleme je nach verwendeter Zielfunktion p-Median- oder p-Center-Probleme. Sollen zu bereits bestehenden Standorten weitere hinzugefügt werden, werden die Probleme (p,q)-Median- und (p,q)-Center-Probleme genannt [Dre89, Dre95b]. Bei Überdeckungsproblemen (engl. set-covering-problems) wird die Anzahl der Standorte erst während der Planung bestimmt. Weiterhin wird noch zwischen Single- und Multi-Facility-Problemen unterschieden, wobei die algorithmische Komplexität der Single-Facility-Probleme geringer ist und sie oft durch Algorithmen mit polynomieller Laufzeit exakt gelöst werden können. Den einzelnen Facilities kann eine Kapazität zugewiesen werden. In diesem Fall muss durch die Zuweisung des Bedarfs sichergestellt werden, dass die Kapazität nicht überschritten wird. Bei einem hierarchischen Modell werden die Facilities nach ihrer Art in verschiedene Ebenen gegliedert. Zur Komplexitätsreduktion kann dieser Fall durch eine mehrstufige Planung behandelt werden. Desweiteren können Wechselwirkungen zwischen den einzelnen Standorten auftreten.
Neben der verwendeten Metrik zur Bestimmung von Entfernungen unterscheiden sich Standortprobleme auch die Art und Anzahl der Ziele. Bei mehreren Zielen muss im all-

gemeinen zwischen diesen abgewogen werden. Die Art der Ziele ist häufig implizit durch die Modellstruktur gegeben. So wird in einem Median-Problem immer die (gewichtete) Summe der Entfernungen zu den Standorten minimiert. Die verschiedenen Zielfunktionen werden im nächsten Abschnitt ausführlich behandelt.

2.1.3 Zielfunktionen in der Standortplanung

Die vielfältigen Anwendungen der Standortplanung erfordern Modelle für die unterschiedlichsten Arten von Facilities, da der optimale Standort vom Zweck der Einrichtung abhängt. Diese Unterschiede werden in den Modellen durch die Zielfunktionen abgebildet. [Dre95a] gibt einen Überblick über Zielfunktionen, die in der Standortplanung verwendet werden. Grundsätzlich wird dort zwischen nützlichen (erwünschten) Facilities und schädlichen (unerwünschten) Facilities unterschieden. Nützliche Einrichtungen erbringen eine Dienstleistung, welche die Kunden möglichst in ihrer Nähe haben wollen. Der Standort einer nützlichen Facility sollte daher möglichst nahe bei den Kunden liegen, diese ziehen die Einrichtung in ihre Richtung (pull objective). Von einer unerwünschten Einrichtung gehen dagegen Störungen aus, welche die Kunden nicht in ihrer Nähe haben möchten. Ein solcher Standort sollte daher so weit entfernt von den Kunden wie möglich platziert werden, die Kunden stoßen eine entsprechende Einrichtung ab (push objective)[Dre95a]. Auch Kombinationen aus nützlicher und schädlicher Einrichtung sind denkbar (push-pull objective). Solche Einrichtungen sind in der Praxis häufiger zu finden als rein schädliche, wobei sich der Kundenkreis, für den die Facility attraktiv ist von dem Kundenkreis unterscheiden kann, für den die Facility unattraktiv ist[Kal06]. Eine Push-Pull-Zielfunktion wird im einfachsten Fall durch eine gewichtete Summe aus Push- und Pull-Zielfunktion gebildet. Eine andere Möglichkeit ist die Einführung von Mindest- bzw. Maximalentfernungen, um die zulässigen Standorte einzugrenzen, und nur eine Zielfunktion zu optimieren.
[Dre95a] erwähnen noch eine weitere Klasse von Zielfunktionen. Diese balancierenden Funktionen dienen dazu, die Distanzen zwischen Kunden und Standort möglichst gleichmäßig zu gestalten. Hinter solchen Zielfunktionen steht der Gedanke einer gerechten Anordnung der Einrichtungen. Allerdings ist der Gerechtigskeitsbegriff schwer zu quantifizieren [Dre95a]. Daher wird in dieser Arbeit nicht näher auf Zielfunktionen eingegangen, die Distanzen balancieren sollen. Aufgrund der Modularität des Recursive-Partitioning-Algorithmus stellt die Integration solcher Ziele keine große Herausforderung dar.
Push- und Pull-Zielfunktionen sind Funktionen der Distanzen zwischen Kunden und Standorten und werden für das Location-Problem, die Auswahl eines oder mehrerer Standorte, verwendet. Bei der Auswahl mehrerer Standorte, ordnen die betrachteten Zielfunktionen implizit jeden Kunden der nächstgelegenen Einrichtung zu. Für die Berechnung der Zielfunktion sind nur die Distanzen zwischen den Kunden und dem Standort, dem sie zugeordnet werden, von Bedeutung. Wie viele Kunden oder wie viel Bedarf den ein-

Name	Beschreibung		Zielkategorie
Median	Punkt minimaler Abstandssumme	Minisum	Pull
Center	Punkt mit minimaler Maximaldistanz	Minimax	Pull
Centdian	Kombination aus Center und Median		Pull
k-Center	Center der k weitest entfernten Kunden		Pull
Maxian	Punkt mit maximaler Abstandssumme	Maxisum	Push
Anticenter	Punkt mit maximaler Minimaldistanz	Maximin	Push

Tabelle 1: Wichtige Zielfunktionen der Standortplanung

zelnen Standorten zugeordnet wird, spielt hingegen keine Rolle. Den gesamten Bedarf möglichst gleichmäßig unter den Einrichtungen aufzuteilen, ist jedoch ein wichtiges Ziel des Allocation-Problems. Auf Zielfunktionen für eine möglichst gerechte Aufteilung des Bedarfs wird im Abschnitt über Gebietsplanung näher eingegangen.

Nützliche Facilities

Nützliche Einrichtungen werden so nahe wie möglich an die Kunden platziert. Hierfür werden die Distanzen zwischen Basisgebieten und und zugeordneten Standorten minimiert. Die älteste in der Standortplanung verwendete Zielfunktion in der Standortplanung ist die zur Berechnung des gewichteten geometrischen Medians.

$$f_{median} = \min_{x \in X} \{ \sum_{i=1}^{n} w_i \cdot d(v_i, x) \}, \quad n = |V| \tag{2.1}$$

Der Faktor w_i ist ein für jedes Basisgebiet individuelles Gewicht. Mediane werden häufig zur Kostenminimierung verwendet. Der Wert w_i entspricht in diesem Fall den Kosten der Versorgung des i-ten Kunden (Transportkosten etc.). Da solche Zielfunktionen die Summe der Distanzen zur Facility minimieren, wird diese Klasse auch Minisum genannt[Dre95a]. Die p-Mediane X_p einer Menge sind die p Punkte X_p, für welche die Summe der gewichteten Abstände der Kunden zum jeweils nächsten Punkt in X_p minimal ist:

$$f_{p-median} = \min_{X_p \subset X, |X_p| = p} \{ \sum_{i=1}^{n} \min_{x \in X_p} (w_i \cdot d(v_i, x)) \} \tag{2.2}$$

Bei der Berechnung der (p,q)-Mediane existieren bereits q Facilities und sollen um p neue Einrichtungen erweitert werden. Die entsprechende Zielfunktion ist:

$$f_{(p,q)-median} = \min_{X_p \subset X, |X_p| = p} \{ \sum_{i=1}^{n} \min\{ \min_{y \in Y_q} (w_i \cdot d(v_i, y)), \min_{x \in X_p} (w_i \cdot d(v_i, x)) \} \} \tag{2.3}$$

Standortprobleme mit einer Minisum-Zielfunktion werden p-Median-Probleme genannt. Während private Akteure für ihre Einrichtungen in der Regel kostenminimale Standorte suchen, werden für öffentliche Einrichtungen gerechte Standorte gesucht. In diesem Zusammenhang gilt eine Planung als gerecht, wenn der Nutzer, der am schlechtesten gestellt wird, so gut wie möglich behandelt wird. Dieses Ziel wird durch das geometrische Zentrum erreicht.

$$f_{center} = \min_{x \in X}\{\max_{i \in V}(w_i \cdot d(v_i, x)\} \tag{2.4}$$

Entsprechend den p-Medianen sind die p-Zentren die p Punkte, so dass die größte Entfernung eines Kunden zum nächsten Punkt in X_p minimal wird.

$$f_{p-center} = \min_{X_p \subset X}\{\min_{x \in X_p}(w_i \cdot d(v_i, x))\} \tag{2.5}$$

Entsprechend den Gleichungen (2.3) und (2.5) lautet die Zielfunktion für die (p,q)-Center:

$$f_{(p,q)-center} = \min_{X_p \subset X}\{\max[\min_{y \in Y_q}(w_i \cdot d(v_i, y), \min_{x \in X_p}(w_i \cdot d(v_i, x))]\} \tag{2.6}$$

Diese Zielfunktionsklasse wird auch Minimax genannt, die entsprechenden Probleme heißen p-Center-Probleme[Dre95a].
Der Centdian [Hal76] ist eine konvexe Kombination aus Median und Zentrum. Sowohl die mittlere Entfernung als auch die maximale Entfernung gehen in die Zielfunktion ein [Kal06]

$$f_{centdian} = \alpha \cdot f_{median} + (1 - \alpha) \cdot f_{center} \tag{2.7}$$

Diese Zielfunktion ist eine Verallgemeinerung der Zielfunktionen von Median und Zentrum: $\alpha = 1 \Rightarrow f_{centdian} = f_{median}$ und $\alpha = 0 \Rightarrow f_{centdian} = f_{center}$. Hier wird ein kostengünstiger Standort gesucht, wobei die maximale Distanz zu einem Kunden nicht zu groß sein sollte. Eine weitere Kombination aus Median und Zentrum ist das k-Zentrum. Hier wird ein Standort mit minimaler Abstandssumme zu den k Kunden mit der größten Entfernung gesucht [Kal06].

$$f_{k-center} = \min_{x \in X}\{\sum_{i=n-k+1}^{n} w_i \cdot d(v_i, x)\} \tag{2.8}$$

Auch das k-Zentrum verallgemeinert die Konzepte von Median und Zentrum. Wird $k = n$ gewählt, entspricht die Zielfunktion der des Medians, für $k = 1$ entspricht sie der des Zentrums.
Eine weitere Klasse von Zielfunktionen der Standortplanung sind die Abdeckungsprobleme (covering problems). Hier existieren zwei Varianten, das Max-Cover-Problem und das Min-Cover-Problem. Beim Max-Cover-Problem sollen mit einer festen Anzahl Facilities möglichst viele Kunden oder möglichst viel Bedarf abgedeckt werden. Die Reichweite einer Einrichtung wird etwa durch einen Radius bestimmt, indem alle Kunden bedient werden, oder durch eine Kapazität, wobei nur die nächstliegenden Kunden bedient werden, bis die Kapazität ausgelastet ist. Beim Min-Cover-Problem soll das gesamte Planungsgebiet durch möglichst wenig Facilities bedient werden. In diesem Fall ist die Anzahl der gewählten Standorte endogen und wird erst im Planungsverlauf festgelegt.

Schädliche Facilities

Schädliche oder unerwünschte Einrichtungen werden so weit wie möglich von den Kunden entfernt platziert. Eine mögliche Zielfunktion eines solchen Standortproblems ist die des Maxians. Der Maxian ist der Punkt, der die Summe der gewichteten Distanzen maximiert.

$$f_{maxian} = \max_{x \in X}\{\sum_{i=1}^{n} w_i \cdot d(v_i, x)\} \tag{2.9}$$

Der Maxian oder ist das Gegenstück zum Median und wird daher auch Antimedian genannt, diese Zielfunktionsklasse heißt Maxisum[Dre95a]. Offensichtlich muss die Menge der möglichen Standorte begrenzt werden, damit diese Funktion eine eindeutige Lösung liefert, da sonst der Standort unendlich weit in jede beliebige Richtung verschoben werden kann[Dre95a].
Eine weitere Zielfunktion für schädliche Einrichtungen maximiert die minimale Entfernung eines Kunden zum Standort.

$$f_{anticenter} = \max_{x \in X}\{\min_{i \in V}(w_i \cdot d(v_i, x)\} \tag{2.10}$$

Diese Problemstellung entspricht der Suche nach dem größtmöglichen leeren Kreis innerhalb der Punktemenge. Der Mittelpunkt dieses Kreises ist dann der optimale Standort für die Einrichtung. Da diese Zielfunktion das Gegenteil des Zentrums berechnet, wird der entsprechende Punkt Antizentrum genannt, die Zielfunktion wird Maximin genannt. [Dre95a]
Auch wenn nur wenige Arbeiten mehr als eine schädliche unerwünschte Einrichtung gleichzeitig betrachten, ist es möglich, entsprechend den p-Medianen und p-Zentren auch die p-Maxiane und p-Antizentren zu berechnen. Die entsprechenden Zielfunktionen sind

$$f_{p-maxian} = \max_{X_p \subset X, |X_p|=p}\{\sum_{i=1}^{n} \min_{x \in X_p}(w_i \cdot d(v_i, x))\} \tag{2.11}$$

$$f_{p-anticenter} = \max_{X_p \subset X}\{\min_{x \in X_p}[w_i \cdot d(v_i, x)]\} \tag{2.12}$$

2.1.4 Komplexität

[KH79b] liefern einen Beweis dafür, dass das p-Median-Problem für allgemeine Graphen NP-schwer ist. Dies gilt auch für einfache Netzwerkstrukturen mit einem maximalen Grad von 3. [KH79a] erbringen den Beweis der NP-Schwere auch für das p-Center-Problem auf einfachen Netzwerken (maximaler Knotengrad wieder 3). Die NP-schwere für diskrete p-Median- und p-Center-Probleme (ohne Netzwerkstruktur) wird durch [MS84] bewiesen. Set-Covering-Probleme gehören zu den 21 Problemen, für die Richard Karp die NP-Vollständigkeit nachwies [Kar72]. Gilt $P \neq NP$ existiert für Standortprobleme mit mehreren Standorten kein Algorithmus der eine optimale Lösung in polynomieller Zeit berechnen kann.
Im Gegensatz zu den Multi-Facility-Problemen ist die Berechnung des 1-Medians bzw. des

1-Centers einer Menge in polynomieller Zeit möglich. Betrachte dazu folgendes simples Verfahren. Zuerst werden in quadratischer Laufzeit alle paarweisen Distanzen der Punkte berechnet. Anschließend müssen für jeden einzelnen Punkt die Summe der Entfernungen zu allen anderen Punkten bzw. die maximale Distanz zu einem anderen Punkt berechnet werden. Dieser Schritt erfordert für jeden einzelnen Knoten lineare und daher für die gesamte Menge quadratische Laufzeit. Die Berechnung des Minimalwertes kann in linearer Zeit absolviert werden. Mit diesem Vorgehen können sowohl 1-Median als auch 1-Center mit quadratischem Aufwand berechnet werden.

2.1.5 Anwendungen

Die Standortplanung hat zahlreiche betrieblichswirtschaftliche Anwendungen, von denen die Wahl eines Produktionsstandortes die historisch bedeutendste ist. Eng verwandt mit dieser Problemstellung ist die Platzierung von Außenlagern und Zweigstellen. Auch die Planung betrieblicher Abläufe kann mit den Methoden der Standortplanung durchgeführt werden. Beispiele hierfür sind das Airline Crew Scheduling oder die Produktionsprogrammplanung in einer Fabrik [Das95]. Im öffentlichen Bereich sind nahezu alle öffentlichen Einrichtungen wie Schulen, Bibliotheken und Freizeiteinrichtungen Gegenstand der Standortplanung. Besondere Beachtung in der Literatur fand die Planung von Notfalleinrichtungen. Hierzu zählen Polizei- und Feuerwehrstationen sowie Krankenhäuser und Rettungswagen. Die Planung öffentlicher Einrichtungen wird häufig als Abdeckungsproblem formuliert, da hier eine möglichst kostengünstige Versorgung der gesamten Bevölkerung angestrebt wird.

An der Schnittstelle zwischen privater und öffentlicher Planung liegen Infrastrukturprobleme. Versorgungseinrichtungen wie Kraft- und Wasserwerke erbringen Dienste, wirken aber aufgrund der Emission von Lärm oder Abgasen unattraktiv auf Wohngebiete. Während hier zwischen kostengünstigen Standorten und den Bedürfnissen der Bürger abgewogen wird, überwiegen bei Standorten für die Müllentsorgung Push-Ziele. [Hak64] untersuchte die optimalen Standorte von Switches in Kommunikationsnetzwerken. Dies ist sowohl für öffentliche als auch betriebliche Netze von Interesse.

2.2 Gebietsplanung

Gebietsplanung bezeichnet die Aggregation kleiner Einheiten (Basisgebiete) zu übergeordneten Einheiten, den Distrikten, Territorien oder Gebieten, wobei jene vorher definierten Kriterien genügen sollen [Kal06]. Jedem Basisgebiet ist die sogenannte Aktivität zugeordnet, welche die Größe des Bedarfs dieses Basisgebiets angibt. Ein Distrikt kann ein ausgewiesenes Zentrum enthalten, welches die Versorgung dieses Distrikts übernimmt, dies muss aber nicht der Fall sein. Im Unterschied zur Standortplanung sind die Zentren, falls vorhanden, vorher bekannt und werden im Laufe der Planung nicht verändert. Die Gebietsplanung mit Zentren kann daher als Entsprechung der Allocation-Phase der

Standortplanung betrachtet werden.
Nach [HS71] wird zu Beginn der Planung jene Eigenschaft der Basisgebiete als Aktivität definiert, für die das Balancekriterium erfüllt sein soll. Der Zusammenhang eines Distriktes ist bei der Einteilung von Service- oder Verkaufsgebieten besonders wichtig, um die Erfüllung des Bedarfs durch das Zentrum zu ermöglichen oder zumindest Fahrzeiten durch andere Gebiete zu vermeiden, wird aber auch bei Wahlkreisen eingefordert.
Eine große Bedeutung für die weitere Planung hat die Frage, ob Verwaltungsgrenzen beachtet oder ignoriert werden. Sollen Verwaltungseinheiten intakt bleiben, kann dies eine mehrstufige Planung erfordern. So dürfen in einigen Staaten Wahlkreise keine Grenzen der Gliedstaaten überschreiten. In diesem Fall wird den Gliedstaaten eine Anzahl an Wahlkreisen zugewiesen, die ihrem Bevölkerungsanteil entspricht und die Einteilung der Wahlkreise erfolgt separat in jedem Gliedstaat. Auch für die Planung von Verkaufsgebieten kann ein solches Vorgehen sinnvoll sein, etwa wenn es sich bei der Verwaltungsgrenze auch um eine Sprachgrenze handelt.

