

Dokumentation – Gebietsverteilung, Greenfield, Whitespot

Inhalt

1. Allgemeine Informationen	2
2. Klassenaufbau und Beziehungsdiagramm.....	4
3. Algorithmus Gebietsverteilung	6
a. Algorithmus Aufbau.....	6
b. Erzeugung der Gleichverteilung	8
c. Beachtung von Kompaktheit und Gleichverteilung	9
4. Algorithmus Greenfield	10
5. Algorithmus Whitespot	12
6. Algorithmus zur Erreichung der Gleichverteilung	14
7. Notwendige Parameter	15
8. Limitierungen	15
9. Perspektiven.....	16
a.) Ermittlung der erlaubten Abweichung.....	16
b.) Abbruch beim Greenfieldalgorithmus.....	16
c.) Verbesserung der Serverrechenzeit	16
d.) Verbesserung der Rechenzeit bei großen Datenmengen	17
e.) Sonderfälle	17
a. Inseln	17
b. PLZ in PLZ.....	17
f.) Erweiterbarkeit des Algorithmus	18
a. Zentrenoptimierung	18
b. STRAB.....	18

1. Allgemeine Informationen

Alle drei implementierten Algorithmen stammen aus dem Feld des Geomarketings und beschäftigen sich mit einer geeigneten Verteilung von Geometrien unter der Berücksichtigung bestimmter Kriterien. Das Ziel ist es gegebene Geometrien, wie Postleitzahlgebiete, auf geeignete Art und Weise zu einer bestimmten Anzahl übergeordneter Gebiet zusammenzufassen. Dabei gibt es drei verschiedene Ansätze:

1. Die Standorte, zu denen übergeordnete Gebiete gebildet werden sollen, sind gegeben. Dies wird als Gebietsverteilung bezeichnet.
2. Es sind keine Standorte gegeben, zu denen Gebiete erzeugt werden sollen. Es müssen somit Standorte UND die zugehörigen übergeordneten Gebiete erzeugt werden. Dies wird als Greenfieldanalyse bezeichnet.
3. Es ist eine bestimmte Anzahl an Standorten gegeben, jedoch sollen zusätzlich n weitere Standorte erzeugt werden. Somit muss für die gegebenen Standorte, sowie für neue Standorte je ein übergeordnetes Gebiet erzeugt werden. Dies wird als Whitespotanalyse bezeichnet.

Die Erzeugung eines solchen Gebietes als Anwendung einer Gebietsverteilung, kann wie in Abbildung 1 dargestellt aussehen. In diesem Fall sind 10 Sparkassenstandorte im Raum Dresden gegeben, zu denen Einzugsgebiete ermittelt wurden. Als Basis dienen dabei PLZ5 Gebiete.

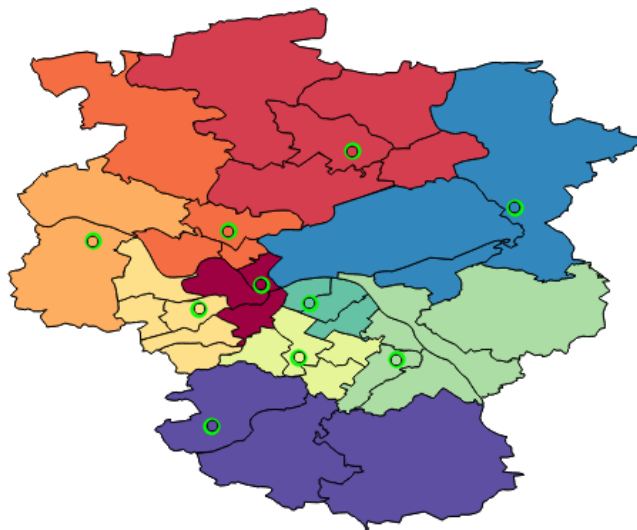


Abbildung 1 Gebietsverteilung am Beispiel von Sparkassenstandorten im Raum Dresden. Die mit grün umrahmten Punkte stellen die Sparkassenfilialen dar.

Bei der Erstellung der übergeordneten Gebiete gibt es je nach Anwendungsfall verschiedenen Kriterien, die beachtet werden können. Unter <http://dol.iwr.kit.edu/Gebietsplanung.php> lassen sich eine Vielzahl an Beispielen für die Anwendung von Gebietsverteilungen und eine Übersicht möglicher Planungskriterien einsehen. In den hier implementierten Algorithmen werden folgende Kriterien berücksichtigt:

- Zusammenhang: Gebiete eines übergeordneten Gebietes hängen zusammen. Von jeder zugehörigen Geometrie, kann also jede andere zugehörige Geometrie erreicht werden, ohne ein anderes übergeordnetes Gebiet durchkreuzen zu müssen. Es werden somit keine „Inseln“

erzeugt. Ein Beispiel ist in Abbildung 2 zu finden. Der Zusammenhang der Gebiete ist optional und kann vom Nutzer aktiviert und deaktiviert werden.

- **Ausgeglichenheit:** Jedes erzeugte übergeordnete Gebiet soll ausgeglichen bezüglich eines gegebenen Kriteriums verglichen zu den anderen Gebieten sein. Solch ein Kriterium können Haushaltszahlen, Einwohner oder andere Daten sein. Die Ausgeglichenheit kann mittels eines Schwellwerts angepasst werden. Dieser gibt die maximal erlaubte Abweichung von der Ausgeglichenheit an. Je höher dieser Schwellwert gesetzt ist, umso kürzer ist die Berechnungszeit. Dabei kann es jedoch sein, dass dennoch eine Verteilung der Gebiete erreicht wird, die einen geringeren Schwellwert aufweist.
- **Zugehörigkeit der Zentren:** Jeder gegebene Standort liegt in einer der Geometrien, die seinem übergeordneten Gebiet zugeordnet sind. Es wird somit nicht der Fall auftreten, dass ein Zentrum abseits seines Gebietes liegt.
- **Kompaktheit:** Es wird versucht, die erzeugten Gebiete möglichst kompakt zu gestalten. Dabei besteht die Möglichkeit einer Nutzereingabe für die Wichtung der Kompaktheit und der Balance. Diese beiden Wichtungsfaktoren werden während der Umverteilung, die zum Erreichen der Gleichverteilung durchgeführt wird, berücksichtigt. Zur Ermittlung der Kompaktheit wird das Cox Maß verwendet. Eine Erklärung davon ist weiter unten zu finden. Die Beachtung der Kompaktheit findet lediglich bei der Aufteilung nach Luftlinie statt.

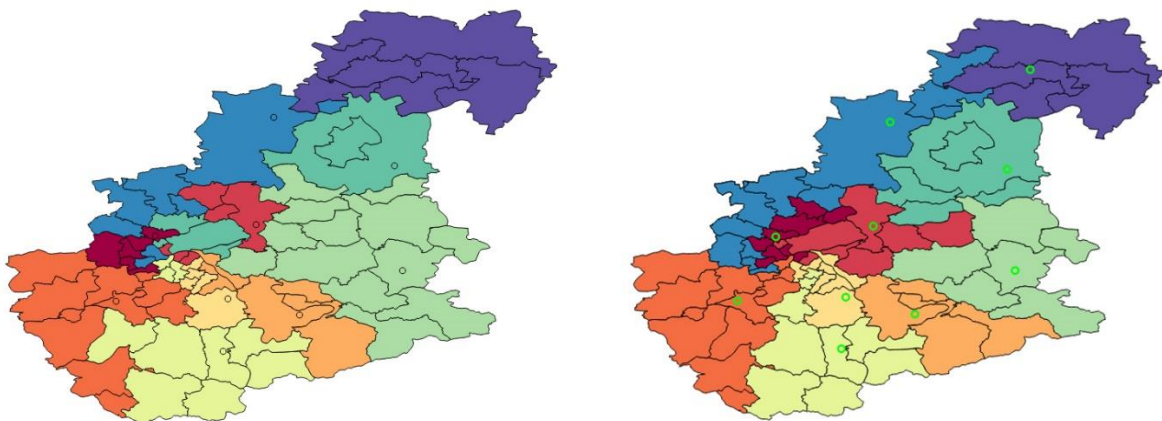


Abbildung 2 Links: Nicht zusammenhängende Verteilung der Gebiete, rechts: Zusammenhängende Verteilung der Gebiete

Um die Verteilung der Gebiete durchführen zu können, beinhaltet jeder der drei Algorithmen drei Grundelemente, die um eventuell notwendige Erweiterungen (bei der Greenfield- und Whitespotanalyse) ergänzt werden. Die Grundelemente sind die Initialisierung der Daten, die Gebietsverteilung an sich und die Visualisierung der Ergebnisse.

