



*Julius-Maximilians-Universität Würzburg*  
*Institut für Informatik*  
*Lehrstuhl für verteilte Systeme (Informatik III)*

# **Untersuchung von Clusterungsverfahren zur diskreten Charakterisierung des Verkehrsaufkommens in mobilen Kommunikationssystemen**

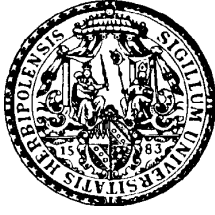
Diplomarbeit im Fach Informatik  
vorgelegt von

**Marius Heuler**

Angefertigt am



Lehrstuhl für verteilte Systeme (Informatik III)  
Bayerische Julius-Maximilians-Universität Würzburg



*Julius-Maximilians-Universität Würzburg*  
*Institut für Informatik*  
*Lehrstuhl für verteilte Systeme (Informatik III)*

# **Untersuchung von Clusterungsverfahren zur diskreten Charakterisierung des Verkehrsaufkommens in mobilen Kommunikationssystemen**

Diplomarbeit im Fach Informatik  
vorgelegt von

**Marius Heuler**

geboren am 3. November 1972 in Werneck

Angefertigt am



Lehrstuhl für verteilte Systeme (Informatik III)  
Bayerische Julius-Maximilians-Universität Würzburg

Betreuer:

Prof. Dr.-Ing P. Tran-Gia  
Dipl.-Inform. K. Tutschku

Abgabe der Arbeit:

24. April 1998

# Erklärung

Ich versichere, die vorliegende Diplomarbeit selbständig und unter ausschließlicher Verwendung der angegebenen Literatur angefertigt zu haben.

Würzburg, den 24. April 1998

---

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>5</b>
<b>2</b>	<b>Planung mobiler Kommunikationssysteme</b>	<b>1</b>
2.1	Bedarfsorientierte Planung .....	7
2.2	Das Konzept der Demand-Nodes .....	9
2.2.1	Definition eines Demand-Nodes.....	11
2.2.2	Zielkriterien für die Erzeugung von Demand-Nodes.....	12
2.2.3	Diskretisierung der Verkehrsdichte.....	14
2.2.4	Gewinnung der Daten über die Verkehrsdichte.....	14
2.2.5	Erzeugung der Demand-Nodes.....	17
2.3	Testgebiete für die Demand-Node Erzeugung.....	18
2.3.1	Würzburger Gebiet.....	18
2.3.2	Dallas Gebiet.....	18
2.3.3	Synthetisches Gebiet .....	19
2.3.4	Konstantes Gebiet.....	21
<b>3</b>	<b>Clusterungsverfahren</b>	<b>22</b>
3.1	Verfahren des Rekursiven Partitionierens .....	24
3.1.1	Algorithmus .....	24
3.1.2	Ergebnisse des Algorithmus.....	28
3.1.3	Nachteile des Algorithmus.....	31
3.1.4	Merkmale des Algorithmus.....	32
3.2	Verfahren des Zusammenfassens von Knoten.....	33
3.2.1	Algorithmus .....	33
3.2.1.1	Zufällige Knotenauswahl.....	35
3.2.1.2	Prüfung der benachbarten Knoten.....	35
3.2.1.3	Zusammenfassen des Knotens mit einem Nachbarn ...	38
3.2.1.4	Erzeugung der Demand-Nodes aus den Knoten .....	38
3.2.2	Ergebnisse des Algorithmus.....	40
3.2.3	Bewertung des Algorithmus .....	43
3.2.4	Verkehrsverteilung der Demand-Nodes .....	44

3.3	Optimierung des Zusammenfassens von Knoten .....	46
3.3.1	Optimierungsphase I .....	46
3.3.1.1	Algorithmus.....	47
3.3.1.2	Verkehrsverteilung der Demand-Nodes.....	50
3.3.2	Optimierungsphase II .....	52
3.3.2.1	Algorithmus.....	52
3.3.3	Ergebnis nach beiden Optimierungsphasen.....	55
3.3.4	Verbesserte Optimierung .....	57
3.3.5	Ideal erreichbare Verkehrsverteilung.....	60
3.3.6	Ergebnisse des Algorithmus mit Optimierung .....	62
3.3.7	Merkmale des Algorithmus.....	64
<b>4</b>	<b>Vergleich verschiedener Verfahren</b> .....	<b>65</b>
4.1	Testverfahren.....	67
4.1.1	Messung einer Testfläche .....	67
4.1.2	Verwendung einer quadratischen Testfläche .....	69
4.1.3	Meßreihe einer gegebenen Testfläche.....	70
4.1.4	Meßergebnisse abhängig vom Radius .....	71
4.1.5	Kriterien für geeignete Parameter der Meßkurven .....	72
4.1.6	Parameter der Meßkurven .....	72
4.2	Meßergebnisse des Verkehrsvergleichs.....	74
4.2.1	Auswahl der Meßkurven.....	75
4.2.2	Vergleich beider Verfahren.....	77
4.2.3	Ergebnisse mit quadratischer Testfläche .....	79
4.2.4	Untersuchung der Optimierung .....	81
4.2.5	Vergleich verschiedener Anzahlen an Demand-Nodes.....	85
4.2.6	Untersuchung der Auswirkung der Flächenbegrenzung.....	88
4.3	Repräsentation der Landnutzungsklassen.....	93
4.3.1	Meßverfahren .....	93
4.3.2	Meßergebnisse.....	96
<b>5</b>	<b>Anwendungsbeispiele der Clusterung</b> .....	<b>98</b>
5.1	Einfluß der Clusterung der Teilnehmer auf die subjektive Dienstgüte .....	98
5.1.1	Modellierung der subjektiven Dienstgüte.....	98

5.1.2	Beschreibung des verwendeten Modells.....	99
5.1.3	Definition der subjektiven Dienstgüte .....	100
5.1.4	Modelle der Verteilung der Teilnehmer .....	101
5.1.5	Messung der Verteilung der Teilnehmer .....	102
5.1.6	Ergebnisse der subjektiven Dienstgüte .....	104
5.1.7	Modellierung mit Wiederholungsversuchen.....	105
5.1.8	Bewertung der Ergebnisse .....	107
5.1.9	Stochastische Punktprozesse als Planungsmodell .....	107
5.2	Integriertes Planungstool <i>ICEPT</i> .....	108
5.2.1	Struktur heutiger mobiler Kommunikationsnetze .....	108
5.2.2	Planung mobiler Kommunikationssysteme mit <i>ICEPT</i> .....	109
5.2.3	Planungsergebnisse von <i>ICEPT</i> .....	111
5.3	Bedarfsabschätzungswerkzeug <i>CUTE</i> .....	112
5.4	Senderpositionierung anhand der Demand-Nodes.....	114
<b>6</b>	<b>Zusammenfassung und Ausblick</b>	<b>116</b>
<b>7</b>	<b>Literaturverzeichnis</b>	<b>118</b>
<b>8</b>	<b>Abbildungsverzeichnis</b>	<b>121</b>

# 1 Einleitung

In der heutigen Zeit ist eines der vordringlichsten Ziele die optimale Ausnutzung der vorhandenen Ressourcen und die Minimierung des Verbrauchs an Ressourcen für zukünftige Anwendungen. Dies trifft nicht nur bei den natürlichen Rohstoffen zum Schutz der Umwelt zu, sondern auch in allen Bereichen der Ökonomie. Natürlich muß dabei auch die Qualität der Anwendung sichergestellt werden. Die verfügbaren Ressourcen sind sehr häufig ein beschränktes Gut. In Kommunikationssystemen z.B. ist die verfügbare Kapazität im mobilen Bereich durch den zur Verfügung stehenden Frequenzbereich und in Festnetzen durch die Bandbreite der verlegten Leitungen eingeschränkt.

Für eine bestimmte Anwendung hängen die dazu nötigen Ressourcen maßgeblich vom Bedarf danach ab, der oft stark mit der Position schwankt. Um dieser Verteilung des Bedarfs Rechnung zu tragen, werden meist verteilte Systeme eingesetzt, bei der die Versorgung der Gesamtfläche auf einzelne Standorte aufgeteilt ist. Jeder Standort bedient dabei nur den Bedarf in seinem Bereich. Die Position einer Versorgungseinrichtung legt auch die nötigen Ressourcen für sie fest. Im Wettbewerb hängt der Erfolg einer Anwendung entscheidend von der gewinnbringenden Wahl der Standorte ab, da diese im nachhinein nicht mehr revidiert werden können. Das zweite Kriterium für eine erfolgreiche Anwendung ist die optimale Verteilung der verfügbaren Ressourcen, bzw. eine ausreichende Bereitstellung dieser.

Bei der Standortplanung wird hinsichtlich einer möglichst geringen Anzahl von Versorgungseinrichtungen und einem möglichst kleinen Verbrauch an Ressourcen optimiert. Zusätzlich muß eine vorher festgelegte Güte des Versorgungsgrades sichergestellt werden. Der Versorgungsgrad wird normalerweise durch die Höhe des abgedeckten Bedarfs definiert. Dazu wird seine Verteilung im Planungsgebiet modelliert. Aus Gründen der Komplexität kann für die Planung nicht jeder einzelne Punkt mit einem bestimmten Bedarf herangezogen werden. Oft sind auch keine genauen Informationen über einzelne Bedarfspunkte vorhanden, sondern es ist nur eine Abschätzung des ungefähren Bedarfs in einem Gebiet möglich. Die räumliche Verteilung des Bedarfs ist zusätzlich meist sehr ungleichmäßig. Der Bedarf nach Kommunikation z.B. ist im urbanem Gebiet um Größenordnungen höher als im freien Gelände. Aus diesen Gründen wurde das Konzept der Demand-Nodes entwickelt, siehe dazu [27] und [30]. Die Demand-Nodes stellen dabei Punkte dar, die einen vorgegebenen Bedarf auf ihrer Fläche repräsentieren. Sie sind der Repräsentant des Gesamtverkehrs, der durch die einzelnen Punkte mit gegebenem Bedarf in ihrer Umgebung gebildet wird. Diese Punkte können auch allein virtuell vorhanden sein, falls z.B. nur eine Abschätzung über die Summe des Kommunikationsverkehrs möglich ist. Die Demand-Nodes werden durch ein sogenanntes Clusterungsverfahren erzeugt, wobei mit Clusterung die Ungleichmäßigkeit der Bedarfsverteilung bezeichnet wird. Die Approximationsgüte des Bedarfs durch die Demand-Nodes ist dabei das wich-

tigste Qualitätskriterium. Die Clusterungsverfahren haben aus diesem Grund einen entscheidenden Einfluß auf den Erfolg der Planung. Daher ist die Intention dieser Arbeit die Entwicklung dieser Clusterungsverfahren und ihre Bewertung im Hinblick auf ihre Eignung für gegebene Problemstellungen. Der entwickelte Algorithmus sollte zuerst für die Planung von mobilen Kommunikationssystemen verwendet werden, weshalb in dieser Arbeit nur die Anwendung für diesen Fall dargelegt wird. Die Grundlagen und der entwickelte Algorithmus sind jedoch universell für alle Probleme der Standortplanung einsetzbar.

Ziel der Arbeit ist es, Algorithmen zur Modellierung des Verkehrsaufkommens in mobilen Kommunikationssystemen zu entwerfen und zu untersuchen, wozu das Konzept der Demand-Nodes verwendet wird. In diesem Rahmen wird ein neues Verfahren zur Erzeugung der Demand-Nodes entwickelt. Der Ansatz dieses Verfahrens des Zusammenfassens von Knoten ist vollkommen verschieden von dem des bekannten rekursiven Partitionierens. Die Nachteile des rekursiven Partitionierens, die sich bei der Bewertung dieses Algorithmus gezeigt haben, sind bei dem neuen Verfahren nicht mehr vorhanden. Beide in dieser Arbeit beschriebenen Algorithmen sind in das am Lehrstuhl entwickelte Mobilfunkplanungssystem *ICEPT*, siehe dazu Kapitel 5.2, integriert worden und können wahlweise zur Planung eingesetzt werden. Auf Grundlage des neuen Algorithmus wird das Werkzeug *CUTE* zur Abschätzung des Bedarfs entwickelt. Damit kann der Bedarf durch die Demand-Nodes für vielfältige Zwecke modelliert werden. Im Rahmen der Darstellung der beiden Verfahren wird auch eine qualitative Bewertung der Eignung durchgeführt. Zusätzlich wird ein Gütemaß definiert, auf dessen Grundlage die Clusterungsverfahren auch quantitativ bewertet werden können. Die Qualität der Approximation kann hiermit in Zahlen gefaßt werden.

In der vorliegenden Arbeit wird zuerst in Kapitel 2 eine Einführung in die Planung von mobilen Kommunikationsnetzen gezeigt, wobei in Kapitel 2.2 das Konzept der Demand-Nodes näher beleuchtet wird. In Kapitel 3 werden nach einer kurzen Einführung in das Konzept der Clusterung zwei Verfahren dazu beschrieben. Nach der Darstellung des Verfahrens des rekursiven Partitionierens wird das Verfahren des Zusammenfassens von Knoten entwickelt, wobei eine qualitative Untersuchung beider Verfahren enthalten ist. Ein quantitatives Gütemaß wird in Kapitel 4 definiert und zur Untersuchung verschiedener Aspekte der Approximationsgüte der Ausgangsdaten verwendet. Anwendungsbeispiele für das neue Clusterungsverfahren und die Auswirkungen der Clusterung werden in Kapitel 5 dargelegt. In Kapitel 6 befindet sich schließlich eine Zusammenfassung der Ergebnisse und ein Ausblick auf zukünftige Anwendungen.



## 2 Planung mobiler Kommunikationssysteme

Für die Planung eines mobilen Kommunikationssystems sind verschiedene Ansätze denkbar. Die klassischen, analytischen Planungsmethoden, die in [6] und [12] beschrieben werden, sind schon in den bekannten Planungswerkzeugen wie *PEGASOS* [11] oder *PLANET* [18] im Einsatz. Die Positionen der Sender werden im wesentlichen durch die Eingaben eines erfahrenen Netzwerkingenieurs festgelegt, der durch diese Programmpakete dabei unterstützt wird. Der große Nachteil dieses Ansatzes ist die starke Abhängigkeit der Planungsgüte von der Ausbildung und Erfahrung des Ingenieurs.

In einem modernen, integrativen Ansatz, wie er in dem Paket *ICEPT*, siehe [27], [29] und [30], verwendet wird, läßt sich der Vorgang durch die Verwendung von Daten über die Verteilung des Kommunikationsbedarfs im Planungsgebiet automatisieren. Das Planungssystem *ICEPT* wird in Kapitel 5.2 beschrieben. Dieser bedarfsorientierte Ansatz ergibt bei der Planung von mobilen Kommunikationssystemen wesentliche Vorteile. Einer ist die Möglichkeit der Automatisierung des Planungsvorgangs. Der wesentlich Nutzen liegt aber darin, daß die Planung nicht mehr von dem subjektiven Urteil eines Ingenieurs abhängt. Die Position und Kapazität der Sender wird statt dessen durch einen definierten Algorithmus festgelegt, der ein objektives Ergebnis liefert. Das Resultat wird deterministisch durch die Wahl der Planungsparameter und der Vorgaben hinsichtlich der Dienstgüte bestimmt.

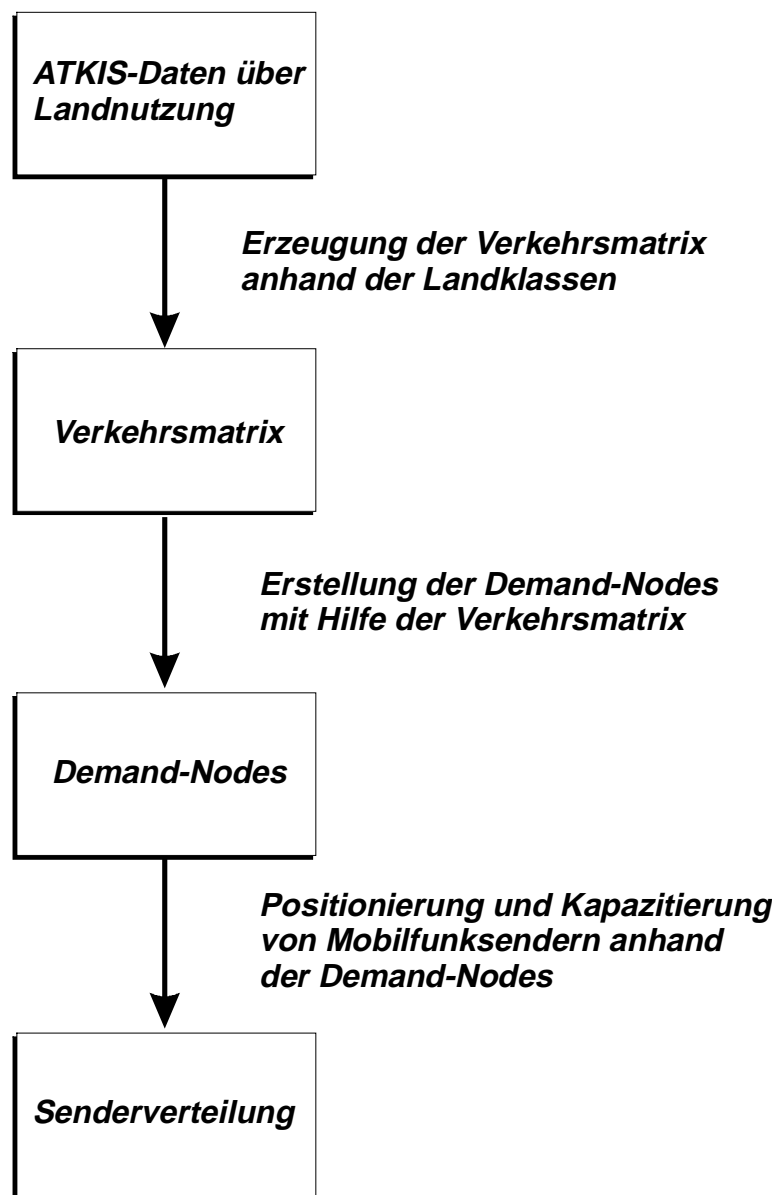
Der wichtigste Aspekt bei der Planung ist die Senderpositionierung, da zum einen die Auswahl eines Standorts durch die Geographie und Morphologie beschränkt ist. Zum anderen kann die Position eines Senders später nicht mehr verändert werden. Andere Parameter, wie die Kanalverteilung oder die Kapazität, können im nachhinein noch adaptiert werden. Nur die bedarfsorientierte Planung ermöglicht eine optimale Positionierung der Sender gemäß vorgegebener Konstanten, wie verfügbare Frequenzen oder physikalische Gegebenheiten, und wählbarer Parameter, wie Dienstgüte oder Versorgungsgrad.

Um Unklarheiten der Darstellung zu vermeiden, werden im folgenden Text die englischen Originalbezeichnungen wie z.B. „Demand-Node“ statt etwa „Bedarfsknotenpunkt“ verwendet.

### 2.1 Bedarfsorientierte Planung

Die verschiedenen Phasen bei der bedarfsorientierten Planung eines mobilen Kommunikationssystems werden in Abbildung 2.1.1 aufgezeigt. In dieser Arbeit wird vor allem der Schritt der Erzeugung der Demand-Nodes aus der Verkehrsmatrix betrachtet. Die Erzeugung der Verkehrsmatrix wird in [14] untersucht.

Verschiedene Verfahren zur Positionierung und Kapazitiering von Sendern werden in [15] verglichen.



*Abbildung 2.1.1: Phasen der Planung eines mobilen Kommunikationssystems*

Bei der bedarfsorientierten Planung ist es besonders wichtig, die Teilnehmeraktivität korrekt zu modellieren. Zur Planung und Dimensionierung von Netzen muß vor allem die Lage und die Intensität von Benutzeraktivität erfaßt werden. Der Bedarf nach Kommunikation ergibt sich durch die Aktivität der Benutzer, z.B. durch die von ihnen geführten Telefongespräche. Durch den Kommunikationsbedarf verschiedener Benutzer im Planungsgebiet wird eine bestimmte Höhe an Kommunikationsverkehr erzeugt. Dieser Kommunikationsverkehr pro Fläche wird als Verkehrsdichte oder Verkehrsintensität bezeichnet und ist stark abhängig von der räumlichen Verteilung der Benutzer.

Die räumliche Verteilung der Verkehrsdichte bestimmt die Standorte der Sender und die Höhe der Verkehrsdichte die Kapazität der Sender. Um z.B. die Anzahl der nötigen Kanäle für einen bestimmten Sender festzulegen, wird der Kommunikationsverkehr in dem Versorgungsgebiet dieses Senders bestimmt. Anhand des Kommunikationsverkehrs kann dann die Anzahl der benötigten Kanäle bei einer vorher festgelegten Dienstgüte berechnet werden.

Durch die Mobilität der Teilnehmer und vor allem durch die je nach Tageszeit unterschiedliche Intensität der Aktivität ist auch die Verkehrsdichte zeitlich abhängig. Zur Planung wird daher die höchste Verkehrsdichte herangezogen, die im Verlauf eines Tages auftritt. Dazu wird die Verkehrsdichte verwendet, die in der Hauptverkehrsstunde, im Englischen „busy hour“ genannt, vorliegt.

Eine andere Möglichkeit ist es, Szenarien für verschiedene Tageszeiten zu erstellen und mit diesen zu planen. Dabei ergeben sich aber mehrere Lösungen, die jeweils zu einem bestimmten Zeitpunkt optimal sind. Die Position der Sender kann jedoch nicht zeitabhängig variiert werden, allenfalls die Kapazität einzelner Sender. In den heute verbreiteten Mobilfunksystemen, wie z.B. nach dem GSM-Standard (Global System for Mobile Communications), sind die Kapazitätszuteilungen zu einzelnen Sendern aber statisch. Für diese Systeme reicht die Planung mit einer statischen Verkehrsdichte aus. Falls die Kapazitätsverteilung zwischen Sendern dynamisch adaptierbar ist, kann mit zeitabhängigen Verkehrsdichten gearbeitet werden. Dies ist z.B. bei den auf dem CDMA-Verfahren (Code Division Multiple Access) basierenden Netzen möglich. Die Repräsentation des Kommunikationsverkehrs im Planungsgebiet durch die Demand-Nodes ist aber in diesem Fall ebenso anwendbar.

## **2.2 Das Konzept der Demand-Nodes**

Die alleinige Grundlage für die Planung von mobilen Kommunikationssystemen sind die Teilnehmer des Dienstes, durch deren Kommunikation Verkehr erzeugt wird. Durch ihre Position und Mobilität wird die Lage der Sender und durch die Gesprächsintensität und Gesprächsdauer die Kapazität der Sender bestimmt. Falls man a priori diese Daten aller Teilnehmer kennen würde, könnte das Mobilfunknetz optimal ausgelegt werden. Zum einen sind nicht alle Daten bekannt, zum anderen würde die Berücksichtigung jedes einzelnen Benutzers bei der Planung die Komplexität und Datenfülle so erhöhen, daß eine Planung schlichtweg unmöglich wäre. Aus diesen Gründen faßt man den von einzelnen Teilnehmern erzeugten Kommunikationsverkehr zu Punkten zusammen, die einen festen, vorher festgelegten, Kommunikationsverkehr repräsentieren.

Diese Punkte werden in der Mobilfunkplanung „Demand-Nodes“ genannt und im nächsten Kapitel exakt definiert. Die Demand-Nodes werden durch ein Clusterungsverfahren aus den vorhandenen Daten über die Verkehrsdichte erzeugt. Die Clusterung wird in Kapitel 3 definiert und verschiedene Verfahren dazu beschrieben.

Durch die Erzeugung der Demand-Nodes wird zum einen eine Diskretisierung des von den Teilnehmern initiierten Kommunikationsverkehrs erreicht. Wie in Kapitel 2.2.1 gezeigt wird, werden statistische Eigenschaften des durch die Kommunikation erzeugten Verkehrs ausgenutzt, um eine Beschreibung des Verkehrs pro Fläche zu erreichen. Dadurch wird eine Funktion der Verkehrsdichte definiert, die die Höhe der Verkehrsdichte an jedem Punkt im Planungsgebiet angibt. Die hiermit gewonnene Diskretisierung des Kommunikationsverkehrs ist essentiell, da keine Daten über die Position und die dort geführten Gespräche einzelner Teilnehmer vorliegen. Die Verkehrsdichtefunktion dient vor allem als theoretisches Modell für die Planung. Meist sind zu wenig Informationen über die Verteilung des Kommunikationsverkehrs vorhanden, um direkt die Verkehrsdichtefunktion angeben zu können.

Zum anderen hat die Aggregation des Teilnehmerverkehrs eine räumliche Diskretisierung des Kommunikationsverkehrs der einzelnen Teilnehmer zum Ergebnis. In der Planung muß nicht mehr die Kommunikation einzelner, sich bewegender Teilnehmer berücksichtigt werden. Statt dessen liegt eine Menge von Punkten vor. Jeder Punkt repräsentiert dabei den Kommunikationsverkehr der Teilnehmer in seiner Umgebung. Die Positionen der Demand-Nodes können auch als virtuelle, bewegungslose Teilnehmer angesehen werden, die einen bestimmten Kommunikationsverkehr repräsentieren.

Die beiden Diskretisierungen sind nötig, um die gewaltigen Datenmengen, die eine Berücksichtigung der Aktivität aller Benutzer erzeugen würde, auf ein für die Planung handliches Maß zu reduzieren. Außerdem sind meist nur unvollständige Daten über den Bedarf an Kommunikation vorhanden. Die diskretisierte Repräsentation des Kommunikationsverkehrs durch die Demand-Nodes kann man wesentlich einfacher abschätzen als die Informationen über unabhängig agierende einzelne Teilnehmer. Zu dieser Abschätzung können z.B. geographische Daten verwendet werden.