2.2.1 Entwicklung der Gebietsplanung

Während politische Parteien in den USA schon seit dem frühen 19. Jahrhundert Wahlkreise einteilten, um die Anzahl ihrer Abgeordneten zu maximieren [Mar08], begann die wissenschaftliche Bearbeitung des Problems in den 1960er Jahren, nachdem das oberste Gericht der Vereinigten Staaten, der United States Supreme Court, in mehreren Urteilen eine Neueinteilung nach bestimmten Kriterien erzwang. Das Gericht forderte, dass Wahlkreise gleich viele Einwohner haben, zusammenhängend und kompakt sein sollten. Vor allem das Kompaktheitskriterium soll das sogenannte Gerrymandering verhindern. [HWS+65] entwickeln ein Verfahren für die objektive Wahlkreiseinteilung, indem sie das Problem als Warehouse-Location-Problem auffassen. Als Kompaktheitsmaß wurde die Summe der quadrierten Abstände der Bevölkerung zum geografischen Mittelpunkt des Wahlkreises verwendet. [GN70] unterscheiden zwischen Bevölkerungs- und geografischer Kompaktheit. Bevölkerungskompaktheit misst den Abstand der Bevölkerung zum Zen-

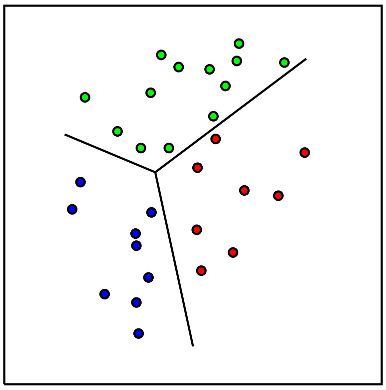


Abbildung 2: Gebietsplan mit Grenzen der Distrikte

trum des Distrikts, während geografische Kompaktheit die Form des Distrikts beschreibt. Geografische Kompaktheit kann wiederum in Distanzkompaktheit, bei der die maximale Distanz zweier Basisgebiete innerhalb eines Distrikts bestimmt wird, und Formkompaktheit, dem Verhältnis von maximaler Distanz zur Fläche, unterteilt werden.
Für die Planung von Verkaufsgebieten führen [HS71] das Aktivitätsmaß ein als die Eigenschaft, welche das Kriterium der Balance erfüllen und somit möglichst gleichmäßig auf die Gebiete aufgeteilt werden sollte. Desweiteren verweisen sie auf die Ähnlichkeiten zwischen der Einteilung von Wahlkreisen und jener von Verkaufsgebieten.
[HOR72] weist darauf hin, dass bereits die Gruppierung der Bevölkerung zu Basisgebieten voreingenommen sein kann, so dass objektive Planung unmöglich werden kann. Ein Distrikt gilt als kompakt, wenn er nahezu konvex ist. Eine kritische Betrachtung beliebter Kompaktheitsmaße findet sich in [You88]. Das Problem der Verwendung euklidischer Distanzen als Kompaktheitsmaß wird von [MJN98] hervorgehoben. Dieses Problem lässt sich auch nicht durch Verwendung gewichteter Entfernungen lösen. Daher definieren [MJN98] den Abstand zweier Basisgebiete als Netzwerkabstand, wobei zwei Gebiete dann benachbart sind, wenn sie eine gemeinsame Grenze teilen.
Die Grenzlänge zwischen den Distrikten wird von [BEL03] als Kompaktheitsmaß eingeführt. Außengrenzen der Planungsregion werden nicht betrachtet. [BLL03] definieren für die Planung von Zuständigkeitsgebieten der häuslichen Pflege als Kompaktheitsmaß den Quotienten der durchschnittlichen Entfernung zweier Basisgebiete eines Distrikts und des Quadrats des Bedarfs der Basisgebiete.
In der Geschichte der Gebietsplanung mangelt es an eigenständigen Lösungsverfahren. Gebietsplanungsprobleme wurden gelöst, indem die Problemstellung auf andere Optimierungsprobleme übertragen wurde und dann mit den entsprechenden Verfahren gelöst wurde. Ein spezifisches Verfahren zur Lösung von Gebietsplanungsproblemen mit Mitteln der algorithmischen Geometrie wird von [KNS05] vorgestellt.

2.2.2 Ziele der Gebietsplanung

Im Gegensatz zur Standortplanung ist die Lage von Standorten, falls vorhanden, während der Gebietsplanung fest. Daher existieren für die Gebietsplanung keine Zielfunktionen, die zwischen attraktiven und unattraktiven Einrichtungen unterscheiden. Wichtige Ziele der Gebietsplanung sind Balance und (geografische) Kompaktheit [Kal06]. Balance bezeichnet dabei das Ziel, dass die Distrikte bezogen auf die Aktivität möglichst gleich groß sind. Kompaktheit hingegen bezieht sich auf die Struktur eines Gebiets und ist wesentlich schwieriger zu definieren. [Kal06] beschreibt ein kompaktes Gebiet als "annähernd rund und nicht verzerrt“. Diese Definition von Kompaktheit stammt aus den beiden wichtigsten Anwendungen der Gebietsplanung, des Designs von Wahlkreisen und Verkaufsgebieten von Außendienstmitarbeitern. Beim Wahlkreisdesign erschwert die Forderung nach kompakten Gebieten das Gerrymandering, die Einteilung der Wahlkreise nach demographischen

Kriterien, um die Sitzanzahl einer Partei zu maximieren. Ein kompaktes Verkaufsgebiet soll die unproduktiven Fahrtzeiten des Mitarbeiters zwischen den Kunden minimieren.

Balance

Die Balance eines Distrikts ist definiert als die Abweichung der Aktivität dieses Distrikts vom Durchschnitt aller Distrikte, wobei die Aktivität eines Distrikts der Summe der Aktivitäten seiner Basisgebiete entspricht. Da die einzelnen Gebiete annähernd gleich groß sein sollten, ist der Zielwert ist die größte relative Abweichung eines Distriktes vom Mittelwert. Für ein Gebietsplanungsproblem mit n Distrikten $D_{i \leq n}$ ist die Balance wie folgt definiert:

$$f_{balance} = \max_i |\frac{w(D_i) - \mu}{\mu}|, \quad \mu = \frac{\sum_i^n w(D_i)}{n} \tag{2.13}$$

Sind den Zentren der Distrikte wie bei der Planung von Verkaufsgebieten Kapazitäten zugeordnet, ist es sinnvoll, Balance als gleichmäßige Auslastung der einzelnen Zentren zu definieren. Die Auslastung u_i des i-ten Zentrums und die durchschnittliche Auslastung u_μ sind nachstehend definiert.

$$u_i = \frac{w(D_i)}{c_i}, \quad u_\mu = \frac{\sum_i^n w(D_i)}{\sum_i^n c_i} \tag{2.14}$$

Daraus ergibt sich die Definition der Balance der Auslastungen zu

$$f_{balance-utilitization} = \max_i |\frac{u_i - u_\mu}{u_\mu}| \tag{2.15}$$

Kompaktheit

Während Balance und Zusammenhang einfach zu definieren und damit bei einer Lösung einfach zu kontrollieren sind, ist Kompaktheit ein schwierig zu definierendes Ziel. [Kal06] beschreibt ein kompaktes Gebiet als "annähernd rund und nicht verzerrt“ [Kal06]. Aufgrund der schwierigen Formalisierung dieses Begriffs existieren entsprechend viele Versuche, Kompaktheit zu definieren. Grundsätzlich können Bevölkerungskompakheit und geografische Kompaktheit voneinander unterschieden werden [GN70]. Bevölkerungskompaktheit bezieht sich auf die (gewichteten) Distanzen der Basisgebiete zum Zentrum des Distrikts. Hat der Distrikt wie bei der Wahlkreiseinteilung kein ausgewiesenes Zentrum, kann etwa das geografische Zentrum verwendet werden. Geografische Kompaktheit misst entweder die Ausdehnung des Distrikts oder dessen Form. Funktionen, mit denen Bevölkerungskompaktheit gemessen werden, sind häufig aus der Standortplanung bekannt. Bevölkerung ist ein spezielles Aktivitätsmaß, die allgemeine Bezeichnung ist Aktivitätskompaktheit.
[You88] beschreibt acht geografische Kompaktheitsmaße und geht insbesondere auf deren Schwächen ein. Von diesen Maßen sind die von [Reo61] sowie [Sch65] eingeführten die bedeutendsten. Das Roeck-Maß [Reo61] vergleicht die Fläche eines Distrikts mit der Fläche des kleinsten Kreises, der den Distrikt enthält. Der Wert des Maßes entspricht dem

Flächeninhalt des Distrikts geteilt durch den Flächeninhalt des umschließenden Kreises mit Radius r.

$$f_{roeck} = \frac{A}{\pi \cdot r^2} \tag{2.16}$$

Der Wert, den das Roeck-Maß dem Distrikt zuordnet, liegt daher zwischen 0 und 1. Ein Distrikt ist umso kompakter, je größer der Wert ist. Ein großes Problem dieses Maßes ist, dass Gebiete einen guten Wert erreichen, obwohl sie "innerhalb einer begrenzten Fläche hin und her mäandrieren“ [You88]. [Sch65] führt ein Maß ein, dass den Umfang eines Distriktes mit dem Umfang eines Kreises mit gleicher Fläche vergleicht. Der Wert, der dem Distrikt zugewiesen wird, ist der Umfang des Distrikts p geteilt durch den Umfang des Kreises.

$$f_{schwartzberg} = \frac{p}{2 \cdot \sqrt{\frac{A}{\pi}}} \tag{2.17}$$

Je kleiner der Wert, desto kompakter ist das bewertete Gebiet. [You88] kritisiert, dass selbst kompakte Distrikte vom Schwartzberg-Maß schlecht bewertet werden, wenn der Umfang zu lang, was aufgrund der Berücksichtigung administrativer Grenzen möglich ist. Letzlich werden alle untersuchten Maße durch [You88] abgelehnt, was zeigt, wie schwierig die Definition eines zweckmäßigen Kompaktheitsmaßes ist. Dabei ist zu beachten, dass Kompaktheit lediglich ein Metaziel ist, dass verwendet wird, um ein anderes Ziel zu erreichen. Bei der Wahlkreiseinteilung ist dieses Ziel die gerechte Zuordnung der Wähler zu den einzelnen Wahlkreisen. Das Ziel der Gestaltung von Außendienstbezirken ist die Minimierung der unproduktiven Fahrtzeit der Mitarbeiter[HS71].

2.2.3 Anwendungen

Die in der Literatur bedeutendsten Anwendungen der Gebietsplanung sind die Einteilung von Wahlkreisen und Bezirken von Außendienstmitarbeitern. Die Wahlkreisplanung ist dabei ein Beispiel für Gebietsplanung ohne ausgewiesene Zentren in den Distrikten. Ein weiteres Beispiel ist die Einteilung eines Stromnetzes in Regionen, die dann privatisiert werden sollen. Im Gegensatz dazu haben die Bezirke von Außendienstmitarbeitern ein solches Zentrum, den Arbeits- oder Wohnort des zuständigen Mitarbeiters. Eine Schule ist ebenso das Zentrum eines Schulbezirks wie Ämter für ihre zuständigen Bezirke. Hier ist eine Gebietsplanung dann notwendig, wenn die Wahl der Schule oder des amtlichen Stelle den Bürgern nicht freigestellt ist. Ein Beispiel hierfür sind Stimmbezirke, in denen die Wähler ein bestimmtes Wahllokal nutzen müssen.

[Kal06] erwähnt die mehrstufige Gebietsplanung von Winterdiensten, Müllentsorgung und den Notfalldiensten Polizei, Feuerwehr und Rettungsdiensten. Hierbei wird auf der oberen Ebene die Planungsregion in Gebiete unterteilt, in denen jeweils ein Depot, eine Polizei- oder Feuerwache oder ein Krankenhaus liegt. Auf dieser Ebene findet eine Gebietsplanung mit Zentren statt. Auf der unteren Ebene werden die Zuständigkeiten innerhalb der Gebiete im Rahmen einer Gebietsplanung ohne Zentrum auf die einzelnen Fahrzeuge verteilt. Sind die Standorte der Zentren festzulegen, handelt es bei der Planung auf der oberen Ebene um Standortplanung.

2.3 Einordnung dieser Arbeit

Gegenstand dieser Arbeit ist die Weiterentwicklung des von [Kal06] entwickelten Verfahrens zur Gebietsplanung, das in einfacher Weise zur Standortplanung verwendet werden kann. Ziel ist ein Verfahren zur Planung zusätzlicher Standorte unter Beachtung bereits vorhandener Einrichtungen. Dabei sollen einerseits die Standorte der neu einzurichtenden Facilities ausgewählt werden. Zum anderen sollen die Distrikte der einzelnen Einrichtung gebildet, die Basisgebiete den Facilities zugewiesen werden. Den einzelnen Facilities ist keine Kapazität zugewiesen.
Dabei werden diskrete Standortprobleme betrachtet. Der Bedarf fällt nur an bestimmten Punkten an und ist im Voraus bekannt und unabhängig von der Entfernung zum nächsten Standort. Da nur eine Planungsperiode betrachtet wird, ist der Bedarf statisch. Weiterhin wird ein Ein-Produkt-Modell verwendet, das Aktivitätsmaß besitzt somit nur eine Dimension. Die Zahl der auszuwählenden Standorte ist exogen vorgegeben und die Menge der möglichen Standorte entspricht der Menge der Basisgebiete. Die Anzahl der durch die Planung verfolgten Ziele ist variabel, wobei für die Bewertung des Algorithmus mehrere Ziele verwendet werden.

3 Recursive Partitioning

Ein Gebietsplanungsproblem besteht aus einer Menge an Basisgebieten und einer Zahl einzuteilender Distrikte. Für eine Lösung müssen Basisgebiete so auf die Distrikte aufgeteilt werden, dass die Kriterien Balance, Zusammenhang und Kompaktheit erfüllt werden. Eine äquivalente Formulierung des Problems besagt dass die Basisgebiete zu größeren Distrikten zusammengefasst werden, wobei ein Distrikt eine Menge an Basisgebieten repräsentiert. Der von [Kal06] entwickelte Recursive-Partitioning-Algorithmus (Rec-Part-Algorithmus) nutzt die zweite Formulierung des Gebietsplanungsproblems und verwendet einen Teile-und-Herrsche-Ansatz zur Lösung.
Die ursprüngliche Probleminstanz wird so lange in kleinere Teilprobleme unterteilt, bis ein solches Teilproblem nur noch aus einem zu bildenden Distrikt und einer Liste an Basisgebieten besteht. Diese Basisgebiete bilden dann in natürlicher Weise einen Distrikt. Für ein Problem mit Zentren muss deren Anzahl mindestens so groß sein muss wie der Anzahl der Distrikte. Die Ausgabe des Algorithmus ist eine Liste an Distrikten, wobei jeder Distrikt eine Liste der ihm zugeordneten Basisgebiete und Einrichtungen enthält. Ein Distrikt besteht dabei immer aus mindestens einem Basisgebiet. Handelt es sich um ein Problem mit Zentren, enthält ein Distrikt auch mindestens eine Einrichtung.
Der Recursive-Partitioning-Algorithmus ist durch die Trennung von Modellierung und

Problemstellung sehr flexibel verglichen mit anderen Verfahren der Gebietsplanung. Die Berechnung und Bewertung der Partitionen ist unabhängig von der Bewertung der Lösung. Hierdurch können für die Aufteilung eines Problems in seine Teilprobleme und die Bewertung der berechneten Lösung verschiedene Kriterien verwendet werden. Weiterhin ist der Algorithmus modular aufgebaut. So können alternative Verfahren für die Berechnung der Partitionen ebenso einfach in den Algorithmus integriert werden wie neue Maße für die Bewertung von Qualität und Gültigkeit der Partitionen. Auch die Reihenfolge, in der die einzelnen Teilprobleme bearbeitet werden, kann von der hier verwendeten Breitensuche abweichen.

3.1 Partitionen

Der Recursive-Partitioning-Algorithmus erzeugt für ein (Teil-)Problem verschiedene Aufteilungen, die Partitionen, die nach verschiedenen Kriterien bewertet werden. Die Bewertungen werden gewichtet und zu einem Wert zusammengefasst, der die Qualität einer Partition beschreibt. Anhand dieser Werte werden die Partitionen sortiert und aus der Aufteilung mit der besten Bewertung werden zwei Teilprobleme erzeugt.

Definition 1 (Partition)**.** Eine Partition $P = (B_l, B_r)$ einer Punktemenge B ist ein 2-Tupel, sodass $B_l, B_r \neq \emptyset$, $B_l \cup B_r = B$, $B_l \cap B_r = \emptyset$.

Die beiden Mengen B_l, B_r werden linke und rechte Seite bzw. Hälfte der Partition genannt. Eine Partition, deren Seiten durch eine Gerade voneinander getrennt werden können, wird als Line Partition bezeichnet. Während die Zahl der möglichen Partitionen einer Punktemenge B exponentiell mit $|B|$ wächst, ist die Anzahl der Line Partitions durch $|B|$ beschränkt [Kal06]. Die Gerade, die eine Line Partition in zwei Seiten trennt, wird Linie oder Trennlinie genannt.

Definition 2 (Problem)**.** Ein Problem $P = (B, q)$ ist definiert durch eine Punktemenge B und einer Zahl $q \in \mathbb{N}^+$, welche die Anzahl der Distrikte beschreibt, in die B aufgeteilt werden soll.

Ein Problem beschreibt eine Gebietsplanungsinstanz. Die Menge B repräsentiert die Menge der Basisgebiete. Jedem Punkt $p \in B$ wird dazu eine Aktivität zugeordnet. Ein Problem $(B, 1)$ stellt einen Distrikt dar.

Definition 3 (Partition eines Problems)**.** Eine Partition $P = (B_l, B_r, q_l, q_r)$ eines Problems $PP = (B, q)$ besteht aus einer Partition (B_l, B_r) sowie zwei Zahlen $q_l, q_r \in \mathbb{N}^+$, $q_l + q_r = q$.

Die Zahl $q_l (q_r)$ gibt an, in wie viele Distrikte die linke Seite $B_l (B_r)$ der Partition aufgeteilt werden soll. Eine Partition unterteilt ein Problem in zwei Teilprobleme, die analog zu den beiden Seiten der Partition linkes und rechtes Teilproblem oder linker und rechter Nachfolger genannt werden. Ist $P_l = (B_l, q_l)$ $[P_r = (B_r, q_r)]$ Nachfolger des Problems $P = (B, q)$, so heißt P Elternproblem von P_l $[P_r]$.

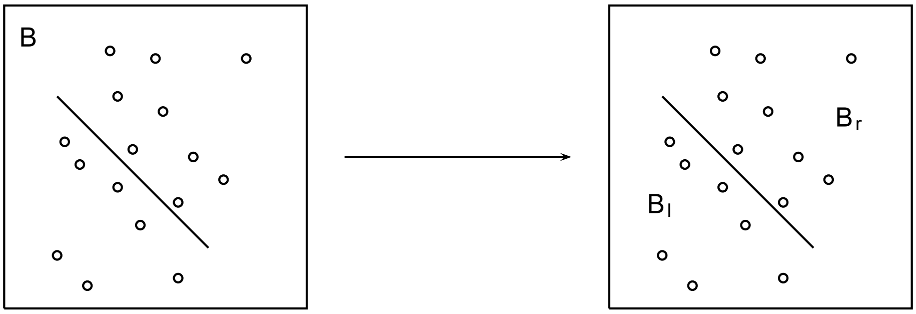


Abbildung 3: Line Partition eines Problems mit Basisgebieten B in zwei Teilprobleme mit Basisgebieten B_l und B_r

Ein Problem gilt als gelöst, wenn seine Nachfolger erzeugt wurden. Dies bedeutet allerdings nicht, dass eine Lösung berechnet werden kann (siehe Abschnitt Backtracking). Durch die Aufteilung der Probleme in je zwei Teilprobleme wird eine Binärbaumstruktur induziert. Die Wurzel dieses Baumes bildet das ursprünglich zu lösende Problem, welches daher auch Root-Problem genannt wird. Der gesamte Baum wird Problem-Baum genannt.

Definition 4 (Suchrichtung)**.** Eine Suchrichtung k ist ein Winkel α, um den die Koordinaten der Punkte einer Punktmenge B gedreht werden.

Nach der Koordinatentransformation werden die Punkte anhand ihrer transformierten x-Koordinaten sortiert.
Für die Aufteilung der Distrikte eines Problems existieren $q - 1$ Möglichkeiten. Der Rec-Part-Algorithmus betrachtet nur Partitionen mit gleich vielen Distrikten auf beiden Seiten. Für eine gerade Anzahl q werden folglich beiden Seiten jeweils die Hälfte der Distrikte zugeordnet ($q_l = q_r = \frac{q}{2}$). Ist die Anzahl der Distrikte eines Problems ungerade, werden zwei Aufteilungen betrachtet. Die erste Aufteilung ordnet der linken Hälfte $q_{l1} = \lfloor \frac{q}{2} \rfloor$ Distrikte zu. Die zweite Aufteilung weist der linken Hälfte $q_{l2} = \lceil \frac{q}{2} \rceil$ Distrikte zu. Dementsprechend werden der rechten Seite $q_{r1} = \lceil \frac{q}{2} \rceil$ bzw. $q_{r2} = \lfloor \frac{q}{2} \rfloor$ zugewiesen. Aufgrund dieser Aufteilung der Distrikte sind die Teilprobleme einer Ebene des Problem-Baums etwa gleich groß.
Der Rec-Part-Algorithmus erzeugt Partitionen für ein Problem balanceorientiert. Zunächst werden die Basisgebiete entsprechend einer Suchrichtung sortiert. Anschließend wird durch die sortierte Liste iteriert und die Aktivität der einzelnen Basisgebiete summiert, bis diese Summe einen Zielwert *thr* überschreitet. Der Zielwert wird so gewählt, dass die durchschnittliche Aktivität pro Distrikt auf beiden Seiten möglichst gleich groß ist. Für gerade

q wird nur ein Zielwert benötigt:

$$thr = \frac{\sum_{i \in B} w(i)}{2}$$

Ist die Anzahl der Distrikte q ungerade, werden zwei Partitionen erzeugt und es werden zwei Zielwerte berechnet:

$$\begin{aligned} thr_1 &= \sum_{i \in B} w(i) \frac{\lfloor \frac{q}{2} \rfloor}{q} \\ thr_2 &= \sum_{i \in B} w(i) \frac{\lceil \frac{q}{2} \rceil}{q} \end{aligned}$$

Sei i der Index des Basisgebiets, für das die Summe der Aktivitäten der betrachteten Basisgebiete den Zielwert thr übersteigt ($i \in B : \sum_{j<i} w(j) < thr, \sum_{j \leq i} w(j) \geq thr$). Dann werden die beiden Seiten der Partition wie folgt gebildet:

$$\begin{aligned} B_l &= \{j | j < i\} \\ B_r &= \{j | j > i\} \end{aligned}$$

Das Basisgebiet i wird dem Teilproblem zugeordnet, dass die Balance der Partition minimiert wird. Falls $\sum_{j \leq i} w(j) - thr > \frac{w(i)}{2}$ wird es der linken Hälfte zugeordnet, andernfalls der rechten.