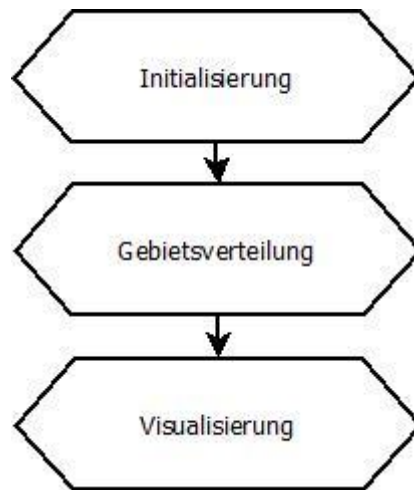


Abbildung 3 Grundelemente jedes Algorithmus

Bei der Initialisierung werden alle notwendigen Daten generiert und in Variablen abgelegt. Dies beinhaltet neben den gegebenen Standorten auch die gegebenen Geometrien zuzüglich ihrer Eigenschaften. Diese beinhalten u.A. die benachbarten Gebiete, der Wert des Kriteriums, das gleich verteilt werden soll und die Form des Gebietes als PostGIS Geometrie. Innerhalb der Initialisierung werden folgende Funktionen aufgerufen:

- **getNrOrSum:** Diese Funktion berechnet die Anzahl an Geometrien in der Datenbank. Der Wert wird für spätere Berechnungen und das Auslesen der Daten aus der Datenbank notwendig sein.
- **initPolygons:** Diese Funktion ermittelt die Basisdaten der gegebenen Geometrien und speichert diese im PolygonContainer (siehe Klassendiagramm) ab. Es sind dabei die ID, die Geometrie und der Wert des Kriteriums enthalten.
- **initNeighbours:** Diese Funktion liest zu jeder Geometrie die Nachbarn aus und speichert diese als Polygon ab (siehe Klassendiagramm). Zuzüglich werden alle Teilumfänge berechnet, die an ein benachbartes Polygon grenzen. Dieser Wert ist während der Umverteilung zur Berechnung des Kompaktheitsmaßes von Nöten.
- **initCentroids:** Diese Funktion ermittelt zu jeder Geometrie das zugehörige Zentroid und speichert dieses. Das Zentroid wird später bei der Gebietsverteilung notwendig sein, wenn die Geometrien bezüglich ihrer Distanz zu den Standorten zugeordnet werden. Es ist notwendig, das Zentroid vorher zu speichern, um eine Vielzahl an Datenbankzugriffen zu vermeiden.
- **initArea:** Diese Funktion berechnet zu jeder Geometrie den Flächeninhalt und speichert diesen ab. Der Flächeninhalt ist bei der Umverteilung zur Bestimmung des Kompaktheitsmaßes notwendig.
- **initCircumferences:** Diese Funktion berechnet den Umfang zu jeder Geometrie und speichert diesen. Der Wert wird während der Umverteilung für die Bestimmung des Kompaktheitsmaßes benötigt.

Während der Gebietsverteilung werden die Gebiete auf die Standorte verteilt. Dies beinhaltet die anfängliche Zuordnung sowie die Umverteilung der zugeordneten Gebiete, um die gewünschte Gleichverteilung zu erzielen. Eine detaillierte Beschreibung dieses Vorgehens ist in Kapitel 3 Algorithmus Gebietsverteilung zu finden. An dieser Stelle werden ebenso die aufgerufenen Funktionen erläutert.

Im Schritt der Visualisierung werden die Ergebnisse anschließend an den Manager zurückgegeben und dargestellt.