Das Konzept der Demand-Nodes stammt ursprünglich aus dem Bereich der Standortplanung, welches in [7] beschrieben wird. In [27] und [30] wurde es für die Planung von mobilen Kommunikationssystemen adaptiert. In [7] werden Strategien zur optimalen Positionierung von Firmen und Geschäften, die ein festgelegtes Gebiet versorgen sollen, untersucht. Dabei wird ermittelt, inwieweit eine bestimmte Position der Versorgungsanlage einzelne Teile im bedürftigen Gebiet bedient. Die Verteilung des Bedarfs im Gebiet, der von der Position und eventuell anderen Faktoren abhängt, wird dabei durch einzelne Punkte mit genau definiertem Verbrauch beschrieben. Dieses Konzept läßt sich leicht auf den Bereich der Planung von mobilen Kommunikationssystemen übertragen. Dort bedient ein Sender als Versorger den Kommunikationsbedarf einer Menge von Teilnehmern.

## 2.2.1 Definition eines Demand-Nodes

*Definition:*

*Ein Demand-Node repräsentiert den Schwerpunkt einer Fläche, die einen vorher festgelegten Bedarf an Kommunikation enthält. Die Höhe des Bedarfs wird dabei durch die Anzahl von Gesprächsanforderungen pro Zeiteinheit definiert.*

Wie bereits dargelegt, wird der Kommunikationsbedarf durch die Demand-Nodes sowohl räumlich als auch hinsichtlich des Verkehrsaufkommens diskretisiert. In der ursprünglichen Definition vertritt jeder Demand-Node im Planungsgebiet eine identische Höhe des Kommunikationsverkehrs. Dadurch repräsentiert die Dichte der Demand-Nodes direkt die Verkehrsdichte, wobei aber ein Quantisierungsfehler durch die Diskretisierung vorhanden ist. Um die Einsatzmöglichkeiten des Konzepts der Demand-Nodes zu erweitern, wird diese strenge Bedingung aufgeweicht. Im folgenden muß der von den Demand-Nodes vertretene Kommunikationsverkehr nicht identisch sein. Trotzdem bleibt das Ideal des gleichen Kommunikationsverkehrs ein wichtiges Ziel. In Kapitel 3 wird diese Zielsetzung bei der Untersuchung von Algorithmen zur Erzeugung der Demand-Nodes genauer beleuchtet.

Ein Demand-Node repräsentiert die Summe des Kommunikationsverkehrs auf seiner Fläche. Dabei geht man davon aus, daß die Zeitpunkte der Verkehrsanforderungen einzelner Teilnehmer sowie die Dauer eines Kommunikationsvorgangs von anderen Verkehrsanforderungen statistisch unabhängig sind. Nach Messungen in der Praxis, siehe dazu [4], wird dieses Kriterium sehr gut erfüllt. Welcher Teilnehmer die Verkehrsanforderungen initiiert ist bedeutungslos, solange die Zeitpunkte voneinander unabhängig sind. Falls die Anzahl der summierten Teilnehmer ausreichend groß ist, können diese statistischen Eigenschaften ausgenutzt werden, und der Vorgang der Verkehrsinitiierung eines Demand-Nodes darf als Poisson-Prozeß modelliert werden. Durch die Zusammenfassung einer größeren Anzahl von Teilnehmern ist es bei den genannten statistischen Voraussetzungen möglich, eine negativ-exponentielle Verteilung der Zeitpunkte der Verkehrsanforderungen anzunehmen. Dies bedeutet eine negativ-exponentielle Verteilung der Zwischenankunftszeit. Die Grundlage zu dieser Annahme liefert das Gesetz der großen Zahlen, durch das sich bei einer unendlichen Anzahl von Teilnehmern genau die negativ-exponentielle Verteilung ergibt. Diese Annahmen sind für das verkehrstheoretische Modell nötig, welches bei der Summation der Verkehrsintensität implizit angewandt wird. Die aus der Verkehrstheorie dazu verwendeten Grundlagen werden teilweise in Kapitel 5.1 und ausführlich in [24] beschrieben.

Der Kommunikationsverkehr und die Position sind für die Beschreibung eines Demand-Nodes bei der Netzplanung oft nicht ausreichend. Für die Modellierung der Wellenausbreitung, die zur Untersuchung von Feldstärke und Interferenz der Funksignale nötig ist, wird eine Charakterisierung der Position des Senders und Mobilteilnehmers benötigt. Die Wellenausbreitungsmodelle, wie

z.B. das Hata-Modell aus [8] oder das Modell nach COST231 aus [19], verwenden dazu sogenannte Landnutzungsklassen. In [19] werden diese Modelle der Wellenausbreitung erläutert. Ein Beispiel für die Landnutzungsklassen ist auf Seite 16 in Tabelle 2.2.2 zu finden. Aus diesem Grund wird jedem Demand-Node bei der Erzeugung zusätzlich eine Landnutzungsklasse zugeordnet, die das Gebiet des Demand-Nodes charakterisiert.

Für die komplette Beschreibung eines Demand-Nodes werden die in Tabelle 2.2.1 genannten Parameter benötigt.

<b>Parameter eines Demand-Nodes</b>
<i>Position</i>
<i>Kommunikationsverkehr</i>
<i>Landnutzungsklasse</i>

*Tabelle 2.2.1: Definition eines Demand-Nodes*

## **2.2.2 Zielkriterien für die Erzeugung von Demand-Nodes**

Bei der Darstellung des Kommunikationsverkehrs im Planungsgebiet durch Demand-Nodes müssen verschiedene Aspekte berücksichtigt werden, um eine möglichst genaue Approximation der Verkehrsverteilung im Planungsgebiet zu erreichen.

Damit die Auffassung als Poisson-Prozeß der Verkehrsinitiierung im Demand-Node korrekt ist, müssen genügend viele Verkehrsanforderungen summiert werden. Im allgemeinen wird diese Annahme bei Telefonverkehr nach [4] sehr gut erreicht.

Für eine genügend genaue Repräsentation der räumlichen Verkehrsverteilung muß die Anzahl der Demand-Nodes ausreichend groß sein. Andernfalls würden die Position und Kapazität der Sender nicht dem wirklichen Kommunikationsverkehr gerecht werden.

Um auch Flächen mit geringem Kommunikationsverkehr im Planungsgebiet genau zu repräsentieren, darf die von einem Demand-Node dargestellte Fläche nicht zu groß sein. Große Teile dieser Gebiete würden sonst bei der Planung nicht berücksichtigt, da in ihnen nur sehr wenige oder gar keine Demand-Nodes positioniert wären.

Die Verteilung der Demand-Nodes in einem bestimmten Teilbereich des Planungsgebietes darf nicht von der Form des Planungsgebietes abhängen. Dazu muß die Verteilung der Demand-Nodes in einem Teilgebiet bei beliebiger Wahl des Planungsgebietes identisch sein, sofern dieses Teilgebiet komplett in ihm

enthalten ist. Durch Randeffekte darf sich höchstens die Verteilung der Demand-Nodes an der Peripherie des Teilgebietes ändern.

Durch den Algorithmus, der die Demand-Nodes erzeugt, darf keine höhere Ordnung dieser Demand-Nodes vorgetäuscht werden, als in der Verkehrsverteilung vorhanden ist. Vor allem darf keine, nur durch den Algorithmus bedingte, Struktur der Demand-Node Positionen erzeugt werden. Die Verteilung der Teilnehmer als Initiatoren der Verkehrsanforderungen ist von Natur aus nicht regelmäßig. Eine nur vorgetäuschte Struktur der Demand-Nodes würde zu falschen Planungsergebnissen führen, wenn z.B. Interferenzen mit den Versorgungsflächen von Sendern auftreten. Vor allem behindert eine feste Form der Demand-Node Flächen eine genaue Repräsentation der Verkehrsverteilung im Planungsgebiet.

Der Kommunikationsverkehr jedes Demand-Nodes sollte gleich sein, damit die Planung vereinfacht wird. Wie bereits bei der Definition der Demand-Nodes erläutert, muß der Verkehr der Demand-Nodes nicht identisch sein. Dieses Ideal sollte aber dennoch möglichst gut erreicht werden. Der von den Demand-Nodes repräsentierte Kommunikationsverkehr wird eventuell durch das Maximum eines Demand-Nodes nach oben abgeschätzt. Der dabei gemachte Fehler wird minimiert, falls möglichst viele Demand-Nodes den gleichen und die restlichen Demand-Nodes einen geringeren Kommunikationsverkehr aufweisen.

Alle oben genannten Aspekte müssen berücksichtigt werden, um eine möglichst genaue Approximation der räumlichen Verkehrsverteilung im Planungsgebiet zu erreichen. Der Algorithmus und die Parameterwahl für die Erstellung der Demand-Nodes haben einen entscheidenden Einfluß auf die Qualität der Planung. Dabei sind die beiden wichtigsten Kriterien für die Beschreibung eines Demand-Nodes der maximal von ihm repräsentierte Kommunikationsverkehr und die maximal von ihm dargestellte Fläche.

Bei der Erzeugung der Demand-Nodes ergibt sich das Problem, sich widersprechende Ziele erfüllen zu müssen. Für eine gute räumliche Darstellung ist eine hohe Anzahl von Demand-Nodes vorteilhaft, was aber einen niedrigen Kommunikationsverkehr pro Demand-Node bewirkt. Dies hat einen negativen Einfluß auf die Voraussetzungen der Annahme eines Poisson-Prozesses für die Verkehrsinitiierung. Außerdem erhöht eine große Anzahl von Demand-Nodes die Planungskomplexität. In der Praxis muß deshalb eine mittlere Quantität als Kompromiß gefunden werden, bei der allen obigen Aspekten genüge geleistet wird.

Um die Demand-Nodes mit einem Algorithmus erzeugen zu können, werden vorher Daten über die Verkehrsdichte im Planungsgebiet benötigt.

### 2.2.3 Diskretisierung der Verkehrsdichte

Die Verkehrsdichte im Planungsgebiet wird im Modell durch die zweidimensionale stetige Verkehrsdichtefunktion beschrieben. Im allgemeinen ist diese Funktion aber nicht bekannt und kann auch nicht einfach aus vorhandenen geographischen Daten erstellt werden. Aus diesem Grund behilft man sich mit der Verwendung eines Rasters, wobei man die diskreten Verkehrswerte in einer zweidimensionalen Matrix speichert. Der Informationsverlust durch die räumliche Diskretisierung kann nicht ermittelt werden, da die wirkliche Verkehrsdichtefunktion nicht bekannt ist. Bei der Planung wird direkt eine Matrix mit Näherungswerten erzeugt, für die in der Praxis meist ein Raster von 10 m auf 10 m verwendet wird. Zur Erstellung der Matrix reicht es nun aus, daß die Verkehrsdichte für jedes dieser atomaren Flächenelemente bekannt ist oder durch andere Verfahren abgeschätzt werden kann. Diese Flächenelemente sind die alleinige Grundlage der Beschreibung der Verkehrsdichte, d.h. im ganzen Planungsprozeß sind keine genaueren Daten über die Verkehrsdichte verfügbar. Auch die Approximationsgüte der Demand-Nodes kann nur im Vergleich mit diesen atomaren Flächenelementen bewertet werden. Trotzdem repräsentiert diese Verkehrsmatrix hinreichend genau die Verkehrsdichtefunktion. Vor allem aber kann sie mit den Ergebnissen von [14] aus vorhandenen geographischen Daten über die Landnutzung erstellt werden, wie in Kapitel 2.2.4 beschrieben wird. In Abbildung 2.2.1 wird der Vorgang der Diskretisierung als Grafik dargestellt.

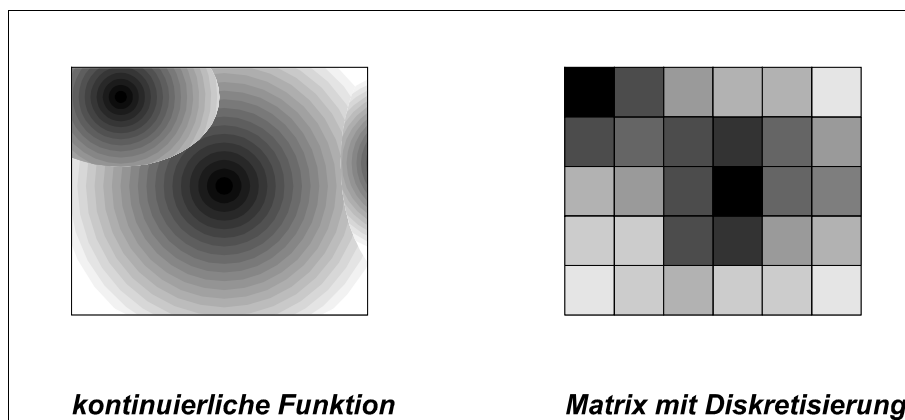


Abbildung 2.2.1: Diskretisierung einer kontinuierlichen Funktion zu einer Matrix

### 2.2.4 Gewinnung der Daten über die Verkehrsdichte

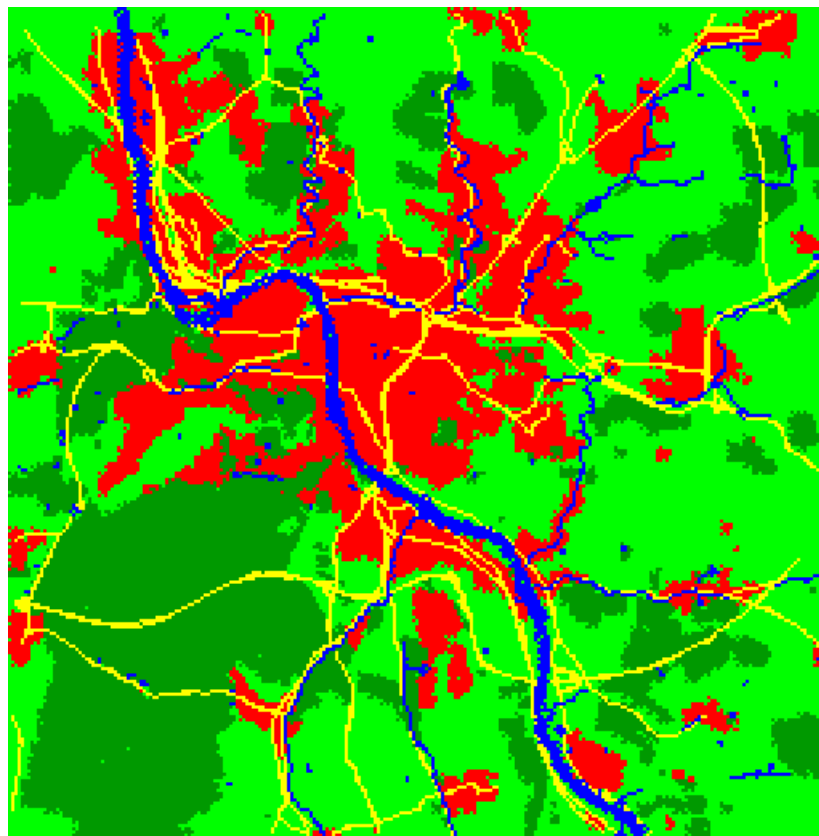
Für eine möglichst genaue Approximation der Verkehrsdichte durch die Demand-Nodes ist die Genauigkeit und Struktur der Rohdaten über die Verkehrsdichte entscheidend. Die Gesamtplanung eines Mobilfunknetzes kann nur so gut sein, wie das schlechteste Glied in der Kette der Planungsphasen, die in



Abbildung 2.1.1 dargestellt werden. Deshalb ist die Wahl geeigneter Rohdaten und eine realitätsnahe Abschätzung des Verkehrs im Planungsgebiet eine essentielle Voraussetzung für korrekte Planungsergebnisse.

Die Gewinnung von Verkehrsdaten aus geographischen Informationen über die Landnutzung wird ausführlich in [14] beschrieben. Aus den Daten über die Landnutzung an einem bestimmten Ort wird eine Aussage über die Verkehrsdichte an dieser Stelle getroffen. In Abbildung 2.2.2 ist die Verteilung der Landnutzungsklassen für das Würzburger Gebiet dargestellt.

Zuerst wird eine Matrix im Raster von  $100\ m$  auf  $100\ m$  über die Landnutzung des Planungsgebietes erzeugt. Zur Gewinnung von Informationen über die Landnutzung wurden Daten aus dem „Amtlichen Topographischen Kartographischen Informations-System“ *ATKIS* [1] verwendet. Dieses Geoinformationssystem ist Grundlage für die Erstellung topographischer Karten. Die Daten beschreiben im Vektorformat die Position und Ausdehnung von Objekten, wie z.B. Siedlung, Straße und Vegetation. Aus diesen Daten kann dann eine Matrix über die Landnutzung erzeugt werden.



*Abbildung 2.2.2: Landnutzungsklassen des Würzburger Gebiets*

Das Würzburger Gebiet<sup>1</sup> hat dabei eine Ausdehnung von 16.1 km auf 15.1 km, wobei jedes atomare Flächenelement die Größe 100 m auf 100 m hat.

Für jede Landnutzungsklasse  $i$  definiert man eine entsprechende Matrix  $E_i(x,y)$ , die nur Einträge der Landklasse  $i$  enthält:

$$E_i(x, y) = \begin{cases} 1 & \text{falls Landklasse } i \text{ an Position } (x, y) \\ 0 & \text{sonst.} \end{cases}$$

Mit dieser Notation läßt sich die Schätzfunktion der Verkehrsdichte definieren als:

$$S(x, y) = \sum_i \eta_i E_i(x, y) .$$

Diese Schätzfunktion gibt für jedes atomare Flächenelement mit dem Mittelpunkt  $(x, y)$  die Verkehrsdichte an. In der Verkehrsmatrix  $T(x, y)$  wird für jedes Matrixelement der Kommunikationsverkehr abgelegt, der sich durch die Verkehrsdichte  $S(x, y)$  mal der Fläche eines atomaren Flächenelements ergibt. Die Faktoren  $\eta_i$  sind Gewichte, die die Verkehrsdichte einer bestimmten Landnutzungsklasse wiedergeben. Sie werden in Tabelle 2.2.2 dargestellt und sind vom Planungsgebiet abhängig. Der Kommunikationsverkehr an der Stelle  $(x, y)$  entspricht also der gewichteten Summe der an dieser Stelle vorhandenen Landklassen.

<b>Land- klasse</b>	<b>Farbe</b>	<b>Beschreibung</b>	<b>Verkehrs- dichte</b>	<b>Verkehrsanteil pro Element</b>
1	Blau	Wasser	$\eta_1$	0.00009000
2	Dunkelgrün	Wald	$\eta_2$	0.00090001
3	Grün	Freies Gelände	$\eta_3$	0.00900009
4	Rot	Siedlung	$\eta_4$	0.09000090
5	Gelb	Straße	$\eta_5$	0.90000900

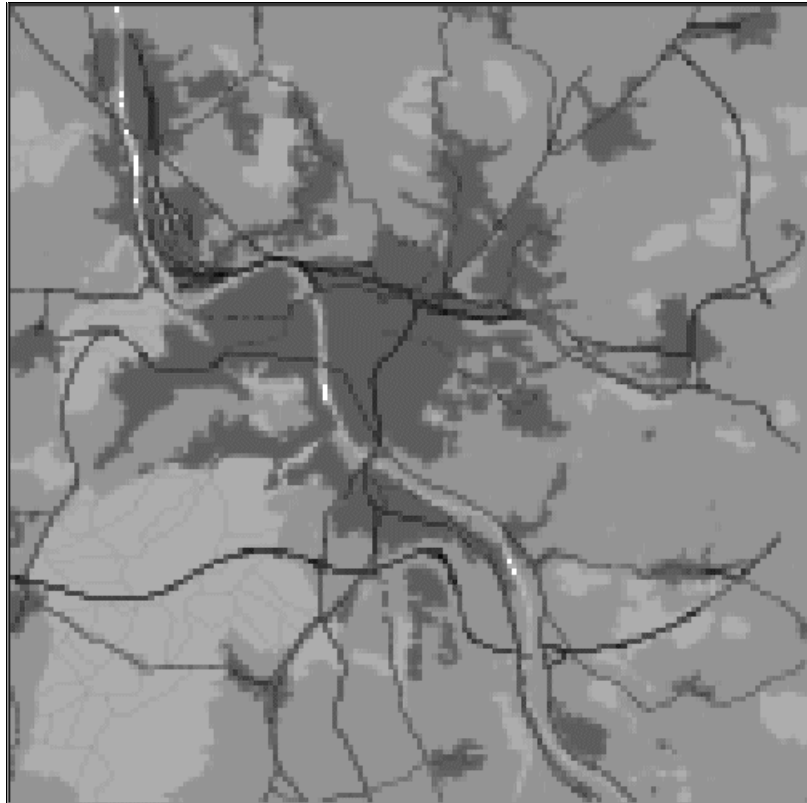
*Tabelle 2.2.2: Landnutzungsklassen*

Um den Kommunikationsverkehr zu bestimmen, den jede Landklasse im Würzburger Gebiet repräsentiert, wird eine exponentielle Verteilung des Kommunikationsverkehrs der Landklassen angenommen. Weiterhin wird der Gesamtverkehr im Planungsgebiet abgeschätzt, der zur Hauptverkehrszeit herrscht. Die Gewichte  $\eta_i$  werden dabei so gewählt, daß die Summe des Kommunikationsverkehrs in der Matrix dem geschätzten Gesamtverkehr entspricht. In Tabelle 2.2.2

<sup>1</sup> Die Lage des Würzburger Gebiets im Gauß-Krüger Format, welches in [14] beschrieben wird, ist: Rechtswert: 4344500 - 4360500, Hochwert: 5510500 - 5525500.

ist der Verkehrsanteil pro Gewicht bei der verwendeten exponentiellen Verteilung mit Basis 10 angegeben.

In Abbildung 2.2.3 ist die fertige Verkehrsmatrix des Würzburger Gebiets dargestellt, die nach obiger Formel berechnet wurde. In der Darstellung gibt die Intensität der Schwarzfärbung die Verkehrsdichte logarithmisch an.



*Abbildung 2.2.3: Verkehrsdichte des Würzburger Gebiets*

Zu beachten ist bei der Verkehrsmatrix, daß Gebiete mit gleichen Landklassen auch den gleichen Kommunikationsverkehr in der Matrix zugeordnet bekommen. Dadurch enthalten große Teilgebiete der Verkehrsmatrix, z.B. in den Waldgebieten oder der Innenstadt, die gleichen Werte. Diese Tatsache muß bei der Demand-Node Erzeugung berücksichtigt werden.

### **2.2.5 Erzeugung der Demand-Nodes**

Zur Erzeugung der Demand-Nodes muß ein geeigneter Algorithmus gefunden werden, der Demand-Nodes mit einer möglichst guten Approximation der Verkehrsverteilung im Planungsgebiet erzeugt. Die Approximationsgüte der Demand-Nodes kann mit den in Kapitel 2.2.2 genannten Kriterien überprüft werden.

Im wesentlichen werden in dieser Diplomarbeit zwei Algorithmen zur Erzeugung von Demand-Nodes dargestellt:

- a) Der Algorithmus des „Rekursiven Partitionierens“ aus [14] wird in Kapitel 3.1 beschrieben.
- b) Der Algorithmus des „Zusammenfassens von Knoten“ ist ein im Rahmen dieser Diplomarbeit entwickeltes Verfahren, welches in den Kapiteln 3.2 und 3.3 dargestellt wird.

Die beiden Algorithmen unterscheiden sich dabei wesentlich in ihrer grundlegenden Arbeitsweise. In Kapitel 4 werden die Verfahren anhand verschiedener Kriterien untersucht und bewertet.

## **2.3 Testgebiete für die Demand-Node Erzeugung**

Als Grundlage für die Untersuchung der Erzeugung von Demand-Nodes werden verschiedene Verkehrsmatrizen verwendet. Diese sind entweder gemessen, aus geographischen Daten berechnet oder synthetisch erzeugt. Bei den folgenden Abbildungen der Verkehrsmatrix stellt die Intensität der Schwarzfärbung die Verkehrsdichte logarithmisch dar. Die Ausdehnung des synthetischen Gebiets und des konstanten Gebiets entsprechen der des Würzburger Gebiets.

### **2.3.1 Würzburger Gebiet**

Das Würzburger Gebiet ist schon oben beschrieben worden. Die Verkehrsdichte ist in Abbildung 2.2.3 dargestellt.

### **2.3.2 Dallas Gebiet**

Die Daten in Abbildung 2.3.1 sind original Planungsdaten von *NORTEL Wireless Networks*, Richardson Texas. Sie beschreiben die Stadt Dallas, USA, mit Umgebung. Das Dallas Gebiet<sup>2</sup> hat eine Ausdehnung von 192.4 km auf 192.4 km.

Die Daten über die Verkehrsdichte sind aus mehreren einzelnen Datensätzen zusammengesetzt, die auf Messungen beruhen. An den weißen Randgebieten liegen keine Daten über den Kommunikationsverkehr vor. Dort wird ein konstanter Kommunikationsverkehr auf sehr niedrigem Niveau angenommen.

---

<sup>2</sup> Die Lage des Dallas Gebiets im Gauß-Krüger Format ist: Rechtswert: 621800 - 814200, Hochwert: 3533900 - 3726300.