Im nächsten Schritt des Algorithmus werden die berechneten Partitionen auf ihre Gültigkeit hin überprüft. Eine Partition ist ungültig, wenn die vorgegebene Toleranz für die Balance einer der beiden Seiten überschritten wird oder wenn einer Seite weniger Basisgebiete als Distrikte zugeordnet wurden. Die Toleranz der Balance bezieht sich auf die durchschnittliche Aktivität aller Distrikte und nicht nur derjenigen des aktuellen Teilproblems.

Gültige Partitionen werden nach Balance und Kompaktheit bewertet. Die gewichtete Summe der einzelnen Maße bildet die Bewertung der Partition. Da die einzelnen Maße Werte stark unterschiedlicher Größenordnung zurückgeben können, werden die Werte auf das Intervall $[0, 1]$ projiziert. Stellt ein Bewertungsmaß eine Maximierungsaufgabe da, z.B. das Roeck-Maß, werden die Werte innerhalb des Intervalls invertiert, so dass eine Partition umso besser bewertet wird, je kleiner die gewichtete Summe der einzelnen Bewertungen ist. Aus der Partition mit dem besten Score werden die beiden Teilprobleme erzeugt. Die Funktion ErzeugeTeilprobleme((B, q)) stellt die Funktionalität für die Berechnung der Partitionen, ihre Bewertung sowie die Erzeugung der Nachfolger bereit.

3.2 Ablauf des Algorithmus

Bevor der Algorithmus das Root-Problem lösen kann, wird überprüft, ob eine Lösung grundsätzlich berechenbar ist. Zu den Fällen, in denen keine Lösung berechnet werden kann, gehören:

```
Funktion ErzeugeTeilprobleme(B, q)
  Eingabe : Basisgebiete B, Anzahl Distrikte q
  Ausgabe : p = (B_l, B_r, q_l, q_r)] : B = B_l ∪ B_r, q = q_l + q_r
  foreach k ∈ K do
  |  Sortiere Basisgebiete
  |  P ← ErzeugePartitionen((B, q))
  |  foreach p ∈ P do
  |  |  berechne Gültigkeit
  |  |  berechne Bewertung
  |  end
  end
  Wähle p_best : ∀p ∈ P : p_b ⪰ p
  Gebe p_best aus
```

- $q = 0$
- $q > |B|$
- Anzahl Suchrichtungen $K = 0$
- $max\_dev = 0$

Die ersten beiden Fälle betreffen schlecht definierte Probleme. Ist die Zahl zu bildenden Distrikte 0, kann nur die leere Menge zurückgegeben werden. Übersteigt die Anzahl q der Distrikte die der Basisgebiete, ist mindestens ein Distrikt leer, enthält kein Basisgebiet. Wird die Anzahl der Suchrichtungen auf 0 gesetzt, können keine Partitionen berechnet werden und somit keine Teilprobleme erzeugt werden. Der Parameter max_dev gibt die relative Abweichung der Balance einer Partition an, die toleriert wird. Für $max\_dev = 0$ werden alle erzeugten Partitionen als ungültig eingestuft, welche die Aktivität nicht perfekt auf beide Seiten aufteilen. Da diskrete Probleme behandelt werden, ist die Wahrscheinlichkeit, dass dies möglich ist, sehr klein.
Algorithmus 3.1 verwaltet zwei Datenstrukturen, den binären Baum der Probleme und ihrer Nachfolger und eine Liste der aktiven Probleme. Ein Problem wird aktiv genannt, wenn es mehr als einen zu bildenden Distrikt enthält und noch keine Nachfolger erzeugt wurden. Die Probleme, die einen Distrikt repräsentieren ($q = 1$), bilden die Menge der Blätter des Baumes. Nachdem alle aktiven Probleme gelöst wurden, wird der Baum traversiert und alle Blätter in die Menge der Distrikte aufgenommen. Die Knoten des Baumes erhalten einen level-order-Index, welcher der Reihenfolge der Traversierung durch eine Breitensuche entspricht (Abbildung 3.4).

Innerhalb der Liste aktiver Probleme werden die Probleme gemäß ihrer Indizes im Baum

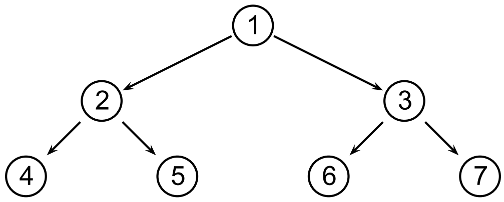


Abbildung 4: Problem-Baum mit Indizes der Teilprobleme

angeordnet, Probleme höherer Ebenen werden vor Problemen niedrigerer Ebenen eingeordnet und bearbeitet. Gelöste Probleme werden aus der Liste entfernt. Erzeugte Teilprobleme werden der Liste entsprechend ihrer Indizes hinzugefügt, es sei denn sie bilden einen Distrikt und sind Teil der Lösung ($q = 1$).
Das Verhalten des Algorithmus wird durch Steuervariablen beeinflusst, die vor dem Start durch den Nutzer festgelegt werden können.

- Anzahl der Suchrichtungen
- maximale Abweichung der Toleranz
- Gewichtsvektor für die Bewertungsmaße

Die Anzahl der Suchrichtungen wird in der Variablen K gespeichert. K muss eine positive ganze Zahl sein. Je mehr Suchrichtungen betrachtet werden, desto größer ist die Wahrscheinlichkeit einer besseren Lösung. Ebenso steigt die Laufzeit des Algorithmus, da mehr Partitionen berechnet werden müssen. Die Suchrichtungen werden gleichmäßig im Intervall $[0, \pi)$ verteilt und von 0 bis $k - 1$ nummeriert. Die Suchrichtung k bildet mit der x-Achse den Winkel $k \cdot \frac{\pi}{K}, k < K$.
Eine Eigenschaft jeder Partition ist, dass die Balance der Teilprobleme mindestens so groß wie die des Problems ist, aus dem sie erzeugt wurden. Somit wächst die Balance monoton im Problem-Baum. Daher ist es wichtig, die Balance bei der Aufteilung eines Problems zu kontrollieren. Hierfür wird die maximal erlaubte Balance max_dev verwendet. Eine Partition ist ungültig, wenn sie die maximale Toleranz überschreitet. Ungültige Partitionen werden verworfen. Kleine Werte von max_dev erhöhen die Qualität der Lösung, vergrößern aber das Risiko, dass der Algorithmus keine gültige Lösung berechnen kann.
Für die Bewertung einer Partition können mehrere Maße verwendet werden. Um aus den verschiedenen Bewertungen einen Wert zu berechnen, müssen sie gewichtet werden. Die Gewichte werden im Gewichtsvektor $wt \in \mathbb{R}^d$ gespeichert. Die Dimension d des Vektors entspricht der Anzahl der verwendeten Maße. Der Vektor wird in der Regel normiert, sodass $\sum_{i=1}^{d} wt[i] = 1$. Das Gewicht $wt[i]$ gibt dann den relativen Anteil des i-ten Maßes an der Bewertung der Partition an.

Name	Beschreibung	Wertebereich
K	Anzahl der Suchrichtungen	$\mathbb{N}^+$
max_dev	maximale Abweichung der Toleranz	$\mathbb{R}_{>0}$
wt	Gewichtsvektor der Maße	$\mathbb{R}^d$

Tabelle 2: Steuervariablen des Rec-Part-Algorithmus

Algorithmus 3.1 : Recursive-Partitioning

Eingabe : Basisgebiete B, Anzahl Distrikte q

Ausgabe : Distrikte $S = \{(B, q) | q = 1\}, \quad |S| = q,$

$A = \{(B, q)\}$ `/* Liste aktiver Probleme */`
$T = \emptyset$ `/* Problem-Baum */`
while $A \neq \emptyset$ **do**
    $(B, q) \in A$ `/* aktuelles Problem */`
    $(B_l, B_r, q_l, q_r) \leftarrow$ ErzeugeTeilprobleme((B, q))
    if $q_l > 1$ **then**
        $A \leftarrow A \cup (B_l, q_l)$ `/* Füge zu aktiven Problemen hinzu */`
    end
    if $q_r > 1$ **then**
        $A \leftarrow A \cup (B_r, q_r)$
    end
    $T \leftarrow T \cup (B_l, q_l) \cup (B_r, q_r)$ `/* Füge zu Baum hinzu */`
    $A \leftarrow A \setminus (B, q)$ `/* Entferne aus aktiver Liste */`
end
$S \leftarrow \{(B, q) | (B, q) \in T, q = 1\}$ `/* Blätter bilden die Lösung */`
Ausgabe S

3.3 Backtracking

Kann für ein Problem P keine gültige Partition berechnet werden, ist dieses Problem mit den vorhandenen Einstellungen nicht lösbar und wird verworfen. Da sich die Lösung eines Problems aus den Lösungen der Unterprobleme zusammensetzt, betrifft die Nichtlösbarkeit eines Problems auch alle übergeordneten Probleme, insbesondere das Elternproblem EP. Damit das Elternproblem dennoch gelöst werden kann, müssen andere Nachfolger erzeugt werden. Dieser Vorgang wird Backtracking genannt.

Die Funktion ErzeugeTeilprobleme() berechnet und sortiert für alle K Suchrichtungen die gültigen Partitionen. Erweist sich ein Teilproblem während Algorithmus als unlösbar, kann auf die Partition mit der nächstbesseren Bewertung zurückgegriffen werden. Aus dieser werden dann zwei neue Teilprobleme generiert. Die neuen Teilprobleme werden in den Problem-Baum aufgenommen und ersetzen die alten Nachfolger, die verworfen werden.

Probleme werden implizit entfernt, indem sie durch die neuen Nachfolger des Elternproblems ersetzt werden.
Es kann der Fall eintreten, dass alle berechneten Partitionen eines Problems zu unlösbaren Teilproblemen führen. Dann ist das Problem selbst unlösbar und löst ein Backtracking aus. Der Algorithmus betrachtet das entsprechende Elternproblem und versucht, dieses mit einer anderen Partition zu lösen.
Mehrere Backtrackings hintereinander führen dazu, dass der Algorithmus zu höheren Ebenen des Baumes zurückkehren muss. Dies führt zu der Entfernung ganzer Teilbäume. Dabei besteht die Möglichkeit, Probleme aus dem Baum zu entfernen, die bereits in der Liste, welche die aktiven Probleme speichert, enthalten sind. Für die Lösung eines Gebietsplanungsproblems ist nur das aktuelle Aussehen des Baums relevant. Verworfene Probleme tragen nicht zur Lösung bei. Die Bearbeitung verworfener Probleme wird verhindert, indem jedes Problem der Liste im Baum gesucht wird. Ist es dort nicht vorhanden, befindet es sich in einem entfernten Teilbaum und muss nicht mehr bearbeitet werden. Aufgrund des Indexes ist der Pfad, der von der Wurzel zum Problem führt bekannt. Das Problem kann daher effizient im Baum gesucht werden.
Abbildung 3.5 zeigt einen Problembaum, in dem zwei Backtracking-Schritte ausgeführt werden. Für das grüne Problem müssen neue Teilprobleme erzeugt werden. Die roten Probleme werden dadurch aus dem Baum entfernt. Die beiden eingerahmten Probleme verbleiben in der Liste der aktiven Probleme. Sie werden jedoch nicht bearbeitet, da sie im Baum nicht mehr aufgefunden werden können.

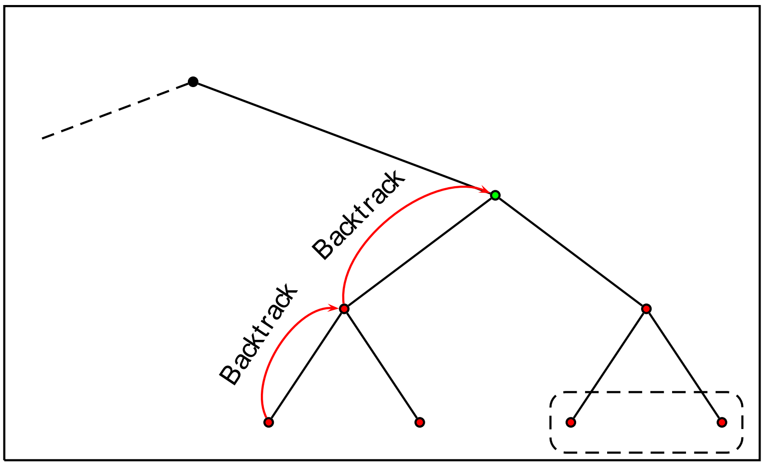


Abbildung 5: Zwei Backtracking-Schritte

Backtracking kann bis zum Root-Problem zurückführen. Löst auch das Root-Problem ein Backtracking aus, ist das Gebietsplanungsproblem mit den aktuellen Parametern nicht lösbar. In diesem Fall werden den Parametern neue, weniger restriktive Werte zugewiesen (relaxiert) und die Ausführung beginnt von neuem. Dieses Verhalten kann zu einer Endlosschleife führen, wenn das Gebietsplanungsproblem nicht gelöst werden kann. Daher wird für jeden Parameter ein Wert definiert, bis zu dem er relaxiert werden kann.

Diese Maximalwerte können durch den Benutzer festgelegt werden. Sind alle Parameter maximal relaxiert, startet der Algorithmus einen letzten Durchlauf. Ist dieser nicht erfolgreich, wird die Ausführung abgebrochen. Das Gebietsplanungsproblem gilt dann als nicht lösbar.
Die zu relaxierenden Parameter sind die Anzahl der Suchrichtungen und die erlaubte Abweichung der Balance einer Partition. Indem beide Parameter erhöht werden, können für jedes Problem mehr gültige Partitionen erzeugt werden und die Wahrscheinlichkeit, dass eine Lösung gefunden werden kann, steigt. Der Parameter *RelaxMax* gibt an, wie häufig relaxiert wird, bis den Parametern ihre Extremwerte zugewiesen werden.

3.4 Einbeziehen von Facilities

Ein Gebietsplanungsproblem kann Zentren der Distrikte enthalten. Diese werden durch den Rec-Part-Algorithmus den einzelnen Distrikten zugeordnet. Grundsätzlich werden Facilities wie Basisgebiete behandelt, jedoch gibt es Unterschiede, die im Folgenden beschrieben werden. Sei nun F eine Menge an Facilities. Die Definitionen können auf Facilities erweitert werden.

Definition 5 (Partition mit Facilities)**.** Eine Partition $P = (B_l, B_r, F_l, F_r)$ einer Punktemenge B und einer Facility-Menge ist ein 4-Tupel, sodass (B_l, B_r) eine Partition von B ist und (F_l, F_r) eine Partition von F.

Wiederum wird eine Partition mit Facilities Line Partition genannt, wenn die beiden Seiten durch eine Gerade voneinander getrennt werden können.

Definition 6 (Problem mit Facilities)**.** Ein Problem mit Facilities $P = (B, F, q)$ ist definiert durch ein Problem (B, q) und eine Facility-Menge F.

Definition 7 (Partition eines Problems mit Facilities)**.** Eine Partition $P = (B_l, B_r, F_l, F_r, q_l, q_r)$ eines Problems mit Facilities $PP = (B, F, q)$ besteht aus der Partition (B_l, B_r, q_l, q_r) des Problems (B, q) und einer Partition der Facilities (F_l, F_r).

Facilities werden wie die Basisgebiete nach der Suchrichtung sortiert. Während der Iteration durch die beiden Mengen werden Basisgebiete und Facilities so lange der linken Hälfte der Partition zugewiesen bis entweder der Zielwert der Aktivität und eine minimale Anzahl an Facilities erreicht oder eine maximale Zahl an Facilities erreicht wird.
Bei der Aufteilung soll das Verhältnis von Facilities zu Distrikten in jeder Hälfte der Partition möglichst gleich groß sein. Im einfachsten Fall gilt $f = q$ und damit $|F_l| = q_l$ sowie $|F_r| = q_r$. Weichen die Anzahl an Facilities und die Zahl der Distrikte voneinander ab, werden eine minimale und eine maximale Anzahl Facilities für das linke Teilproblem definiert.
Allgemein ist der Quotient $\frac{f}{q}$ nicht ganzzahlig und eine Aufteilung mit $\frac{f_l}{q_l} = \frac{f_r}{q_r} = \frac{f}{q}$

unmöglich. In diesem Fall wird Zuordnung der Facilities zu den beiden Seiten der Partition folgendermaßen vorgenommen. Bei einer geraden Anzahl Distrikte wird der Nachkommaanteil des Quotienten $\frac{f}{q}$ ermittelt. Ist dieser größer als 0,5, werden einem Teilproblem maximal $\lceil \frac{f}{q} \rceil \cdot \frac{q}{2}$ Facilities zugewiesen. Dementsprechend ist die minimale Anzahl Facilities in einer Hälfte $f - \lceil \frac{f}{q} \rceil \cdot \frac{q}{2}$. Ist der Nachkommaanteil kleiner als 0,5, werden einem Teilproblem mindestens $\frac{\lfloor \frac{f}{q} \rfloor \cdot q}{2}$ Facilities zugewiesen. Entsprechend ist die maximale Anzahl Facilities einer Hälfte $f - \frac{\lfloor \frac{f}{q} \rfloor \cdot q}{2}$. Tabelle 3 fasst die Berechnung der minimalen und maximalen Anzahl Facilities für die Hälften einer Partition zusammen.

$\frac{f}{q} - \lfloor \frac{f}{q} \rfloor > 0,5$	$maxFac = \frac{\lceil \frac{f}{q} \rceil \cdot q}{2}$
	$minFac = f - maxFac = f - \frac{\lceil \frac{f}{q} \rceil \cdot q}{2}$
$\frac{f}{q} - \lfloor \frac{f}{q} \rfloor \leq 0.5$	$minFac = \frac{\lfloor \frac{f}{q} \rfloor \cdot q}{2}$
	$maxFac = f - minFac = f - \frac{\lfloor \frac{f}{q} \rfloor \cdot q}{2}$

Tabelle 3: Aufteilung der Facilities bei gerader Anzahl Distrikte

Für ein Problem mit ungerader Zahl an Distrikten werden zwei Partitionen erzeugt. Für die erste Partition gilt $q_l = \lfloor \frac{q}{2} \rfloor$, $q_r = \lceil \frac{q}{2} \rceil$, für die zweite $q_l = \lceil \frac{q}{2} \rceil$, $q_r = \lfloor \frac{q}{2} \rfloor$. Die Berechnung der minimalen und maximalen Anzahl der Facilities wird für beide Partitionen analog zur Aufteilung bei gerader Anzahl Distrikte durchgeführt.

$\frac{f}{q} - \lfloor \frac{f}{q} \rfloor > 0,5$	$maxFac[0] = \lceil \frac{f}{q} \rceil \cdot \lfloor \frac{q}{2} \rfloor$
	$maxFac[1] = \lceil \frac{f}{q} \rceil \cdot \lceil \frac{q}{2} \rceil$
	$minFac[0] = f - maxFac[1] = f - \lceil \frac{f}{q} \rceil \cdot \lceil \frac{q}{2} \rceil$
	$minFac[1] = f - maxFac[0] = f - \lceil \frac{f}{q} \rceil \cdot \lfloor \frac{q}{2} \rfloor$
$\frac{f}{q} - \lfloor \frac{f}{q} \rfloor \leq 0,5$	$minFac[0] = \lfloor \frac{f}{q} \rfloor \cdot \lfloor \frac{q}{2} \rfloor$
	$minFac[1] = \lfloor \frac{f}{q} \rfloor \cdot \lceil \frac{q}{2} \rceil$
	$maxFac[0] = f - minFac[1] = f - \lfloor \frac{f}{q} \rfloor \cdot \lceil \frac{q}{2} \rceil$
	$maxFac[1] = f - minFac[0] = f - \lfloor \frac{f}{q} \rfloor \cdot \lfloor \frac{q}{2} \rfloor$

Tabelle 4: Aufteilung der Facilities bei ungerader Anzahl Distrikte

Diese Aufteilung erzielt ein ähnliches Verhältnis von Facilities zu Distrikten in beiden Teilproblemen, erlaubt jedoch kleine Abweichungen.