2. Klassenaufbau und Beziehungsdiagramm

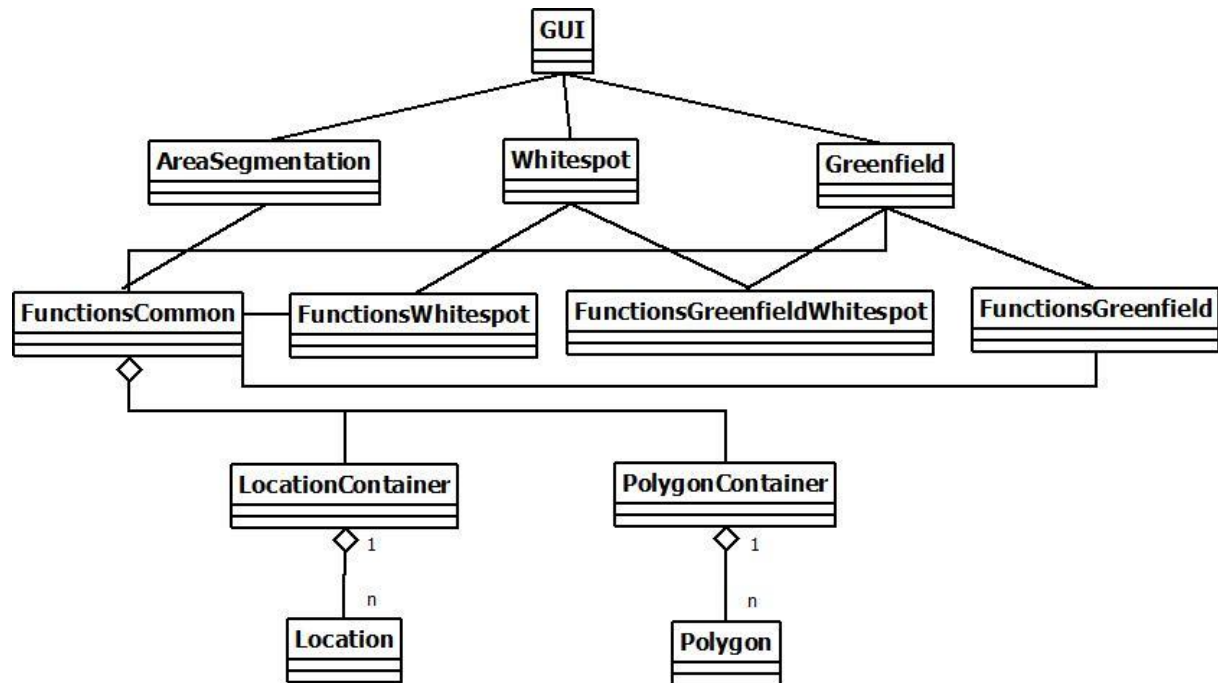


Abbildung 4 Übersicht über die Beziehungen

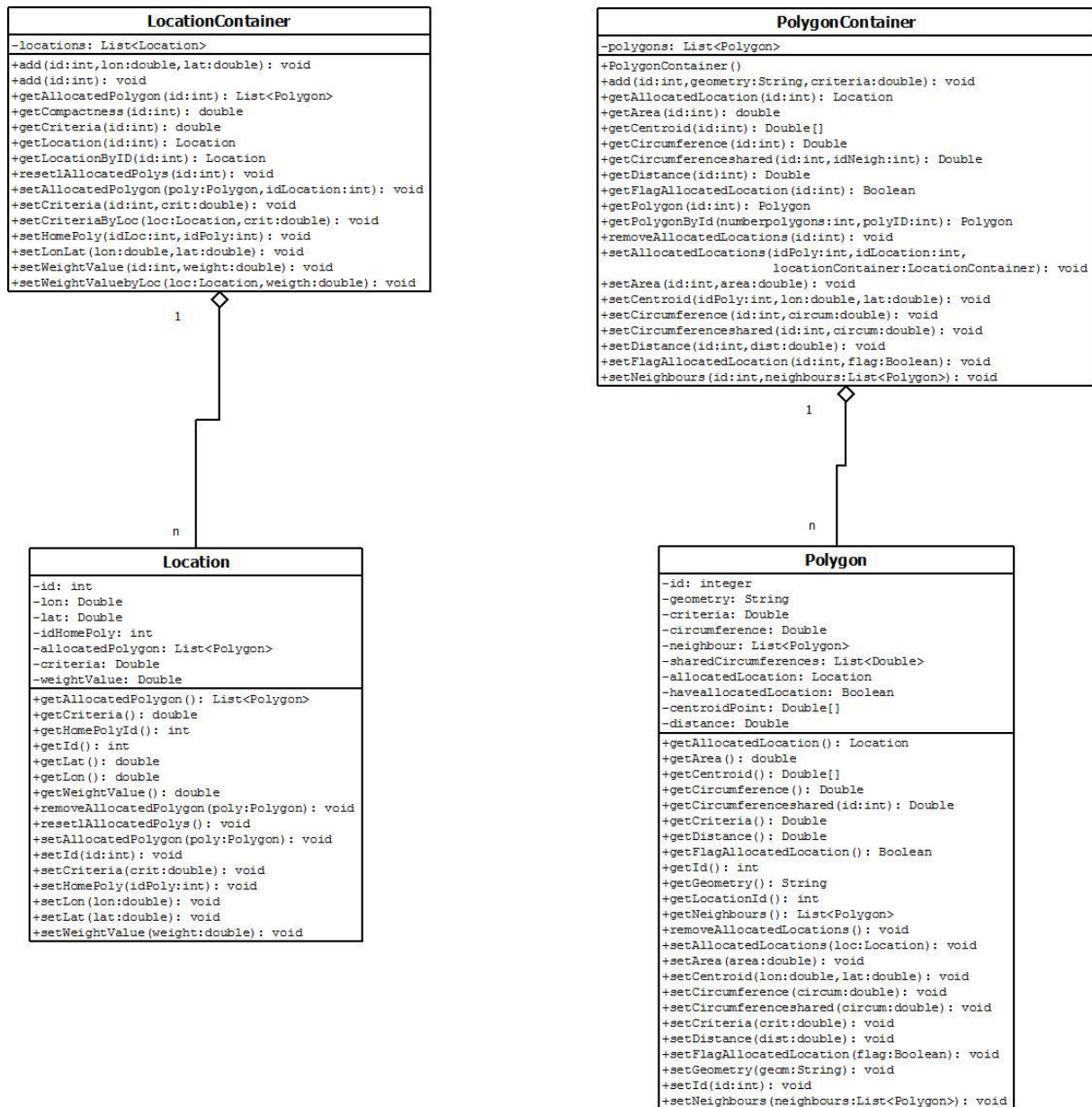


Abbildung 5 Klassendiagramm

3. Algorithmus Gebietsverteilung

a. Algorithmus Aufbau

Der Gebietsverteilungsalgorithmus ist die Basis für alle drei implementierten Algorithmen, jedoch sind für die Durchführung der Greenfield- bzw. Whitespotanalyse Erweiterungen notwendig. Der Gebietsverteilungsalgorithmus verteilt die gegebenen Geometrien auf die gegebenen Standorte unter der Beachtung der Gleichverteilung auf. Das Vorgehen des Algorithmus gliedert sich dabei in die folgenden Schritte:

1. Aufteilung der Gebiete auf die Standorte bezüglich ihrer Distanz, das heißt jedem Standort werden die Gebiete zugeteilt, die ihm und wirklich nur ihm am nächsten liegen
2. Umverteilung der zugeordneten Gebiete unter Berücksichtigung der Wichtungsfaktoren, um eine Gleichverteilung zu erzeugen.

Bei der Verteilung der Geometrien im ersten Schritt, wird keine Rücksicht auf die Gleichverteilung der Haushaltszahlen, etc. genommen. Stattdessen wird lediglich die Distanz der Gebiete zu den Standorten beachtet. Die Distanz ergibt sich entweder aus der direkten Distanz (Aufteilung nach Luftlinie) oder der Fahrtstrecke/Fahrtzeit (Aufteilung nach Fahrtstrecke/-zeit). Die Aufteilung nach Distanz ohne Berücksichtigung der Gleichverteilung hat die zwei Vorteile, dass 1. die Gebiete durch diese Vorgehen von vornherein kompakt gestaltet werden, was eine gute Voraussetzung für die Umverteilung darstellt, und 2. Dieses Verfahren als Basis für eine Erweiterung des Algorithmus genommen werden kann, wenn der Benutzer eine Verteilung der Gebiete nur nach Distanzen erhalten möchte. Dieser Ansatz ist somit bereits implementiert. Die Verteilung nach Distanz ist an einem Beispiel im Raum Dresden auf der folgenden Abbildung dargestellt.

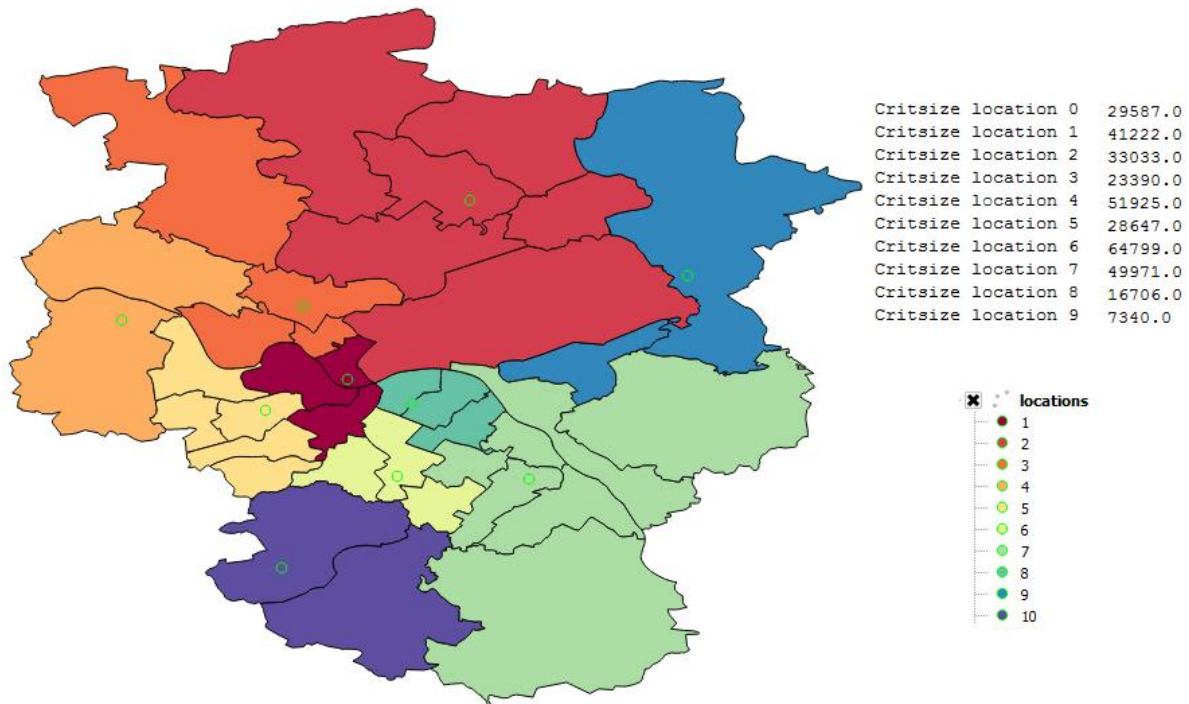


Abbildung 6 Zuordnung der Gebiete nach Distanz am Beispiel von Dresden

Um den Distanzalgorithmus durchführen zu können, werden zunächst die Distanzen initialisiert und gespeichert. Die Distanz wird mittels des Zentroides einer Geometrie bestimmt. Anfänglich wurde die Distanz mit der Funktion `ST_Distance(geom,geom)` von PostGIS bestimmt, doch Tests zeigten sich, dass die Verwendung des Zentroids zu leicht besseren Ergebnissen führt. Zudem wird die Distanz nun über das Orthodrome berechnet, um keine Datenbankzugriffe mehr durchführen zu müssen (im Fall von der Verwendung der Luftlinie). Bei der Verwendung von Fahrtstrecke/-zeit wird die Distanz/Zeit mittels des Routingalgorithmus initialisiert.

Nachdem die Zuordnung nach Distanz erfolgt ist, wird die Umsortierung der Geometrien vollzogen, so dass eine Gleichverteilung erzielt wird. Bei der Umsortierung gibt es 3 wichtige Parameter:

- Wert für die erlaubte Abweichung von der Gleichverteilung
- Wichtungsfaktor für die Kompaktheit (Luftlinie) bzw. Fahrtstrecke/-zeit
- Wichtungsfaktor für die Gleichverteilung

Die genaue Beschreibung des Vorgehens zum Erreichen der Gleichverteilung wird im folgenden Kapitel beschrieben. Das Resultat der Berechnung ist in der folgenden Abbildung aufgeführt.