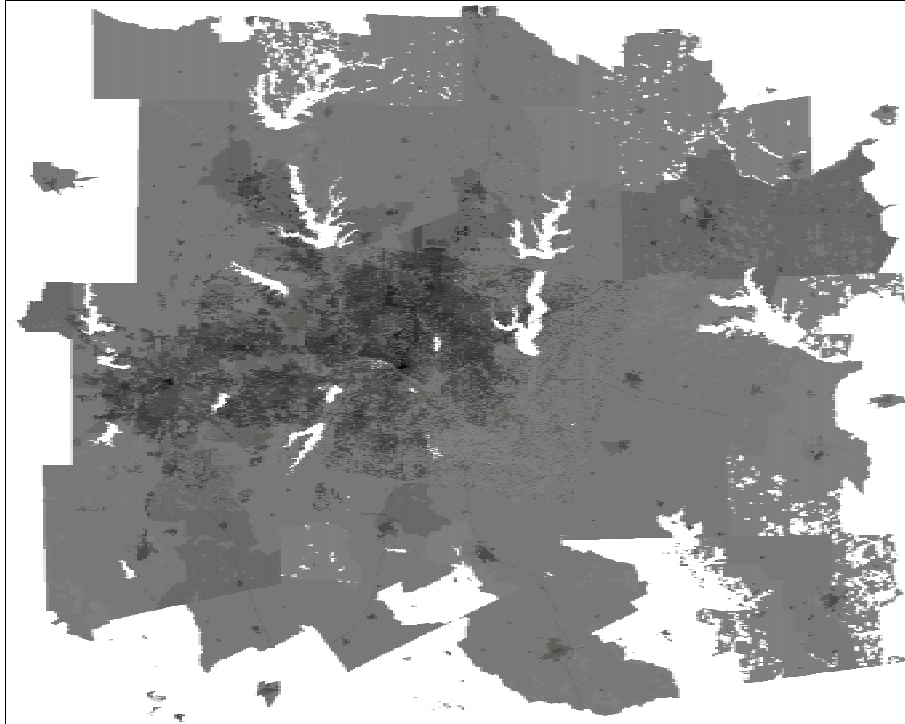


Abbildung 2.3.1: Verkehrsdichte des Dallas Gebiets

### 2.3.3 Synthetisches Gebiet

Der Kommunikationsverkehr des synthetischen Gebiets wurde durch vier überlagerte Funktionen erzeugt, die einen Kommunikationsverkehr gemäß einer Gaußverteilung mit den in Tabelle 2.3.1 genannten Parametern liefern.

Der Kommunikationsverkehr  $T(x,y)$  ergibt sich dabei aus der Summe der Verkehrswerte der einzelnen Gaußverteilungen  $T_i(x,y)$ :

$$T_i(x, y) = \frac{1}{\sqrt{2\pi}} \cdot e^{w_i(x, y)} \cdot s_i, \text{ mit}$$

$$w_i(x, y) = -\frac{1}{2(x_{di} + y_{di})^2} \cdot \left( \frac{(x - x_{ci})^2}{x_{di}} + \frac{(y - y_{ci})^2}{y_{di}} \right),$$

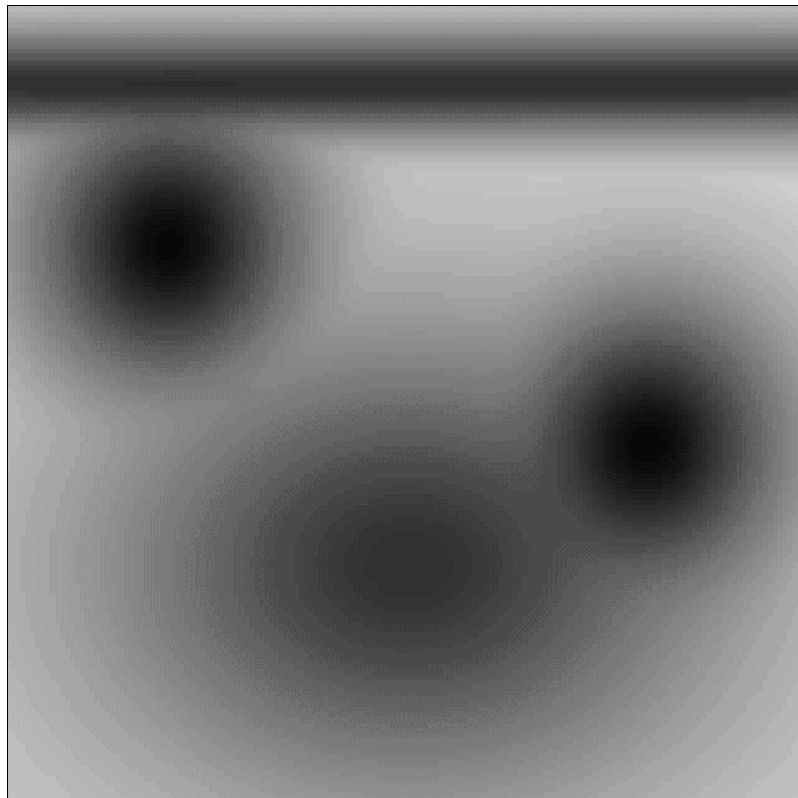
$$T(x, y) = \sum_{i=1}^4 T_i(x, y) .$$

Die  $x$ - und  $y$ -Koordinaten sind in der Tabelle im Bereich  $[0;1]$  angegeben. Der Gesamtverkehr des synthetischen Gebiets wurde dabei so normiert, daß er dem Würzburger Gebiet entspricht.

$i$	$x_{ci}$	$y_{ci}$	$x_{di}$	$y_{di}$	$S_i$
1	0.2	0.7	3.0	3.0	2.0
2	0.8	0.45	3.0	3.0	2.0
3	0.5	0.3	6.0	3.0	1.0
4	0.5	0.9	50.0	0.005	1.0

*Tabelle 2.3.1: Parameter der Gaußverteilungen*

Die beiden Funktionen  $T_1(x,y)$  und  $T_2(x,y)$  können als kreisförmige Ballungszentren angesehen werden, während die Funktion  $T_3(x,y)$  ein elliptisches Zentrum mit niedrigerem Kommunikationsverkehr beschreibt. Mit der Funktion  $T_4(x,y)$  wird ein sehr schmales Band mit hohem Kommunikationsverkehr beschrieben, das eine Straße darstellen soll. In Abbildung 2.3.2 ist die Verkehrsverteilung des synthetischen Gebiets dargestellt.



*Abbildung 2.3.2: Verkehrsdichte des synthetischen Gebiets*

### **2.3.4 Konstantes Gebiet**

In den zum Aufbau der Verkehrsmatrix verwendeten geographischen Daten werden oft größere Objekte, wie z.B. Siedlungs- oder Waldgebiete, beschrieben. Innerhalb dieser Flächen muß eine konstante Verkehrsdichte angenommen werden, sofern nicht zusätzliche Daten darüber vorliegen. Um das Verhalten der Algorithmen in diesem Bereich genauer zu bewerten, wurde eine Verkehrsmatrix mit konstantem Kommunikationsverkehr erzeugt. Anhand dieser Matrix kann untersucht werden, wie sich die Algorithmen in diesem Grenzfall verhalten.

### 3 Clusterungsverfahren

Mit Clusterungsverfahren bezeichnet man die Algorithmen und Methoden für das Gruppieren und Klassifizieren von Objekten, wobei ein Objekt entweder durch eine Menge von Eigenschaften oder durch seine Beziehungen zu anderen Objekten beschrieben wird. Bei der Clusterung werden keine Kategoriebezeichnungen verwendet, in die Objekte mit bestimmten Eigenschaften eingeordnet werden. Dies unterscheidet die Clusterung von den Verfahren der Mustererkennung oder Entscheidungsanalyse, bei denen die Objekte bestimmten namentlich bezeichneten Kategorien zugeordnet werden.

Das Ziel der Clusterung ist es vielmehr, eine geeignete und korrekte Organisation der Daten zu erhalten, anstatt Regeln zum Aufteilen zukünftiger Daten in bestimmte Kategorien. Die Clusterungsverfahren zielen auf das Erkennen von Strukturen in den Daten.

In [3] werden folgende Definitionen von Clustern aufgezählt:

1. „Ein Cluster ist eine Menge von Entitäten, die ähnlich sind, und Entitäten von verschiedenen Clustern sind nicht ähnlich.“
2. „Ein Cluster ist eine Ansammlung von Punkten im Testgebiet, so daß die Entfernung zwischen zwei beliebigen Punkten des Clusters kleiner ist als die Distanz zwischen einem beliebigen Punkt im Cluster und irgendeinem Punkt außerhalb.“
3. „Cluster beschreiben verbundene Gebiete eines mehrdimensionalen Raums, die eine relativ hohe Dichte von Punkten besitzen, wobei sie von anderen Regionen durch eine Region mit einer relativ geringen Dichte von Punkten getrennt sind.“

Die beiden letzten Definitionen beschreiben das Clustern auf räumlichen Strukturen, wie sie in dieser Arbeit benötigt werden.

Bei der bedarfsorientierten Mobilfunkplanung wird der Kommunikationsverkehr im Planungsgebiet zu Demand-Nodes zusammengefaßt, wobei jeder einzelne Demand-Node einem Cluster entspricht. Die Demand-Nodes stellen laut Definition die Ansammlung einer Gruppe von Teilnehmern dar, deren Kommunikationsverkehr sie repräsentieren. Bei dieser Voraussetzung entspricht die Clusterung der Definition 3, wobei die Mobilfunkteilnehmer als Punkte angenommen werden, die dann zu den Demand-Nodes zusammengefaßt werden.

In der Praxis sind die Positionen der einzelnen Teilnehmer nicht bekannt, und für die Planung ist es ausreichend, den von ihnen initiierten Kommunikationsverkehr in einem bestimmten Gebiet zu kennen. Deshalb diskretisiert man den Kommunikationsverkehr im Planungsgebiet, wie in Kapitel 2.2.3 beschrieben wird, und erzeugt aus den Informationen über die Verkehrsdichte an einer Position die Demand-Nodes. Hier entspricht die Demand-Node Erstellung mehr einer Diskretisierung des Kommunikationsverkehrs als einem Clusterungsver-



fahren im eigentlichen Sinn. Die erzeugten Demand-Nodes repräsentieren aber entsprechend der Definition 3 einzelne Cluster der Mobilfunkteilnehmer. Die Grenzen von Clusterung und Diskretisierung sind dabei, je nach Definition, fließend.

Um eine Clusterung eines Gebiets zu erstellen, sind prinzipiell zwei Ansätze möglich. Zum einem kann man, ausgehend vom Gesamtgebiet, dieses weiter aufteilen und so von oben nach unten, im Englischen „top down“ genannt, die Cluster erstellen. Eine andere Möglichkeit ist, ausgehend von einzelnen Punkten, diese zusammenzufassen und dadurch von unten nach oben, auch „bottom up“ genannt, die Cluster zu bilden.

Im folgenden werden Algorithmen für beide Möglichkeiten untersucht, wobei zuerst ein „top down“ Verfahren betrachtet wird.

## 3.1 Verfahren des Rekursiven Partitionierens

Das Verfahren des rekursiven Partitionierens ist ein klassisches „top-down“ Verfahren. Ausgehend von der Gesamtfläche werden bereits vorliegende Gebiete rekursiv weiter aufgeteilt, bis die Anzahl der Teilungen den vorher festgelegten Grenzwert erreicht. Für jedes der so gebildeten Teilgebiete wird ein Demand-Node erzeugt.

Das rekursive Partitionieren entspricht einer hierarchischen Clusterung, auch Partitions-Clusterung genannt. Die Grundidee dazu wird in [3] dargelegt. Der hier benutzte Algorithmus entspricht dem in [14] und [28] verwendeten.

Als Parameter benötigt der Algorithmus die Anzahl der rekursiven Teilungsschritte  $R$ . Diese Zahl bestimmt eindeutig die Anzahl der erzeugten Demand-Nodes und legt damit den von ihnen repräsentierten Kommunikationsverkehr fest. Jeder mit diesem Verfahren erzeugte Demand-Node stellt dabei einen identischen Kommunikationsverkehr dar.

### 3.1.1 Algorithmus

Der Algorithmus arbeitet von oben nach unten, in dem er rekursiv bereits vorliegende Gebiete weiter aufteilt.

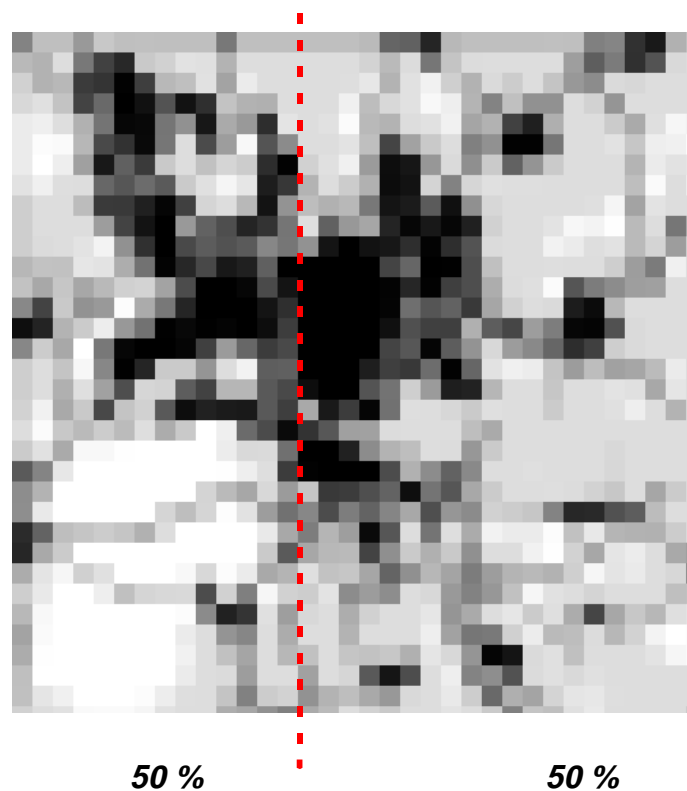


Abbildung 3.1.1: 1. Schritt des Algorithmus

Vor dem Start des Algorithmus wird der Gesamtverkehr der Planungsregion als Summe der Verkehrswerte in der Matrix bestimmt. Danach wird von links beginnend der Kommunikationsverkehr in jeder Spalte summiert, bis die Hälfte des Gesamtverkehrs der Matrix überschritten ist. Da nur ganze Spalten zum Kommunikationsverkehr addiert werden, können im Regelfall nicht genau 50% jeweils dem rechten bzw. linken Gebiet zugeordnet werden. Um eine genaue Aufteilung auf beide Regionen zu erreichen, wird der Kommunikationsverkehr der Spalte, die beide Gebiete trennt, anteilig aufgeteilt. Damit ist der Gesamtverkehr der Matrix genau gleich auf die beiden Flächen verteilt.

Nun wird rekursiv jedes der beiden Gebiete nach dem gleichen Schema weiter aufgeteilt, bis die maximale Anzahl an Teilungen erreicht ist. Dabei wird abwechselnd entweder spaltenweise oder zeilenweise genau die Hälfte des Kommunikationsverkehrs in dem Teilgebiet gesucht. In Abbildung 3.1.2 ist die Verkehrsmatrix nach dem zweiten Teilungsschritt dargestellt. Hierbei wurden die beiden Teilgebiete des ersten Schrittes nun zeilenweise wieder genau zur Hälfte aufgeteilt. Insgesamt sind vier Gebiete entstanden, von denen jedes exakt 25% des Gesamtverkehrs enthält.

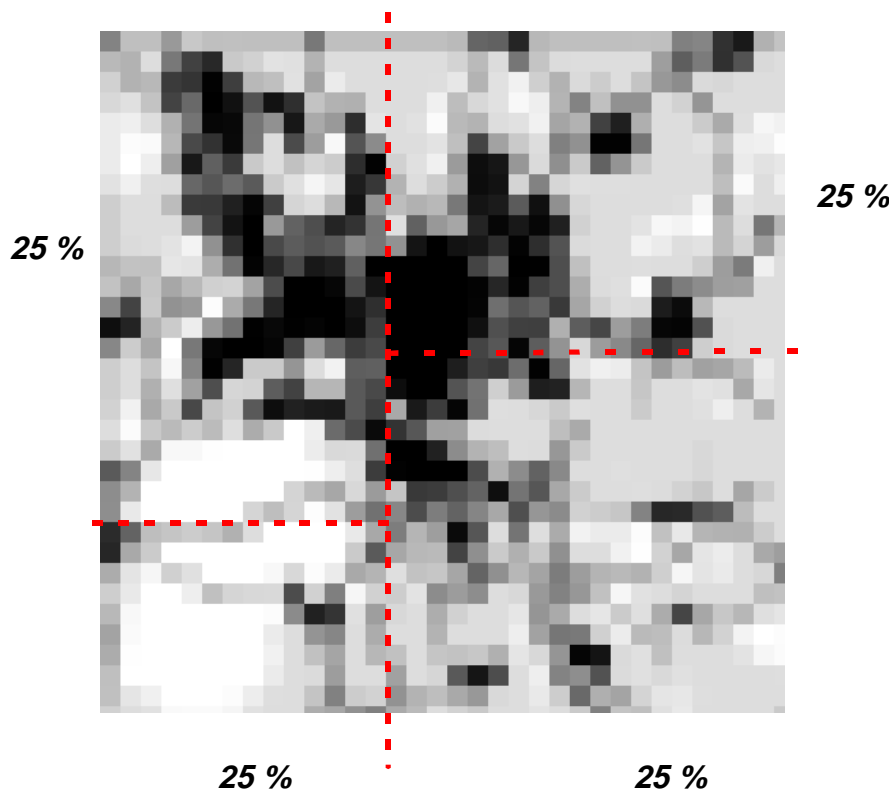
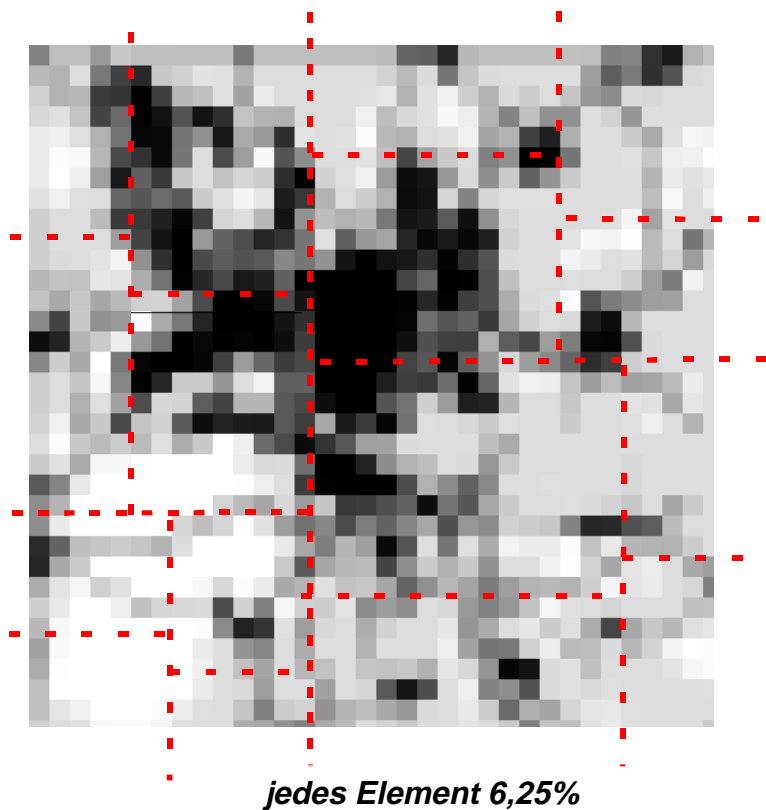


Abbildung 3.1.2: 2. Schritt des Algorithmus

Diese Aufteilung wird solange wiederholt, bis die Anzahl  $R$  der Teilungsschritte erreicht ist. Dann enthält jedes Teilgebiet genau  $\frac{1}{2^R}$  des Gesamtverkehrs. Da für jede Region ein Demand-Node erzeugt wird, ergibt sich deren Anzahl zu

$N = 2^R$ . In Abbildung 3.1.3 wird das Ergebnis des Algorithmus für  $R = 4$  dargestellt. Es sind  $2^4 = 16$  Teilgebiete entstanden, von denen jedes den Kommunikationsverkehr  $\frac{1}{2^4} = 6,25\%$  des Gesamtverkehrs enthält.

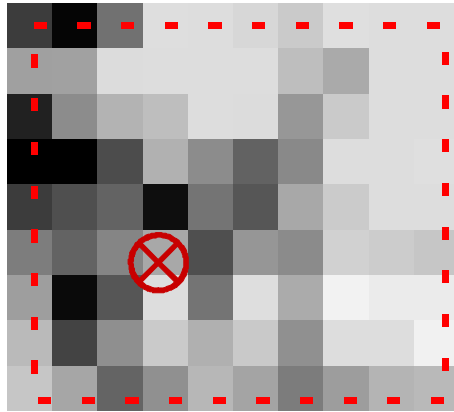
Dabei muß die Obergrenze für die Anzahl der möglichen Regionen beachtet werden. Sie ist begrenzt durch die Bedingung, daß jede Teilregion mindestens vier Matrixelemente enthalten muß. Ein Demand-Node muß offensichtlich mindestens ein komplettes atomares Flächenelement enthalten, damit der Algorithmus sinnvolle Ergebnisse liefern kann. Durch die Aufteilung der Matrixelemente an den Grenzen der Teilgebiete auf mehrere Gebiete ergibt sich diese Mindestzahl an Matrixelementen. Dadurch ist auch die maximale Approximationsgüte der Verkehrsmatrix durch die Demand-Nodes begrenzt, wie in Kapitel 4.2.5 gezeigt wird.



*Abbildung 3.1.3: Ergebnis des Algorithmus für  $T = 4$  Teilungen*

Nachdem der Kommunikationsverkehr der Matrix aufgeteilt wurde, müssen für die einzelnen Teilgebiete Demand-Nodes erzeugt werden. Die Position der Demand-Nodes wird durch die Gewichtung des Kommunikationsverkehrs der Teilfläche ermittelt, wie in Abbildung 3.1.4 dargestellt wird. Der Schwerpunkt dieser Gewichtung ergibt die Position des neuen Demand-Nodes, wobei der

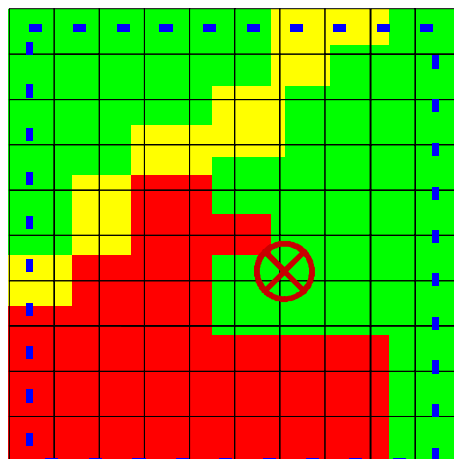
Kommunikationsverkehr für alle Demand-Nodes gleich ist und  $\frac{1}{2^R}$  des Gesamtverkehrs entspricht.



 **nach Verkehr gewichteter  
Schwerpunkt des Knotens**

Abbildung 3.1.4: Position des Demand-Nodes

Die Landklasse des Demand-Nodes wird bestimmt, indem die Landklasse an der Position des Demand-Nodes genommen wird, wie in Abbildung 3.1.5 gezeigt.



**Landnutzungsclassen an  
der Position des Knotens:** 

Abbildung 3.1.5: Bestimmung der Landklasse

### 3.1.2 Ergebnisse des Algorithmus

Im folgenden werden die Ergebnisse des Algorithmus für die vier verschiedenen Vergleichsflächen untersucht. Dazu wird die Verteilung der Demand-Nodes, die aus der Verkehrsmatrix erzeugt wird, betrachtet. Eine genaue Analyse der verschiedenen Algorithmen und ihrer Eignung für die Netzplanung wird in Kapitel 4 durchgeführt.

Die vier Vergleichsgebiete werden in Kapitel 2.3 genau beschrieben, wo auch die Verkehrsdichten der Gebiete grafisch dargestellt werden. In Tabelle 3.1.1 werden die zur Berechnung verwendeten Parameter des Algorithmus für die vier Vergleichsflächen aufgelistet.

<b>Vergleichsfläche</b>	<b>Rekursionstiefe</b>	<b>Anzahl der Demand-Nodes</b>
<i>Würzburger Gebiet</i>	10	1024
<i>Dallas Gebiet</i>	15	32768
<i>Synthetisches Gebiet</i>	10	1024
<i>Konstantes Gebiet</i>	10	1024

*Tabelle 3.1.1: Parameter des Algorithmus bei den Vergleichsflächen*

Die Verteilung der Demand-Nodes muß die Verkehrsverteilung im Planungsgebiet möglichst genau repräsentieren. Dies kann man visuell anhand der Abbildungen der Demand-Node und Verkehrsverteilungen überprüfen.

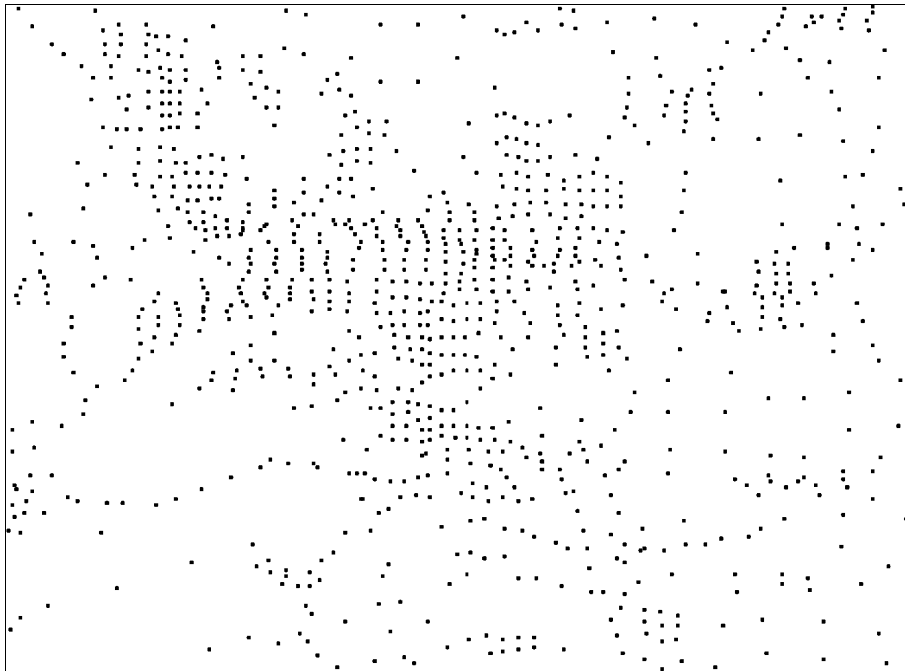
In Abbildung 3.1.6 sind die Demand-Nodes des Würzburger Gebiets abgebildet, welche aufgrund der Verkehrsmatrix aus Abbildung 2.2.3 erzeugt wurden. Bei der Betrachtung zeigt sich, daß die Demand-Nodes den Kommunikationsverkehr plastisch darstellen. Man kann z.B. die Autobahnen durch eine Linie von Demand-Nodes erkennen. In Flächen mit konstanter Verkehrsdichte, wie etwa der Innenstadt von Würzburg, bilden die Demand-Nodes quadratische Muster. Man sieht auch, daß in Flächen mit relativ niedriger Verkehrsdichte nahezu keine Demand-Nodes liegen.