3.5 Standortplanung mit dem Rec-Part-Algorithmus

Der Rec-Part-Algorithmus kann auf einfache Weise so erweitert werden, dass er auch für die Standortplanung verwendet werden kann. Für die Lösung eines Standortproblems mit

p auszuwählenden Standorten werden zunächst mit dem Rec-Part-Algorithmus p Distrikte ohne Zentrum gebildet. Für jeden dieser Distrikte wird dann ein Standort entsprechend der Zielfunktion ausgewählt, somit ein Single-Facility-Standortproblem gelöst. Die Lösung eines p-Median-Problems kann mit Algorithmus 3.2 berechnet werden.

Algorithmus 3.2 : Berechnung der p-Mediane mit Rec-Part-Algorithmus

Eingabe : Basisgebiete B, Anzahl Standorte p

Ausgabe : p Standorte

$M = \emptyset$

$S \leftarrow$ Recursive-Partitioning(B,p)

for $i \leftarrow 1$ **to** p **do**

  $M \leftarrow M \cup m$, m Median von S_i

end

return M

Indem andere Zielfunktionen verwendet werden, können die entsprechenden Standortprobleme wie das p-Center-Problem gelöst werden.
Wird der Algorithmus für die Standortplanung verwendet, müssen lediglich p Single-Facility-Probleme gelöst werden. Diese können in im Gegensatz zum ursprünglichen p-Facility-Problem in polynomieller Zeit exakt gelöst werden.

3.6 Analyse

Die Laufzeit des Recursive-Partitioning-Algorithmus wird durch drei Parameter bestimmt. Diese Parameter sind die Anzahl n der Basisgebiete, die Anzahl p der zu bildenden Distrikte sowie die Anzahl K der Suchrichtungen. Die wesentlichen Operationen sind die Sortierung der Basisgebiete für jede Suchrichtung k sowie das Berechnen und Sortieren der erzeugten Partitionen für jedes Teilproblem. Die Basisgebiete müssen in jeder Ebene des Problem-Baums für jede Suchrichtung sortiert werden. Für jedes Teilproblem werden $O(K)$ verschiedene Partitionen erzeugt. Diese müssen sortiert mit Aufwand $\mathcal{O}(K \log K)$ sortiert werden. Die Anzahl der zu lösenden Teilprobleme liegt in $\mathcal{O}(p)$. Somit beträgt der Aufwand für das Sortieren der Partitionen insgesamt $\mathcal{O}(p \cdot K \log K)$. Das Sortieren der Basisgebiete hat Komplexität $\mathcal{O}(n \log n)$ und wird auf der $\log p$ Ebenen des Baumes für jeweils K Suchrichtungen durchgeführt. Damit liegt der gesamte Aufwand für das Sortieren der Basisgebiete in $\mathcal{O}(\log p \cdot K \cdot n \log n)$. Die Zeitkomplexität des Recursive-Partitioning-Algorithmus ergibt sich damit zu $\mathcal{O}(K \cdot (p \log K + \log p \cdot n \log n))$ [Kal06]. Wie in Abschnitt 6 zu sehen sein wird, werden selbst bei einer sehr harten Balance-Schranke (0.5% erlaubte Abweichung) nur bei etwa der Hälfte aller bearbeiteten Teilprobleme Backtrackings ausgelöst. Dieser Anteil kann als unabhängig von der Anzahl der Suchrichtungen angesehen werden. Er ist jedoch nicht unabhängig von der Anzahl der Basisgebiete. Je weniger Basisgebiete ein Teilproblem enthält, desto größer ist die Wahrscheinlichkeit, dass es zu einem Backtracking kommt. Jedoch ist der Anteil der Backtrackings klein genug, dass sich der asymptotische Aufwand nicht ändert.

4 Maße

Der Teile-und-Herrsche-Ansatz des Recursive-Partitioning-Algorithmus erfordert neben der Bewertung einer Lösung auch die Bewertung der Partition eines Teilproblems. Die Modularität des Algorithmus ermöglicht dabei die Verwendung verschiedener Kriterien für die Bewertung einer Partition und der Lösung. Ein Maß zur Bewertung einer Lösung kann hierbei auch zur Bewertung einer Partition herangezogen werden. Umgekehrt ist dies i.A. nicht der Fall, wenn für die Bewertung einer Partition die Trennlinie mit einbezogen wird, da ein einzelner Distrikt keine Trennlinie enthält. Im folgenden Abschnitt werden die verwendeten Bewertungskriterien vorgestellt.
Wie in Abschnitt 2.3 erwähnt, sind die wichtigsten Ziele der Gebietsplanung Balance, Zusammenhang der einzelnen Gebiete und Kompaktheit. Um die Qualität einer Partition hinsichtlich dieser Ziele bewerten zu können, werden Maße verwendet, welche den Grad der Erfüllung eines Ziels messen. Da der Zusammenhang eines Gebiets entweder gegeben ist oder nicht, wird auf eine weitere Diskussion des Maßes verzichtet, auch wenn der Zusammenhang einer Lösung ihre Bewertung beeinflusst. Balancemaße bewerten Partitionen mit gleichmäßiger Verteilung der Aktivität besser.
Kompaktheitsmaße werden in Aktivitätskompaktheitsmaße und geografische Kompaktheitsmaße unterteilt [GN70]. Erstere berechnen eine Funktion der Distanzen der Aktivität zum Zentrum, während letztere Ausdehnung, Fläche und Umfang eines Distrikts betrachten. Für die richtige Gewichtung mehrerer Maße ist entscheidend, ob es sich bei den einzelnen Maßen um Minimierungsmaße oder Maximierungsmaße handelt. Einer gut bewertete Partition werden durch Minimierungsmaße niedrige Wert, durch Maximierungsmaße höhere Werte zugeordnet.
Neben den bereits implementierten Maße wurden für diese Arbeit Maße neu entwickelt. Die neu entwickelten Maß sind das Utilization-Maß, welches die durchschnittliche Auslastung einer Facility bestimmt, das Diameter-Maß, mit dem das Ausmaß eines Distrikts gemessen wird und das K-Nearest-Neighbor-Maß, welches die durchschnittliche Entfernung eines Basisgebiets zu seinen Nachbarn approximiert. Das Maß Compactness-Epsilon-Area wurde nach einer bei [KNS05] skizzierten Idee implementiert.

4.1 Balance-Maße

4.1.1 Balance

Balance ist definiert als die relative Abweichung der Aktivität eines Distrikts von der durchschnittlichen Aktivität pro Distrikt. Enthält eine Seite einer Partition mehr als einen

Distrikt, wird die relative Abweichung der Aktivität dieser Seite von dem Produkt aus durchschnittlicher Aktivität eines Distriktes und der Anzahl der Distrikte der Seite berechnet. Eine hohe Gewichtung dieses Maßes sorgt für eine möglichst gleichmäßige Aufteilung der Aktivität auf die beiden Seiten einer Partition. Eine Partition ist balancierter, wenn ihr Balance-Wert kleiner ist, es handelt sich um ein Minimierungsmaß.

Definition Mit der durchschnittlichen Aktivität pro Distrikt für die Planungsregion mit q Distrikten $w_{avg} = \frac{\sum w(i)}{q}$ und den durchschnittlichen Aktivitäten für die linke und rechte Seite der Partition

$$\begin{aligned} w_{avg}^{l} &= \frac{\sum w(i)}{q_l}, i \in B_l \\ w_{avg}^{r} &= \frac{\sum w(i)}{q_r}, i \in B_r \end{aligned}$$

ist die Balance einer Partition wie folgt definiert:

$$f_{balance} = \max(|1 - \frac{w_{avg}^{l}}{w_{avg}}|, |1 - \frac{w_{avg}^{r}}{w_{avg}}|) \tag{4.1}$$

Aufwand Für die Berechnung dieses Maßes muss auf beiden Seiten der Partition die Summe der Aktivität aller Basisgebiete berechnet werden. Vorausgesetzt, die durchschnittliche Aktivitätssumme eines Distrikts ist bekannt, liegt der Aufwand für die Berechnung damit in $\mathcal{O}(|B|)$.

4.1.2 Auslastung

Sind in einem Gebietsplanungsproblem den Zentren Kapazitäten zugeordnet, ist das Ziel der Planung nicht unbedingt, die Aktivität gleichmäßig auf die einzelnen Distrikte aufzuteilen. Vielmehr kann es zielführender sein, die Aktivität entsprechend den Kapazitäten der einzelnen Einrichtungen auf die Distrikte aufzuteilen. Das Utilization-Maß (engl. für Auslastung) bestimmt die relative Abweichung der Auslastung eines Distrikts von der durchschnittlichen Auslastung aller Distrikte. Dieses Minimierungsmaß berücksichtigt anders als das absolute Balance-Maß die Kapazität der einzelnen Facilities. Ziel ist die gleichmäßige Auslastung der Einrichtungen, wodurch Distrikten unterschiedliche Aktivitätssummen zugeordnet werden können.

Definition Bezeichne $c(j)$ die Kapazität der Facility mit Index j und seien die durchschnittliche Auslastung u_{total} einer Einrichtung in der gesamten Planungsregion sowie die durchschnittlichen Auslastungen u_l und u_r einer Facility der linken bzw. rechten Seite einer Partition wie folgt definiert:

$$u_{total} = \frac{\sum\limits_{i \in B} w(i)}{\sum\limits_{j \in F} c(j)}, \quad u_l = \frac{\sum\limits_{i \in B_l} w(i)}{\sum\limits_{j \in F_l} c(j)}, \quad u_r = \frac{\sum\limits_{i \in B_r} w(i)}{\sum\limits_{j \in F_r} c(j)}$$

Dann ergibt sich die Definition des Utilization-Maßes einer Partition zu:

$$f_{utilization} = \max(|1 - \frac{u_l}{u_{avg}}|, |1 - \frac{u_r}{u_{avg}}|) \tag{4.2}$$

Aufwand Die Berechnung der Auslastungen erfordert die Summe der Aktivitäten wie auch jene der Kapazitäten eines Teilproblems. Daher liegt der Aufwand in $\mathcal{O}(|B| + |F|)$.

4.2 Aktivitätskompaktheit

4.2.1 Distanzsumme

Das Weber-Problem berechnet den Median einer Punktmenge. Zielfunktion ist das Minimum der gewichteten Entfernungssumme zwischen Kunden (Basisgebieten) und zu wählendem Standort. Diese Minisum-Zielfunktion lässt sich auch auf die Bewertung einer Partition anwenden. Für dieses Maß wird die Summe der Entfernungen der Basisgebiete einer Hälfte zur jeweils nächstgelegenen Einrichtung, die in der selben Hälfte liegt, berechnet. Die Entfernungssummen beider Hälften werden für die Bewertung einer Partition addiert. Dieses Maß wird für die Berechnung der p-Mediane eingesetzt. Partitionen werden schlechter bewertet, wenn sie Basisgebiete von der nächstgelegenen Einrichtung trennen. Jedes Basisgebiet, dass nicht mehr der nächstgelegenen Einrichtung zugeordnet werden kann, erhöht den Wert dieses Minimierungsmaßes. Es können verschiedene Metriken verwendet werden.

Definition

$$f_{distance-sum} = \sum_{i \in B_l} \min_{j \in F_l}(d(i,j)) + \sum_{i \in B_r} \min_{j \in F_r}(d(i,j)) \tag{4.3}$$

Aufwand Die minimalen Entfernungen zu einer gültigen Einrichtung werden in $\mathcal{O}(|B| \cdot |F|)$ berechnet. Das Aufsummieren der Werte erfolgt mit linearem Aufwand, sodass die Komplexität dieses Minimierungsmaßes in $\mathcal{O}(|B| \cdot |F|)$ liegt.

4.2.2 Maximale Distanz zur Facility

Dieses Entscheidungskriterium ist abgeleitet von der Minimax-Zielfunktion, welche die maximale Entfernung eines Basisgebiets zum nächstgelegenen bzw. dem zugeordneten Zentrum minimiert. Es kann je nach Anwendung auch die maximale Entfernung zu einem gemeinsamen Punkt, wie dem Median oder dem Center eines Teilproblems berechnet werden. Der einer Partition zugeordnete Wert ist das Maximum aus beiden Hälften. Bei

Verwendung des Minimax-Kriteriums soll das Basisgebiet, das die maximale Distanz zur zugeordneten Einrichtung hat, möglichst gut gestellt werden. Entscheidend für die Bewertung einer Parition ist nur die maximale Distanz. Ein Basisgebiet, das von der nächstgelegenen Facility getrennt wird, erhöht die Bewertung nur dann, wenn die Distanz größer als die bisherige Maximaldistanz wird.

Definition

$$f_{max-distance-to-facility} = \max(\max_{i \in B_l} \min_{j \in F_l} d(i,j), \max_{i \in B_r} \min_{j \in F_r} d(i,j)) \tag{4.4}$$

Aufwand Für jedes Basisgebiet muss zunächst die Distanz zur nächsten Facility berechnet werden. Der Aufwand für diese Berechnung steigt mit dem Produkt aus der Anzahl an Basisgebieten und Zentren und liegt in $\mathcal{O}(|B| \cdot |F|)$. Die Berechnung der maximalen Entfernung ist linear und erhöht den asymptotischen Aufwand nicht.

4.2.3 Nr-to-best-Facility

Der Rec-Part-Algorithmus berechnet normalisierte Berechnungen der einzelnen Maße, was bei bestimmten Funktionen zu Problemen führen kann. Die maximale Entfernung eines Basisgebiets zur Einrichtung ist anfällig für Ausreißer, die den Wert nach oben verzerren. Hat nur eine Partition eines Problems einen solchen Ausreißer, marginalisiert dieser Wert bei der Normalisierung die Werte der anderen Partitionen, wodurch deren Unterschiede verschwimmen. Dies hat den Effekt, dass für dieses Problem das entsprechende Maß seine Bedeutung verliert und die Reihenfolge der Partitionen nur noch von den anderen verwendeten Kriterien abhängt. Um diesen Effekt zu verhindern, wird das NrToBest-Maß verwendet, das die Zahl der Basisgebiete misst, die nicht mehr ihrer nächstgelegenen Facility zugeordnet werden können, wobei nur die Facilities des aktuellen Teilproblems betrachtet werden.

Definition

$$f_{NrToBest} = |\{x \in B_l | \min_{j \in F_r} d(x,j) < \min_{j \in F_l} d(x,j)\}| + |\{x \in B_r | \min_{j \in F_l} d(x,j) < \min_{j \in F_r} d(x,j)\}| \tag{4.5}$$

Aufwand Für jedes Basisgebiet muss die nächstgelegene Facility bestimmt werden: $\mathcal{O}(|B| \cdot |F|)$.

4.3 Geografische Kompaktheit

4.3.1 Compactness-Basic

[KNS05] definieren zwei Kompaktheitsmaße für eine Partition des Recursive-Partitioning-Algorithmus. Das CompactnessBasic-Maß definiert die Kompaktheit der Partition als Länge des Schnitts der Linie L, welche die beiden Seiten der Partition voneinander trennt, mit der konvexen Hülle C des zu zerlegenden Teilproblems. Die Linie L ist dabei senkrecht zur Suchrichtung k und verläuft durch das letzte Basisgebiet, das noch der linken Seite der Partition zugeordnet wurde. Alternativ kann die Trennlinie so definiert werden, dass das erste Basisgebiet, das der rechten Seite zugeordnet wurde, auf der Linie liegt oder diese zwischen den beiden genannten Basisgebieten verläuft. Werden andere Arten von Basisgebieten als punktförmige verwendet, muss die Berechnung angepasst werden. Da der Teil der Trennlinie, der zwischen den Schnittpunkten liegt, die Länge der Grenzen der entsprechenden Distrikte erhöht, sollte er möglichst gering ausfallen [KNS05]. Dieses Maß kann nicht zur Bewertung einer Lösung eingesetzt werden.

Definition

$$f_{compact-basic} = d(S_1, S_2) \tag{4.6}$$

mit den Schnittpunkten S_1, S_2 von Trennlinie L und konvexer Hülle der Basisgebiete C sowie der Distanz $d(x, y)$ zwischen zwei Punkten x,y. Die Wahl der Metrik ist dabei frei.

Aufwand Aus der Suchrichtung und dem letzten Punkt der linken Hälfte kann die Trennlinie wie die Distanz der Schnittpunkte in konstanter Zeit berechnet werden. Die Berechnung der konvexen Hülle einer Punktmenge B hat die Laufzeit $\mathcal{O}(|B| \log |B|)$. Die konvexe Hülle besteht schlimmstenfalls aus $|B|$ Punkten. Der Aufwand für die Berechnung der Schnittpunkte von Hülle und Linie ist dann $\mathcal{O}(|B|)$. Somit ist der Gesamtaufwand für dieses Maß $\mathcal{O}(|B| \log |B|)$.

4.3.2 Compactness-Epsilon-Area

In manchen Fällen misst das Compactness-Basic-Maß die Grenze zwischen den zwei Seiten einer Partition nur ungenau. Sind die Basisgebiete nicht gleichmäßig in der konvexen Hülle verteilt, besteht die Möglichkeit, dass Bereiche existieren, in denen die Punkte weit von der Trennlinie entfernt liegen. Die effektive Grenze der Distrikte ist dann deutlich kleiner als vom Compactness-Basic-Maß berechnet. [KNS05] skizzieren daher noch ein anderes Kompaktheitsmaß, welches für diese Arbeit implementiert wurde. Hierzu werden alle Punkte mit einem Abstand zur Trennlinie L unterhalb eines Grenzwertes orthogonal auf die Linie projiziert. Der größte Abstand zweier Abbildpunkte ergibt den Wert des Compactness-Epsilon-Area-Maßes. Genau wie beim Compactness-Basic-Maß ist es nicht möglich, mit diesem Maß Lösungen zu bewerten.

Definition

$$f_{compact-epsilon} = \max_{d(x,L),d(y,L)<\epsilon} d(x', y'), \quad x' = \pi(x), y' = \pi(y) \tag{4.7}$$

wobei π die Orthogonalprojektion auf die Linie L ist.

Aufwand Für jedes einzelne Basisgebiet ist der Aufwand für die Berechnung von Abstand zur Linie sowie für die Projektion auf diese konstant und linear für die gesamte Menge B. Der Aufwand zur Berechnung des maximalen Abstands der Abbilder ist wiederum linear für die betrachteten Punkte, welche eine Teilmenge von B darstellen. Insgesamt ergibt sich ein Aufwand von $\mathcal{O}(|B|)$.

4.3.3 Durchmesser

Das Durchmesser-Maß (engl. Diameter) wurde für diese Arbeit entwickelt, um die maximale Ausdehnung eines Distrikts zu messen. Der Durchmesser einer Punktmenge ist der maximale Abstand zweier Punkte dieser Menge und ein Maß für die Ausdehnung eines Distriktes oder Teilproblems. Daher kann er auch als Kompaktheitskriterium verwendet werden. Dabei haben kompaktere Teilprobleme und Distrikte kleinere Durchmesser. Der Wert für eine Partition entspricht dem Maximum beider Seiten.

Definition

$$f_{diameter} = \max(\max_{i,j \in B_l} d(i,j), \max_{i,j \in B_r} d(i,j)) \tag{4.8}$$

Ein ähnliches Maß berechnet für beide Seiten einer Partition die Summe aller paarweisen Distanzen der Basisgebiete. Damit wird nicht mehr nur die Ausdehnung in eine Richtung berücksichtigt.

$$f_{pairwise-distance-sum} = \sum_{i,j \in B_l} d(i,j) + \sum_{i,j \in B_r} d(i,j) \tag{4.9}$$

Aufwand Für beide Maße müssen alle paarweisen Entfernungen zwischen den Basisgebieten berechnet werden. Daher ist die Laufzeit beider Maße $\mathcal{O}(n^2)$.

4.3.4 K-Nearest-Neighbor

Dienstleistungen, die am Kunden erbracht werden, erfordern oft, dass ein Außendienstmitarbeiter mehrere Kunden in einer Tour besucht. Die Arbeitszeit eines Außendienstmitarbeiters wird maximiert, wenn seine Fahrtzeit zwischen den Kunden minimiert wird. Zusätzlich müssen noch die Fahrtzeiten vom Arbeitsplatz des Mitarbeiters zum ersten Kunden der Tour und vom letzten Kunden der Tour zurück zum Arbeitsplatz berücksichtigt werden. Das K-Nearest-Neighbor-Maß geht davon aus, dass ein Kunde während einer Tour wahrscheinlich von einem der nächstgelegenen Kunden (seinen Nachbarn) angefahren wird. Dazu werden jedem Kunden Durchschnittskosten zugeordnet, die der durchschnittlichen Entfernung zu seinen k nächstgelegen Nachbarn entspricht. Es werden jedoch nur solche Nachbarn berücksichtigt, die sich auf der selben Seite der Partition wie der Kunde selbst befinden. So sollen Aufteilungen besser bewertet werden, die möglichst wenige Basisgebiete von ihren nächsten Nachbarn trennen. Implementiert wird eine Approximation, bei der die für jedes Basisgebiet die $2k$ nächsten Nachbarn bestimmt werden. Für die Bewertung eines Teilproblems werden dann für jedes Basisgebiet nur diese Nachbarn betrachtet. Finden sich weniger als k dieser Nachbarn in der gleichen Hälfte der Partition, wird mit einem Strafterm aufgefüllt.