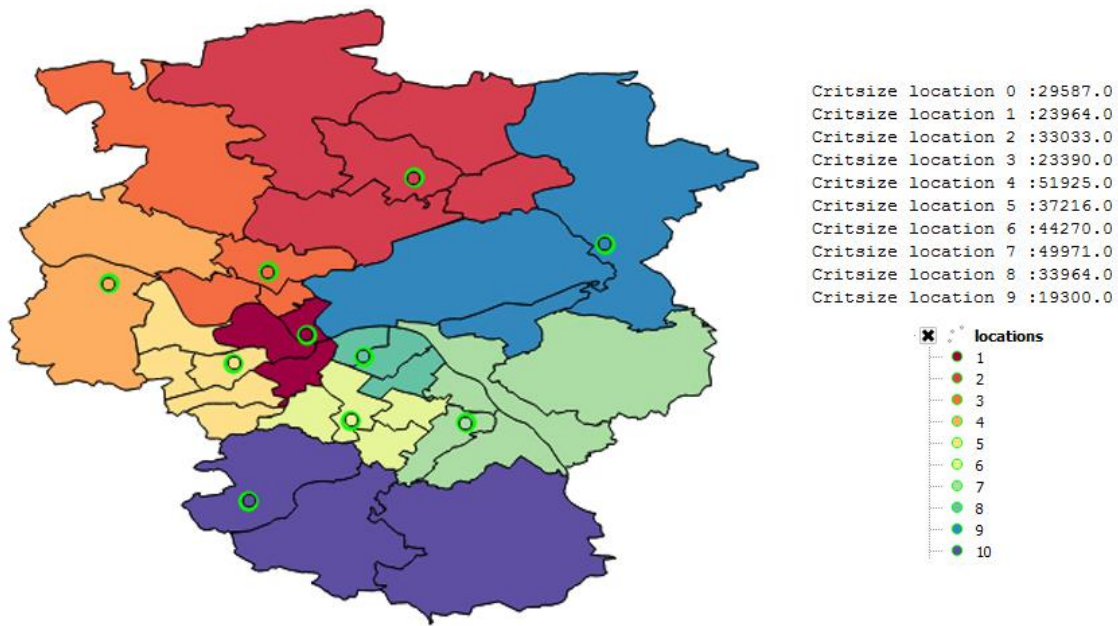


Abbildung 7 Gebietsverteilung unter Verwendung folgender Parameter: Erlaubte Abweichung 50, Wichtungsfaktor Kompaktheit 100, Wichtungsfaktor Gleichverteilung 0

b. Erzeugung der Gleichverteilung

- Bei der Umsortierung wird jeder Standort schleifenartig durchgegangen. Es wird geprüft, ob die zugehörige Summe der Haushaltszahlen, etc. innerhalb der erlaubten Abweichung liegt. Ist dies der Fall, wird mit der Prüfung des nächsten Standortes fortgesetzt. Liegt der Wert jedoch außerhalb des erlaubten Wertes, wird eine Umsortierung vorgenommen. Für die Umsortierung werden zunächst alle Nachbargebiete des Gebietes ermittelt, welches die erlaubte Abweichung nicht erfüllt. Anschließend werden von diesen Gebieten alle Geometrien ermittelt, die unmittelbar benachbart zu dem Gebiet sind. Nun wird die Geometrie ermittelt, deren Umverteilung die beste Veränderung bezüglich Kompaktheit und Verbesserung der Gleichverteilung erzielt. Das Vorgehen ist in dem folgenden Ablaufdiagramm noch einmal detaillierter beschrieben:

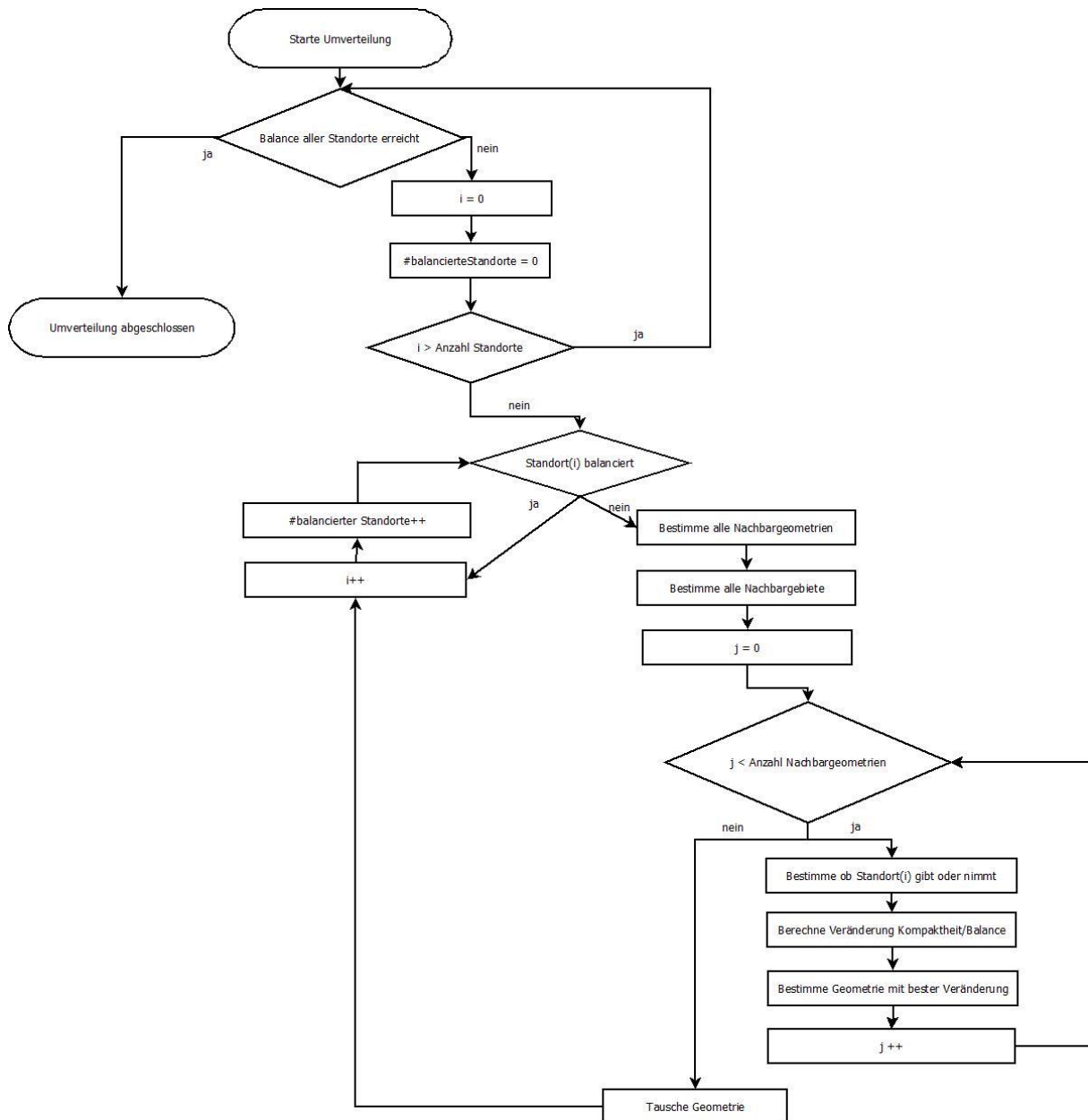


Abbildung 8 Ablauf der Gleichverteilung

c. Beachtung von Kompaktheit und Gleichverteilung

Zur Ermittlung der Veränderung von Kompaktheit und der Gleichverteilung wird die folgende Wichtungsfunktion verwendet:

$$\vartheta = \Delta c * \omega_1 + \Delta crit * \omega_2 \rightarrow \min$$

In dieser sind die Veränderung der Kompaktheit (Δc), sowie die Veränderung der Balance ($\Delta crit$) integriert. Diese werden mittels der gegebenen Faktoren gewichtet, so dass der Nutzer Einfluss auf die endgültige Form seiner Gebiete nehmen kann.

Die Veränderung der Kompaktheit wird mithilfe des Cox Maßes bestimmt. Dieses betrachtet das Verhältnis der Fläche des Gebietes zur Fläche eines Kreises mit gleichem Umfang wie das Gebiet. Je näher dieses Verhältnis dem Wert 1 liegt, umso kompakter ist das Gebiet. Schreibt man dieses in folgende Formel um:

$$\Delta c = 1 - c \rightarrow \min$$

So gilt es die Abweichung Δc zu minimieren.

Wird eine Aufteilung nach Fahrtstrecke/-zeit vorgenommen, wird die Kompaktheit nicht beachtet. Stattdessen soll die Summe aller Fahrtwege innerhalb eines übergeordneten Gebietes minimal werden. Es werden somit alle Summen verglichen und das Gebiet ausgewählt, was entsprechend der Wichtungsfaktoren den besten Wert liefert.