Für das Dallas Gebiet in Abbildung 3.1.7 kann im Prinzip das gleiche ausgesagt werden, wobei die Flächen ohne Demand-Nodes hier wesentlich größer sind. Außerdem ist die Spannweite der Verkehrsdichte dort beträchtlich höher.

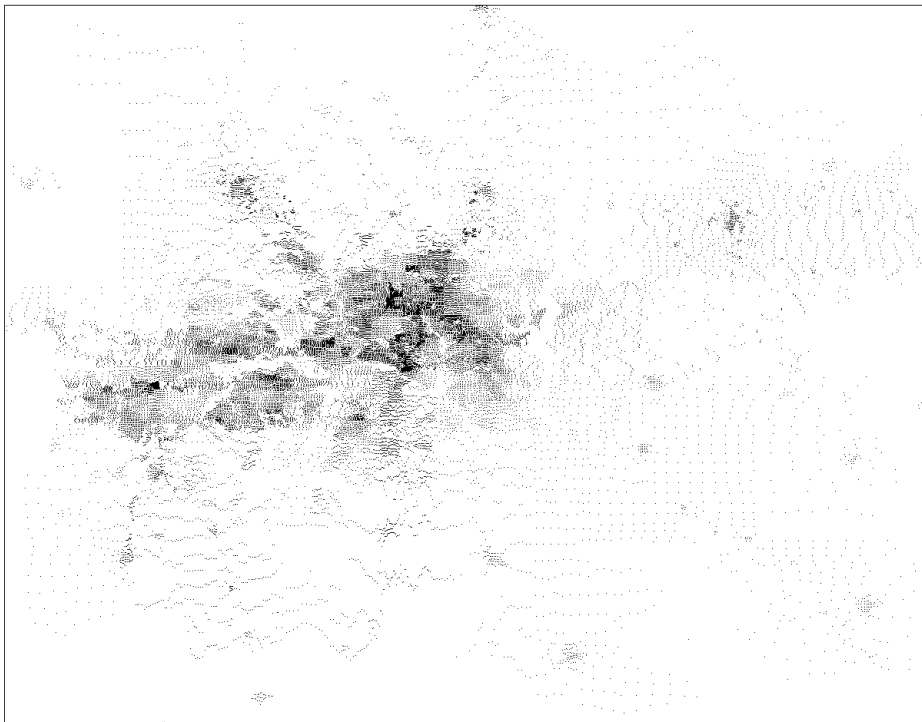
Interessant sind die Demand-Nodes des synthetischen Gebiets in Abbildung 3.1.8, die in quadratischen Mustern auftreten. Der Kommunikationsverkehr liegt aber gemäß radialer Gauß-Verteilungen vor, die in Abbildung 2.3.2 gezeigt werden.

In Abbildung 3.1.9 sind die Demand-Nodes des konstanten Gebiets abgebildet. Man sieht deutlich, daß ein regelmäßiges quadratisches Muster gebildet wird.

Diese Musterbildung des Algorithmus wird noch für sehr problematische Effekte sorgen.



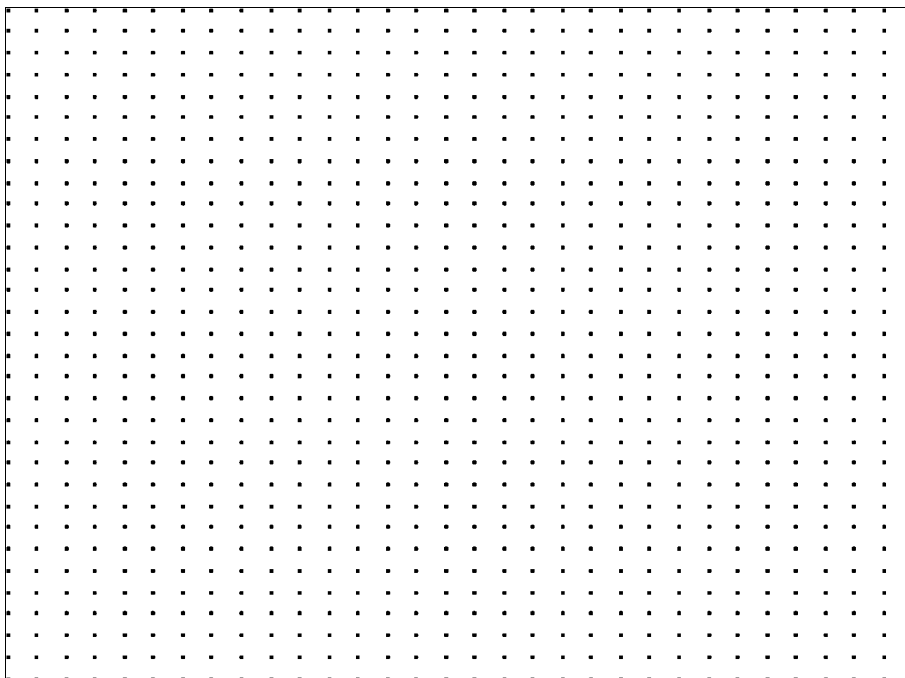
*Abbildung 3.1.6: Demand-Nodes des Würzburger Gebiets*



*Abbildung 3.1.7: Demand-Nodes des Dallas Gebiets*



*Abbildung 3.1.8: Demand-Nodes des synthetisches Gebiets*



*Abbildung 3.1.9: Demand-Nodes des konstanten Gebiets*



### 3.1.3 Nachteile des Algorithmus

Im folgenden werden die Ergebnisse des Algorithmus qualitativ bewertet. In Kapitel 4 wird eine quantitative Untersuchung der Algorithmen durchgeführt, die mit Meßwerten die folgenden Erkenntnisse untermauert.

Auffällig bei diesem Algorithmus ist die quadratische Struktur der Demand-Node Verteilung, die durch die Restriktion der Demand-Node Fläche auf eine rechteckige Form hervorgebracht wird. Diese Beschränkung kann zu einer sehr schlechten Repräsentation des Kommunikationsverkehrs durch die Demand-Nodes führen, wie in den obigen Beispielen zu sehen ist.

Eine viel wesentlichere Einschränkung wird durch diese Beschränkung hervorgerufen. Die Approximationsgüte der Verkehrsdichtefunktion durch die Demand-Nodes hängt hier stark von der Form der Testfläche ab. Da diese die Versorgungsgebiete von Sendern darstellen sollen, die im allgemeinen eine unregelmäßige Form aufweisen, führt diese Einschränkung zu unbrauchbaren Ergebnissen bei der Netzplanung. In einem Gebiet mit konstanter Verkehrsdichte sollte der Kommunikationsverkehr innerhalb einer beliebigen Testfläche nach der Formel  $T(r) = A \cdot c$  nur von der Größe der Testfläche  $A$  und der konstanten Verkehrsdichte  $c$  abhängen. Die Form der Testfläche darf dabei keinen Einfluß haben. Bei diesem Algorithmus hängt der Kommunikationsverkehr aber auch von der Form der Testfläche und ihrer Lage ab. Eine minimale Änderung, z.B. der Größe oder der Form, führt zu nicht vorhersehbaren Abweichungen des durch die Demand-Nodes innerhalb der Fläche repräsentierten Kommunikationsverkehrs. Es treten dann Interferenzen zwischen den rechteckigen Demand-Node Flächen und der Form der Versorgungsgebiete auf. In den Kapiteln 4.2.2 und 4.2.3 wird dieses Ergebnis durch quantitative Resultate verifiziert.

Ein weiteres Problem ist die fehlende Begrenzung der Größe der Demand-Node Flächen. Dadurch kommt es in Bereichen mit niedrigem Kommunikationsverkehr zu einer ungenauen Darstellung des Kommunikationsverkehr, da nur sehr wenige Demand-Nodes in diesem Bereich liegen. Ein einzelner Demand-Node repräsentiert dort eine große Fläche. Im Dallas Gebiet z.B. werden Demand-Nodes mit einer Fläche von über  $400 \text{ km}^2$  erzeugt. Abgesehen davon, daß kein Sender eine Fläche dieser Größe versorgen könnte, wird auch die gesamte Modellierung der Wellenausbreitung sehr ungenau. Als gravierender erweist sich das Problem, daß Bereiche in dieser Größenordnung vollkommen frei von Demand-Nodes sind. Diese Gebiete können später bei der Planung überhaupt nicht berücksichtigt werden, obwohl auch hier ein niedriger Kommunikationsverkehr vorhanden ist. Vor allem kann keine auch nur annähernd flächendeckende Versorgung erreicht werden.

Durch die Abhängigkeit der Demand-Node Verteilung von der Form und Größe des Planungsgebietes ergibt sich eine Restriktion, die eine Planung größerer Gebiete unmöglich macht. Es ist zwar prinzipiell möglich, ein großes Gebiet in mehrere Bereiche aufzuteilen und in jedem Teilgebiet die Demand-Nodes ein-

zeln erzeugen zu lassen. Die dabei entstehende Verteilung der Demand-Nodes ist aber nicht identisch mit einer Verteilung, die bei einer anderen Aufteilung des Gesamtgebietes entstehen würde. Die erzeugten Demand-Nodes sind auch nicht invariant bezüglich einer Drehung des Planungsgebietes. Durch das mit dem Gesamtgebiet startende rekursive Aufteilen des Algorithmus ergeben selbst minimale Änderungen der Gebietsgrenzen eine andere Verteilung der Demand-Nodes.

### 3.1.4 Merkmale des Algorithmus

Im folgenden werden die bereits beschriebenen charakteristischen Merkmale des rekursiven Partitionierens prägnant in Form einer Aufzählung zusammengefaßt:

- Die Fläche eines Demand-Nodes ist immer rechteckig. Die Position des Demand-Nodes ist dabei der verkehrsmäßige Schwerpunkt.
- Die Fläche, die ein Demand-Node überdeckt, kann sich maximal über das gesamte Planungsgebiet erstrecken. Es kann kein Grenzwert für die Größe vorgegeben werden.
- Die Anzahl der Demand-Nodes  $N$  ist immer eine Zweierpotenz der Teilungsschritte  $R$ :  $N = 2^R$ . Die maximale Anzahl der Demand-Nodes ist begrenzt durch die Bedingung, daß jeder Demand-Node mindestens vier Flächenelemente der Verkehrsmatrix repräsentieren muß.
- Der Kommunikationsverkehr aller Demand-Nodes ist gleich und entspricht  $N^{-1}$  des Gesamtverkehrs. Die Grenzen der Demand-Nodes liegen nicht unbedingt auf den Rändern der Matrixelemente. Vielmehr werden einem Demand-Node an seinen Grenzen anteilige Matrixelemente zugeordnet, wie in Abbildung 3.1.4 zu sehen ist.
- Die Landklasse eines Demand-Nodes ergibt sich aus der Landklasse, die an der Position des verkehrsmäßigen Schwerpunktes des Demand-Nodes vorliegt.
- Die Struktur der erzeugten Demand-Nodes ist streng deterministisch. In Gebieten mit ähnlichem Kommunikationsverkehr werden regelmäßige quadratische Muster erzeugt.
- Die Verteilung der Demand-Nodes hängt von der Form und Größe des Planungsgebietes ab.
- Die Approximationsgüte der Verkehrsdichtefunktion durch die Demand-Nodes hängt innerhalb eines Versorgungsgebietes eines Senders von dessen Form ab.

## 3.2 Verfahren des Zusammenfassens von Knoten

Das Verfahren des Zusammenfassens von Knoten ist ein typisches „bottom-up“ Verfahren. Zu Beginn wird jedem Flächenelement der Verkehrsmatrix ein Knoten zugeordnet. In einer Schleife werden dann zufällig Knoten ausgewählt und mit einem geeigneten Nachbarknoten zusammengefaßt, falls bestimmte Voraussetzungen erfüllt sind. Die Schleife endet, falls kein geeigneter Knoten mehr gefunden werden kann. Für jeden der verbleibenden Knoten wird ein Demand-Node erzeugt.

Als Parameter benötigt der Algorithmus:

- Maximaler Kommunikationsverkehr  $T_{max}$ . Dieser Wert bestimmt den maximalen Kommunikationsverkehr, den ein Demand-Node repräsentieren kann.
- Maximale Fläche  $A_{max}$ . Dieser Wert gibt die maximale Größe der Fläche an, die ein Demand-Node überdeckt.

Da der Kommunikationsverkehr aller Demand-Nodes in diesem Algorithmus nicht identisch sein muß, haben auch einige einen kleineren Kommunikationsverkehr als  $T_{max}$ . Ein Ziel des Algorithmus ist es daher, für alle Demand-Nodes ungefähr gleichen Kommunikationsverkehr zu erreichen. Um auch in Bereichen des Planungsgebietes mit geringem Verkehrsaufkommen eine genügend hohe Approximationsgüte zu erzielen, darf die Fläche eines Demand-Nodes nicht zu groß sein. Deswegen wird die Flächenbegrenzung  $A_{max}$  eingeführt.

Diese beiden Parameter entscheiden direkt über die Anzahl der erzeugten Demand-Nodes, wobei durch die zufallsgesteuerte Arbeitsweise des Algorithmus die Anzahl der erzeugten Demand-Nodes nicht absolut vorhergesagt werden kann. Da sich der Zufall aber auf die willkürliche Wahl des nächsten zu bearbeitenden Objekts beschränkt, unterscheiden sich die Ergebnisse bei verschiedenen Läufen des Algorithmus beinahe nicht. Die Anzahl der Demand-Nodes schwankt nur im Promillebereich.

### 3.2.1 Algorithmus

In diesem Algorithmus werden aus bereits vorhandenen, kleineren Knoten durch geeignetes Zusammenfassen neue Knoten gebildet. Dazu liegt als Grundlage die Verkehrsmatrix vor, die in Abbildung 3.2.1 dargestellt wird.

Für jedes Flächenelement in der Verkehrsmatrix wird nun ein Knoten erzeugt, der genau dieses Flächenelement enthält. Ein Knoten wird dabei durch folgende Teile festgelegt:

- zum Knoten gehörende Flächenelemente der Matrix
- Gesamtverkehr der zugeordneten Flächenelemente

- Gesamtfläche der zugeordneten Flächenelemente
- Suchrichtung bei der Ermittlung eines geeigneten Nachbarn

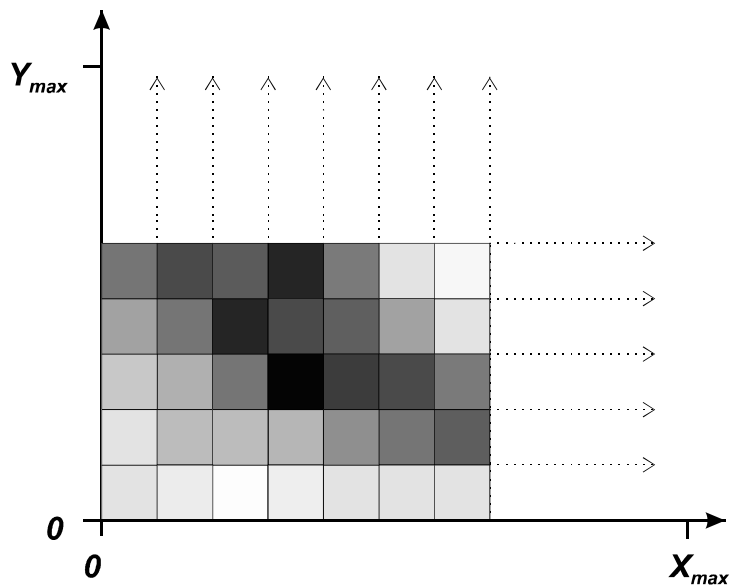


Abbildung 3.2.1: Verkehrsmatrix

Nach dieser Initialisierung der Knoten läuft das in Abbildung 3.2.2 gezeigte Schema ab, bis kein Knoten mehr mit einem Nachbarn zusammengefaßt werden kann.

<i>Wiederhole:</i>				
	1. Wähle zufällig einen Knoten aus.			
	2. Prüfe die benachbarten Knoten, ob ein geeigneter Kandidat zum Vereinigen vorhanden ist:			
	<table border="1"> <tr> <td><i>Nein</i></td> <td><i>Ja</i></td> </tr> <tr> <td></td> <td>3. Fasse den ausgewählten Knoten mit diesem Nachbarn zusammen.</td> </tr> </table>	<i>Nein</i>	<i>Ja</i>	
<i>Nein</i>	<i>Ja</i>			
	3. Fasse den ausgewählten Knoten mit diesem Nachbarn zusammen.			
<i>Bis keine Knoten mehr zusammengefaßt werden können</i>				
4. Erzeuge die Demand-Nodes aus den verbliebenen Knoten.				

Abbildung 3.2.2: Algorithmus des Zusammenfassens von Knoten

### 3.2.1.1 Zufällige Knotenauswahl

Aus der Menge der Knoten wird zufällig ein Knoten  $K$  zum Prüfen ausgewählt, wie in Abbildung 3.2.3 gezeigt wird.

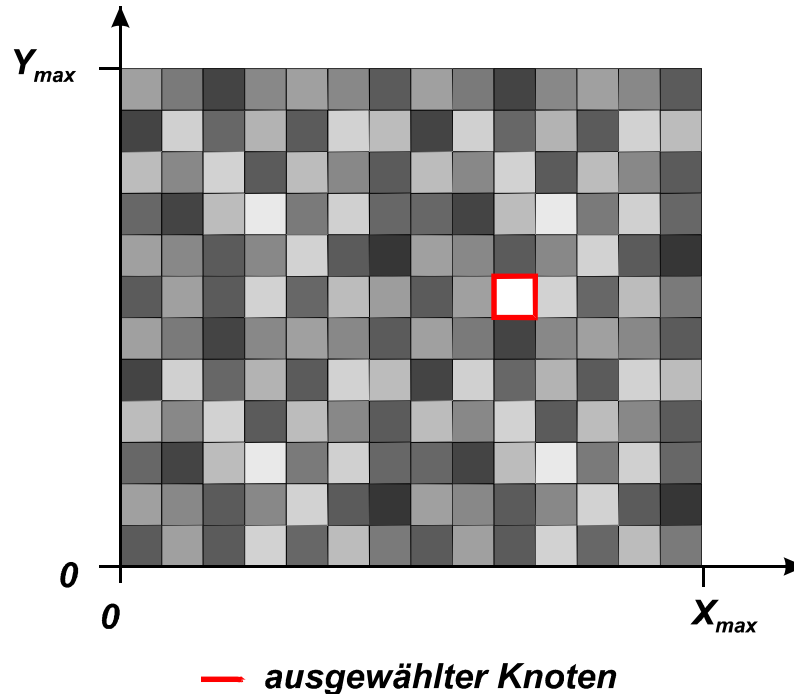
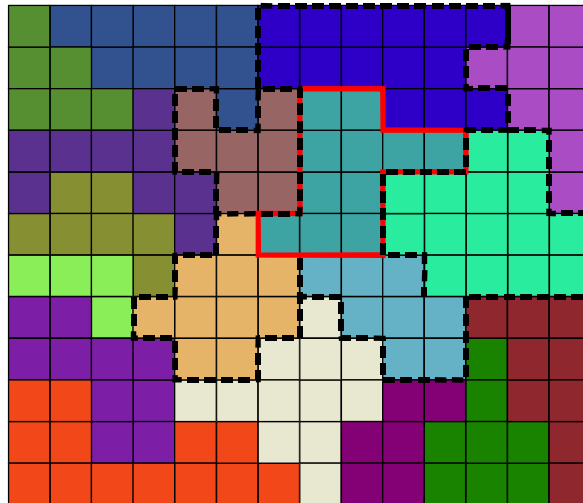


Abbildung 3.2.3: Zufällige Auswahl eines Knotens

### 3.2.1.2 Prüfung der benachbarten Knoten

Die Nachbarknoten des Knotens  $K$  werden nach einem geeigneten Kandidaten für das Zusammenfassen mit ihm durchsucht. Als Nachbarn eines Knotens werden alle anderen Knoten angesehen, die ihn horizontal oder vertikal mit der Kante mindestens eines Flächenelements berühren. In Abbildung 3.2.4 ist ein Beispiel für einen Knoten mit seinen Nachbarn dargestellt.

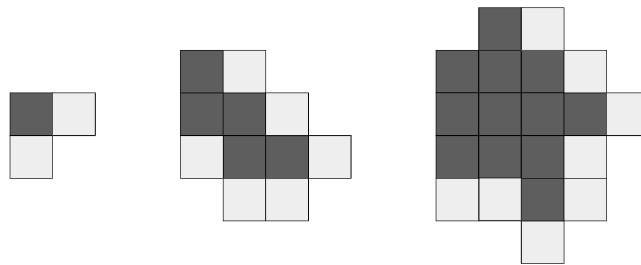
Damit die Knoten beim Zusammenfassen gleichmäßig in alle Richtungen anwachsen, werden nicht alle Nachbarn eines Knotens bei der Suche nach einem Kandidaten durchsucht. Vielmehr wird entweder nur rechts und unterhalb, oder nur links und oberhalb nach geeigneten Nachbarn gesucht. Dazu enthält jeder Knoten eine Variable, die die aktuelle Suchrichtung anzeigt. Diese wird nach jeder Nachbarsuche für diesen Knoten gewechselt. In Abbildung 3.2.5 sind die Flächenelemente, auf denen nach Nachbarn gesucht wird, für beide Suchrichtungen anhand von drei Beispielknoten aufgetragen.



— **Knoten K**    ..... **zu testende Nachbarn**

Abbildung 3.2.4: Nachbarn eines Knoten

**Suche rechts und unterhalb:**



**Suche links und oberhalb:**

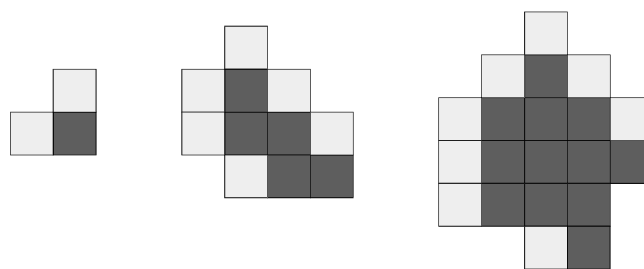


Abbildung 3.2.5: Suchfolge

Diese abwechselnde Suche ist nötig, damit die Knoten nicht bevorzugt in eine Richtung wachsen, wenn der Kommunikationsverkehr der Nachbarelemente identisch ist. In der Verkehrsmatrix besteht, bedingt durch ihre Erzeugung aus den Landnutzungsklassen, in größeren Gebieten der gleiche Kommunikationsverkehr. Dies führt zu Flächenelementen mit identischem Kommunikationsverkehr. Um regelmäßige Muster beim Zusammenfassen von Knoten zu vermeiden,

ist eine zufällige Auswahl des Knotens  $K$  für eine ausgewogene Form der Knotenflächen notwendig.

Unter den Nachbarn der aktuellen Suchfolge wird der Nachbar gesucht, der folgende Bedingungen erfüllt:

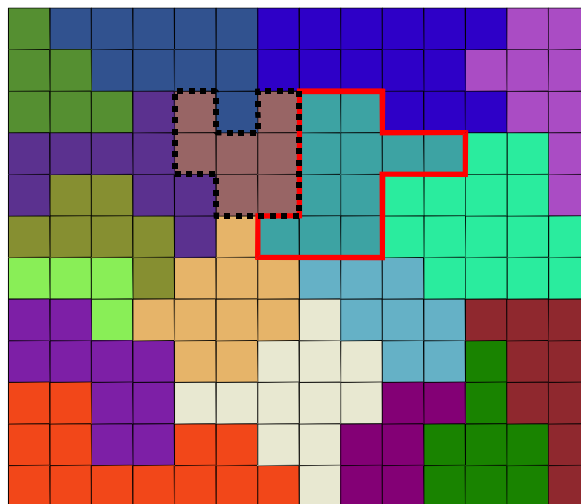
$$T_K + T_N < T_{max} ,$$

$$A_K + A_N \leq A_{max} ,$$

$$T_N \leq T_i \quad \forall i \in Nh(K) .$$

Hierbei bezeichnet  $N$  den zum Zusammenfassen ausgewählten Nachbarknoten,  $T_K$  den Kommunikationsverkehr des Knotens  $K$ ,  $A_K$  die Fläche des Knotens  $K$  und  $Nh(K)$  die Menge der Nachbarn des Knoten  $K$  gemäß der aktuellen Suchfolge.  $T_{max}$  gibt den maximalen Kommunikationsverkehr und  $A_{max}$  die maximale Fläche für einen Demand-Node an. Beides sind Parameter des Algorithmus.