Definition Sei N_i die Menge der k nächsten, gültigen Nachbarn für das Basisgebiet i. Dann ist das k-nearest-neighbor-Maß definiert als:

$$f_{knn} = \sum_{i \in B} \frac{\sum_{j \in N_i} d(i,j)}{k} \tag{4.10}$$

Aufwand Zur Berechnung der k nächstgelegenen Basisgebiete müssten für jedes Basisgebiet die Distanzen zu allen anderen Basisgebieten berechnet werden. Daher steigt der Aufwand für dieses Maß quadratisch mit der Anzahl mit der Anzahl der Basisgebiete. Dieser Aufwand für die einzelnen Teilprobleme wird reduziert, indem für jedes Basisgebiet die $2 \cdot k$ nächsten Nachbarn vorberechnet werden. Bei der Bewertung einer Partition werden dann nur noch diese $2k$ Nachbarn in der jeweiligen Hälfte gesucht. Mit binärer Suche liegt der Aufwand für jeden Nachbarn in $(O)(\log n)$. Da k eine Konstante ist, die unabhängig von der Zahl der Basisgebiete in einem Teilproblem ist ist der Gesamtaufwand des Maßes $(O)(n \log n)$.

4.3.5 Distanz zur Trennlinie

Das LineDistance-Maß misst die minimale Distanz zwischen einer Einrichtung und der Trennlinie, wobei eine möglichst große Distanz angestrebt wird. Auf diese Weise soll verhindert werden, dass Distrikte so eingeteilt werden, dass die zugehörigen Zentren an ihren Rändern liegen. Ein Zentrum am Rand eines Distriktes ist nicht zielführend, da sowohl die durchschnittliche Entfernung zu den Basisgebieten als auch die maximale Entfernung unnötig hoch wäre. Die Bewertung einer Aufteilung ist die minimale Distanz einer Einrichtung einer der beiden Seiten zur Trennlinie L. Da dieses Maß eine Trennlinie zur Berechnung benötigt, kann es keine Lösungen bewerten. Als Variante können die Entfernungen der Facilities zu den Grenzen des jeweiligen Distrikts verwendet werden.

Definition

$$f_{line-distance} = \min_{j \in F} d(j, L) \tag{4.11}$$

Aufwand Der Aufwand zur Berechnung der Distanzen wächst linear mit der Anzahl der Facilities. Für dieses Maß bedeutet ein hoher Wert eine gute Beurteilung.

4.3.6 Reock-Maß

Der Reock-Test[Reo61] berechnet in einem ersten Schritt den kleinsten Kreis, der den Distrikt enthält. In einem zweiten Schritt wird die Fläche des Distrikts durch die des enthaltenden Kreises geteilt, wobei die Fläche eines Distrikts durch die Fläche der konvexen Hülle des Distrikts entspricht. Je größer der so berechnete Wert, desto kompakter ist der Distrikt. Dieses Verfahren kann auch auf Teilprobleme angewendet werden, die mehr als einen Distrikt enthalten. Für die Bewertung der Kompaktheit einer Partition wird das Minimum der Werte beider Seiten genannt, da es sich bei diesem Kriterium um ein Maximierungsmaß handelt.

Definition Seien r_l, r_r die Radien der kleinsten umschließenden Kreise des linken und rechten Teilproblems sowie A_l und A_r die Flächen der beiden Teilprobleme. Dann ist das Roeck-Maß für die Partition

$$f_{roeck} = \min(\frac{A_l}{\pi \cdot {r_l}^2}, \frac{A_r}{\pi \cdot {r_r}^2}) \tag{4.12}$$

Aufwand Die Berechnung des kleinsten Kreises, der die Punktmenge enthält, erfolgt nach der Methode von Elzinga-Hearn[EH72], die beliebige Genauigkeit in $\mathcal{O}(|B|)$ erreicht [DS87]. Der asymptotische Aufwand für die Berechnung der konvexen Hüllen ist $\mathcal{O}(|B| \log |B|)$. Die Fläche eines konvexen Polygons kann in linearer Zeit berechnet werden, wobei die Eingabe die Anzahl der Eckpunkte ist. Der gesamte Aufwand für die Berechnung des Roeck-Tests ist somit $\mathcal{O}(|B| \log |B|)$.

4.3.7 Schwartzberg-Maß

Der Schwartzberg-Test vergleicht den Umfang eines Distriktes mit dem Umfang eines Kreises gleicher Fläche. Die Bewertung eines Distrikts ergibt sich durch den Quotienten aus dem Umfang des Distriktes und dem des Kreises. Umfang und Fläche eines Teilproblems entsprechen Umfang und Fläche seiner konvexen Hülle.

Definition Das Schwartzberg-Maß für ein Teilproblem mit Umfängen der konvexen Hüllen p_l, p_r und Flächen A_l, A_r ist

$$f_{schwartzberg} = \max(\frac{p_l}{2 \cdot \sqrt{\frac{A_l}{\pi}}}, \frac{p_r}{2 \cdot \sqrt{\frac{A_r}{\pi}}}) \tag{4.13}$$

Aufwand Der aufwendigste Schritt für die Berechnung des Schwartzberg-Maßes ist die Berechnung der konvexen Hülle mit Aufwand in $\mathcal{O}(|B| \log |B|)$. Die anderen Schritte haben lineare Laufzeit.

5 Erweiterungen des Algorithmus

In diesem Abschnitt werden die Erweiterungen des Recursive-Partitioning-Algorithmus beschrieben, die im Rahmen dieser Arbeit entwickelt wurden. An eine Erweiterung wurde die Anforderung gestellt, die asymptotische Laufzeit des Algorithmus nicht zu erhöhen. Die Erweiterungen können in drei Kategorien eingeteilt werden:

- Sortierung der Basisgebiete
- Aufteilungsverfahren zur Berechnung von Partitionen
- Berechnung von Dummy-Facilities für Maße, die Facilities einbeziehen

Die Sortierung der Basisgebiete bestimmt die Reihenfolge, in der über diese bei der Berechnung einer Partition iteriert wird. Das verwendete Aufteilungsverfahren bestimmt die Zuordnung der Basisgebiete zu den Seiten einer Partition und somit wie die Teilprobleme erzeugt werden. Dummy-Facilities werden für Maße benötigt, die Funktionen der Distanzen zu den Facilities einer Partition berechnen. Die Standorte einiger Facilities werden erst festgelegt, nachdem der Algorithmus alle Distrikte berechnet hat, und sind daher während des Algorithmus unbekannt. Somit können sie bei der Bewertung einer Partition nicht berücksichtigt werden, was zu Ungenauigkeiten bei der Bewertung der Partitionen führt. Desweiteren wird die Möglichkeit vorgestellt, die verwendeten Maße während des Algorithmus zu wechseln. Die Vorberechnung bestimmter Werte zur Beschleunigung des Algorithmus wurde nicht im Rahmen dieser Arbeit implementiert, wird aber dennoch beschrieben.

5.1 Sortierungen

Die spezifische Eigenschaft einer Line Partition ist, dass die beiden Seiten der Partition durch eine Linie voneinander getrennt werden können. Diese Eigenschaft bedingt, dass die Basisgebiete entlang einer Suchrichtung sortiert werden müssen. Werden außer den Line Partitions auch andere Partitionen betrachtet, ermöglicht dies andere Sortierungen der Basisgebiete. Grundsätzlich wird jedoch an dem Verfahren der Berechnung einer Partition festgehalten. Der Algorithmus iteriert weiterhin durch die sortierten Basisgebiete und summiert die Aktivität, bis ein Zielwert erreicht wird.

5.1.1 Winkelsortierung

Die Basisgebiete eines Teilproblems können nach den Winkeln, die sie mit dem Mittelpunkt dieses Teilproblems bilden, sortiert werden. Der Mittelpunkt eines Teilproblems ist der Punkt, dessen rotierte x-Koordinaten den Mittelwert von minimaler und maximaler rotierter x-Koordinate und dessen rotierte y-Koordinaten den Mittelwert der beiden extremen y-Koordinaten bilden.

$$\begin{aligned} x_m &= \frac{\min_{i \in B}(x_{rot}(i)) + \max_{i \in B}(x_{rot}(i))}{2} \\ y_m &= \frac{\min_{i \in B}(y_{rot}(i)) + \max_{i \in B}(y_{rot}(i))}{2} \end{aligned}$$

Der Winkel α , den ein Basisgebiet i mit dem Mittelpunkt bildet, ist definiert als

$$\alpha = \arctan \frac{y_{rot}(i) - y_m}{x_{rot}(i) - x_m} \tag{5.1}$$

Für diese Sortierung wird nur der jeweilige Winkel α, den ein Basisgebiet mit dem Mittelpunkt bildet, nicht aber die Distanz zum Mittelpunkt betrachtet. Das Teilproblem wird als Kreis betrachtet und durch diese Sortierung werden gut balancierte Kreisausschnitte berechnet. Ziel dieser Sortierung sind natürlichere Grenzverläufe zwischen den einzelnen Distrikten. Die Grenze zwischen zwei Partitionen kann nicht durch eine Linie, sondern durch zwei Linien beschrieben werden. Diese Trennlinien beginnen im Mittelpunkt und bilden mit diesem den Winkel des ersten betrachteten Basisgebiets, sowie des letzten, mit dem der Zielwert überschritten wurde. Jede Suchrichtung kann als Offset verwendet werden, von dem aus die Iteration durch die nach Winkel sortierten Basisgebiete startet. So können ebenso viele Partitionen erzeugt werden, wie durch das ursprüngliche Verfahren.

5.1.2 Netzwerkdistanzen

Liegt dem Gebietsplanungsproblem ein Netzwerk zugrunde, etwa ein Straßennetzwerk, können die Basisgebiete nach ihren Netzwerkentfernungen zu einem ausgewählten Startpunkt sortiert werden. Als Startpunkte können die Basisgebiete gewählt werden, die für eine Suchrichtung die minimale rotierte x-Koordinate besitzen. Die Sortierung der Basisgebiete kann anhand gewichteter und ungewichteter Netzwerkentfernungen erfolgen. Für gewichtete Netzwerkentfernungen muss ein Algorithmus zur Berechnung der Entfernungen verwendet werden, wie der Algorithmus von Dijkstra. Die Sortierung anhand ungewichteter Netzwerkentfernung entspricht einer Breitensuche ausgehend vom Startpunkt. In beiden Fällen werden die nächsten Basisgebiete solange der linken Hälfte zugeordnet, bis der Zielwert erreicht wurde.

5.2 Aufteilungen

Die oben beschriebene Berechnung der Partitionen berücksichtigt nur die Balance. Die Aktivität wird entsprechend dem Verhältnis der Anzahl an Distrikten auf die beiden Seiten aufgeteilt. Dieses Vorgehen erzielt gut-balancierte Lösungen, Kompaktheit spielt jedoch bei der Berechnung keine Rolle. Aufteilungsverfahren, die sich an der Kompaktheit orientieren, könnten kompaktere Lösungen erzeugen. Für ein solches Verfahren muss entschieden werden, welche Definition der Kompaktheit verwendet wird. Hier werden folgende kompaktheitsorientierte Aufteilungsverfahren vorgestellt:

- BestLineDistance
- BestCompactnessBasic
- BestCompactnessEpsilon

Für jedes dieser Aufteilungsverfahren werden die Basisgebiete $|B|$ und Facilities $|F|$ eines Teilproblems nach einer Suchrichtung sortiert. Entsprechend wird für jede Suchrichtung mindestens eine Partition erzeugt. Wie bei der balanceorientierten Aufteilung werden für eine ungerade Anzahl q der zu bildenden Distrikte zwei Aufteilungen für eine Suchrichtung berechnet.
Auch die kompaktheitsorientierten Verfahren müssen die Beschränkungen bei der Zuordnung der Facilities einhalten. Seien l die Mindestanzahl sowie r die maximale Zahl an Facilities, die der linken Hälfte der Partition zugeordnet werden. Dann verläuft die Trennlinie zwischen den beiden Hälften durch ein Basisgebiete $i \in B$, für das gilt:

$$x_{rot}(F[l]) < x_{rot}(i) < x_{rot}(F[r])$$

Dabei bezeichnen $x_{rot}(i)$ die rotierte x-Koordinate des Basisgebiets mit Index i oder Facility und $F[l]$ sowie $F[r]$ die Facilities mit den Indizes l und r in der nach der aktuellen Suchrichtung sortierten Menge F. Es müssen für alle Verfahren nur die Trennlinien betrachtet werden, welche diese Eigenschaft aufweisen.

5.2.1 BestLineDistance

Das BestLineDistance-Verfahren basiert auf der Überlegung, die Facility eines Distrikts möglichst nahe dessen geografischen Zentrums zu platzieren. Die Ausmaße eines Distriktes sind während des Algorithmus nicht bekannt und es ist auch nicht festgelegt, welche Facility welchem Distrikt zugeordnet wird. Daher kann der Abstand der Facility zum geografischen Zentrum eines Distrikts nicht approximiert werden. Für konvexe Distrikte gilt jedoch, dass das geografische Zentrum von den Grenzen des Distrikts entfernt ist. Daher ist es sinnvoll, Facilities möglichst weit von allen Grenzen entfernt zu platzieren bzw. die Grenzen so zu wählen, dass die Distanz zu den Facilities maximal wird. Das Verfahren BestLineDistance ermittelt daher die Trennlinie L, die den größten minimalen Abstand zu einer der Facilities des Teilproblems hat.

$$f_{BestLineDistance} = \max_{L} \min_{j \in F} d(L, f)$$

Für jede mögliche Trennlinie wird der minimale Abstand zu einer Facility berechnet. Da die Linie L orthogonal zur Suchrichtung verläuft, müssen nur die beiden Facilities berücksichtigt werden, zwischen deren rotierten x-Koordinaten die rotierte x-Koordinate des Basisgebiets, das die Trennlinie induziert, liegt.

5.2.2 BestCompactness

Jede Trennlinie L einer ausgewählten Partition vergrößert die Länge der (inneren) Grenzen der Distrikte [KNS05]. Dabei ist nur derjenige Teil der Linie relevant, der innerhalb der Grenzen bzw. der Hülle des Teilproblems verläuft. Um die Grenzlänge zu minimieren, müssen daher solche Trennlinien gewählt werden, für welche die Distanz der beiden Schnittpunkte mit der Hülle eines Teilproblems minimal ist. Seien H die Hülle eines aufzuteilenden Teilproblems und L die Trennlinie. Dann wird für die BestCompactness-Aufteilung folgende Funktion minimiert:

$$f_{BestCompactness} = \min_{\{S_1,S_2\}=L\cap H} d(S_1, S_2)$$

S_1 und S_2 bilden hierbei die Schnittpunkte von Hülle und Trennlinie. Eine Trennlinie L, welche mehr als zwei Schnittpunkte mit der Hülle hat, erzeugt unzusammenhängende Partitionen und wird als nicht gültig bewertet. Dieser Fall kann nur eintreten, wenn die verwendete Hülle nicht konvex ist.
Die Aufteilung BestCompactnessBasic berechnet für jede der möglichen Aufteilungen den Schnitt von Trennlinie L und konvexer Hülle. Der Nachteil ist, dass die konvexe Hülle die Ausmaße eines Distriktes überschätzen kann. Eine BestCompactnessEpsilon-Aufteilung berücksichtigt daher nur die Basisgebiete innerhalb eines Streifens vorgegebener Breite beiderseits von L.

5.3 Dummy-Facilities

Einige der beschriebenen Maße berechnen Funktionen von Distanzen zu Facilities. Das Ziel des hier vorgestellten Verfahrens ist die Auswahl von Standorten zusätzlicher Facilities. Da diese Standorte erst berechnet werden, nachdem alle Distrikte eingeteilt wurden, sind sie während des Algorithmus unbekannt. Die Aussagekraft eines Maßes, welches Distanzen zu den Einrichtungen berechnet, ist durch die Nichtberücksichtigung der noch zu wählenden Standorte eingeschränkt. Daher werden nun zwei Verfahren vorgestellt, mit denen die noch zu wählenden Standorte geschätzt werden können. Beide Verfahren identifizieren Regionen eines Teilproblems, in denen relativ zu der vorhandenen Aktivität wenige Facilities vorhanden sind. In diesen Regionen werden dann sogenannte Dummy-Facilities platziert. Diese Dummy-Facilities werden anschließend für die Berechnung von Distanzen zur jeweils nächstgelegenen Facility verwendet. Die minimale Entfernung d eines Basisgebiets i zur nächstgelegenen Facility wird unter Einbeziehung der berechneten Dummy-Facilities zu

$$d = \min(\min_{j\in F}(d(i,j), \min_{j\in DF}(d(i,j))$$

Dabei bezeichnen F die Menge der (vorhandenen) Facilities und DF die Menge de berechneten Dummy-Facilities. Dummy-Facilities müssen nur für die Teilprobleme berechnet werden, für die $|F| < q$. Hierbei bezeichnet q wieder die Anzahl der zu bildenden

Distrikte. In einem Teilproblem mit $|F| = q$ werden keine neuen Standorte ausgewählt. Dementsprechend muss die Berechnung der Distanz zur nächsten Facility nicht angepasst werden.

Dummy-Facilities nach Winkel Das erste Verfahren für die Berechnung von Dummy-Facilities vergleicht die Aktivitätssummen der Basisgebiete in den Kreissektoren, die von jeweils zwei Facilities und dem Mittelpunkt eines Teilproblems gebildet werden. Im ersten Schritt werden Basisgebiete und Facilities nach den Winkeln, die sie mit dem Mittelpunkt des Teilproblems bilden, sortiert. Diese Winkel werden nach Gleichung 5.1 berechnet. Die Winkel der Facilities mit dem Mittelpunkt implizieren eine Aufteilung des Teilproblems in Kreisausschnitte. Die Aktivität jedes Basisgebiets wird nun dem entsprechenden Kreisausschnitt zugeordnet. Den Kreisausschnitten werden die Dummy-Facilities nun so zugeordnet, dass Aktivität pro Dummy-Facility in der maximal wird. Die Aktivität pro Dummy-Facility ergibt sich dabei aus der gesamten Aktivität eines Kreisausschnitts geteilt durch die Anzahl der diesem Ausschnitt zugeordneten Dummy-Facilities.

Dummy-Facilities nach Zellen Für diese Berechnung der Dummy-Facilities wird zunächst ein Raster über das Teilproblem gelegt. Für dieses Raster besteht für eine Partitionshälfte mit q Distrikten aus jeweils r Zeilen und Spalten, mit $r = \max(q, 10)$. Die Aktivität eines Basisgebiets wird der Zelle des Rasters zugeordnet, die es enthält. Für jede Facility, die in einer Zelle liegt, wird die durchschnittliche Aktivität eines Distrikts dieses Teilproblems, von der Gesamtaktivität der Zelle subtrahiert. Die Bewertung einer Zelle wird aufgrund einer gewichteten Summe der Aktivitäten einer Zelle sowie ihrer Nachbarzellen berechnet. Die Aktivität einer Zelle geht dabei mit dem vierfachen Gewicht in die Bewertung ein. Die an den vier Seiten angrenzenden Zellen erhalten das doppelte Gewicht, während die an den Ecken einer Zelle anschließenden Zellen einfach gewichtet werden. Im nächsten Schritt werden die Zellen nach ihrer Bewertung sortiert. Solange weitere Dummy-Facilities berechnet werden müssen, werden folgende Schritte durchgeführt. Die Zelle mit der besten Bewertung wird ausgewählt und eine Dummy-Facility an ihrem Mittelpunkt erstellt. Da diese Dummy-Facilitiy das Zentrum eines Distrikts repräsentiert, wird die Aktivität dieser Zelle um die durchschnittliche Aktivität eines Distriktes verringert und die Bewertung aller betroffenen Zellen entsprechend der neuen Aktivität der Zelle aktualisiert.