Für die Veränderung der Balance wird die Abweichung vom zu erreichenden Mittelwert bestimmt. Diese Abweichung sollte ebenso minimiert werden, was sich in folgender Formel ausdrücken lässt:

$$\Delta_{crit} = \left| 1 - \frac{crit_n}{\sum_1^n crit} \right| \rightarrow \min$$

Um nun die Veränderung des Wertes vergleichen und den besten Tauschpartner identifizieren zu können, muss die Veränderung der Werte nach dem Tausch zu denen vor dem Tausch verglichen werden. Der Wert mit der geringsten Abweichung, entspricht dem am meisten geeignetsten Tauschpartner.

$$\Delta \vartheta = \vartheta_{neu} - \vartheta_{old} \rightarrow \min$$

$\Delta \vartheta$ kann dabei Werte >0 sowie Werte <0 annehmen (bei Luftlinie):

- $\Delta \vartheta > 0$ kennzeichnet, dass eine Verschlechterung des Wertes der Kompaktheit und Balance vorliegt
- $\Delta \vartheta < 0$ kennzeichnet, dass eine Verbesserung des Wertes der Kompaktheit und Balance vorliegt

4. Algorithmus Greenfield

Der Greenfieldalgorithmus nutzt als Basis den Algorithmus der Gebietsverteilung. Jedoch müssen zunächst Standorte erzeugt werden, so dass die Gebietsverteilung durchgeführt werden kann. Um diese erstellen zu können, werden zunächst so viele übergeordneten Gebiete erstellt wie neue Standorte erzeugt werden. Dabei wird mit einem Randpolygon gestartet. Von diesem aus werden so viele benachbarte Gebiete zu diesem Polygon hinzugefügt, bis deren gemeinsame Summe der Haushaltszahl, etc. einen gewissen Schwellwert erreicht. Dieser Schwellwert berechnet sich aus der Anzahl aller Haushalte, etc. des gesamten Untersuchungsgebietes dividiert durch die Anzahl zu erzeugender Standorte. Um Fehler zu vermeiden, weil bereits alle Geometrien verteilt sind, wird dieser Schwellwert um 10% gesenkt. Sobald ein übergeordnetes Gebiet erzeugt wurde, wird das nächste Gebiet erzeugt. Dafür wird ein benachbartes Randpolygon als neuer Start verwendet. Falls kein Randpolygon zur Verfügung steht, wird entweder ein abseits gelegenes Randpolygon oder ein beliebiges noch nicht zugeteiltes Polygon als Startpunkt zur Erzeugung des nächsten übergeordneten Gebietes ausgewählt. Der Ablauf ist vereinfacht in der folgenden Grafik dargestellt:

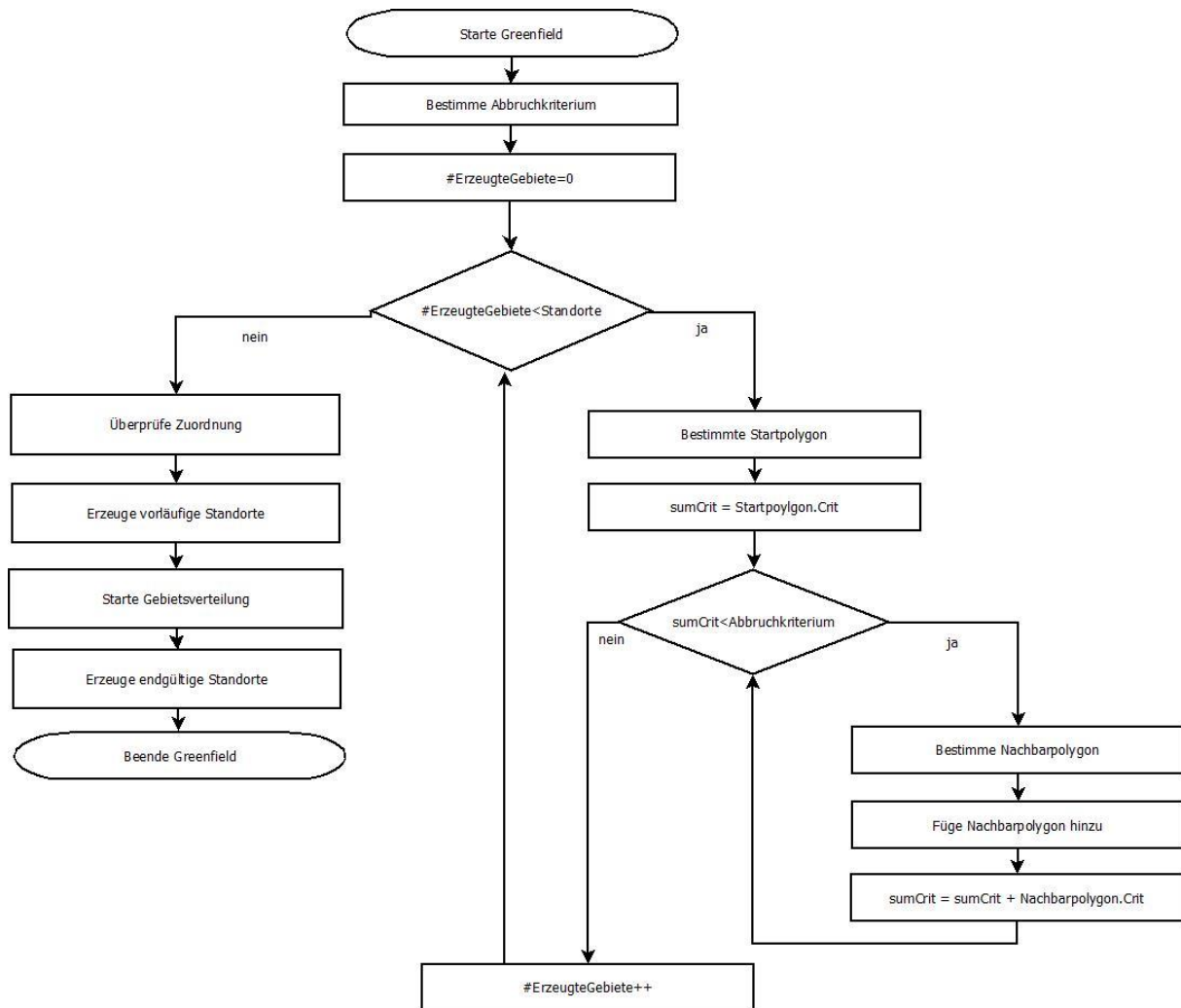


Abbildung 9 Ablauf der Greenfieldanalyse

Untersuchungsgebietes auf ein Gebiet zugewiesen sind. Durch die ungenaue Festlegung der Abbruchgrenze zur Erzeugung der anfänglichen übergeordneten Gebiete, kann es vorkommen, dass nicht alle Gebiete verteilt sind. Die nicht zugeteilten Gebiete werden somit dem übergeordneten benachbarten Gebiet zugeteilt, das die geringste Summe an Haushaltszahlen, etc. aufweist.

Sobald alle Geometrien zu übergeordneten Gebieten zusammengefasst worden sind, werden in die Mitte dieser Gebiete vorläufige Standorte gesetzt. Anschließend wird der bereits bekannte Gebietsverteilungsalgorithmus ausgeführt, mit Initialisierung, Gebietsverteilung und Visualisierung.

Nachdem diese Schritte ausgeführt worden sind, werden nun die endgültigen Standorte nach demselben Prinzip berechnet, wie es bei den vorläufigen Standorten der Fall war. Zu beachten ist, dass diese nun aufgrund der Umverteilung an einer anderen Position liegen können, als die vorläufigen Standorte. Weiterhin ist wichtig anzumerken, dass die Festlegung der „HomePolys“, also der Polygone, die eindeutig einem Standort zugeordnet werden müssen, da dieser sich innerhalb dieser Geometrie befindet, bei diesem Algorithmus nicht vorgenommen wird, da am Anfang keinerlei Standorte gegeben sind und die Bestimmung der endgültigen Position dieser nicht an eine bestimmte Geometrie gebunden ist.

Ein Ergebnis der Greenfieldanalyse ist in der folgenden Abbildung dargestellt.

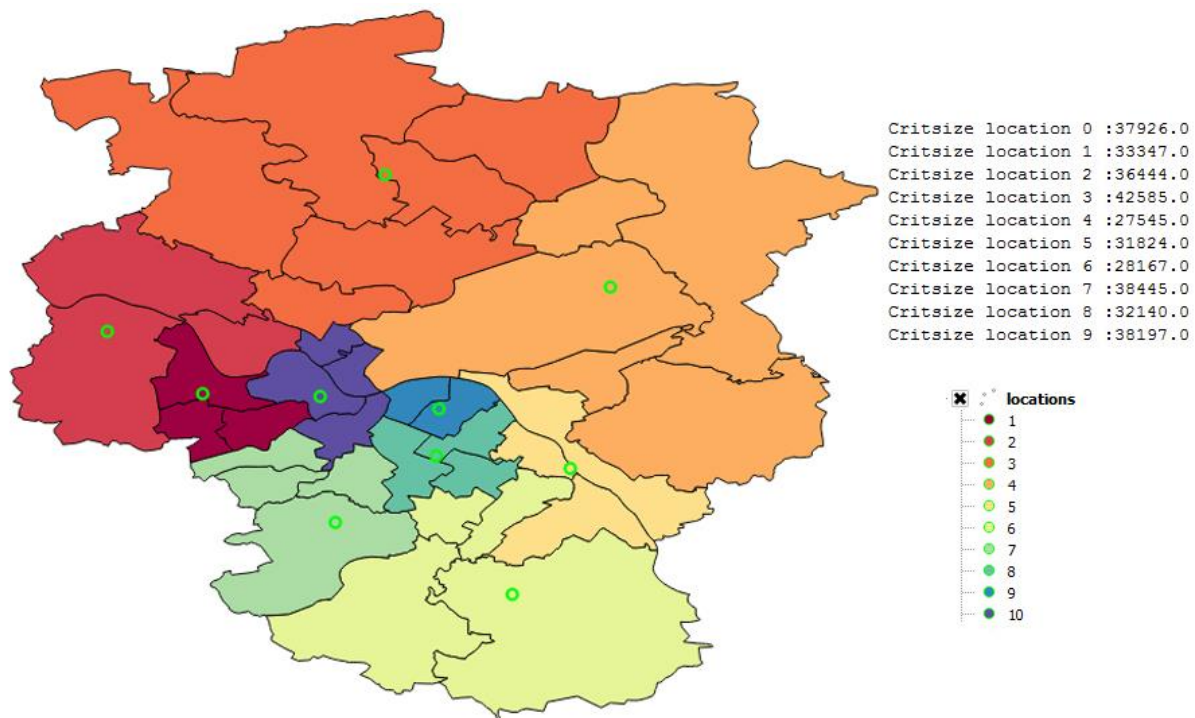


Abbildung 10 Ergebnis der Greenfieldanalyse im Raum Dresden. 10 neue Standorte wurden erzeugt, unter Verwendung folgender Parameter: Erlaubte Abweichung 30, Wichtungsfaktor Kompaktheit 10, Wichtungsfaktor Balance 90