Der Nachbar mit dem geringsten Kommunikationsverkehr wird ausgewählt, mit dessen Verschmelzung der neue Knoten die maximalen Werte für Kommunikationsverkehr und Fläche noch nicht überschreitet. In Abbildung 3.2.6 ist der zum Zusammenfassen ausgewählte Knoten dargestellt. Falls kein Nachbar alle obigen Bedingungen erfüllt, wird ein anderer Knoten  $K$  zufällig gewählt und die Schleife wiederholt. Die Schleife wird beendet und mit der Erzeugung der Demand-Nodes begonnen, falls für keinen Knoten ein geeigneter Nachbar gefunden werden kann.



— **Knoten K**    ..... **ausgewählter Knoten N**

Abbildung 3.2.6: Zum Vereinigen ausgewählter Knoten  $N$

### 3.2.1.3 Zusammenfassen des Knotens mit einem Nachbarn

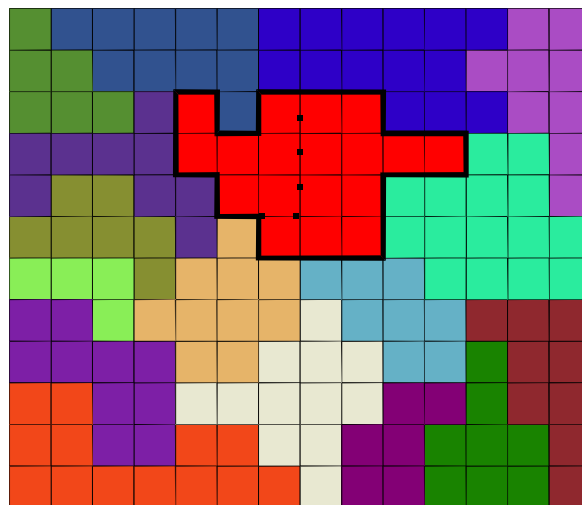
Der Knoten  $K$  wird mit seinem ausgewählten Nachbarn  $N$  zu dem neuen Knoten  $Z$  kombiniert:

$$T_Z = T_K + T_N ,$$

$$A_Z = A_K + A_N ,$$

$$E_Z = E_K \cup E_N .$$

Wobei  $E_X$  die Menge der Flächenelemente angibt, die zum Knoten  $X$  gehören.



— *neuer Knoten*

Abbildung 3.2.7: Aus der Vereinigung hervorgegangener neuer Knoten  $Z$

Die beiden alten Knoten  $K$  und  $N$  werden entfernt. In Abbildung 3.2.7 ist das Ergebnis des Zusammenfassens dargestellt.

### 3.2.1.4 Erzeugung der Demand-Nodes aus den Knoten

Falls keine Knoten mehr mit ihren Nachbarn kombiniert werden können, endet die Schleife. Für jeden noch verbliebenen Knoten wird ein Demand-Node erstellt. Die Position des Demand-Nodes ergibt sich aus dem nach Kommunikationsverkehr gewichteten Schwerpunkt der einzelnen Flächenelemente des Knotens, wie in Abbildung 3.2.8 gezeigt:

$$x_G = \frac{1}{\sum_{\forall e \in E(Z)} T_e} \cdot \sum_{\forall e \in E(Z)} x_e \cdot T_e ,$$

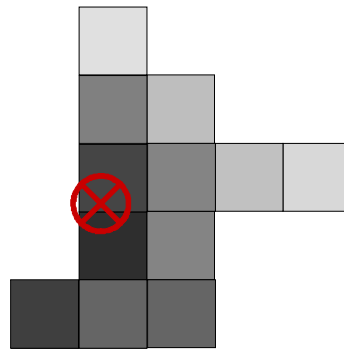


$$y_G = \frac{1}{\sum_{\forall e \in E(Z)} T_e} \cdot \sum_{\forall e \in E(Z)} y_e \cdot T_e .$$

Hierbei wird mit  $Z$  der Knoten bezeichnet, für den der Demand-Node erzeugt wird.  $T_e$  gibt den Kommunikationsverkehr des Flächenelements  $e$  an, wobei  $e$  für ein Flächenelement des Knotens  $Z$  steht, und  $E(Z)$  die Menge der Flächenelemente des Knotens  $Z$  angibt.

Der Schwerpunkt  $G(x_g, y_g)$  ist dann die Position des neuen Demand-Nodes. Der Kommunikationsverkehr des erzeugten Demand-Nodes entspricht dem Kommunikationsverkehr  $T_Z$  des Knotens.

### Verkehrswerte



 **nach Verkehr gewichteter Schwerpunkt des Knotens**

Abbildung 3.2.8: Position des Demand-Nodes

Für die Landklasse des Demand-Nodes wird die am häufigsten auftretende Landklasse in der Fläche des Knotens genommen, wie in Abbildung 3.2.9 illustriert wird. In der Planung wird die Landklasse nur für die Berechnung der Wellenausbreitung verwendet. Für diese ist die ganze Umgebung des Demand-Nodes und nicht nur seine Position ausschlaggebend. Die Modelle der Wellenausbreitung werden deshalb nach der vorherrschenden Landnutzung in der Umgebung eines Demand-Nodes gewählt. Genauere Untersuchungen in Kapitel 4.3 zeigen, daß die Wahl der Landklasse in dieser Weise eine bessere Darstellung der Landnutzung des Planungsgebietes ergibt.

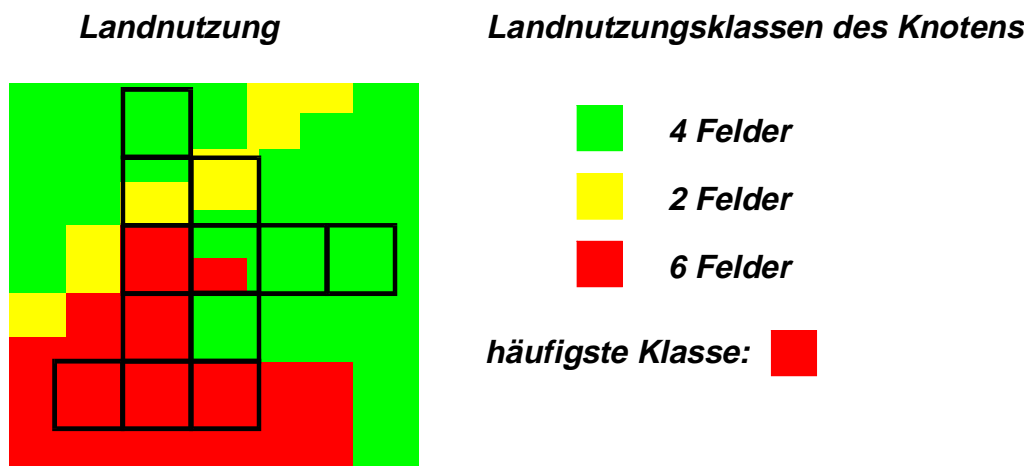


Abbildung 3.2.9: Wahl der Landklasse

### 3.2.2 Ergebnisse des Algorithmus

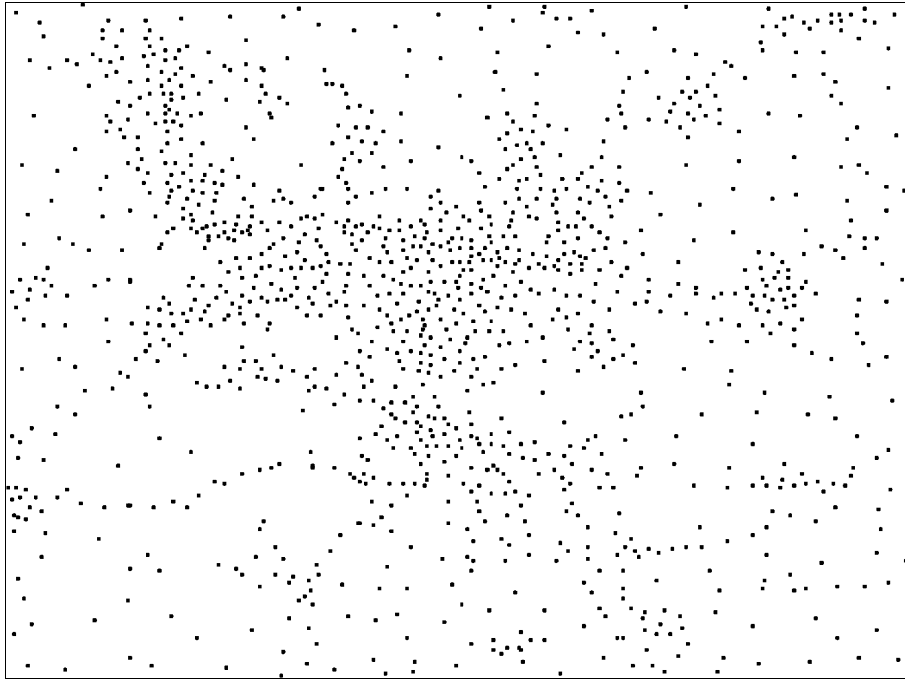
Im folgenden werden die Ergebnisse des Algorithmus für die vier verschiedenen Vergleichsflächen, die in Kapitel 2.3 beschrieben werden, untersucht. In Tabelle 3.2.1 werden die zur Berechnung verwendeten Parameter aufgelistet.

<b>Vergleichsfläche</b>	<b>Maximaler Verkehr<sup>3</sup></b>	<b>Maximale Fläche</b>	<b>Anzahl der Demand-Nodes</b>
<i>Würzburger Gebiet</i>	0.08	0.75 km <sup>2</sup>	1289
<i>Dallas Gebiet</i>	0.0002	25 km <sup>2</sup>	40589
<i>Synthetisches Gebiet</i>	0.08	0.75 km <sup>2</sup>	1381
<i>Konstantes Gebiet</i>	0.08	0.75 km <sup>2</sup>	1281

Tabelle 3.2.1: Parameter des Algorithmus bei den Vergleichsflächen

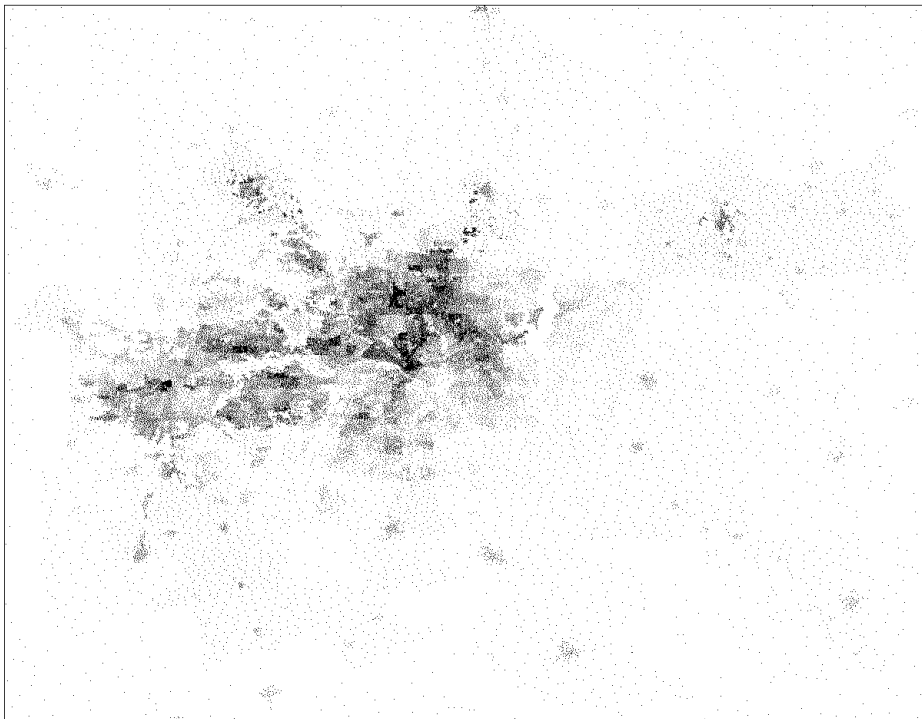
In Abbildung 3.2.10 sind die für das Würzburger Gebiet erzeugten Demand-Nodes abgebildet. Die Demand-Nodes stellen die Verkehrsverteilung aus Abbildung 2.2.3 gut dar, die einzelnen Objekte, wie z.B. Autobahnen oder Dörfer, sind erkennbar. Im Vergleich zum Verfahren des rekursiven Partitionierens fällt auf, daß die Demand-Nodes keine Muster bilden und viel gleichmäßiger verteilt sind. Auch werden die Bereiche mit niedrigem Kommunikationsverkehr durch die Demand-Nodes genügend repräsentiert, was durch die Begrenzung ihrer Fläche bei diesem Verfahren erreicht wird.

<sup>3</sup> Der Kommunikationsverkehr ist in der Pseudoeinheit der jeweiligen Vergleichsfläche angegeben. Um auf *Erlang* zu kommen, muß dieser Wert gemäß des wirklichen Gesamtverkehrs skaliert werden.

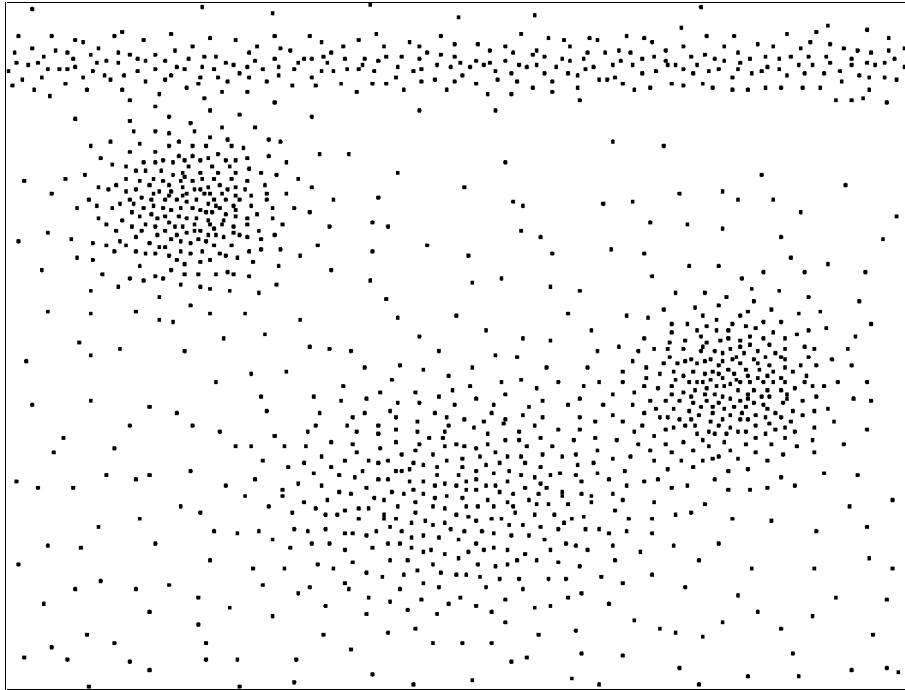


*Abbildung 3.2.10: Demand-Nodes des Würzburger Gebiets*

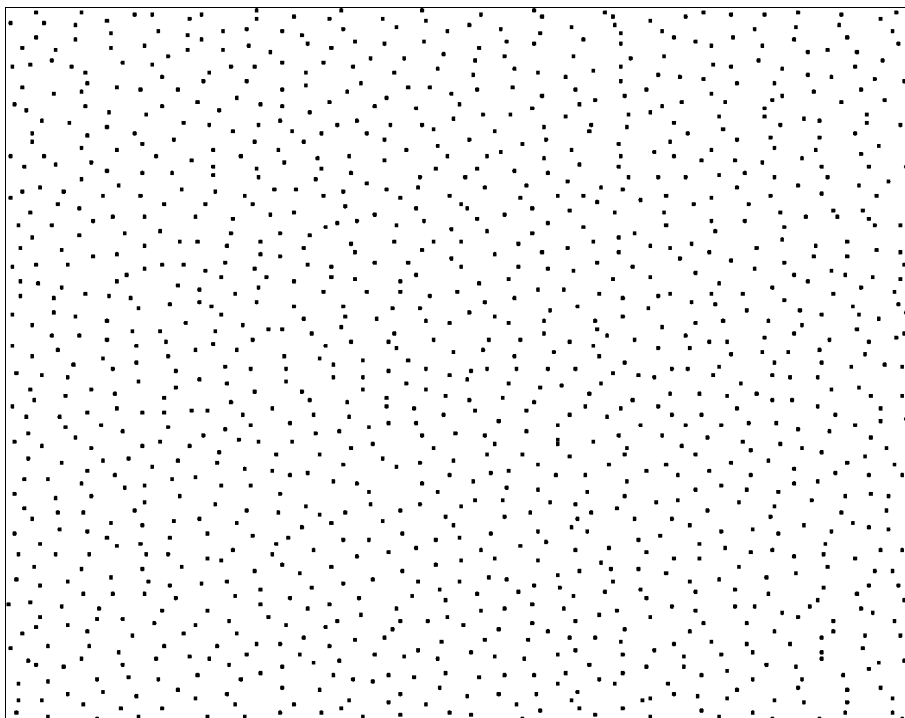
Vor allem im Dallas Gebiet in Abbildung 3.2.11 wird der Vorteil der Flächenbegrenzung deutlich. Selbst der Außenbereich mit sehr niedrigem Kommunikationsverkehr wird durch Demand-Nodes dargestellt. In diesem Bereich erzeugt das Verfahren des rekursiven Partitionierens nahezu keine Demand-Nodes.



*Abbildung 3.2.11: Demand-Nodes des Dallas Gebiets*



*Abbildung 3.2.12: Demand-Nodes des synthetischen Gebiets*



*Abbildung 3.2.13: Demand-Nodes des konstanten Gebiets*

Die radialen Gaußverteilungen im synthetischen Gebiet aus Abbildung 2.3.2 werden durch die Demand-Nodes in Abbildung 3.2.12 korrekt wiedergegeben.

In diesem Gebiet erzeugt das rekursive Partitionieren eine mangelhafte Repräsentation des Kommunikationsverkehr, wie in Abbildung 3.1.8 sichtbar ist.

Bei der Darstellung des konstanten Gebiets in Abbildung 3.2.13 ist ein wesentlicher Vorteil des Zusammenfassens von Knoten sichtbar. Statt ein Muster, wie in Abbildung 3.1.9 durch das Verfahren des Partitionierens erzeugt, wird eine absolut gleichmäßige Zufallsverteilung der Demand-Nodes erstellt.

### 3.2.3 Bewertung des Algorithmus

Bei der rein qualitativen Untersuchung dieses Algorithmus zeigt sich, daß alle Nachteile des Verfahrens des rekursiven Partitionierens aufgehoben sind.

Durch die Begrenzung der Fläche eines Demand-Nodes wird auch in Gebieten mit sehr niedrigem Kommunikationsverkehr eine genügende Repräsentation der Verkehrsverteilung erreicht. Die unbeschränkte Demand-Node Fläche war ein großer Kritikpunkt aus der Praxis der Mobilfunkplanung. Die maximale Fläche kann nun beliebig durch den Parameter des Algorithmus festgelegt werden.

Die Fläche eines Demand-Nodes kann bei diesem Verfahren eine beliebige Form annehmen und ist nicht mehr auf eine rechteckige Gestalt beschränkt. Dies hat mehrere Vorteile. Die Demand-Nodes können den Kommunikationsverkehr besser approximieren, da durch eine frei wählbare Form eine genauere Darstellung möglich ist.

Weiterhin erzeugt der Algorithmus keine störenden Muster, die eine Struktur höherer Ordnung des Kommunikationsverkehrs vortäuschen, als in der Verkehrsmatrix vorhanden ist. Die erzeugten Demand-Nodes bilden vielmehr eine Verteilung mit maximaler Entropie. Bei Flächen mit konstanter Verkehrsdichte wird eine absolut zufällige, homogene Verteilung der Demand-Nodes erstellt. Der Kommunikationsverkehr innerhalb einer beliebigen Testfläche, die z.B. das Versorgungsgebiet eines Senders darstellt, hängt damit nur von ihrer Größe und Position ab und nicht von ihrer Form. In einem Gebiet mit konstanter Verkehrsdichte ergibt sich der Kommunikationsverkehr innerhalb der Testfläche nach der Formel  $T(r) = A \cdot c$  in Abhängigkeit von der Größe der Testfläche  $A$  und der konstanten Verkehrsdichte  $c$ .

Die gerade genannten wesentlichen Vorteile dieses Verfahrens erkaufte man sich durch einen Nachteil: Der Kommunikationsverkehr, den ein Demand-Node repräsentiert, ist nicht mehr für alle Demand-Nodes gleich, was vor allem durch die Beschränkung der maximalen Fläche verursacht wird. Vielmehr besteht eine Verteilung des Kommunikationsverkehrs der Demand-Nodes, die vom Planungsgebiet und den gewählten Parametern abhängt. Diese Verteilung des Kommunikationsverkehrs der Demand-Nodes gilt es nun zu untersuchen und zu optimieren.

### 3.2.4 Verkehrsverteilung der Demand-Nodes

In Abbildung 3.2.14 ist ein Histogramm des Kommunikationsverkehrs der Demand-Nodes des Würzburger Gebiets dargestellt. Für den Parameter  $T_{max}$ , der den maximalen Kommunikationsverkehr eines Demand-Nodes angibt, wurde der Wert 0.08 genommen. Man sieht, daß im Bereich von 0.0024 bis  $T_{max}$  nahezu alle Verkehrswerte vorkommen, wobei die höheren Werte mit größerer Wahrscheinlichkeit angenommen werden. Der Wert 0.0024 stellt dabei das Minimum des von den Demand-Nodes repräsentierten Verkehrs dar.

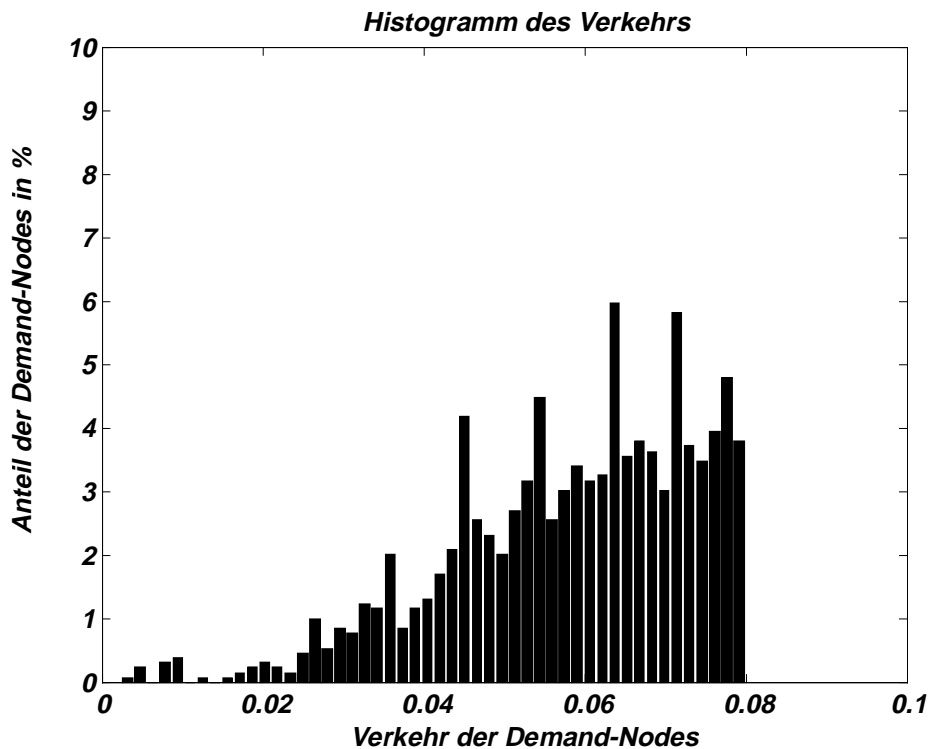


Abbildung 3.2.14: Verkehrshistogramm des Würzburger Gebiets

In Abbildung 3.2.15 sind dieselben Ergebnisse für das Dallas Gebiet, bei einem Wert von 0.0002 für  $T_{max}$ , aufgetragen. Die Verteilung entspricht im Wesentlichen der des Würzburger Gebiets. Durch die größere Fläche dieses Gebiets ist die Verteilung gleichmäßiger.

Für das synthetische und konstante Gebiet sehen die Verteilungen den hier abgebildeten ähnlich und werden deshalb nicht dargestellt.

Die Verteilungen entsprechen nicht dem Ziel, daß alle Demand-Nodes den gleichen Kommunikationsverkehr aufweisen. Vor allem sollte der maximale Kommunikationsverkehr von den meisten Demand-Nodes erreicht werden. Das Ergebnis des Algorithmus muß deshalb noch verbessert werden, um auch diesen Zielen zu genügen.