5.4 Maßwechsel und Vorberechnung

5.4.1 Maßwechsel

Einige der verwendeten Maße entsprechen Zielen der Standort- bzw. Gebietsplanung. Für ein Teilproblem mit nur zwei zu bildenden Distrikten ist es daher sinnvoll, die erzeugten Partitionen entsprechend einem Maß zu bewerten, dass dem jeweils verfolgten Ziel der

Planung entspricht, unabhängig davon, welche Maße vorher verwendet wurden. So kann das DistanceSum-Maß, dessen Zielfunktion der des Medians entspricht, für die Bewertung der letzten Aufteilungen verwendet, wenn ein p-Median-Problem gelöst werden soll. Ein solcher Maßwechsel könnte allgemein zu besseren Ergebnissen führen, wenn ab einer gewissen Tiefe des Baumes ein anderes Maß sich als besser herausstellt.

5.4.2 Vorberechnung

Bestimmte Berechnungen werden im Verlauf des Rec-Part-Algorithmus wiederholt verwendet. Um die Laufzeit des Algorithmus zu verkürzen, können die Ergebnisse dieser Berechnungen gespeichert und im späteren Verlauf abgerufen werden. Zu den wiederkehrenden Berechnungen gehören:

- Rotieren der Koordinaten von Basisgebieten und Facilities
- Berechnen der Distanzen von Basisgebieten und Facilities
- Berechnen der nächstgelegenen Facility eines Basisgebiets

Das Rotieren der Koordinaten erfordert die Berechnung von trigonometrischen Funktionen. Dieser Zeitaufwand kann durch einmaliges Berechnen und Speichern der Lösung reduziert werden, da die rotierten Koordinaten für jede Suchrichtung konstant sind. Da die Anzahl der Facilities in der Regel deutlich geringer ist als die der Basisgebiete ist es sinnvoll, die entsprechenden Distanzen vorzuberechnen und während des Algorithmus nur die Ergebnisse abzufragen. Zu beachten ist hierbei, dass die Anzahl der Facilities durch den Logarithmus der Basisgebiete beschränkt sein sollte, da ansonsten die asymptotische Laufzeit des Algorithmus beeinträchtigt würde. Aus diesem Grund werden auch die Distanzen der Basisgebiete zueinander nicht vorberechnet. Dies erfordert quadratischen Aufwand und verlangsamt den Algorithmus erheblich. Die asymptotische Komplexität des Algorithmus kann durch die Vorberechnungen nicht verringert werden, der Aufwand wird lediglich um einen konstanten Faktor geringer.

6 Tests

Getestet werden vier wesentliche Optionen des Algorithmus, die Anzahl der Suchrichtungen, Aufteilungsverfahren nach verschiedenen Kriterien, unterschiedliche Sortierung der Basisgebiete und die einzelnen Maße. Die verschiedenen Testkonfigurationen werden an einem Satz realistischer Problemen getestet. Dieser Satz besteht aus 43 verschiedenen Problemen mit wenigen hundert bis mehr als 20.000 Basisgebieten (siehe Tabelle 6.1). Die Anzahl der vorhandenen Facilities reicht von drei bis 130. Für jedes Problem werden fünf neue Standorte gesucht. Hiermit wird beabsichtigt, das Verhältnis von neuen zu alten Standorten möglichst stark zu variieren. Während der Test wird gefordert, dass

	Minimum	Maximum
Basisgebiete	284	21315
Facilities	3	130
Basisgebiete pro Facility (alt)	85	540
Basisgebiete pro Facility (neu)	35	417

Tabelle 5: Kennzahlen der Testprobleme

eine Facility innerhalb eines Teilproblems liegt, dem sie zugeordnet wird. Das Testsystem besteht aus einem PC mit Intel Core i5-3570K-Prozessor mit 3,4 GHz und 4 GB RAM. Das installierte Betriebssystem ist die 32-Bit-Version von Windows 7 Professional.
Die berechneten Lösungen werden nach den folgenden Kriterien bewertet:

- Balance
- Summe der Distanzen zur zugeordneten Facility
- maximale Distanz zur Facility
- maximale paarweise Entfernung innerhalb eines Distrikts
- durchschnittliche Entfernung zu den k nächsten Nachbarn
- Schwartzberg-Test

Aufgrund der stark unterschiedlichen Größen der Probleme werden bis auf die Balance und den Schwartzberg-Test alle Bewertungen für jeweils ein Problem skaliert, so dass der besten berechneten Lösung der Wert 0 und der schlechtesten Lösung der Wert 1 zugeordnet wird. Der skalierte Wert v^i_{scaled} einer Zielfunktion für Problem i ergibt sich demnach wie folgt aus dem Wert der Zielfunktion v und dem Minimum min^i und Maximum max^i:

$$v^i_{scaled} = \frac{(v^i - min^i)}{max^i - min^i} \tag{6.1}$$

Um zwei Konfigurationen miteinander zu vergleichen, können verschiedene Ansätze verfolgt werden. Sind alle Bewertungen einer Konfiguration besser als die einer Referenz, so wird die Referenz durch diese Konfiguration dominiert. Dominanz ist ein eindeutiges Kriterium: eine dominierte Konfiguration ist in jedem (getesteten) Fall schlechter als die Referenz. Dominiert keine Konfiguration die andere, müssen andere Kriterien wie Mittelwert, Median, Varianz oder Spannweite herangezogen werden.

6.1 Suchrichtungen

Die Anzahl der Suchrichtungen bestimmt, wie viele Partitionen für ein Teilproblem erzeugt werden. Wie in Abschnit 3.2 beschrieben wird für ein Teilproblem mit gerader

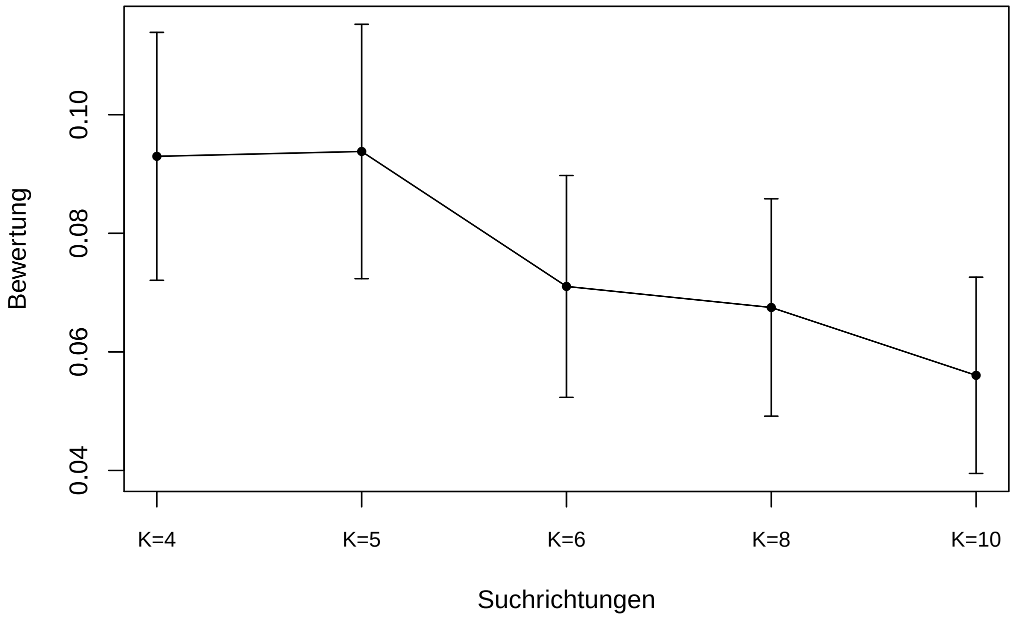


Abbildung 6: Mittelwerte und Varianzen der Balance der Lösungen für verschiedene K

Anzahl an Distrikten für jede Suchrichtung eine Partition erzeugt, bei ungerader Anzahl Distrikte werden zwei Partitionen je Suchrichtung erzeugt. Die Laufzeit des Algorithmus sollte daher ebenso mit höher Anzahl Suchrichtungen steigen wie die Qualität der berechneten Lösungen. Für die Tests werden 4, 5, 6, 8 und 10 Suchrichtungen verwendet. Für jede Suchrichtung werden alle beschriebenen Maße getestet. Als Aufteilungsverfahren wird die bestbalancierte Linie verwendet. Für alle Tests wird die maximale Toleranz zu Beginn auf 0.005 gesetzt.

6.1.1 Qualität der Lösungen

Die Balancewerte der Lösungen nehmen mit steigender Anzahl Suchrichtungen leicht ab. Abbildung 5.1 zeigt Mittelwerte und Varianzen der Balance der berechneten Lösungen für die verschiedenen Werte von K. Die fallende Tendenz der Mittelwerte ist schwach ausgeprägt, so dass lediglich der Mittelwert für $K = 10$ signifikant von den Mittelwerten für $K = 4$ und $K = 5$ unterscheidet. Zwischen den anderen Mittelwerten kann zum Signifikanzniveau von 95% kein Unterschied festgestellt werden. Auffallend ist auch das insgesamt niedrige Niveau der Balancewerte. Grund hierfür ist neben der balanceorientierten Aufteilung der Balance-Grenzwert für gültige Partitionen, so dass schlecht balancierte Partitionen abgelehnt werden. Diese Schranke wird zunächst auf 0.5% gesetzt, kann aber durch Relaxierung während des Algorithmus erhöht werden.

Suchrichtungen	K=4	K=5	K=6	K=8	K=10
Mittelwert	0.08599	0.09230	0.07193	0.06490	0.05540
Varianz	0.04879	0.05666	0.04310	0.03809	0.03285

Tabelle 6: Mittelwerte und Varianzen

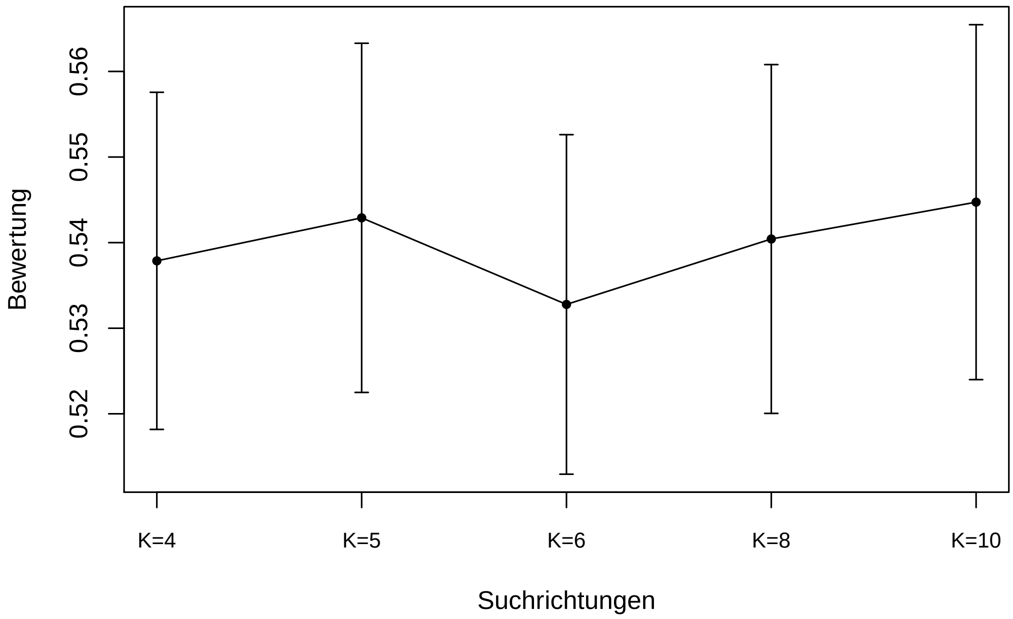
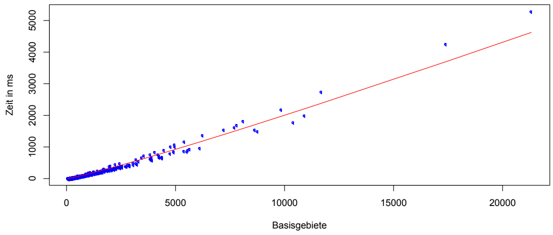


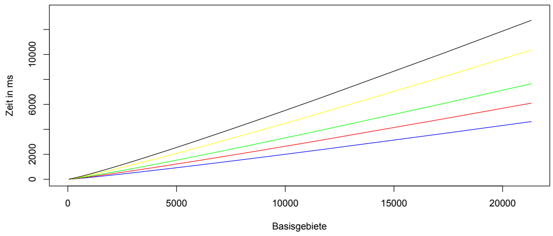
Abbildung 7: Normierte Mittelwerte und Varianzen der Median-Zielfunktion für verschiedene K

Für die anderen Bewertungen ist hingegen keine Tendenz zu erkennen. Teilweise sind die mit $K = 10$ verschiedenen Suchrichtungen berechneten Lösungen im Mittel sogar schlechter als die Lösungen für andere K.

Interessant ist in diesem Zusammenhang der Vergleich mit $K = 2$ und $K = 1$. Während sich die Bewertungen von $K = 4$ und $K = 2$ geringfügig, aber signifikant unterscheiden, können für $K = 1$ nur für drei Probleme überhaupt Lösungen berechnet werden. Um eine Lösung berechnen zu können, ist es wichtig, die Basisgebiete nach unterschiedlichen Suchrichtungen sortieren zu können. Jedoch ist die Anzahl der Suchrichtungen für die Qualität der Lösung von untergeordneter Bedeutung.



(a) Rechenzeit für $K = 4$ und Balance-Maß



(b) Rechenzeit für $K \in \{4, 5, 6, 8, 10\}$

Abbildung 8: Rechenzeit für Teilprobleme nach Anzahl der Basisgebiete

6.1.2 Details

Neben der Qualität der berechneten Lösungen hat die Anzahl der Suchrichtungen Einfluss auf die Laufzeit des Algorithmus. Der asymptotische Aufwand für den Algorithmus ist $\mathcal{O}(k \cdot \log q \cdot n \log c)$ mit k verschiedenen Suchrichtungen und den Anzahlen q der Distrikte und n der Basisgebiete. Damit ist der Aufwand zur Lösung eines einzelnen Teilproblems in $\mathcal{O}(n \log n)$. Abbildung 5.3(a) zeigt für $K = 4$ die Zeit, die für die Lösung eines Teilproblems benötigt wird, in Abhängigkeit der Anzahl der Basisgebiete des Teilproblems. Als Bewertungsmaß wurde das Balance-Maß verwendet. Die rote Regressionskurve wurde mit der Methode der kleinsten Quadrate berechnet und genügt der Gleichung $f(x) = 0.02175 \cdot n \log n$. Für alle getesteten Werte von K ergeben sich ähnliche Bilder, lediglich um den jeweiligen Faktor k verschoben (Abbildung 5.3(b)).
Ein weiteres Merkmal einer Testkonfiguration, welche neben der Rechenzeit auch die Qualität der Lösung beeinflusst, ist die Anzahl der nicht lösbaren Teilprobleme. Jedes Teilproblem, für das keine Lösung berechnet werden konnte, löst ein Backtracking aus. Tabelle zeigt die Anzahl der Backtrackings und deren Anteil an der Gesamtzahl der Probleme für jeden Wert von K. Die hohen Werte erklären sich aus der extrem niedrigen Anfangstoleranz, welche für die Partitionen nur eine Balance von maximal 0,005 erlaubt. Dieser Wert ist sehr niedrig gewählt, so dass vor allem bei Problemen mit vielen Basisgebieten Backtrackings auftreten. Dennoch kann für viele Probleme mit weniger als 1000 Basisgebieten diese Toleranz eingehalten werden. Der Anteil der Backtrackings zeigt keine größeren Unterschiede und erreicht bei $K = 8$ sein Maximum. Hier kann kein großer Unterschied zwischen den einzelnen Tests festgestellt werden. Trotz des etwa gleichbleibenden Anteils an Backtrackings sinkt die Zahl der bearbeiteten Teilprobleme mit zunehmendem K. Auch der Anteil gültiger Partitionen erhöht sich. Dies ist vor allem auf einen größeren

K	4	5	6	8	10
Teilprobleme	40232	39240	35372	33611	30082
Backtrackings	20712	20882	18935	18207	15094
Anteil Backtrackings	51,5%	53,2%	53,5%	54,2%	50,2%
Anteil gültiger Partitionen	65,3%	64,6%	68,9%	67,6%	69,4%

Tabelle 7: Anteil Teilprobleme, die Backtrack verursachen

Anteil gültiger Partitionen bei kleineren Teilproblemen zurückzuführen: je kleiner die Anzahl der Basisgebiete, desto größer ist das relative Gewicht jedes einzelnen Basisgebiets. Da die Partitionen balanceorientiert erzeugt werden, kann genau dann keine gültige Partition berechnet werden, wenn die Aktivität des trennenden Basisgebiets so groß ist, dass die Balancegrenzen nicht erfüllt werden. Da mit steigendem K mehr Sortierungen der Basisgebiete durchlaufen werden, sinkt die Wahrscheinlichkeit, dass dieser Fall eintritt.
Zusammenfassend lässt sich über die Anzahl der Suchrichtungen sagen, dass ihre Wahl den Ablauf des Algorithmus zwar beeinflusst, dieser Einfluss jedoch gering ausgeprägt ist und lediglich die Balance der Lösung beeinflusst. Der vielleicht wichtigste Grund für die Wahl eines höheren Werts für K ist daher ästhetisch-psychologisch. Aufgrund der stärker variierenden Ausrichtung der Trennlinien sehen die durch höhere K erzeugten Distrikte runder aus.

6.2 Maße

Die erzeugten gültigen Partitionen eines Teilproblems werden nach ihrer Bewertung geordnet. Die Unterprobleme werden aus der Partitition mit der besten Bewertung erstellt. Die Bewertung setzt sich aus den gewichteten Bewertungen der verwendeten Maße zusammen. Auf diese Weise beeinflusst die Wahl der verwendeten Maße den Verlauf des Algorithmus. In Abschnitt 4 wurde eine Einteilung der verwendeten Maße in vier Kategorien vorgenommen. Neben dem Balance-Maß, welches eine Kategorie für sich bildet, werden Kompaktheitsmaße in folgende Kategorien eingeteilt:

- Aktivitätskompaktheit
- Formkompaktheit
- Ausdehnungskompaktheit

Form- und Ausdehnungskompaktheit werden zur geografischen Kompaktheit zusammengefasst. Es ist zu erwarten, dass Maße einer Kategorie Lösungen erzeugen, die in der jeweiligen Kategorie gut bewertet werden.

6.2.1 Balance

Anders als bei den Suchrichtungen unterscheiden sich die Ergebnisse der verschiedenen Maße deutlich voneinander. Abbildung 5.4 zeigt die Verteilung der normierten Balance-Bewertungen für die einzelnen Maße. Erwartungsgemäß liefert das Balance-Maß die deutlich besten Ergebnisse. Sowohl Mittelwert als auch Varianz sind signifikant geringer als bei allen anderen verwendeten Maßen. Die anderen Maße unterscheiden sich nicht signifikant voneinander. Für alle Maße ist das Balance-Niveau niedrig, lediglich beim CompactnessBasic-Maß liegt der Mittelwert über 10%. Die erzeugten Lösungen sind insgesamt gut balanciert, was auf die balancorientierte Aufteilung und den niedrigen Grenzwert für gültige Partitionen zurückzuführen ist.