5. Algorithmus Whitespot

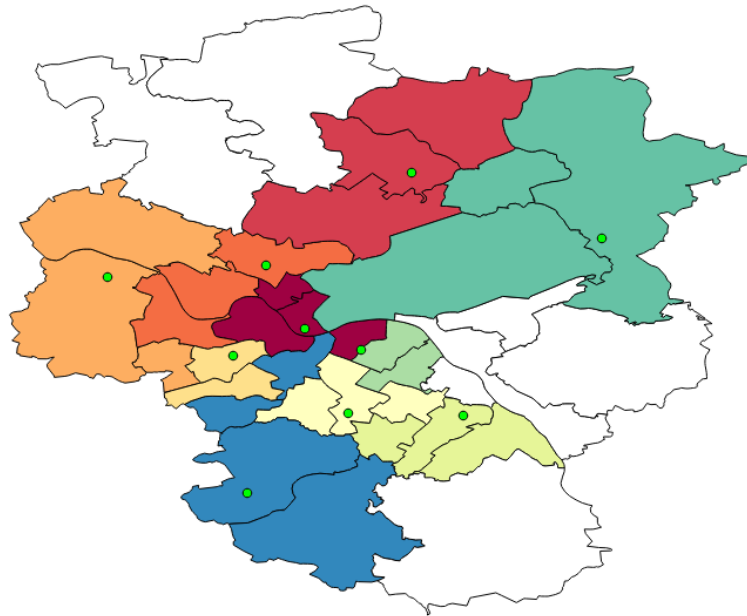
Ebenso wie der Greenfieldalgorithmus baut der Algorithmus der Whitespotanalyse ebenfalls auf den der Gebietsverteilung auf. Jedoch lassen sich hier erneut notwendige Erweiterungen finden. Die Idee der Whitespotanalyse ist es den „weißen Fleck“ auf der Landkarte zu finden. Dafür ist eine bestimmte Anzahl an bestehender Standorte vorgegeben. Zu diesen sollen eine bestimmte Anzahl weiterer Standorte hinzugefügt werden. In der Realität sind dies meist 1-2 Standorte. Um diesen Schritt vornehmen zu können, muss eine Stelle bestimmt werden, an der ein neuer Standort platziert werden kann. Zur Bestimmung dieser werden zunächst zugehörige Geometrien der gegebenen Standorte zu diesen zugeteilt. Das Abbruchkriterium ergibt sich wie bei der Greenfieldanalyse aus dem Mittelwert aller Haushaltszahlen, etc. Anschließend wird aus den noch nicht zugeteilten Gebieten die Geometrie ausgewählt, die das beste Potential für einen neuen Standort aufweist. Dafür werden die Haushaltszahlen, etc. aller möglichen Teilgebiete addiert. Das Teilgebiet mit der höchsten Summe wird das Gebiet, in das der Standort gesetzt wird. Von diesem Teilgebiet wird die Geometrie bestimmt, die die höchste Haushaltszahl, etc. besitzt. Diese wird nun als Startpolygon genommen. Weitere Geometrien werden zu diesem Startpolygon hinzugefügt, bis entweder:

- Das Abbruchkriterium erreicht ist oder
- Keine angrenzenden Geometrien mehr zur Verfügung stehen, die noch nicht einem Standort zugeteilt worden sind

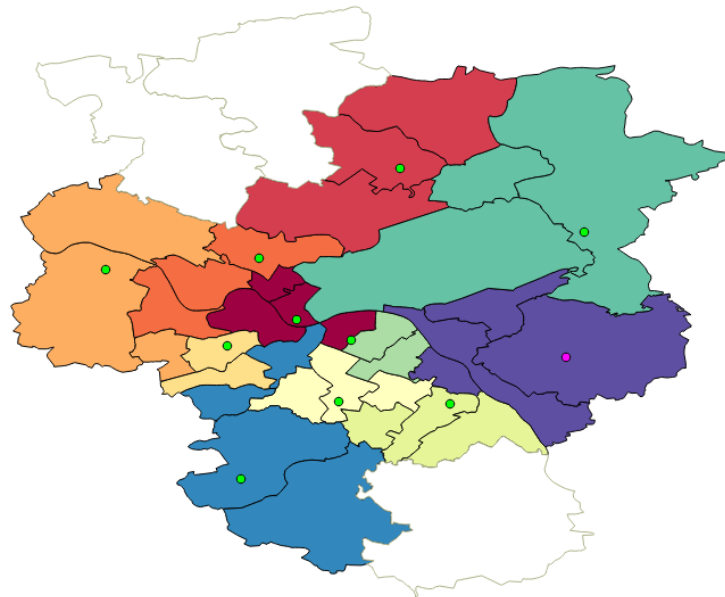
Nachdem diese Schritte vollzogen worden sind, wird wie bei der Greenfieldanalyse geprüft, ob nun alle Geometrien auf Standorte aufgeteilt sind. Falls nicht, werden diese noch verteilt. Anschließend wird der vorläufige Standort des neu zu erzeugenden Standortes bestimmt, so dass die eigentliche Gebietsverteilung starten kann. Nachdem diese durchlaufen ist, wird der endgültige neue Standort bestimmt. Dabei wird dieser aktuell wie bei der Greenfieldanalyse in die Mitte des übergeordneten Gebietes gesetzt.

Um die einzelnen Schritte bei der Whitespotanalyse zu verdeutlichen, sind die Ergebnisse im Folgenden als Grafiken aufgeführt.

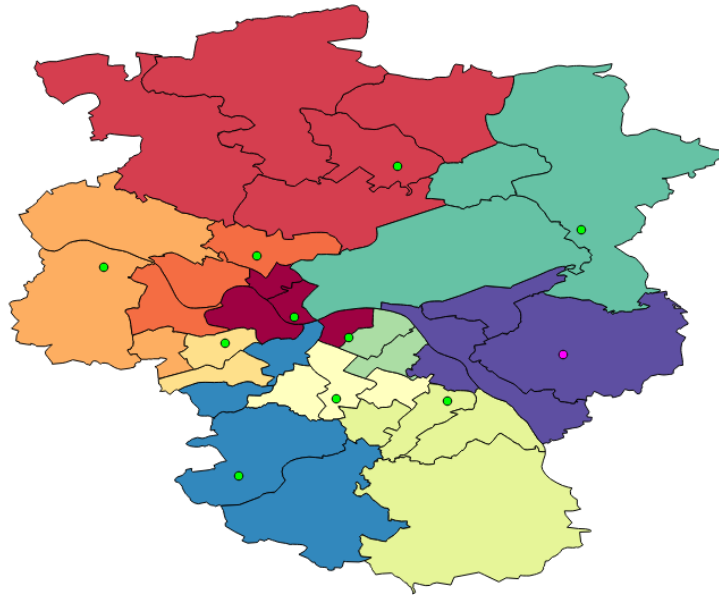
1. Zuordnung der Gebiete zu den bestehenden Standorte



2. Ermittlung des Teilgebietes mit höchster Haushaltszahl. Bestimmung des Startpolygons. Zuordnung angrenzender Gebiete und Setzen des vorläufigen Standorts