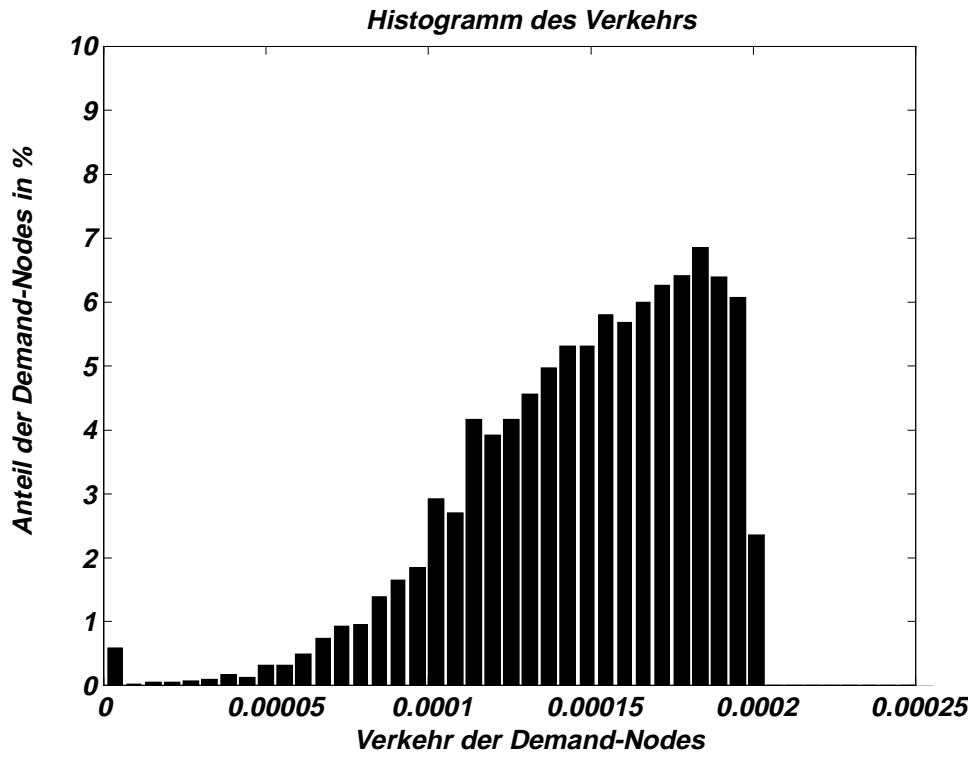


Abbildung 3.2.15: Verkehrshistogramm des Dallas Gebiets

### 3.3 Optimierung des Zusammenfassens von Knoten

Um eine bessere Verteilung des Kommunikationsverkehrs der Demand-Nodes zu erreichen, wurde ein Optimierer entwickelt, der nach dem Ende der Hauptschleife des Algorithmus arbeitet. Die Knoten werden dann nach bestimmten Gesichtspunkten optimiert, und erst anhand der endgültigen Menge der Knoten werden die Demand-Nodes erzeugt.

Der gesamte Ablauf des Algorithmus mit Optimierungsphasen ist in Abbildung 3.3.1 dargestellt.

<i>Wiederhole:</i>				
	1. Wähle zufällig einen Knoten aus.			
	2. Prüfe die benachbarten Knoten, ob ein geeigneter Kandidat zum Vereinigen vorhanden ist:			
	<table border="1"> <tr> <td><i>Nein</i></td> <td><i>Ja</i></td> </tr> <tr> <td></td> <td>3. Kombiniere den ausgewählten Knoten mit diesem Nachbarn.</td> </tr> </table>	<i>Nein</i>	<i>Ja</i>	
<i>Nein</i>	<i>Ja</i>			
	3. Kombiniere den ausgewählten Knoten mit diesem Nachbarn.			
<i>Bis keine Knoten mehr kombiniert werden können</i>				
4. a) <i>Optimierungsphase I</i> b) <i>Optimierungsphase II (optional)</i>				
5. <i>Erzeuge die Demand-Nodes aus den restlichen Knoten.</i>				

Abbildung 3.3.1: Algorithmus mit Optimierung

#### 3.3.1 Optimierungsphase I

In der ersten Optimierungsphase wird versucht, Knoten mit geringem Kommunikationsverkehr zu entfernen, indem man ihre Flächenelemente auf benachbarte Knoten aufteilt. Zu diesem Zweck müssen die Nachbarknoten noch Flächenelemente aufnehmen können, d.h. weder der maximale Kommunikationsverkehr  $T_{max}$ , noch die maximale Fläche  $A_{max}$  dieser Knoten darf überschritten werden.

Schon vor der Optimierungsphase haben viele Knoten den maximalen Kommunikationsverkehr. Durch das Hinzufügen von Flächenelementen aus Knoten, die gelöscht werden, erhöht sich der Kommunikationsverkehr anderer Knoten. Damit erreichen zusätzliche Knoten den maximalen Kommunikationsverkehr.



Aus diesem Grund können nur wenige Knoten überhaupt Flächenelemente aufnehmen. Um die Möglichkeit zur Verteilung von Flächenelementen zu verbessern, wird erlaubt, daß der maximale Kommunikationsverkehr  $T_{max}$  um einen bestimmten Faktor  $R_{T_{max}}$  überschritten werden kann.

Die Parameter der Optimierungsstufe I sind also:

- Minimaler Kommunikationsverkehr  $T_{min}$ . Alle Knoten, die einen Kommunikationsverkehr kleiner als  $T_{min}$  haben, versucht man zu entfernen, indem ihre Flächenelemente auf Nachbarn verteilt werden.
- Faktor  $R_{T_{max}}$ , um den der maximale Kommunikationsverkehr  $T_{max}$  überschritten werden kann.

Die Bedingungen für Knoten lassen es teilweise nicht zu, daß alle Flächenelemente eines zu löschenden Knotens auf Nachbarknoten verteilt werden. Deshalb bleiben am Ende des Algorithmus auch einige Knoten übrig, die einen Kommunikationsverkehr kleiner als  $T_{min}$  haben. Der Wert  $R_{T_{max}} \cdot T_{max}$  gibt den absolut maximalen Kommunikationsverkehr an, der auch bei der Optimierung nicht überschritten werden darf, d.h. nach der Optimierung haben alle Demand-Nodes einen Kommunikationsverkehr kleiner als  $R_{T_{max}} \cdot T_{max}$ . Im Regelfall wird  $R_{T_{max}} = 1.1$  gesetzt, d.h. der maximale Kommunikationsverkehr  $T_{max}$  darf um höchstens 10% überschritten werden.

### 3.3.1.1 Algorithmus

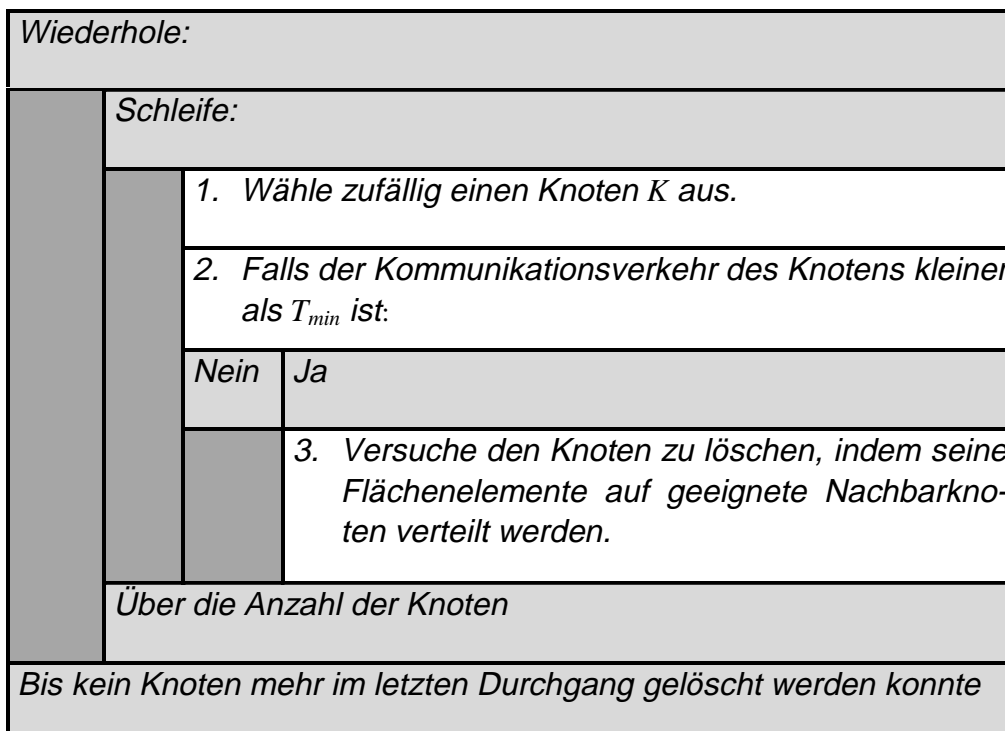


Abbildung 3.3.2: Algorithmus des Löschens von verkehrsarmen Knoten

Der grundsätzliche Aufbau des Algorithmus ist in Abbildung 3.3.2 skizziert. Die Hauptschleife mit der Auswahl eines Knotens zur Optimierung wird so oft wiederholt, wie insgesamt Knoten vorhanden sind. Der Algorithmus wird solange durchgeführt, wie ein Knoten durch den Algorithmus dieser Schleife gelöscht werden konnte.

Für jeden innerhalb der Schleife zufällig ausgewählten Knoten  $K$  läuft der in Abbildung 3.3.3 gezeigte Algorithmus ab. Falls dieser Knoten einen niedrigeren Kommunikationsverkehr als  $T_{min}$  hat, werden für jedes seiner Flächenelemente  $E$  gemäß der in Abbildung 3.2.5 gezeigten Suchfolge die Nachbarknoten untersucht. Dabei wird der Nachbarknoten  $N$  mit dem kleinsten Kommunikationsverkehr genommen, der noch dieses Flächenelement aufnehmen kann, ohne die maximale Anzahl von Flächenelementen oder den bei der Optimierung maximalen Kommunikationsverkehr  $R_{Tmax} \cdot T_{max}$  zu überschreiten:

$$T_K < T_{min} ,$$

$$T_N + T_E < R_{Tmax} \cdot T_{max} ,$$

$$A_N + A_E \leq A_{max} ,$$

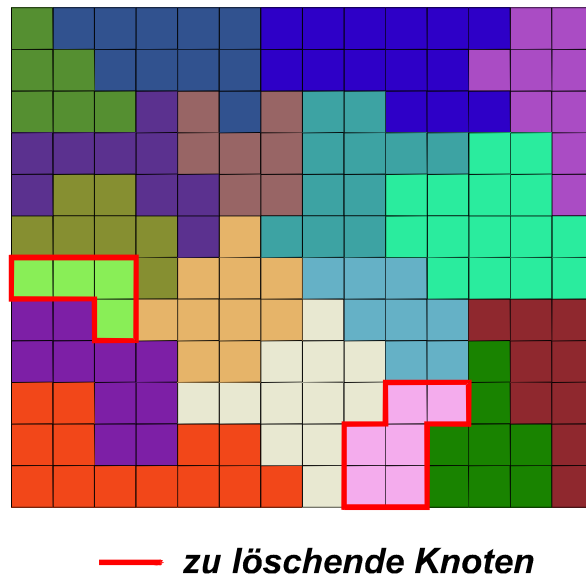
$$T_N \leq T_i \forall i \in Nh(K) .$$

Schleife:				
	1. Suche gemäß der aktuellen Suchfolge nach einem geeigneten Nachbarknoten, der ein Flächenelement aufnehmen kann.			
	2. Falls geeigneter Nachbarknoten gefunden:			
	<table border="1"> <thead> <tr> <th>Nein</th> <th>Ja</th> </tr> </thead> <tbody> <tr> <td></td> <td></td> </tr> </tbody> </table>	Nein	Ja	
Nein	Ja			
	3. Entferne das Flächenelement vom Knoten und füge es beim gefundenen Nachbarknoten ein.			
Über alle Flächenelemente des ausgewählten Knotens				

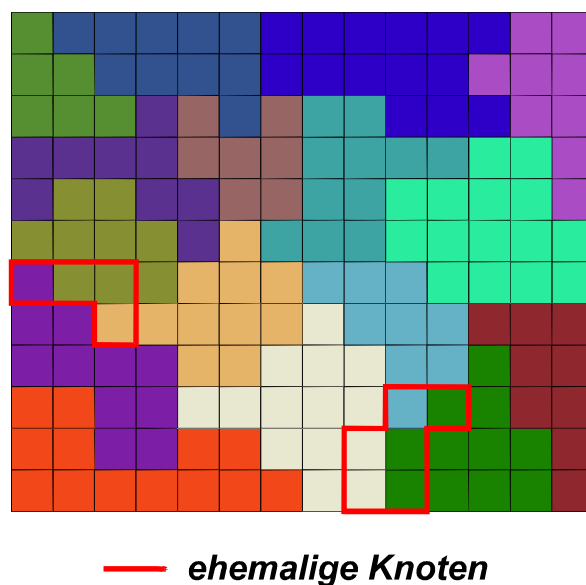
Abbildung 3.3.3: Algorithmus für zu optimierenden Knoten

Falls kein geeigneter Nachbar für das aktuelle Flächenelement gefunden werden kann, wird das Flächenelement zurückgestellt und mit dem nächstem Flächenelement des Knotens fortgefahren. Bei der Auswahl der Flächenelemente müssen zuerst die überprüft werden, die einen anderen Knoten als Nachbarn haben. Falls diese Randelemente zu anderen Knoten wechseln konnten, kom-

men weiter innen gelegene Flächenelemente an den Rand und können dann auch zu einem Nachbarknoten wechseln. Falls alle Flächenelemente eines Knotens zu anderen bewegt werden konnten, wird der Knoten komplett entfernt.



*Abbildung 3.3.4: Beispiel für zwei zu löschende Knoten*



*Abbildung 3.3.5: Gebiet nach der Entfernung der zwei Knoten*

In Abbildung 3.3.4 sind als Beispiel zwei Knoten dargestellt, die einen zu geringen Kommunikationsverkehr besitzen, und damit zum Löschen ausgewählt wurden. In Abbildung 3.3.5 ist das Gebiet nach dem Löschen der beiden Knoten

dargestellt. Die Flächenelemente der beiden zu löschenden Knoten sind gleichmäßig auf die Nachbarknoten verteilt worden, solange diese die genannten Voraussetzungen erfüllen.

Durch die Auswahl des Nachbarn mit dem geringsten Kommunikationsverkehr zur Aufnahme eines Flächenelements wird eine gleichmäßige Verteilung der Flächenelemente auf die Nachbarn sichergestellt. Da sich der Kommunikationsverkehr eines Knotens erhöht, wenn er ein Flächenelement aufnimmt, wird er später nicht mehr minimal in der Nachbarschaft sein. Dann wird ein anderer Knoten aus der Nachbarschaft das nächste Flächenelement aufnehmen.

Die Bedingungen der maximalen Fläche und des maximalen Kommunikationsverkehrs können bewirken, daß kein Nachbarknoten mehr Flächenelemente aufnehmen kann. In diesem Fall können die übrigen Flächenelemente dieses Knotens nicht mehr entfernt werden. Durch das Optimieren anderer Knoten können sie eventuell später noch komplett auf andere Knoten verteilt werden. Falls dies nicht möglich ist, da z.B. mehrere benachbarte Knoten die maximale Fläche erreicht haben, bleibt dieser Knoten mit niedrigem Kommunikationsverkehr übrig. Er bildet dann einen Demand-Node, der einen geringeren Kommunikationsverkehr wie die übrigen Demand-Nodes repräsentiert.

In den folgenden Histogrammen der Testgebiete wird das Ergebnis der Optimierungsphase I im Hinblick auf die Verteilung des Kommunikationsverkehrs der Demand-Nodes untersucht.

### **3.3.1.2 Verkehrsverteilung der Demand-Nodes**

Durch das Löschen von Knoten mit niedrigem Kommunikationsverkehr wird die Verkehrsverteilung der Demand-Nodes dem Ideal gleichen Kommunikationsverkehrs aller Demand-Nodes angenähert. Vor allem ist dieser im Bereich von  $T_{min}$  bis  $R_{T_{max}} \cdot T_{max}$  gehäuft, wobei größere Werte mit höherer Wahrscheinlichkeit angenommen werden. Viele Knoten haben zudem den maximalen Kommunikationsverkehr, da sie durch das Löschen verkehrsarmer Nachbarknoten Flächenelemente dazu gewinnen.

In Abbildung 3.3.6 ist die Verkehrsverteilung der Demand-Nodes des Würzburger Gebiets bei einem minimalen Kommunikationsverkehr  $T_{min}$  von 60% vom maximalen und einer erlaubten Überschreitung von  $T_{max}$  um 10% dargestellt.

In Abbildung 3.3.7 wird die Verkehrsverteilung bei gleichen Parametern für das Dallas Gebiet gezeigt.

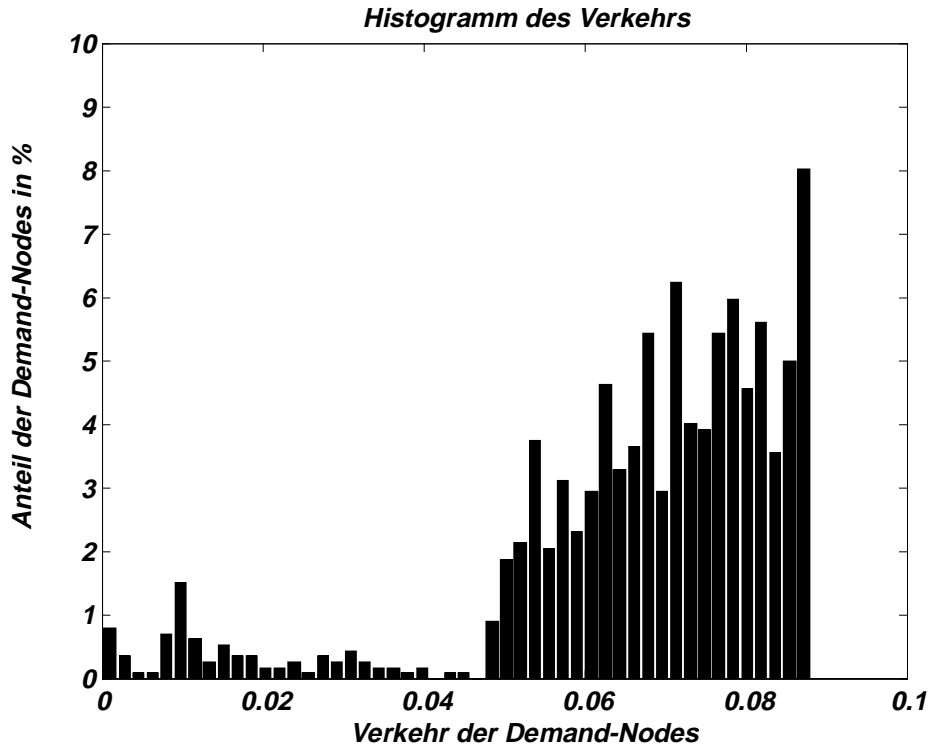


Abbildung 3.3.6: Würzburger Gebiet, mit  $T_{min} = 0.6 \cdot T_{max}$ ,  $R_{Tmax} = 1.1$

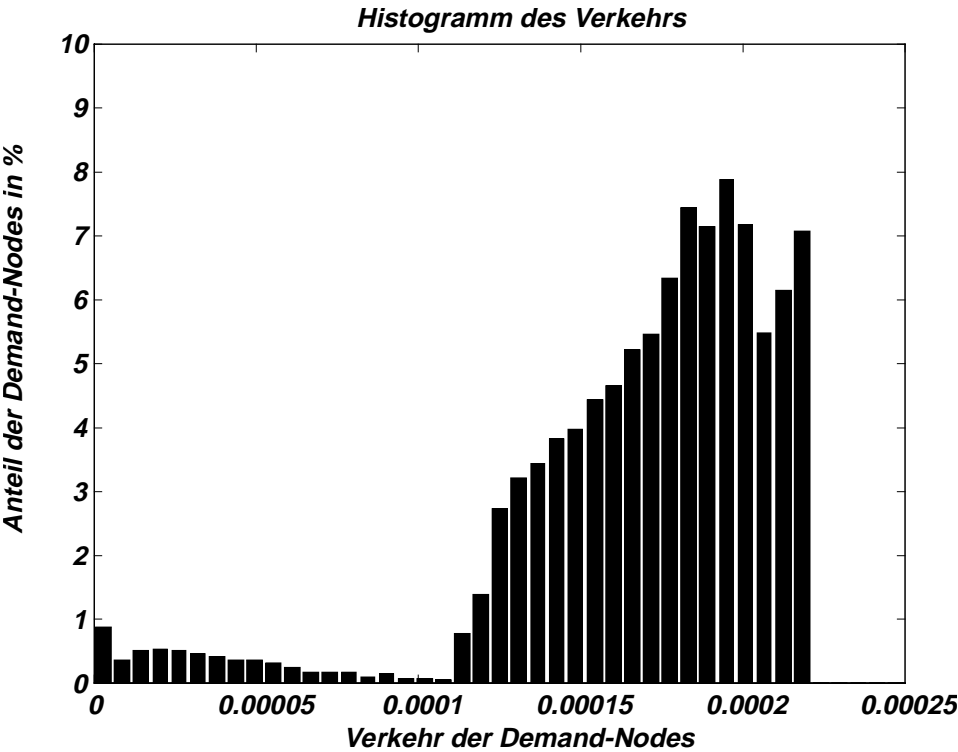


Abbildung 3.3.7: Dallas Gebiet, mit  $T_{min} = 0.6 \cdot T_{max}$ ,  $R_{Tmax} = 1.1$

### 3.3.2 Optimierungsphase II

Obwohl schon eine Verbesserung der Verkehrsverteilung der Demand-Nodes erreicht wurde, ist die Verteilung noch weit vom Ideal entfernt. Deshalb wird ein weiterer Optimierungsschritt eingeführt, der direkt auf der Basis einzelner atomarer Flächenelemente der Verkehrsmatrix arbeitet. Dabei wird versucht, die Flächenelemente zwischen angrenzenden Knoten zu wechseln, falls dadurch die Verkehrsverteilung insgesamt verbessert wird. Für den Algorithmus werden zwei Schwellwerte definiert, die festlegen, in welchem Bereich der Kommunikationsverkehr der Demand-Nodes liegen muß, damit ein Flächenelement von einem Knoten zum anderen wechseln kann.

Die beiden Parameter der Optimierungsstufe II sind:

- Maximaler Kommunikationsverkehr  $T_{select\_max}$ , den ein Knoten haben kann, damit Flächenelemente von ihm entfernt und zu anderen Knoten gewechselt werden können.
- Minimaler Kommunikationsverkehr  $T_{select\_min}$ , den ein Knoten besitzen muß, damit er Flächenelemente aufnehmen kann.

Durch den Parameter  $T_{select\_max}$  wird sichergestellt, daß nur Flächenelemente von Knoten mit geringem Kommunikationsverkehr entfernt werden. Der Parameter  $T_{select\_min}$  ist hingegen nötig, damit nur Knoten Flächenelemente aufnehmen, die bereits einen hohen Kommunikationsverkehr erreicht haben. Mit den beiden Parametern werden die Knoten in zwei Klassen aufgeteilt: In eine Klasse von Knoten mit niedrigem Kommunikationsverkehr, die Flächenelemente abgeben und dabei eventuell komplett entfernt werden können. Und eine zweite, die Flächenelemente hinzugewinnen, bis der maximale Kommunikationsverkehr oder die maximale Fläche erreicht wird.

Durch die Erhöhung der Anzahl der Knoten mit maximalem Kommunikationsverkehr verbessert sich die Verkehrsverteilung der Knoten. Anzumerken ist, daß Knoten auch in beiden Klassen liegen können, falls  $T_{select\_max}$  größer als  $T_{select\_min}$  ist. Im Laufe des Algorithmus werden die Knoten aber einer Klasse zugeordnet, da sich ihr Kommunikationsverkehr durch den Verlust von Flächenelementen unter die Schranke  $T_{select\_min}$  erniedrigt, oder durch Zugewinn über  $T_{select\_max}$  erhöht wird. Meist wird sogar  $T_{select\_max}$  größer als  $T_{select\_min}$  gewählt, da die Ergebnisse der Optimierung durch die erhöhte Flexibilität verbessert werden.

#### 3.3.2.1 Algorithmus

In Abbildung 3.3.8 ist der Grundaufbau des Algorithmus dargestellt. In einer Schleife wird so oft zufällig ein Flächenelement der Matrix ausgewählt, wie Flächenelemente in der Matrix vorhanden sind. Für jedes Flächenelement wird geprüft, ob der Knoten, dem das Flächenelement angehört, durch einen Wechsel dieses Flächenelements zu einem anderen Knoten optimiert werden kann. So-

lange in einem Durchgang der Schleife ein Knoten optimiert werden konnte, wird sie wiederholt.

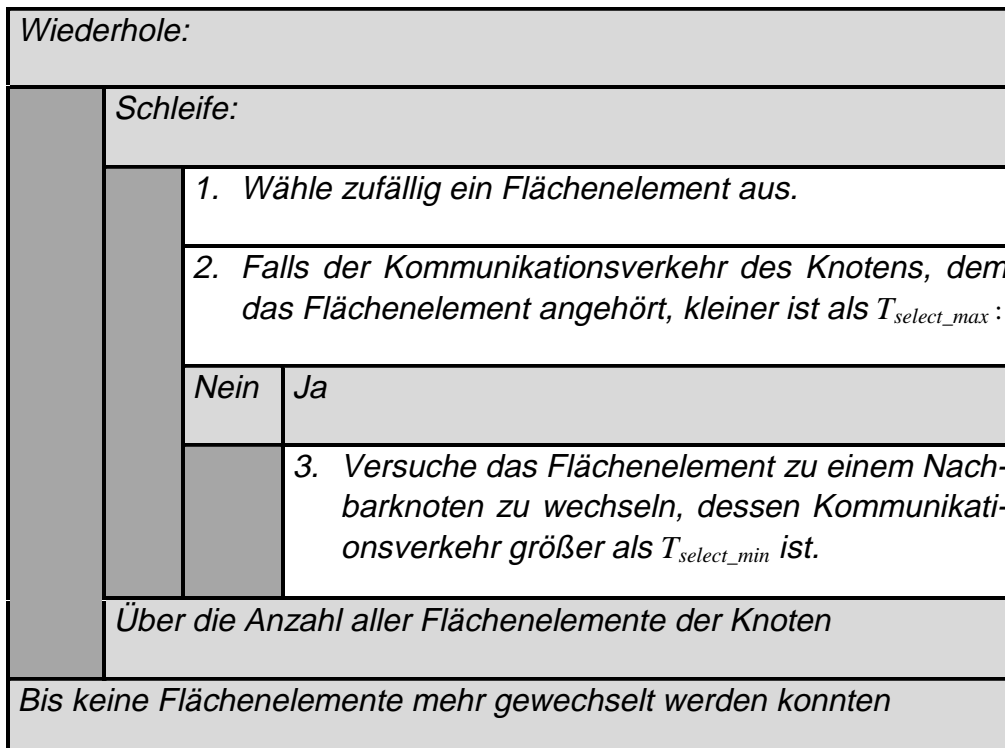


Abbildung 3.3.8: Algorithmus der Optimierung von Flächenelementen

Der Knoten, dem das zufällig ausgewählte Flächenelement  $E$  angehört, wird mit  $K$  bezeichnet. Falls der Knoten  $K$  einen Kommunikationsverkehr kleiner als  $T_{select\_max}$  hat, wird nach einem geeigneten Nachbarknoten gesucht, der dieses Flächenelement aufnehmen kann. Es wird dann dem Nachbarknoten  $N$  mit dem größten Kommunikationsverkehr zugewiesen, der es noch aufnehmen kann. Ein Knoten kann nur ein Flächenelement aufnehmen, falls weder die maximale Anzahl von Flächenelementen, noch der bei der Optimierung maximale Kommunikationsverkehr  $R_{Tmax} \cdot T_{max}$  überschritten wird. Außerdem muß der Kommunikationsverkehr des Knotens größer als  $T_{select\_min}$  sein, damit er als Kandidat zur Aufnahme des Flächenelements  $E$  in Frage kommt.