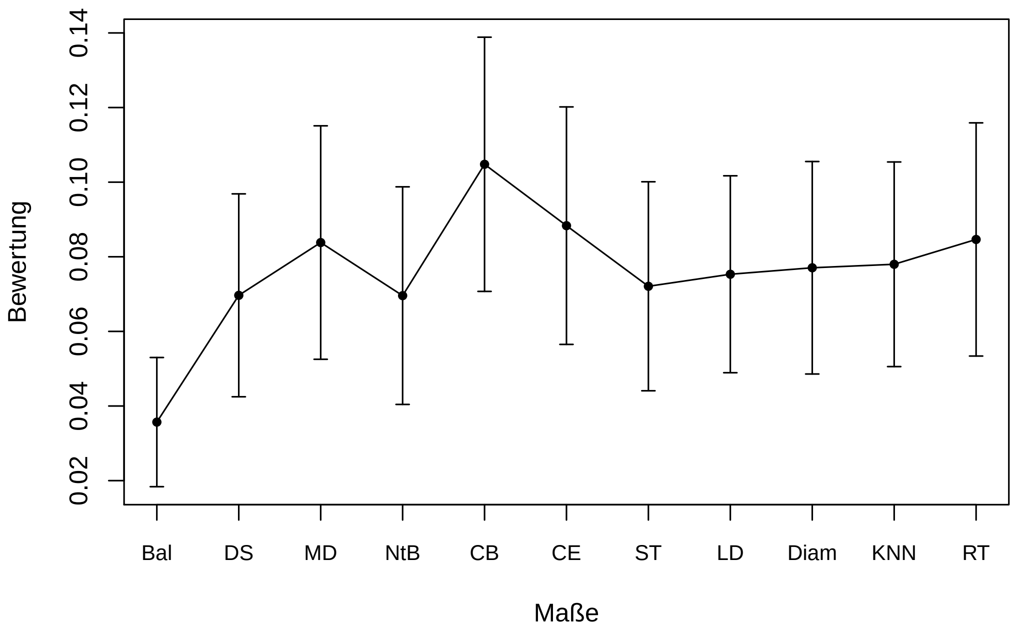
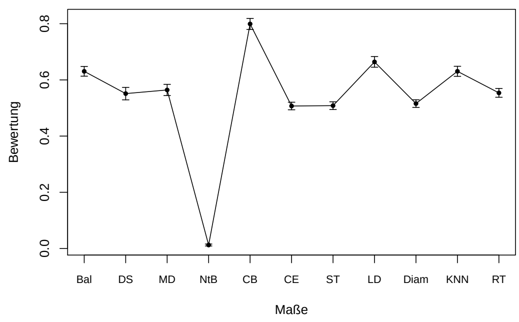


Abbildung 9: Mittelwerte und Varianzen der Balance für die getesteten Maße

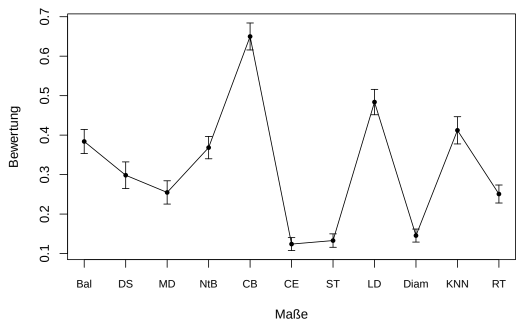
6.2.2 Aktivitätskompaktheit

Aktivitätskompaktheit bezeichnet die Eigenschaft, dass die Aktivität eines Distrikts nahe des Zentrums konzentriert ist. Die verwendeten Zielfunktionen sind (2.1) und (2.2) für den Median eines Distrikts, bzw. die p-Mediane der Lösung und (2.3) sowie (2.4) für das Zentrum eines Distrikts bzw. die p-Zentren der Lösung. Für beide Zielfunktionen existieren entsprechende Maße, Distanzsumme und MaxDistanceToFacility. Ein weiteres Aktivitätskompaktheitsmaß ist das Nr-to-Best-Maß, welches für eine Partition die Anzahl der Basisgebiete misst, die nicht der bestmöglichen Facility zugeordnet werden.

Distanzsumme Abbildung 5.5 zeigt die skalierten Distanzsummen für die getesteten Maße. Die größte Auffälligkeit bietet das Nr-to-Best-Maß, welches mit Abstand die besten Bewertungen liefert und alle anderen Maße dominiert. Im Durchschnitt beträgt die mit diesem Maß berechnete Distanzsumme nur etwa ein Fünftel des Minimums der anderen Maße.



(a) Distanzsumme zu den Facilities(p-Median-Problem)



(b) Maximale Entfernung eines Basisgebiets zur Facility (p-Center-Problem)

Abbildung 10: Skalierte Bewertungen für Distanzsummen und maximale Entfernung zur Facility

Anhand der Ergebnisse können die getesteten Maße in verschiedene Qualitätsstufen eingeteilt werden, deren Lösungen sich ihrer Güte signifikant voneinander unterscheiden. Das Compactness-Basic-Maß wird bis auf das LineDistance-Maß von allen anderen Gruppen dominiert. Das beste Ergebnis eines Maßes dieser Gruppen ist für jedes untersuchte Problem besser als das Ergebnis des CB-Maßes. Die Maße Compactness-Epsilon, Schwartzberg-Test und Diameter dominieren zusammen auch das LineDistance-Maß.

Maximaldistanz zum Zentrum Die Ergebnisse für die maximale Distanz eines Basisgebiets zu der jeweils zugeordneten Facility bieten ein ähnliches Bild wie bei der Distanz-

Rang	Maße
1	Nr-to-Best
2	Compactness-Epsilon, Schwartzberg-Test, Diameter
3	Distanzsumme, MaxDistance, Reock-Test
4	Balance, K-Nearest-Neighbor
5	LineDistance
6	Compactness-Basic

Tabelle 8: Qualitätsstufen für die Distanzsumme

Rang	Maße
1	Compactness-Epsilon, Schwartzberg-Test, Diameter
2	Distanzsumme, MaxDistance, Reock-Test
3	Balance, Nr-to-Best, K-Nearest-Neighbor
4	Line-Distance
5	Compactness-Basic

Tabelle 9: Güteklassen für die p-Center-Probleme

summe, allerdings ohne den Ausreißer NtB. Wieder gruppieren sich Maße zu Qualitätsniveaus. Diese Gruppen sind im wesentlichen die gleichen wie bei den Distanzsummen. Allerdings entspricht die Güte der NtB-Lösungen in diesem Fall etwa derjenigen von Balance- und KNN-Maß.
Die beiden Maße CB und LD liefern wiederum die schlechtesten Ergebnisse. An der Reihenfolge der einzelnen Gruppen ändert sich nichts. Allerdings sind die Unterschiede deutlicher. Die Ergebnisse der CE-, ST- und Diameter-Maße werden nur in Einzelfällen von anderen Maßen übertroffen.

6.2.3 Geografische Kompaktheit

Geografische Kompaktheit umfasst solche Kompaktheitsmaße, die nicht auf Funktionen der Distanzen zwischen Basisgebieten und Zentren eines Distrikts basieren. Als Zielfunktionen werden die Ausdehnung eines Distrikts (4.8), die durchschnittliche Entfernung eines Basisgebiets zu den k nächstgelegenen Nachbarn (4.10) und der Schwartzberg-Test (2.14). Diese Zielfunktionen werden auch für die entsprechenden Maße verwendet. Weiterhin werden die geografischen Kompaktheitsmaße Compactness-Basic, Compactness-Epsilon-Area, Line-Distance und Reock-Test getestet.
Abbildung 5.6 zeigt die Mittelwerte der skalierten Ergebnisse für den maximalen und den durchschnittlichen Durchmesser eines Distriktes. In beiden Fällen bilden die Maße CE, ST und Diameter das höchste Qualitätsniveau. Während diese Maße für den maximalen Durchmesser bis auf drei Probleme die beste Lösung berechnen, wird der durchschnittli-

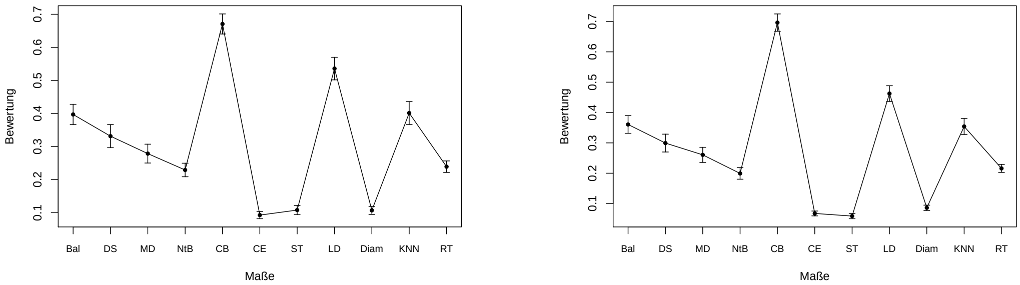


Abbildung 11: Skalierte maximale und durchschnittliche Durchmesser eines Distrikts

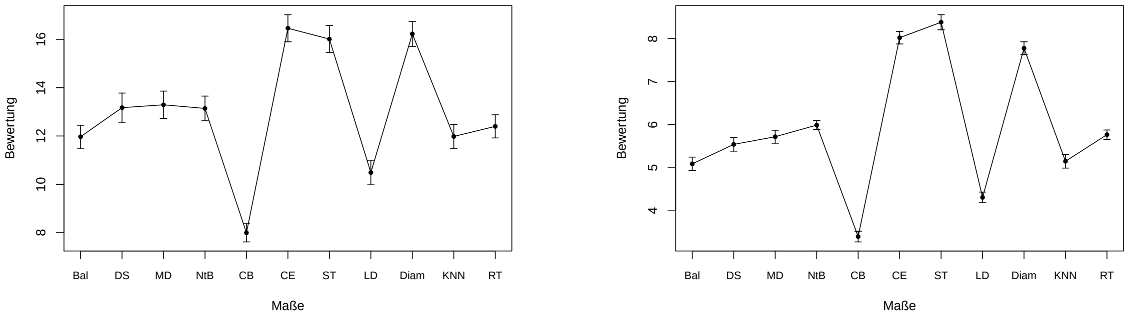
Rang	Maße
1	Compactness-Epsilon, Schwartzberg-Test, Diameter
2	Nr-to-Best, Reock-Test
3	Distanzsumme, MaxDistance
4	Balance, Nr-to-Best, K-Nearest-Neighbor
5	Line-Distance
6	Compactness-Basic

Tabelle 10: Güteklassen für den Durchmesser und die Entfernung zu den k nächsten Nachbarn

che Durchmesser von diesen Maßen sogar dominiert.
Die K-Nearest-Neighbor-Bewertungen unterscheiden sich kaum von denen des Durchmessers. Die Güte der Diameter-Lösungen ist ein wenig, aber messbar, schlechter als die der Lösungen der CB- und ST-Maße. Das KNN-Maß liefert nur unterdurchschnittliche Ergebnisse, welche sich, wie auch beim Durchmesser der Distrikte nicht signifikant von den Ergebnissen des Balance-Maßes unterscheidet. Tabelle 5.5 zeigt die verschiedenen Klassen der Maße für den Durchmesser und die k nächsten Nachbarn.
Die durchschnittlichen Maximalwerte des Schwartzberg-Tests weisen das umgekehrte Muster zu den anderen Bewertungen auf. Die besten Bewertungen werden unter Verwendung des Compactness-Basic erzielt, während die schlechtesten Bewertungen mit den Maßen Compactness-Epsilon, Diameter und Schwartzberg-Test erzeugt werden. Die durchschnittlichen Werte des Schwartzberg-Test liefern das gleiche Ergebnis.

6.2.4 Dummy-Facilities

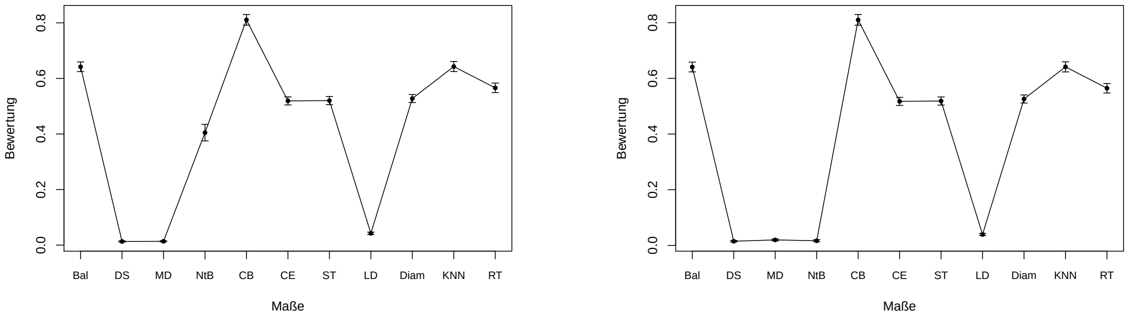
Die Maße Distanzsumme, MaxDistance, Nr-to-Best und LineDistance berechnen Funktionen der Distanz der Facilities zu den Basisgebieten bzw. zur Trennlinie der Partition. In Teilproblemen, in denen die Zahl der Distrikte die der vorhandenen Facilities übersteigt, kann die Einbeziehung von Dummy-Facilities zu besseren Ergebnissen führen. Abbildung 5.8(a) zeigt die Ergebnisse für Dummy-Facilities, die nach Winkeln berechnet



(a) Maximale Werte

(b) durchschnittliche Werte

Abbildung 12: Ergebnisse des Schwartzberg-Tests



(a) nach Winkeln

(b) nach Zellen

Abbildung 13: Distanzsummen mit Dummy-Facilities

wurden. Die Ergebnisse des NtB-Maßes haben sich im Vergleich zu den Ergebnissen ohne Dummy-Facilities verschlechtert, während die anderen drei Maße erheblich profitieren. Die mit den Maßen DS und MD erzielten Distanzsummen erreichen mit Dummy-Facilities diejenigen des NtB-Maßes ohne Dummy-Facilities. Die Distanzsummen werden bei diesen Maßen durch die Verwendung von Dummy-Facilities um 80% reduziert. Auch das Line-Distance-Maß erzielt eine ähnliche Verbesserung, allerdings sind die Distanzsummen weiterhin signifikant größer als bei den anderen beiden Maßen.

Die Ergebnisse mit Dummy-Facilities, die nach Zellen berechnet wurden, sind, wie in Abbildung 5.8(b) zu sehen, denen mit Winkel-Dummies sehr ähnlich. Distance-Sum, Max-Distance und Line-Distance profitieren genau so von den Dummy-Facilities. Der größte Unterschied ist, dass das Nr-to-Best-Maß nun wieder hervorragende Ergebnisse erzielt, die sich nicht signifikant von denen der Maße DS und MD unterscheiden. Abbildung 5.9 zeigt die Methoden, Dummy-Facilities zu berechnen, im Vergleich mit den Ergebnissen ohne diese. Die schwarze Linie zeigt die Ergebnisse ohne Dummy-Facilities, die rote Linie und grüne Linie zeigen die Ergebnisse mit Winkel-Dummies und Zellen-Dummies.

In geringerem Umfang verbessern sich die maximalen und durchschnittlichen Durchmesser

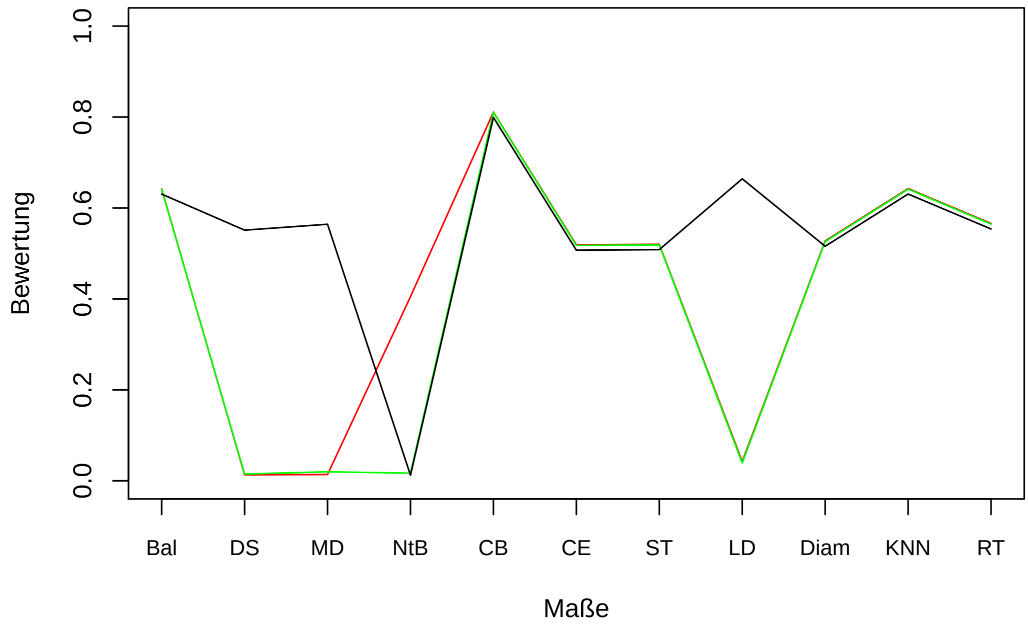


Abbildung 14: Vergleich der Distanzsummen mit und ohne Dummy-Facilities

eines Distrikts sowie die KNN-Bewertung für die Maße DS und MD. Für das DistanceSum-Maß ist die Verbesserung für beide Arten von Dummy-Facilities in etwa gleich, während das MaxDistance-Maß stärker von der Berechnung nach Winkeln profitiert. Für die Maße Nr-to-Best und LineDistance sind die Verbesserungen für diese Zielfunktionen nicht signifikant. Für den Schwartzberg-Test zeigt sich eine Verschlechterung, die für Winkel-Dummies stärker ausgeprägt ist als für Zellen-Dummies.

6.3 Sortierungen

Neben der Sortierung nach Suchrichtung bzw. rotierten x-Koordinaten können die Basisgebiete auch nach dem Winkel, den sie mit dem Mittelpunkt eines Teilproblems bilden, und der Netzwerkentfernung zu einem Startpunkt sortiert werden. Der Winkel, den ein Basisgebiet mit dem Mittelpunkt bildet, ist wie in Gleichung () definiert. Die Netzwerkentfernungen zwischen den einzelnen Basisgebieten sowie den Basisgebieten und den bereits vorhandenen Facilities liegen vorberechnet vor, und werden eingelesen bevor der Algorithmus gestartet wird. Da die Netzwerkentfernungen nur für 25 Probleme vorliegen, werden im folgenden die Tests auf diese Auswahl beschränkt. Diese 25 Probleme gehören hierbei zu den kleinsten der getesteten Probleme. Weiterhin können die Maße CompactnessBasic, CompactnessEpsilon und LineDistance nicht mit den Sortierungen nach Winkeln und Netzwerkentfernungen verwendet werden, da sie Funktionen der Trennlinie berechnen. Eine solche Trennlinie existiert aber für diese Sortierungen nicht.

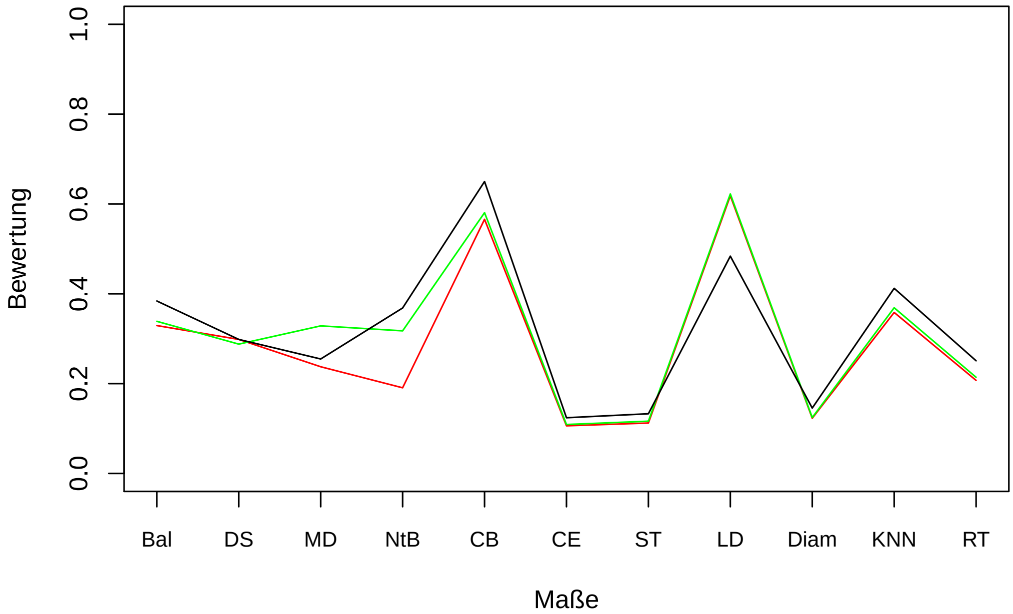
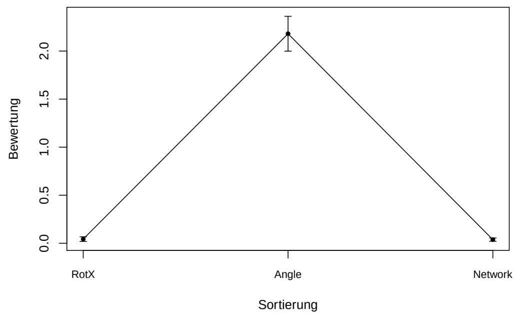


Abbildung 15: Vergleich der maximalen Distanz zur Facility mit und ohne Dummy-Facilities

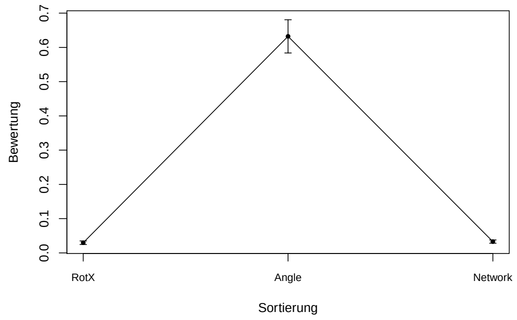
6.3.1 Ergebnisse

Die Mittelwerte der Balance für die getesteten Sortierungen werden in Abbildung 5. a gezeigt. Die Sortierungen nach rotierten x-Koordinaten und Netzwerkentfernungen unterscheiden sich nicht signifikant voneinander. Dagegen ist die Balance der Winkel-Sortierung deutlich schlechter. Durch die notwendigen Relaxierungen, um überhaupt eine Lösung berechnen zu können, sind die Ergebnisse sehr schlecht balanciert. Eine solch hohe Balance wie sie die Winkelsortierung im Mittel zeigt, erlaubt Distrikte, die dreimal so viel Aktivität enthalten wie der Durchschnitt aller Distrikte. Abweichungen nach unten können sogar noch extremer werden: selbst ein leerer Distrikt würde die Balance nicht weiter erhöhen, da ihm eine Balance von 1 zugeordnet würde. Wie in Abbildung 5. b) zu sehen, stimmen die Ergebnisse für die Distanzsummen im wesentlichen mit denen der Balance überein. Die Ergebnisse der Netzwerk-Sortierung sind lediglich geringfügig schlechter als diejenigen der rotierten x-Koordinaten, während die Winkelsortierung die schlechtesten Ergebnisse erzielt.