3. Zuordnung noch nicht aufgeteilter Geometrien



4. Rückgängigmachen der Zuordnungen. Anwendung des Gebietsverteilungsalgorithmus auf alle 11 Standorte.

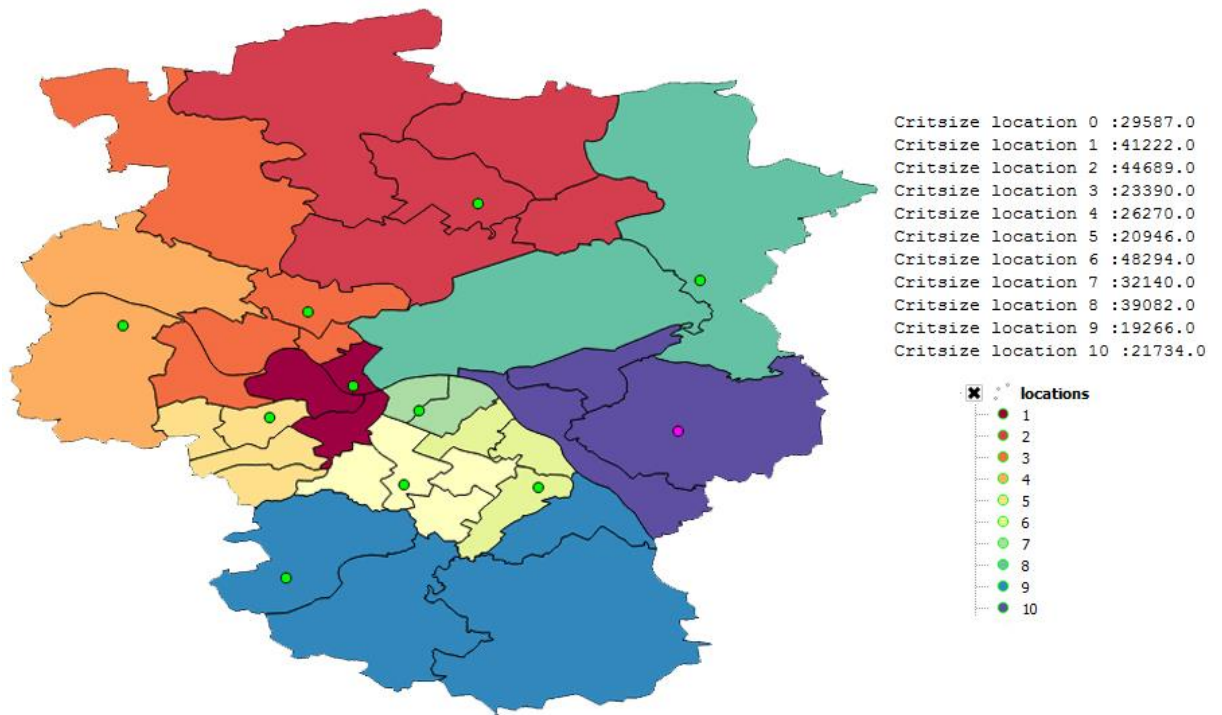


Abbildung 11 Ergebnis der Whitespotanalyse am Beispiel Raum Dresden. 10 gegebene Standorte (grün) ein neu hinzugefügter (pink), unter Verwendung folgender Parameter: Erlaubte Abweichung 50, Wichtungsfaktor Kompaktheit 100, Wichtungsfaktor Balance 0

6. Gleichverteilungsalgorithmus und Prüfung auf Zusammenhang

Innerhalb des Gebietsverteilungsalgorithmus wird nach der initialen Aufteilung der Geometrien auf Standorte eine Umverteilung der Gebiete durchgeführt, um eine möglichst gute Gleichverteilung des betrachteten Kriteriums zu erhalten. Dafür wird die Summe des Kriteriums bei jedem Standort geprüft, ob diese innerhalb der erlaubten Toleranz liegt. Der Toleranzwert wird durch den Nutzer eingegeben. Erzielt diese Prüfung ein negatives Ergebnis, so wird die Umverteilung

initiiert. Dabei werden zunächst alle möglichen Gebiete zur Verteilung bestimmt. Zu diesen Gebieten zählen lediglich die Geometrien, die direkt benachbart sind, zu dem übergeordneten Gebiet, für das eine Umverteilung notwendig ist. Anschließend wird für jedes Gebiet die Veränderung des Verhältnisses der Kompaktheit/Fahrtzeit/-strecke zur Gleichverteilung bestimmt. Anschließend wird die Geometrie ausgewählt, deren Veränderung die besten Werte des Kriteriums erfüllen. Tests zeigten nun, dass dieser komplexe Algorithmus einige Schwachstellen aufweist. Auf diese soll im Folgenden eingegangen werden.

Tests zeigten, dass der Umsortierungsalgorithmus zur Erzeugung der Gleichverteilung mitunter bereits nach mehreren Durchläufen keine Verbesserung erzielt. Stattdessen werden immer wieder die gleichen Gebiete hin und her verteilt, bis der Algorithmus abgebrochen wird. Aus diesem Grund sollte an der Stelle ein verbesserter Umsortierungsalgorithmus implementiert werden, so dass wirklich eine Verbesserung der Werte erzielt wird. Als Beispiel lässt sich da die Berechnung von 200 Standorten in Hamburg aufführen. Die Initialisierung der Parameter und die Aufteilung der Gebiete nach Distanz benötigt lediglich ca. 1-2 Minuten. Der Umsortierungsalgorithmus läuft weitere 3 Minuten, wenn man ihn nach 2000 Durchläufen abbricht. Bricht man ihn erst nach einer höheren Anzahl an Durchläufen ab, so findet trotz erhöhter Rechenzeit keine wirkliche Verbesserung der Ergebnisse statt, da nur dieselben Gebiete hin und her verteilt werden.

Des Weiteren zeigt der Umsortierungsalgorithmus Schwächen, sobald die Polygonanzahl pro Standort sehr groß wird. Die dabei bestehende Hauptursache ist die Überprüfung der Gebiete auf Zusammenhang. Für eine Vielzahl von Möglichkeiten muss geprüft werden, ob das resultierende Gebiet zusammenhängt. Dadurch müssen eine Vielzahl an Schleifendurchläufen absolviert werden. Ein Test an Gesamtdeutschland auf PLZ5-Ebene mit 10 Standorten zeigt das Problem deutlich auf. Sobald die Anzahl an Standorten erhöht wird und die Anzahl an zugeordneten PLZ5-Gebieten pro Standort deutlich reduziert wird, tritt dieses Problem nicht mehr auf. Ein weiteres Beispiel wurde an Hamburg durchgeführt (ca. 2500 PLZ8-Gebiete). Bei der Berechnung von 200 neuen Standorten, liegt eine Rechenzeit von 5 Minuten vor. Wird die Anzahl an Standorten jedoch auf 20 herabgesetzt, dauert die Berechnung aufgrund der Umverteilung mehrere Stunden. Es gibt mehrere Ansätze, dieses Problem zu beheben. Zum einen könnte ein anderer Startalgorithmus dabei helfen, die Anzahl an Umsortierung zu reduzieren. Anstatt der Aufteilung nach Distanzen könnte zum Beispiel ein kombinierter Ansatz aus der Masterarbeit verwendet werden. Ein weiterer Ansatz ist eine Optimierung der Einschränkung möglicher Tauschgeometrien. Da eine sinnvolle Verbesserung an dieser Stelle erst noch implementiert werden muss, wurde eine Möglichkeit eingebaut, dass der Nutzer entscheiden kann, ob die Geometrien zusammenhängen sollen oder nicht. Wird der Zusammenhang der Geometrien deaktiviert, findet die zeitintensive Überprüfung bei der Umverteilung nicht statt. Somit läuft der Algorithmus signifikant schneller. Jedoch kann es passieren, dass einzelne Geometrien eines übergeordneten Gebietes etwas abseits von diesem liegen. Tests zeigten jedoch, dass dies nur auf eine geringe Anzahl an Gebieten zutrifft.