Im Gegensatz zum Verfahren der Optimierungsphase I wird hier der Nachbar mit dem größtem Kommunikationsverkehr gewählt. Dadurch werden die Knoten zu einem maximalem Kommunikationsverkehr hin optimiert. Die Verkehrsverteilung der zu erzeugenden Demand-Nodes wird so dem Ziel des gleichen Kommunikationsverkehrs für alle Demand-Nodes angenähert. Die Knoten mit hohem Kommunikationsverkehr gewinnen, solange es die Begrenzung der Fläche erlaubt, Flächenelemente hinzu, bis sie den maximalen Kommunikationsverkehr erreicht haben. Wenn ein Knoten wegen dieser Bedingungen keine Flächenelemente mehr aufnehmen kann, gewinnt der Knoten mit dem nächst

niedrigerem Kommunikationsverkehr Flächenelemente hinzu. So erreichen im Laufe des Algorithmus immer mehr Knoten den maximalen Kommunikationsverkehr. Die Begrenzung der maximalen Fläche wirkt sich hier kaum aus, da die Knoten mit großer Fläche in einem Bereich des Planungsgebietes mit niedriger Verkehrsdichte liegen. Sie repräsentieren dort meist einen Kommunikationsverkehr kleiner als  $T_{select\_min}$  und können damit kein Flächenelement aufnehmen.

Die Auswahlbedingung für das Wechseln eines Flächenelements  $E$  lautet also:

$$T_K < T_{select\_max} ,$$

$$T_N > T_{select\_min} ,$$

$$T_N + T_E < R_{Tmax} \cdot T_{max} ,$$

$$A_N + A_E \leq A_{max} ,$$

$$T_N > T_i \forall i \in Nh(K) .$$

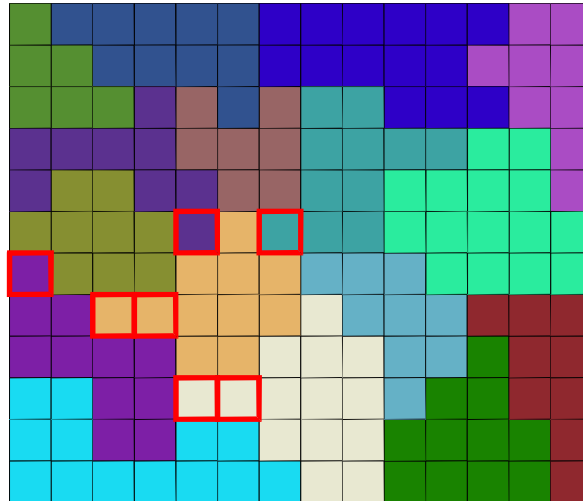
Falls kein geeigneter Nachbar gefunden wird, wird die innere Schleife fortgeführt und ein neues Flächenelement zufällig zum Optimieren ausgewählt. Der gesamte Algorithmus wird solange wiederholt, bis bei einem kompletten Durchlauf der Schleife kein Flächenelement mehr optimiert werden konnte.

Als Beispiel für die Optimierung sind in Abbildung 3.3.9 mehrere Felder dargestellt, die zur Optimierung ausgewählt wurden. Für diese Felder wird ein geeigneter Nachbarknoten gesucht, dem sie zugeordnet werden können, um insgesamt eine bessere Verkehrsverteilung der Knoten zu erreichen.

In Abbildung 3.3.10 werden die optimierten Felder dargestellt, nachdem sie zu ihren neuen Knoten gewechselt sind.

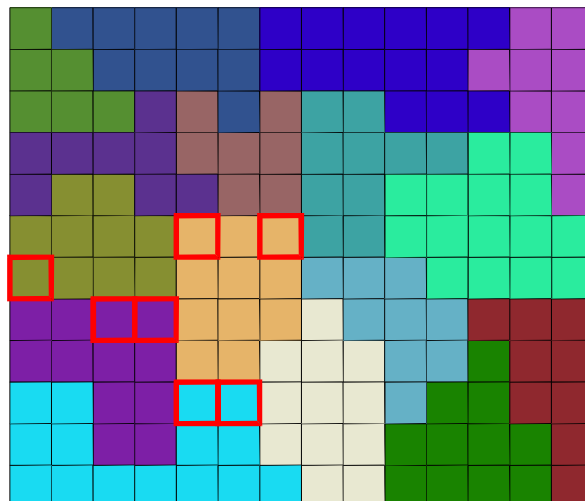
Für ein gutes Ergebnis der Optimierungsphase II müssen möglichst wenige Knoten einen Kommunikationsverkehr kleiner als  $T_{select\_min}$  haben. Diese Knoten können durch den Algorithmus nicht den maximalen Kommunikationsverkehr erreichen. Der Wert  $T_{select\_min}$  darf aber nicht zu klein gewählt werden, da die Optimierung zum maximalen Kommunikationsverkehr um so schwieriger wird, je größer die Differenz zwischen  $T_{select\_min}$  und  $T_{max}$  ist. Für die Vergrößerung des Kommunikationsverkehrs müssen nämlich Flächenelemente von anderen Knoten herangezogen werden. Im vorhergehenden Durchlauf der Optimierungsphase I wird versucht, alle Knoten mit einem Kommunikationsverkehr niedriger als  $T_{min}$  zu löschen. Bei einer Wahl von  $T_{min}$  kleiner als  $T_{select\_min}$  werden die obigen Bedingungen erfüllt. Aus diesem Grund kann die Optimierungsphase II nur nach einem Durchlauf der Optimierungsphase I verwendet werden. Die Optimierungsphase II kann daher nicht einzeln bewertet werden.





— zur Optimierung ausgewählte Felder

Abbildung 3.3.9: Beispiel zur Feldauswahl der Optimierung



— optimierte Felder

Abbildung 3.3.10: Felder nach der Optimierung

### 3.3.3 Ergebnis nach beiden Optimierungsphasen

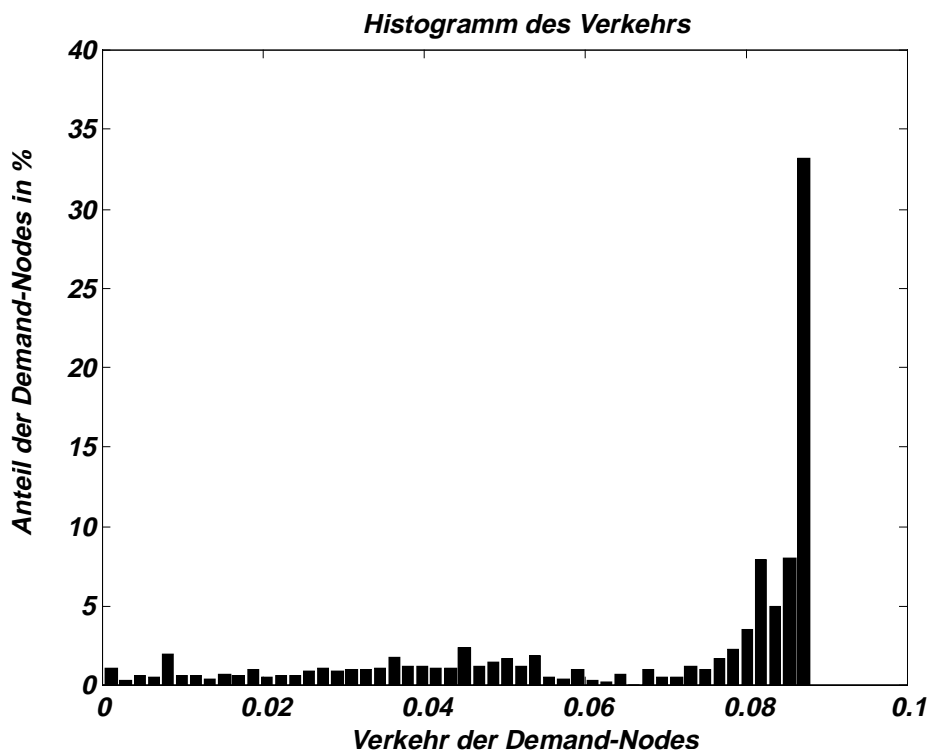
Für die Untersuchung der Verkehrsverteilung nach beiden Optimierungsphasen werden die in Tabelle 3.3.1 genannten Parameter verwendet. Diese Parameter ergaben bei verschiedenen Meßreihen eine Verkehrsverteilung, die dem Ideal am nächsten lag.

<b>Parameter</b>	<b>Wert</b>
$T_{min}$	$0.6 \cdot T_{max}$
$R_{T_{max}}$	1.1
$T_{select\_min}$	$0.7 \cdot T_{max}$
$T_{select\_max}$	$0.9 \cdot T_{max}$

*Tabelle 3.3.1: Für die Optimierung verwendete Parameter*

Die Optimierung der Demand-Nodes durch das Wechseln von einzelnen Flächenelementen innerhalb der Knoten verbessert die Verkehrsverteilung stark, wie man in Abbildung 3.3.11 für das Würzburger Gebiet und in Abbildung 3.3.12 für das Dallas Gebiet erkennen kann.

Die Demand-Nodes häufen sich am maximalen Kommunikationsverkehr und in einem kleinen Bereich unterhalb. Nur wenige Demand-Nodes haben einen niedrigeren Kommunikationsverkehr. Diese sind entweder durch die maximale Fläche beschränkt und können damit nicht mehr Kommunikationsverkehr repräsentieren, oder sie konnten in der Optimierungsphase II nicht optimiert werden, da keine geeigneten Nachbarn gefunden werden konnten.



*Abbildung 3.3.11: Würzburger Gebiet, mit Parametern gemäß Tabelle 3.3.1*

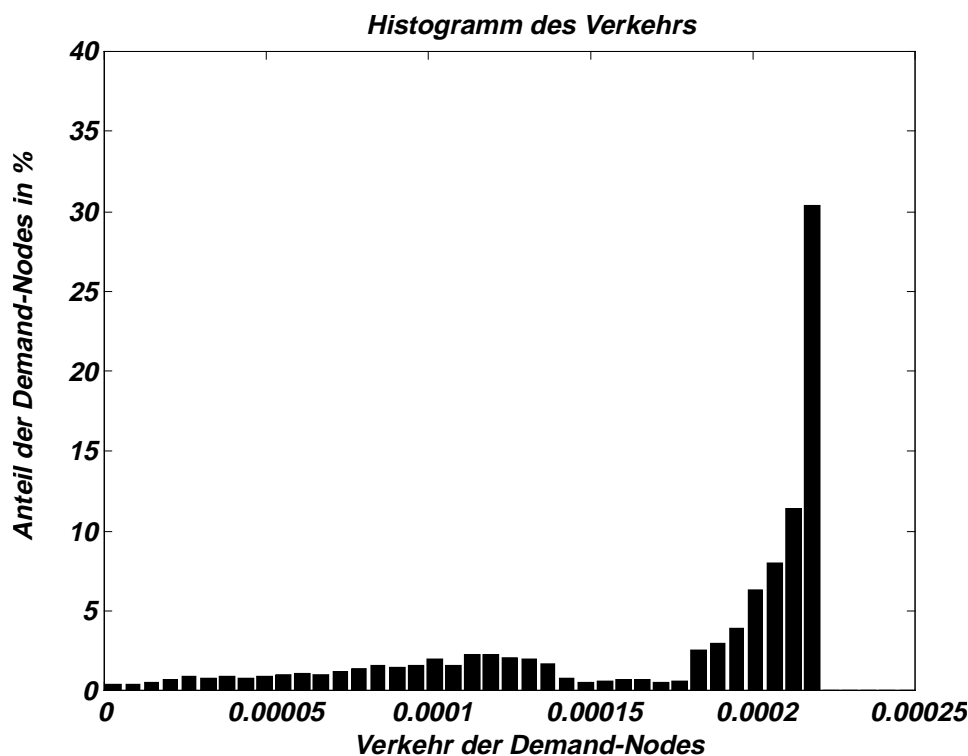


Abbildung 3.3.12: Dallas Gebiet, mit Parametern gemäß Tabelle 3.3.1

Die Verkehrsverteilung der Demand-Nodes ist nach beiden Optimierungsphasen dem Ideal schon sehr nahe. Es ergibt sich aber das Problem, daß bei der Anwendung beider Optimierungsphasen insgesamt fünf Parameter bestimmt und bestmöglich eingestellt werden müssen. Dies ist ein klarer Nachteil gegenüber dem Verfahren ohne Optimierung, bei dem man nur die zwei essentiellen Parameter festgelegt muß.

Bei der Entwicklung der Algorithmen der beiden Optimierungsphasen hat sich gezeigt, daß die besten Ergebnisse mit einer verbesserten Version der Optimierungsphase I erzielt werden. Die Verwendung der Optimierungsphase II ist dann nicht mehr nötig. In Kapitel 3.3.1 wird die verbesserte Version des Algorithmus der Optimierungsphase I beschrieben, die sich nach Tests mit verschiedenen Verfahren als optimal herausgestellt hat.

### 3.3.4 Verbesserte Optimierung

Die verbesserte Optimierungsphase I liefert bei einer bestimmten Wahl der Parameter die besten Ergebnisse. Dazu wird für den Parameter  $T_{min}$  ein sehr hoher Wert genommen, der nur wenig unterhalb des Maximalwertes  $T_{max}$  liegt. In der Praxis werden Werte zwischen 90% und 99% von  $T_{max}$  verwendet. Dies hat zur Folge, daß für fast jeden Knoten versucht wird, ihn zu löschen.

Bei obiger Wahl von  $T_{min}$  werden sogar von Knoten mit sehr hohem Kommunikationsverkehr, z.B. bis zu 99% von  $T_{max}$ , Flächenelemente entfernt. Dadurch

stehen auch Knoten mit hohem Kommunikationsverkehr zur Verfügung, die noch Flächenelemente aufnehmen können, ohne  $T_{max}$  zu übersteigen. Aus diesem Grund ist eine Überschreitung von  $T_{max}$  durch den Faktor  $R_{T_{max}}$  hier nicht nötig. Deshalb wird  $R_{T_{max}}$  auf 1.0 gesetzt, d.h. der absolut maximale Kommunikationsverkehr  $R_{T_{max}} \cdot T_{max}$ , den ein Demand-Node repräsentieren kann, entspricht dem Parameter  $T_{max}$  des Verfahrens ohne Optimierung.

Der Kommunikationsverkehr der Knoten häuft sich dabei im Bereich von  $T_{min}$  bis  $T_{max}$ . Da dieser Bereich bei obiger Parameterwahl sehr schmal ist, ergibt sich eine starke Konzentration des Knotenverkehrs in der Nähe des maximalen Verkehrswertes. Bei der Wahl von  $T_{min} = 0.99 \cdot T_{max}$  konzentriert sich der Kommunikationsverkehr der meisten Knoten in einem Bereich von 1% Breite unterhalb von  $T_{max}$ . Offensichtlich können nicht alle Knoten mit einem Kommunikationsverkehr kleiner als  $T_{min}$  gelöscht werden. Der Algorithmus optimiert diese Knoten beim Versuch, sie zu löschen, aber in der Art, daß am Ende der Kommunikationsverkehr sehr vieler Knoten in diesem Bereich liegt.

Für das Verfahren des Zusammenfassens von Knoten mit verbesserter Optimierung sind insgesamt nur die essentiellen Parameter des Verfahrens ohne Optimierung notwendig. Der Parameter  $T_{min}$  wird auf den Wert  $0.99 \cdot T_{max}$  unabhängig vom verwendeten Gebiet gesetzt.

Insgesamt benötigt der Algorithmus mit verbesserter Optimierung folgende Parameter:

- Maximaler Kommunikationsverkehr  $T_{max}$ . Dieser Wert bestimmt den maximalen Kommunikationsverkehr, den ein Demand-Node repräsentieren kann. Der Parameter entspricht dem des Verfahrens ohne Optimierung.
- Maximale Fläche  $A_{max}$ . Dieser Wert gibt die maximale Größe der Fläche an, die ein Demand-Node beschreibt. Der Parameter entspricht dem des Verfahrens ohne Optimierung.
- Minimaler Kommunikationsverkehr  $T_{min}$ . Alle Knoten, die einen Kommunikationsverkehr kleiner als  $T_{min}$  haben, versucht man zu entfernen, indem ihre Flächenelemente auf Nachbarn verteilt werden. Der Parameter  $T_{min}$  wird im Regelfall auf  $0.99 \cdot T_{max}$  gesetzt.

Da der Parameter  $T_{min}$  unabhängig vom Gebiet gleich bleibt, müssen für das Verfahren mit verbesserter Optimierung keine zusätzlichen Werte eingestellt werden. Die beiden Grundparameter des Verfahrens ohne Optimierung reichen vollkommen aus.

In der Abbildung 3.3.13 ist die Verkehrsverteilung für das Würzburger Gebiet und in Abbildung 3.3.14 für das Dallas Gebiet dargestellt, wobei die gleichen Parameter wie beim Verfahren ohne Optimierung verwendet wurden. Die Ergebnisse entsprechen voll den theoretischen Erwartungen und sind deutlich besser als die mit beiden Optimierungsphasen erzielten. In den Testgebieten ist die Anzahl der Demand-Nodes, die den maximalen Kommunikationsverkehr

erreichen, bei verbesserter Optimierung um etwa 50% bis 90% höher als nach Durchlauf beider Optimierungsphasen.

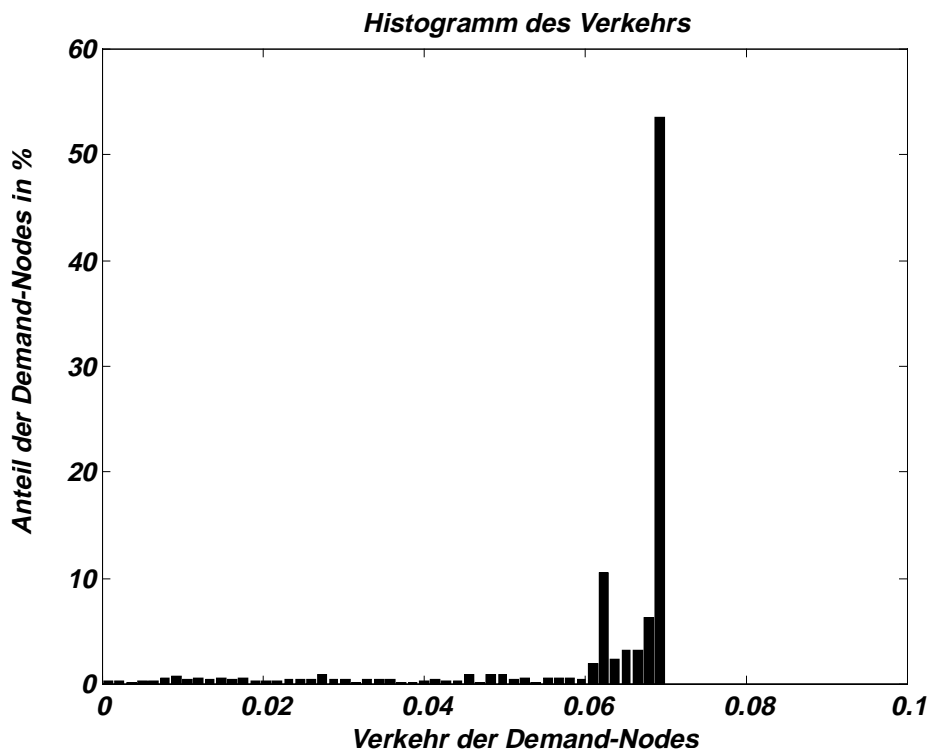


Abbildung 3.3.13: Würzburger Gebiet, verbesserte Optimierung

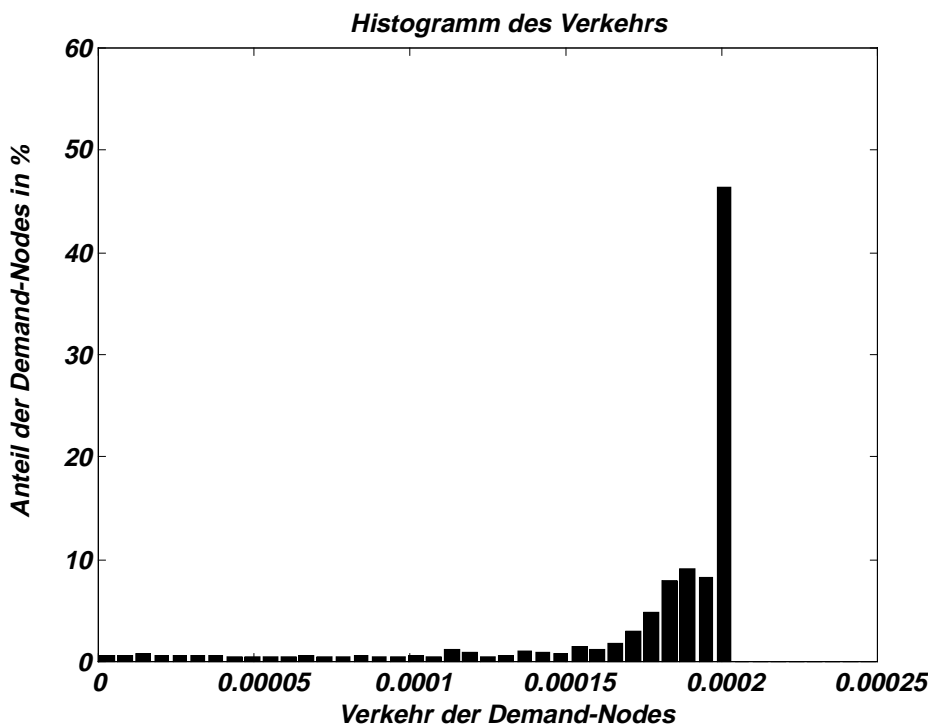


Abbildung 3.3.14: Dallas Gebiet, verbesserte Optimierung

Bei Verwendung der genannten Parameter für die verbesserte Optimierungsphase I ist die Optimierungsphase II nicht mehr nötig. Durch sie kann keine Verbesserung der Verkehrsverteilung mehr erreicht werden. In Kapitel 3.3.5 wird dazu theoretisch untersucht, wie die bestmögliche Verteilung bei den gegebenen Voraussetzungen aussieht.

Die im Histogramm erkennbaren Demand-Nodes mit niedrigerem Kommunikationsverkehr sind durch die Flächenbeschränkung bedingt und hängen vor allem von der Struktur des Gebiets ab. Im idealen Gebiet mit konstanter Verkehrsdichte entspricht die Verkehrsverteilung genau der nach dem Algorithmus theoretisch zu erwartenden Verteilung. Die Verteilung des konstanten Gebiets ist in Abbildung 3.3.15 dargestellt. In diesem Fall ist der Kommunikationsverkehr nahezu aller Demand-Nodes gleich.