Die Ergebnisse für die anderen Zielfunktionen weisen ein jeweils ähnliches Muster auf: Mit der RotX-Sortierung werden die besten Ergebnisse erzielt, gefolgt von der Netzwerksortierung. Die Winkelsortierung berechnet durchgehend die schlechtesten Lösungen.



(a) Balance



(b) Distanzsumme

Abbildung 16: Balance und skalierte Distanzsummen der verschiedenen Sortierungen

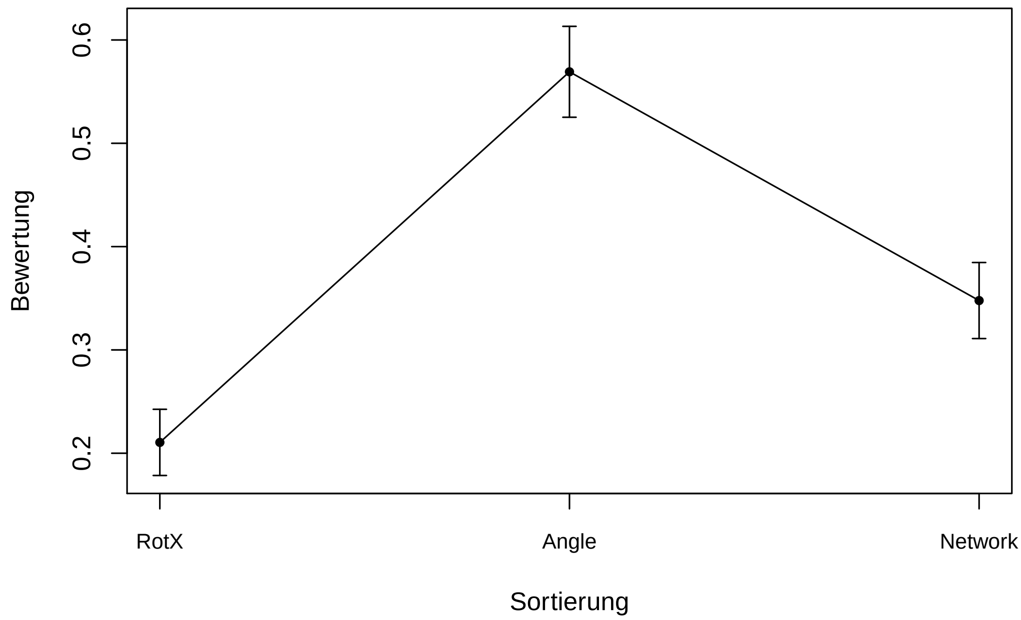


Abbildung 17: Mittelwerte der skalierten maximalen Distanz eines Basisgebiets zur ihm zugeordneten Facility

6.4 Aufteilungen

Das ursprüngliche Aufteilungsverfahren des Recursive-Partitioning-Algorithmus ist balanceorientiert. Es wird die Trennlinie gewählt, durch welche die erzeugte Partition die bestmögliche Balance erzielt (BestBalancedLine = BBL). Alternative Partitionierungen berechnen Trennlinien, welche andere Maße minimieren. So berechnet die Aufteilung BestLineDistance (BLD) diejenige Trennlinie, welche die Distanz zu einer der vorhandenen Facilities maximiert. Das BestCompactnessBasic-Verfahren (BCB) sucht nach einer Trennlinie mit minimalem Schnitt mit der konvexen Hülle der Basisgebiete. Zum Vergleich werden auch die Ergebnisse für das CompactnessEpsilonArea-Aufteilungsverfahren (BCE) angegeben, obwohl dessen Laufzeit quadratisch in der Anzahl der Basisgebiete ist und somit die asymptotische Laufzeit des Rec-Part-Algorithmus erhöht.

6.4.1 Ergebnisse

Die Auswertung der Lösungen zeigt, dass die Suche nach der bestbalanciertesten Trennlinie die besten Ergebnisse erzielt, welche sich für Balance, Distanzsumme und Maximaldistanz auch signifikant von allen anderen unterscheiden (siehe Abbildung 5.). Für die durchschnittliche Distanz zu den k nächsten Nachbarn sind die mit dem BBL-Verfahren berechneten Lösungen im Mittel die besten, unterscheiden sich aber nicht mehr messbar von denen, die mit dem BCB-Verfahren erzeugt wurden. Lediglich bei maximalem und durchschnittlichem Durchmesser der Distrikte schneidet das BBL-Verfahren schlechter ab als die anderen Partitionierungen.
Unter den alternativ entwickelten Partitionierungsverfahren erzielt das BestCompactnessBasic-Verfahren die besten Bewertungen. In Abbildung 5. ist zu sehen, dass dieses Verfahren signifikant bessere Balance- und p-Center-Ergebnisse berechnet, während für andere Zielfunktionen die Ergebnisse zwar schlechter sind, die Differenz zu den anderen Partitionierungen jedoch nicht signifikant ist. Bei der Balance werden von allen drei Verfahren hohe Werte erreicht. Lediglich die BCB-Partitionierungen erreichen Balancewerte unter 20%. Unerwartet sind die schlechten Ergebnisse der BestCompactnessEpsilon-Aufteilung. Während das entsprechende Maß neben dem Schwartzberg-Test und dem Diameter-Maß noch die besten Ergebnisse erzielte, sind die Ergebnisse des Aufteilungsverfahrens deutlich schlechter als die der anderen Partitionierungen.

6.5 Analyse der Ergebnisse

Im Rahmen dieser Arbeit wurden vier Einstellungen getestet. Neben der Anzahl der Suchrichtungen waren dies verschiedene Maße zur Bewertung der erzeugten Partitionen, die Sortierung der Basisgebiete und die Verfahren zur Erzeugung der Partitionen. Es wurde gezeigt, dass diese Parameter einen jeweils unterschiedlich starken Einfluss auf die Qua-

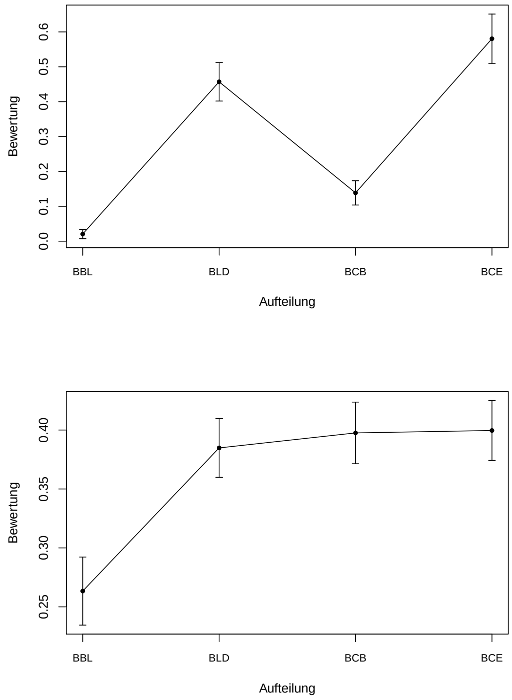
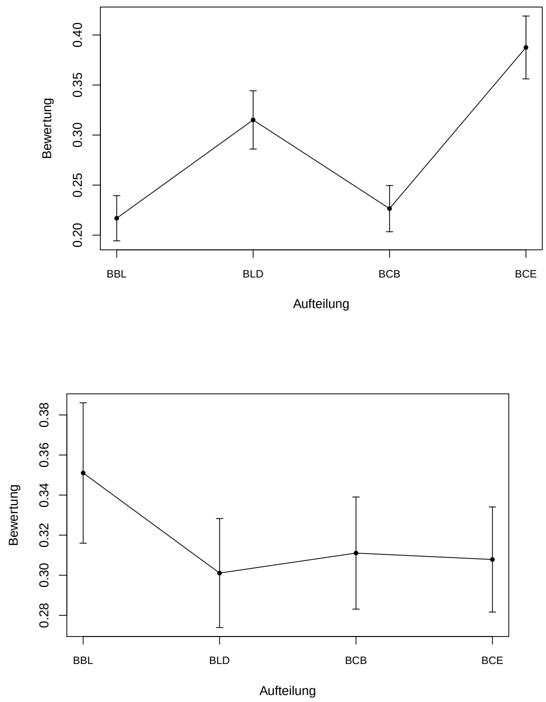


Abbildung 18: Balance und skalierte p-Center-Funktionswerte für die getesteten Aufteilungsverfahren



lität der berechneten Lösung haben.
Die Anzahl der Suchrichtungen übt hierbei den geringsten Einfluss auf das Ergebnis aus. Die Ergebnisse bei vier verwendeten Suchrichtungen unterschieden sich nur bei der Balance messbar von den Ergebnissen mit zehn Suchrichtungen. Signifikant schlechter werden die Ergebnisse erst, wenn zwei Suchrichtungen verwendet werden. Hingegen sind bei nur einer Suchrichtung viele Probleme nicht oder nur mit schlechter Balance lösbar. Offensichtlich ist für eine (gute) Lösung die Möglichkeit essenziell, die Basisgebiete verschieden sortieren zu können, wodurch bei den einzelnen Partitionen eine gleichmäßigere Aufteilung der Aktivität erzielt werden kann. Jedoch kann die Qualität der erzeugten Partitionen ab vier verwendeten Suchrichtungen nur noch wenig erhöht werden, da mit zunehmender Anzahl Suchrichtungen die Differenz der Winkel zwischen den Suchrichtungen sinkt. Dies bestätigen auch die schlechten Ergebnisse, die mit der Winkelsortierung erzielt werden. Die Sortierung nach Winkeln ist letzlich für die verwendeten Offsets die gleiche, nur startet die Iteration durch die Basisgebiete an verschiedenen Stellen.
Mit der Sortierung der Basisgebiete nach Winkeln werden aus diesem Grund deutlich schlechtere Ergebnisse erzielt als mit der Sortierung nach rotierten x-Koordinaten. Besonders aufgrund der schlechten Balance ist die Winkelsortierung kein gleichwertiger Ersatz. Weiterhin ist ein Teilproblem, welches mithilfe dieser Sortierung erzeugt wurde, im Allgemeinen nicht konvex, was im späteren Verlauf des Algorithmus zu unzusammenhängenden Teilproblem bzw. Distrikten führen könnte. Dies gilt auch für Sortierung nach Netzwerkdistanzen. Diese Art, die Basisgebiete zu sortieren, erzielt für Balance und Distanzsumme, qualitativ ähnliche Ergebnisse wie die RotX-Sortierung. Diese Zielfunktionen beziehen alle Basisgebiete in die Berechnung des Funktionswertes mit ein und sind daher für den Recursive-Partitioning-Algorithmus einfacher zu optimieren als Zielfunktionen wie der maximale Durchmesser eines Distrikts oder die maximale Entfernung eines Basisgebiets zum zugeordneten Zentrum. Letzere werden durch die Extremwerte bestimmt und sind daher anfällig für Ausreißer. Mit der RotX-Sortierung gelingt die Optimierung dieser Zielfunktionen noch am besten. Aus diesen Gründen sollten die Netzwerk- und in noch stärkerem Maße die Winkelsortierung vor allem als Ergänzung zur klassischen Sortierung der Basisgebiete verwendet werden.
Auch die alternativen Aufteilungsverfahren können nicht überzeugen. Offenbar ist die balanceorientierte Aufteilung notwendig, um ein akzeptables Balanceniveau zu erhalten. Die ursprüngliche Partitionierung berücksichtigt als einzige der getesteten Varianten die Basisgebiete bzw. etwa die Hälfte der Basisgebiete. Aufgrund des Aufteilungsverfahrens wird jeder Seite einer Partition ungefähr gleich viel Aktivität zugeordnet. Sind sowohl die Verteilung der Basisgebiete als auch die Zuordnung der Aktivität auf die einzelnen Basisgebiete in beiden Hälften der Partition ähnlich, werden auch die Basisgebiete zu ähnlichen Teilen auf die Teilprobleme aufgeteilt, was sich wiederum positiv auf die Zielfunktionen auswirkt, die eine Funktion der Distanzen berechnen.
Die mit den getesteten Maßen berechneten Lösungen unterscheiden sich deutlich vonein-

ander. Die Maße Balance, NrToMin und DistanceSum erfüllen jeweils ihren Zweck, auch wenn für letzteres die Einbeziehung von Dummy-Facilities entscheidend ist. Mit diesen Maßen werden die besten Lösungen bezüglich Balance bzw. Distanzsumme zu den Facilities berechnet. Auch das NrToBest-Maß erzielt gute Ergebnisse bezüglich der Distanzsumme. Maße wie MaxDistanceToFacility und K-Nearest-Neighbor sind hingegen nicht zweckmäßig, da sie für die jeweils entsprechende Zielfunktion nur unterdurchschnittliche Werte erzielen. Dies gilt auch für das Schwartzberg-Test-Maß, welches dafür bei anderen Bewertungen gut abschneidet. Das gleiche gilt für die Maße Compactness-Epsilon-Area und Diameter. Diese Maße bewerten solche Aufteilungen gut, bei denen die Ausdehnung der beiden Teilprobleme möglichst gering ist. Dieses Vorgehen scheint für die meisten Zielfunktionen zweckmäßig zu sein. Für die Minimierung der Distanzsumme ist jedoch ein Maß zu bevorzugen, welches die Distanzen der Basisgebiete zu den Facilities und Dummy-Facilities berücksichtigt.

7 Zusammenfassung

In dieser Arbeit wurde ein heuristisches Verfahren zur kombinierten Standort- und Gebietsplanung vorgestellt. Im Rahmen der Standortplanung waren einige Standorte fest, während andere im Rahmen der Restriktionen frei gewählt werden konnten. Das Verfahren baut auf dem von [Kal06] entwickelten Recursive-Partitioning-Algorithmus auf, einem Algorithmus zur Gebietsplanung. Es wurde gezeigt, wie der Algorithmus zur Standortplanung eingesetzt werden kann. Für die Bewertung der Partitionen des Rec-Part-Algorithmus wurden neue Maße entwickelt, die bei Tests mit den bereits vorhandenen Maßen verglichen wurden. Ebenso wurden Erweiterungen des Algorithmus vorgestellt und getestet. Die Tests zeigten, dass die Anzahl der Suchrichtungen ab einer bestimmten Größe ($K = 4$) keinen nennenswerten Einfluss auf die Qualität der berechneten Lösung hat.

Die neu implementierten Aufteilungsverfahren konnten in den Tests ebenso wenig überzeugen wie die entwickelten Sortierungsverfahren. Die Einführung von Dummy-Facilities für Maße, die eine Funktion der Distanzen der Basisgebiete zu den Facilities berechnen, hat dagegen die Ergebnisse deutlich verbessert. Bei den Tests der Maße zeigten sich verschiedene Qualitätsstufen, die für die einzelnen Zielfunktionen herausgearbeitet wurden.

Literatur

Burcin Bozkaya, Erhan Erkut, and Gilbert Laporte. A tabu search heuristic and adaptive memory procedure for political districting. *European Journal of Operational Research*, 144(1):12–26, 2003.

Marko Blais, Sophie D Lapierre, and Gilbert Laporte. Solving a home-care districting problem in an urban setting. *Journal of the Operational Research Society*, 54(11):1141–1147, 2003.

Richard L. Church and Robert S. Garfinkel. Locating an obnoxious facility on a network. *Transportation Science*, 12(2):107–118, 1978.

Mark S. Daskin. *Network and discrete location : models, algorithms, and applications.* Wiley-interscience series in discrete mathematics and optimizationA Wiley-interscience publication. Wiley, New York, 1995.

Zvi Drezner, Kathrin Klamroth, Anita Schöbel, and George O. Wesolowsky. The weber problem. 1995.

Zvi Drezner. Conditional p-center problems. *Transportation Science*, 23(1):51–53, 1989.

Zvi Drezner, editor. *Facility location a survey of applications and methods.* Springer series in operations research. Springer, Berlin, 1995.

Zvi Drezner. On the conditional¡ i¿ p¡/i¿-median problem. *Computers & operations research*, 22(5):525–530, 1995.

Zvi Drezner and Saharon Shelah. On the complexity of the elzinga-hearn algorithm for the 1-center problem. *Mathematics of operations research*, 12(2):255–261, 1987.

Jack Elzinga and Donald W Hearn. Geometrical solutions for some minimax location problems. *Transportation Science*, 6(4):379–394, 1972.

AJ Goldman and PM Dearing. Concepts of optimal location for partially noxious facilities. *Bulletin of the Operational Research Society of America*, 23(1):B85, 1975.

Robert S Garfinkel and George L Nemhauser. Optimal political districting by implicit enumeration techniques. *Management Science*, 16(8):B–495, 1970.

S Louis Hakimi. Optimum locations of switching centers and the absolute centers and medians of a graph. *Operations Research*, 12(3):450–459, 1964.

S Louis Hakimi. On locating new facilities in a competitive environment. *European Journal of Operational Research*, 12(1):29–35, 1983.

J Halpern. The location of center-median convex combination on an undirected tree. *Journal of Regional Science*, 16:237–245, 1976.

Robert E Helbig, Patrick K Orr, and Robert R Roediger. Political redistricting by computer. *Communications of the ACM*, 15(8):735–741, 1972.

Sidney W. Hess and Stuart A. Samuels. Experiences with a sales districting model: Criteria and implementation. *Management Science*, 18(4):pp. P41–P54, 1971.

Sidney Wayne Hess, JB Weaver, HJ Siegfeldt, JN Whelan, and PA Zitlau. Nonpartisan political redistricting by computer. *Operations Research*, 13(6):998–1006, 1965.

Jörg Kalcsics. *Unified approaches to territory design and facility location*. Operations-Research. Shaker, Aachen, 2006.

Richard M Karp. *Reducibility among combinatorial problems*. Springer, 1972.

Oded Kariv and S Louis Hakimi. An algorithmic approach to network location problems. i: The p-centers. *SIAM Journal on Applied Mathematics*, 37(3):513–538, 1979.

Oded Kariv and S Louis Hakimi. An algorithmic approach to network location problems. ii: The p-medians. *SIAM Journal on Applied Mathematics*, 37(3):539–560, 1979.

Jörg Kalcsics, Stefan Nickel, and Michael Schröder. Towards a unified territorial design approach - applications, algorithms and gis integration. *Top*, 13(1):1–56, 2005.

Kenneth C. Martis. The original gerrymander. *Political Geography*, 27(8):833 – 839, 2008.

Edward Minieka. Conditional centers and medians of a graph. *Networks*, 10(3):265–272, 1980.

Anuj Mehrotra, Ellis L Johnson, and George L Nemhauser. An optimization based heuristic for political districting. *Management Science*, 44(8):1100–1114, 1998.

Nimrod Megiddo and Kenneth J Supowit. On the complexity of some common geometric location problems. *SIAM journal on computing*, 13(1):182–196, 1984.

Jr. Reock, Ernest C. A note: Measuring compactness as a requirement of legislative apportionment. *Midwest Journal of Political Science*, 5(1):pp. 70–74, 1961.

Joseph E Schwartzberg. Reapportionment, gerrymanders, and the notion of compactness. *Minn. L. Rev.*, 50:443, 1965.

George O Wesolowsky. The weber problem: History and perspectives. *Computers & Operations Research*, 1993.

H Peyton Young. Measuring the compactness of legislative districts. *Legislative Studies Quarterly*, 13(1):105–115, 1988.