7. Notwendige Parameter

Die vom Nutzer einzugebenden Parameter sind:

- Angabe der Planungsebene: PLZ5, PLZ8, ...
- Auswahl des Planungsraumes: Selektion von Gebieten der Planungsebene
- Auswahl Verfahren: Gebietsverteilung, Whitespot oder Greenfield

- Auswahl der Planungsvariable: Haushaltszahlen, ...
- Auswahl der Distanzberechnungsmethode: Luftlinie, Fahrtstrecke, Fahrtzeit
- Angabe der Anzahl neuer Standorte bei Greenfield und Whitespot
- Wert der erlaubten Abweichung
- Wichtungsfaktoren für Kompaktheit/Fahrtzeit und Gleichverteilung
- Deaktivierung/Aktivierung der Prüfung auf zusammenhängende Gebiete

8. Perspektiven

Die implementierten Algorithmen weisen an einigen Stellen Verbesserungspotential auf. Auf diese soll im Folgenden eingegangen werden.

a.) Ermittlung der erlaubten Abweichung

Zur Ermittlung der erlaubten Abweichung wird aktuell die Abweichung vom Mittelwert als Prozentwert angenommen. Wird zb als erlaubte Abweichung 50 angegeben und ist der zu erreichende Wert des Potentials 30.000, so dürfen die Ergebniswerte zwischen 15.000 und 45.000 liegen. Der Prozentwert wird in negative sowie positive Richtung beachtet. Anstatt eines Prozentwertes kann an dieser Stelle auch mithilfe von Varianzen gerechnet werden, um ein genaueres Ergebnis zu erzielen.

b.) Umverteilung der Gebiete zur Erzielung der Gleichverteilung

Der größte Verbesserungsbedarf liegt wohl bei der Umverteilung der Gebiete. Wie bereits im Kapitel 6 beschrieben, werden alle übergeordneten Gebiete bezüglich ihrer Kriteriensumme überprüft, ob sie innerhalb oder außerhalb der erlaubten Abweichung liegen. Falls sie außerhalb liegt, wird eine Umverteilung initiiert. Es wurde bereits im Kapitel 6 auf die Schwäche des aktuell implementierten Algorithmus eingegangen, die darin liegt, dass trotz einer Vielzahl von Durchläufen keine Verbesserung der Werte auftritt und zT die erlaubte Abweichung nicht erreicht wird (es gibt hingegen auch Rechenbeispiele, wo diese weit unterschritten wird). In diesem Fall wird ein Abbruchkriterium genutzt, um den Algorithmus nicht in einer Endlosschleife festhängen zu lassen. Bei diesem Algorithmus sollte eine Weiterentwicklung stattfinden, um dieses Problem zu beheben. Es wurde bereits versucht Bedingungen einzuarbeiten, dass ein Standort nicht das gleiche Gebiet zuordnen darf, was in dem Durchlauf zuvor bereits von einem Gebiet umverteilt wurde. Doch dies führte dennoch zu dem gleichen Problem, nur dass nun nicht mehr eine Geometrie zwischen mehreren Standorten hin und her getauscht wird, sondern es eine endliche Anzahl an Geometrien sind, die der Reihe um getauscht werden. Aus diesem Grund muss eine komplexere Lösung gefunden werden.

c.) Abbruch beim Greenfieldalgorithmus

Wie im Abschnitt Greenfieldalgorithmus beschrieben, wird bei der Zuordnung der Gebiete der Durchschnittswert der Haushaltszahlen als Abbruchkriterium genommen. Dieser wird um eine gewisse Prozentzahl verringert, um weniger Probleme zu verursachen, falls bereits zu viele Geometrien verteilt worden sind. An dieser Stelle ist eine sinnvollere Anwendung des Abbruchkriteriums ratsam.

Lösungsvorschlag:

Unabhängig von der verwendeten Planungsebene wird der Mittelwert und somit das Abbruchkriterium besteht, indem die Haushaltszahlensumme (oder ein anderes verwendetes Potential) durch die Anzahl an Standorten dividiert wird. Anschließend werden die Gebiete aufgeteilt, jedoch wird bei jeder Hinzunahme einer neuen Geometrie geprüft, ob sich durch diese an das Abbruchkriterium annähert wird oder wieder weiter entfernt wird, als die vorherige Differenz war. Nur wenn eine Annäherung an das Abbruchkriterium stattfindet, wird die Zuordnung durchgeführt.

d.) Verbesserung der Serverrechenzeit

Durch einen besseren Server ließe sich die Rechenzeit des Algorithmus verringern. Eine weitere Option sind bereits vorgefertigte Tabellen, aus denen Daten ausgelesen werden können, um die rechenaufwendigen Abfragen bei der Initialisierung nicht jedes Mal erneut ausführen zu müssen. Vor allem die Berechnung der Nachbargebiete mit den zugehörigen gemeinsamen Kanten benötigt den Großteil der Rechenzeit. Bei der Whitespot- und Greenfieldanalyse kommt zusätzlich noch die Bestimmung der Randpolygone hinzu, die bei einem großen Datensatz wie Gesamtdeutschland einen erheblichen Teil der Rechenzeit in Anspruch nimmt.

e.) Verbesserung der Rechenzeit bei großen Datenmengen

Tests zeigten, dass es mehrere Stellen zur Verbesserung der Rechenzeit gibt. Vor allem betrifft dies verwendete for-Schleifen, bei denen über alle Geometrien iteriert wird. Bei der Greenfieldanalyse wurde bereits versucht, solche Schleifen zu optimieren, jedoch bestehen weitere Optimierungsstellen innerhalb des Gleichverteilungs- und Whitespotalgorithmus.

Ein Beispiel: Wie im Abschnitt Greenfieldalgorithmus beschrieben, werden zunächst Gebiete erzeugt, um die vorläufigen Standorte bestimmen zu können. Beim Erzeugen der Gebiete werden von der Startgeometrie aus neue Geometrien, die eine geringe Distanz zur Startgeometrie aufweisen, hinzugefügt bis das Abbruchkriterium erreicht ist. Ursprünglich wurde dabei mehrfach über die gesamte Geometriemenge iteriert, um nach und nach die Geometrien mit den kürzesten Distanzen zu bestimmen. Bei der Performanceoptimierung wurden alle Geometrien zunächst nach ihrer Distanz sortiert. Anschließend wurden so viele Geometrien hinzugenommen, wie nötig waren, um das Abbruchkriterium zu erfüllen. Dafür brauchte lediglich zunächst die erste, dann die zweite,... Geometrie aus dem sortierten Feld entnommen werden. Solche Verbesserungen sind auch bei der Gebietsverteilung und der Whitespotanalyse möglich.

f.) Sonderfälle

a. Inseln

Aktuell lassen sich nur Gebiete aufteilen (wenn die Prüfung auf Zusammenhang aktiviert ist), bei denen keine Inseln vorhanden sind. Das Problem von Inseln ist, dass sie keine Nachbar haben, somit endet der Algorithmus in einer Endlosschleife, da die erzeugten Gebiete zusammen hängen müssen, dies jedoch bei Inseln nicht möglich ist. Somit muss ein Ansatz gefunden werden, der das Wasser zwischen Insel und Festland herausrechnet. Da es dabei jedoch eine Vielzahl von Sonderfällen gibt (z.B. Insel beherbergt mehrere PLZ-Gebiete), muss der Algorithmus einen komplexen Sachverhalt abdecken.

b. PLZ in PLZ

Es ist möglich, dass eine Anordnung von Gebieten besteht, bei denen ein Gebiet in einem anderen enthalten ist. Tests zeigen, dass dieser Fall keine Probleme bei der Aufteilung der Gebiete darstellt.

g.) Erweiterbarkeit des Algorithmus

a. Zentrenoptimierung

Aktuell werden neu generierte Standorte bei der Greenfield- und der Whitespotanalyse in die Mitte des erzeugten Gebietes gesetzt. Eine sinnvolle Erweiterung dieses Vorgehens ist eine Standortplatzierung unter Beachtung von potentiell geeigneten Standorten, die mit Hilfe der microm Potentialanalyse erzeugt worden sind.

b. STRAB

Es war zu prüfen, ob die STRAB-Ebene ebenso als Planungsebene genutzt werden kann. Die Betrachtung der Geometrien der STRAB zeigen, dass der Algorithmus nicht für Punktgeometrien genutzt werden kann. Mehrere Aspekte schließen dies aus:

- Die Initialisierung der Geometrien ist auf Polygone ausgelegt, da u.a. der Flächeninhalt und Umfang ausgelesen wird
- Die Verteilung und Umsortierung der Gebiete nutzt das Kompaktheitsmaß als Qualitätskriterium, welches Flächeninhalte und Umfänge von Geometrien zur Berechnung benötigt. Diese sind durch die Punktgeometrien nicht vorhanden.

Es sind somit Ergänzungen notwendig, um den Algorithmus auf STRAB-Ebene nutzen zu können. Es gilt die Initialisierung in der Art und Weise zu verändern, dass keine Flächeninhalte und Umfänge ausgelesen werden. Zudem müssen Nachbarpunkte mit Hilfe von Umkreisen bestimmt werden. Die aktuell implementierte Funktion kann für Punktgeometrien also nicht genutzt werden. Bei der Gebietsaufteilung müssen auch Veränderungen vorgenommen werden. So fällt die Bestimmung der „HomePolys“ weg und die Überprüfung auf zusammenhängende Gebiete muss anders berechnet werden, zum Beispiel indem aus den Punkten, die einem Standort zugeordnet sind, Flächen berechnet werden und geprüft wird, ob die Fläche keine Lücke aufweist. Des Weiteren muss das Setzen neuer Standorte geklärt werden, da diese aktuell in die Mitte des erzeugten Polygons gesetzt werden. Jedoch gilt zu prüfen, ob der Standort auf einem Straßenabschnitt platziert werden muss oder im Raum liegen darf. Weitere notwendige Veränderung lassen sich bei der Implementierung des Algorithmus auf STRAB-Ebene erkennen, dies soll jedoch ein Auszug von notwendigen Anpassungen sein.