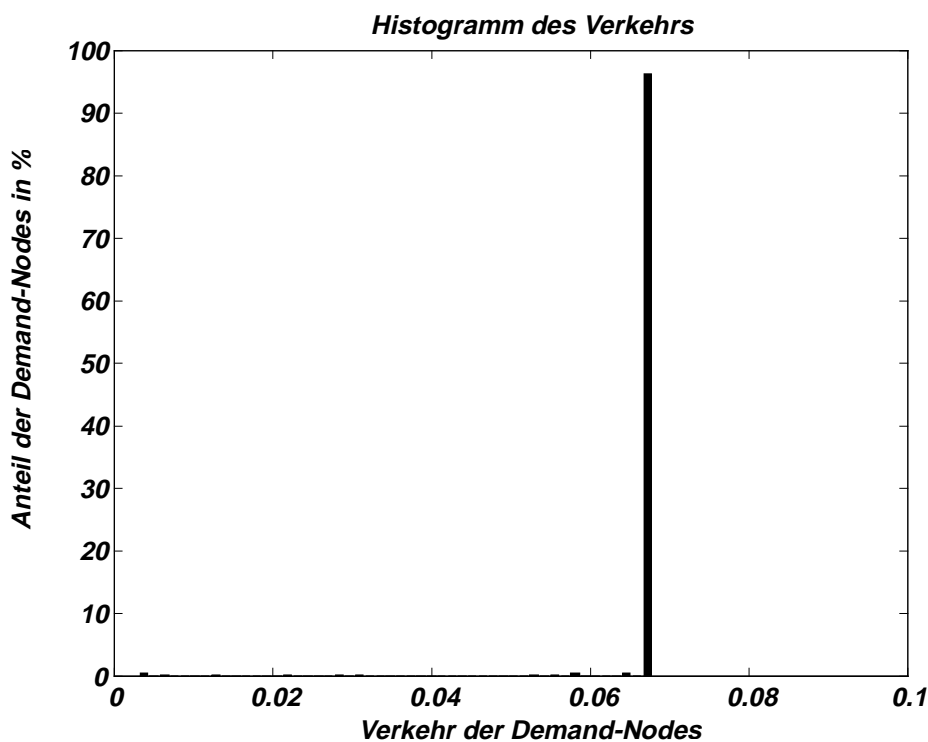


Abbildung 3.3.15: Konstantes Gebiet

### 3.3.5 Ideal erreichbare Verkehrsverteilung

Im folgenden Abschnitt wird untersucht, inwieweit das Ideal des gleichen Kommunikationsverkehrs aller Demand-Nodes in der Theorie erreicht werden kann. Die Verwendung einer Verkehrsmatrix führt schon bei den Grunddaten zu einer Quantisierung des Kommunikationsverkehrs. Da den Demand-Nodes nur ganze Flächenelemente der Verkehrsmatrix zugeordnet werden können, wird der maximale Kommunikationsverkehr mindestens um den Quantisierungsfehler verfehlt. Dieser Fehler hängt vom Planungsgebiet ab und kann mit dem ma-

ximalen Kommunikationsverkehr eines Matrixelements nach oben abgeschätzt werden. Bei einem auf einer Verkehrsmatrix beruhenden Verfahren liegt der Kommunikationsverkehr eines Demand-Nodes im Idealfall in einem Intervall zwischen  $T_{max} - Q$  und  $T_{max}$ . Mit  $Q$  wird dabei der maximale Quantisierungsfehler bezeichnet. Dieser entspricht dem größten Kommunikationsverkehr  $T_{Emax}$  eines Matrixelements  $E_i$  in der Verkehrsmatrix  $T(x,y)$ :

$$T_{Ki} \in [T_{max} - Q, T_{max}], \forall (1 \leq i \leq N) ,$$

$$T_{Emax} \geq T_{Ei}, \forall E_i \in T(x, y) ,$$

$$Q = T_{Emax} .$$

Hierbei gibt  $N$  die Gesamtzahl der Demand-Nodes und  $T_{Ki}$  den Kommunikationsverkehr des Demand-Nodes  $K_i$  an.

Diese einfache Beziehung ist aber nur für den Fall von Demand-Nodes ohne Begrenzung der Fläche korrekt. Im Fall einer Flächenbeschränkung liegt der Kommunikationsverkehr der Demand-Nodes  $K_j$ , deren Fläche den maximalen Wert  $A_{max}$  erreicht, unterhalb dieses Intervalls. Mit  $M_A$  wird die Menge der Demand-Nodes bezeichnet, deren Fläche begrenzt ist:

$$M_A = \{1 \leq j \leq N \mid A_{Kj} = A_{max}\} .$$

Den minimalen Kommunikationsverkehr eines flächenbeschränkten Demand-Nodes kann man nach unten abschätzen durch den minimalen Kommunikationsverkehr eines Flächenelements  $T_{Kmin}$  mal der maximalen Anzahl der Flächenelemente, die ein Demand-Node besitzen darf:

$$T_{Kmin} = T_{Emin} \cdot \frac{A_{max}}{A_E} ,$$

$$T_{Emin} \leq T_{Ei}, \forall E_i \in T(x, y) .$$

Hierbei gibt  $A_E$  die Fläche eines Matrixelements an, die für alle Flächenelemente gleich ist.

Insgesamt ergibt sich folgende ideale Verkehrsverteilung bei einem maximalen Kommunikationsverkehr von  $T_{max}$  und maximaler Fläche von  $A_{max}$  für die Demand-Nodes  $K_i$ :

$$T_{Km} \in [T_{Kmin}, T_{max}], \forall (1 \leq m \in M_A \leq N) ,$$

$$T_{Kn} \in [T_{max} - Q, T_{max}], \forall (1 \leq n \notin M_A \leq N) ,$$

$$M_A = \{1 \leq j \leq N \mid A_{Kj} = A_{max}\} .$$

Das theoretische Ideal von

$$T_{Ki} = T_{max}, \forall (1 \leq i \leq N)$$

kann aus obigen Gründen nicht erreicht werden. Um diese Abweichung zu kompensieren, kann entweder bei der Planung der Kommunikationsverkehr der einzelnen Demand-Nodes berücksichtigt werden. Eine andere Möglichkeit ist, den Kommunikationsverkehr der Demand-Nodes mit  $T_{max}$  nach oben abzuschätzen und diesen Wert für alle Demand-Nodes anzunehmen. Der Fehler hierbei ist sehr klein, da der Kommunikationsverkehr der allermeisten Demand-Nodes nur gering von  $T_{max}$  abweicht. Diese Ergebnisse können in den Histogrammen der Verkehrsverteilung verifiziert werden.

### 3.3.6 Ergebnisse des Algorithmus mit Optimierung

In Abbildung 3.3.16 und Abbildung 3.3.17 werden die Demand-Nodes des Würzburger und des Dallas Gebiets dargestellt, die bei dem Verfahren mit verbesserter Optimierung erzeugt werden. Die verwendeten Parameter entsprechen denen des Verfahrens ohne Optimierung. Deshalb sind die Demand-Node Verteilungen direkt mit den in der Abbildung 3.2.10 und Abbildung 3.2.11 gezeigten vergleichbar, die ohne Optimierung erzeugt wurden.

In Tabelle 3.3.2 wird die Anzahl der Demand-Nodes vor und nach der Optimierung für jeweils einen Durchlauf des Verfahrens gegenübergestellt. Die abgebildete Anzahl der Demand-Nodes variiert bei verschiedenen Läufen, da der Algorithmus nicht rein deterministisch arbeitet. Bei vorgegebener, gleicher Initialisierung des Zufallsgenerators werden exakt identische Demand-Nodes erzeugt. Andernfalls schwankt die Anzahl der Demand-Nodes im Promillebereich. Durch die Optimierung werden abhängig vom Gebiet eine bestimmte Anzahl von Demand-Nodes entfernt.

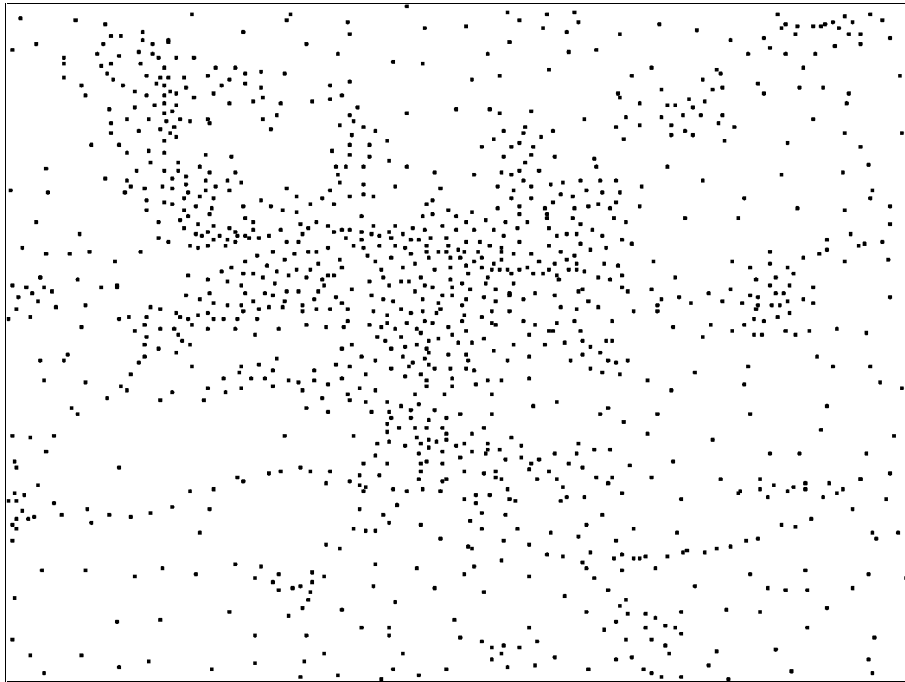
<b>Vergleichsfläche</b>	<b>Anzahl der Demand-Nodes</b>	<b>Anzahl Demand-Nodes nach der Optimierung</b>
<i>Würzburger Gebiet</i>	1289	1226
<i>Dallas Gebiet</i>	40589	34849
<i>Synthetisches Gebiet</i>	1381	1273
<i>Konstantes Gebiet</i>	1281	1125

*Tabelle 3.3.2: Anzahl der Demand-Nodes vor und nach der Optimierung*

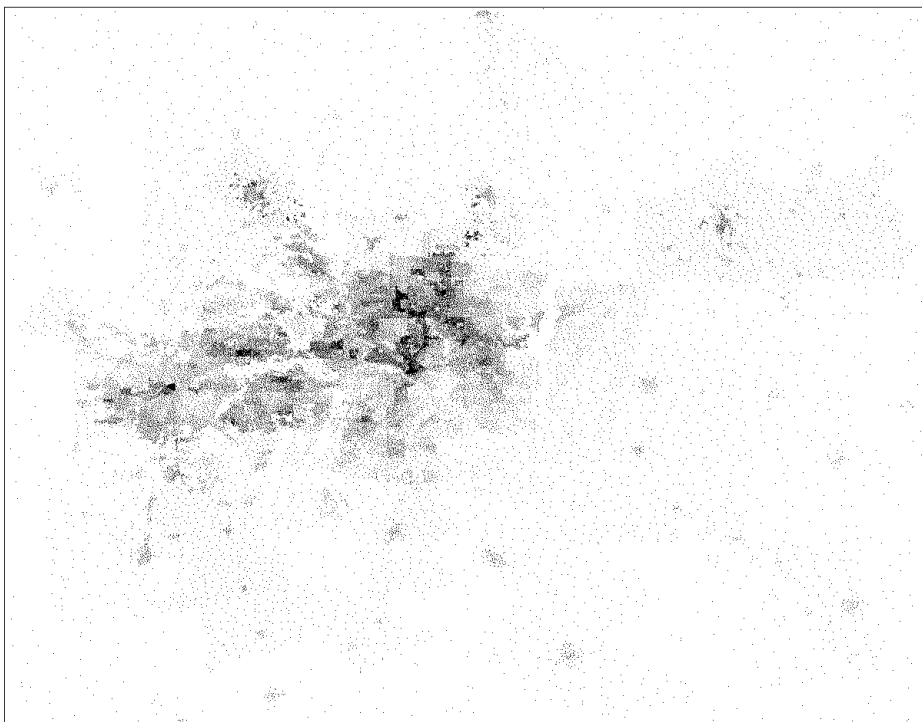
Bei der qualitativen Untersuchung der Demand-Nodes in den Testgebieten zeigt sich kein signifikanter Unterschied zur Verteilung der Demand-Nodes ohne Optimierung. Dies wird in einer quantitativen Analyse der Approximationsgüte in Kapitel 4 verifiziert. Dennoch ist die Verkehrsverteilung der Demand-Nodes nach der verbesserten Optimierung um Größenordnungen dem Idealbild näher.



Vor allem gibt die Dichte der Demand-Nodes in einem Gebiet den wirklichen Kommunikationsverkehr dort an.



*Abbildung 3.3.16: Demand-Nodes des Würzburger Gebiets*



*Abbildung 3.3.17: Demand-Nodes des Dallas Gebiets*

### 3.3.7 Merkmale des Algorithmus

Im folgenden werden nochmals die charakteristischen Merkmale des Algorithmus des Zusammenfassens von Knoten prägnant dargestellt:

- Die Fläche eines Demand-Nodes kann eine beliebige Form haben. Die Position des Demand-Nodes ist dabei der verkehrsmäßige Schwerpunkt dieser Fläche.
- Für die Größe der Fläche, die von einem Demand-Node repräsentiert wird, kann mit einem Parameter ein Grenzwert festgelegt werden.
- Die Anzahl der Demand-Nodes ist frei wählbar. Sie wird implizit durch die Wahl der Parameter des Algorithmus bestimmt. Minimal kann ein Demand-Node erzeugt werden, der den Kommunikationsverkehr des gesamten Planungsgebietes repräsentiert. Maximal kann für jedes Flächenelement der Verkehrsmatrix ein Demand-Node erzeugt werden.
- Der Kommunikationsverkehr der Demand-Nodes ist nicht unbedingt identisch. Vielmehr ergibt sich eine bestimmte Verteilung der Verkehrswerte, die von den Demand-Nodes repräsentiert werden. Die Güte der Verteilung kann dabei durch eine Optimierung der Demand-Nodes wesentlich verbessert werden.
- Für die Landklasse eines Demand-Nodes wird die Landklasse genommen, die den größten Bereich der vom Demand-Node repräsentierten Fläche darstellt.
- Die Struktur der erzeugten Demand-Nodes ist durch die zufällige Wahl der Objekte im Algorithmus nicht streng deterministisch.
- In Gebieten mit ähnlichem Kommunikationsverkehr wird ein homogenes Punktmuster mit maximaler Entropie erzeugt. Dies entspricht dem elementaren Bestreben nach maximaler Entropie in der Natur.
- Die Verteilung der Demand-Nodes ist bis auf Randeffekte unabhängig vom verwendeten Planungsgebiet.

## 4 Vergleich verschiedener Verfahren

Neben der qualitativen Bewertung der Demand-Nodes und der Auswertung der Histogramme beim Verfahren des Zusammenfassens von Knoten, ist auch eine mathematische Untersuchung nötig. Mit der qualitativen Methode kann die Qualität verschiedener Verteilungen von Demand-Nodes hinsichtlich einer Musterbildung oder der Repräsentation geographischer Objekte, wie z.B. Straßen oder Siedlungsgebieten, bewertet werden. Dabei ist aber nur eine rein qualitative Aussage möglich, d.h. es kann kein quantitatives Maß der Güte angegeben werden. Damit die Approximationsgüte der Verkehrsverteilung durch die Demand-Nodes ermittelt werden kann, ist ein quantitatives Fehlermaß notwendig. Dieses kann auch als Vergleichsmaß dienen, um verschiedene Parametersätze für die Verfahren zu bewerten.

Der Zweck der Demand-Nodes ist die möglichst genaue Approximation der Verkehrsverteilung im Planungsgebiet, die hier in Form einer Verkehrsmatrix vorliegt. Die Güte der Verkehrsmatrix kann nicht beurteilt werden, da sie direkt aus den geographischen Daten erzeugt wird. Die Verkehrsmatrix ist deswegen die einzige Vergleichsgrundlage. Ein Qualitätskriterium für Demand-Nodes ist die möglichst genaue Wiedergabe des Kommunikationsverkehrs in der Matrix. Die Abweichung des Kommunikationsverkehrs kann dabei als Fehlermaß dienen.

Zur Messung der Abweichung sind verschiedene Ansätze denkbar. In der räumlichen Statistik wird häufig das Konzept von Testflächen verwendet. Diese werden an unterschiedlichen Positionen über das Gebiet gelegt. Zum Vergleich von zwei Verteilungen werden dann die Daten innerhalb der Testflächen ausgewertet und miteinander verglichen. Außerdem können durch mehrere Meßreihen statistische Aussagen über die Momente der Verteilungen gewonnen werden. Interessant ist auch, inwieweit die Ergebnisse von der Form der Testfläche abhängen. Insgesamt können alle Eigenschaften einer Verteilung im zweidimensionalen Raum durch systematische Meßreihen mit Testflächen untersucht werden.

Bei der Bewertung der verschiedenen Verfahren zur Clusterung sind zum einen die Approximationsgüte der Verkehrsverteilung durch die Demand-Nodes von Interesse. Zum anderen ist auch die Genauigkeit der Repräsentation der Landnutzungsklassen wichtig, da sie für die Wellenausbreitungsmodelle benötigt werden.

Die Wahl der Form der Testflächen ist fundamental, da die Ergebnisse von ihr abhängen können. Die verwendeten Testflächen sollen z.B. die Versorgungsgebiete von Mobilfunksendern darstellen und eine ähnlich Form wie diese aufweisen. Deshalb wird eine kreisförmige Testfläche gewählt, was zudem den Vorteil hat, daß der Flächeninhalt nur durch den Radius  $r$  als einzigen Parameter bestimmt wird. Um die Auswirkung verschiedener Formen zu untersuchen, wird auch eine quadratische Testfläche verwendet. Da das Clusterungsverfahren des

rekursiven Partitionierens nur eine rechteckige Form der Demand-Node Flächen erzeugen kann, ist die Quadratform auch deshalb besonders interessant.

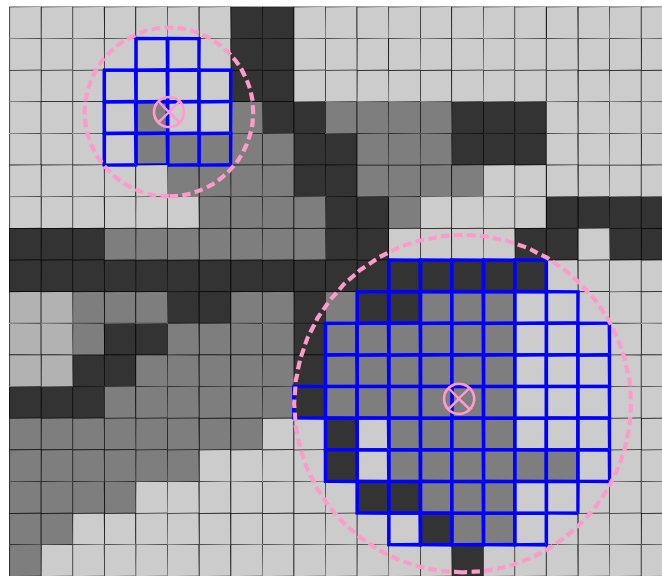


Abbildung 4.1: Kommunikationsverkehr eines Testkreises auf der Matrix

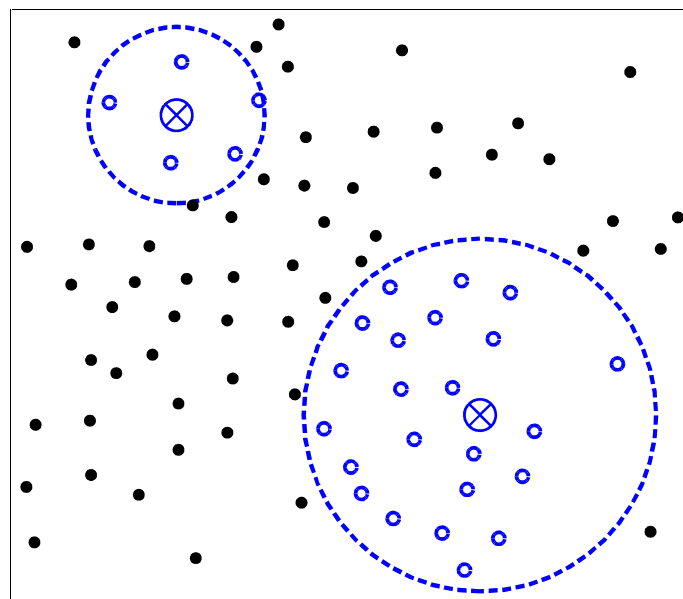


Abbildung 4.2: Kommunikationsverkehr der Demand-Nodes eines Testkreises

Im folgenden wird die Abweichung des Kommunikationsverkehrs untersucht. Dazu werden Testflächen über die Demand-Nodes und die Verkehrsmatrix gelegt, und der Kommunikationsverkehr innerhalb dieser Testflächen gemessen. Falls die gleiche Testfläche an die selbe Position gelegt wird, ist der Kommunikationsverkehr innerhalb im Idealfall identisch. Die Abweichung des Kommuni-

kationsverkehrs wird deswegen als Fehlermaß genommen. Die Vorgehensweise wird in Abbildung 4.1 und Abbildung 4.2 gezeigt.

## 4.1 Testverfahren

Um die Approximationsgüte der Verkehrsverteilung durch die Demand-Nodes zu beurteilen, werden verschiedene Meßreihen durchgeführt. Dazu wird jeweils ein bestimmter Radius  $r$  für die Testkreise vorgegeben. Der Testkreis wird  $M$  mal zufällig innerhalb des Planungsgebietes positioniert, wobei der Kommunikationsverkehr innerhalb des Testkreises in der Verkehrsmatrix gemessen wird. Dieser wird verglichen mit der Summe des Kommunikationsverkehrs der Demand-Nodes innerhalb desselben Testkreises. Zuerst wird bei gegebenem Radius  $r$  die Messung eines Testkreises  $K_i$  beschrieben.

### 4.1.1 Messung einer Testfläche

Für eine Vergleichsmessung wird eine zufällige Position  $(x_i, y_i)$  des Testkreises  $K_i$  gewählt, wobei zu beachten ist, daß der gesamte Testkreis innerhalb des Gebiets liegt. Die Testkreise müssen komplett innerhalb liegen, da ansonsten der Kommunikationsverkehr im außerhalb liegenden Bereich nicht gemessen werden könnte und damit das Fehlermaß verfälschen würde:

$$x_i = \text{rnd}(r, X_{\max} - r) ,$$

$$y_i = \text{rnd}(r, Y_{\max} - r) .$$

Hierbei gibt  $r$  den vorgegebenen Radius des Kreises an.  $X_{\max}$  und  $Y_{\max}$  stellen die Ausdehnung des Gebiets in  $X$ - und  $Y$ -Richtung dar. Die Funktion  $\text{rnd}(a, b)$  liefert einen zufälligen Wert im Bereich  $[a, b]$ .

Da die Testkreise komplett innerhalb des Planungsgebietes liegen müssen, hängen die möglichen Positionen von deren Radius ab. In Abbildung 4.1.1 sind zwei unterschiedlich große Testkreise dargestellt, wobei die erlaubten Positionen für die Kreiszentren innerhalb der dargestellten Rechtecke liegen.

Der Testkreis mit Radius  $r$  wird an der Position  $(x_i, y_i)$  auf die Matrix gelegt. Anschließend wird der Kommunikationsverkehr der Flächenelemente in der Matrix, die komplett innerhalb des Kreises liegen, zu  $T_M$  summiert. Dies wird in Abbildung 4.1 für zwei Testkreise gezeigt:

$$T_M = \sum_{j \in I_M(K_i)} T(x_j, y_j) ,$$

$$I_M(K_i) = \left\{ l \mid \sqrt{|x_i - x_l|^2 + |y_i - y_l|^2} \leq r \right\} .$$

Hierbei gibt  $T(x_j, y_j)$  den Kommunikationsverkehr des Flächenelements  $j$  an, wobei mit  $x_j$  und  $y_j$  seine Position bezeichnet wird.  $x_l$  und  $y_l$  gibt die Position des Flächenelements  $l$  an. Mit  $x_i$  und  $y_i$  wird die Position des Zentrums des Testkreises angegeben.  $I_M(K_i)$  gibt die Menge der Flächenelemente an, die komplett innerhalb des Testkreises liegen.

Die Flächenelemente, die nicht komplett innerhalb liegen, werden nicht berücksichtigt. Dies ist unnötig, da die Position der Testkreise beliebig ist, d.h. sie ist nicht auf die Grenzen der Flächenelemente beschränkt. Durch die Mittelung über viele Positionen der Testkreise wird auch der Kommunikationsverkehr der Flächenelemente berücksichtigt, die nur teilweise überdeckt werden.

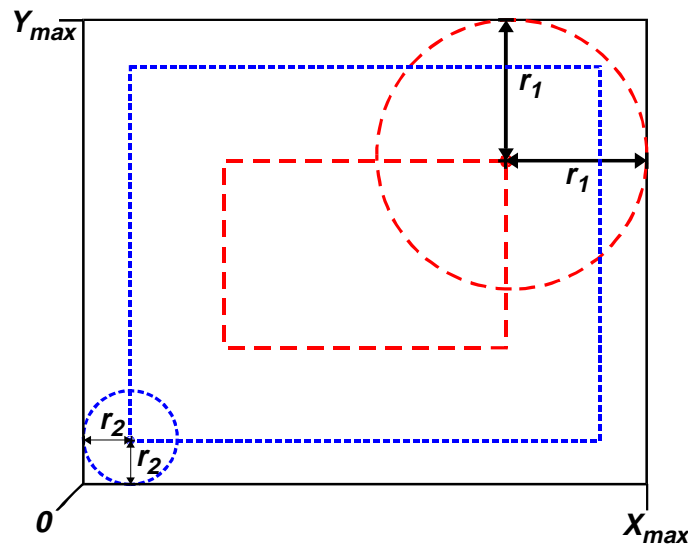


Abbildung 4.1.1: Mögliche Positionen von Testkreisen

Der gemessene Kommunikationsverkehr  $T_M$  gibt den wirklichen Kommunikationsverkehr an, den ein Sender mit einem dem Testkreis entsprechenden kreisförmigen Versorgungsbereich an der Position  $(x_i, y_i)$  versorgen müsste. Wie in Kapitel 2.2.3 beschrieben, liegen keine genaueren Daten über den Kommunikationsverkehr vor. Deshalb wird dieser Wert als Vergleichsmaßstab genommen, dem der Kommunikationsverkehr der Demand-Nodes innerhalb desselben Versorgungsbereiches möglichst gut entsprechen muß.

Nun wird, wie in Abbildung 4.2 dargestellt, der Kommunikationsverkehr der Demand-Nodes bestimmt, die sich innerhalb des Testkreises befinden:

$$T_D = \sum_{j \in I_D(K_i)} T_j ,$$

$$I_D(K_i) = \left\{ l \mid \sqrt{|x_i - x_l|^2 + |y_i - y_l|^2} \leq r \right\}$$

Hierbei gibt  $T_j$  den Kommunikationsverkehr des Demand-Nodes  $j$  an, und mit  $x_l$  und  $y_l$  wird die Position des Demand-Nodes  $l$  bezeichnet.  $I_D(K_i)$  ist die Menge der Demand-Nodes, die komplett innerhalb des Testkreises liegen.

Der Wert  $T_D$  gibt den von den Demand-Nodes im Testkreis repräsentierten Kommunikationsverkehr an, der sich möglichst wenig vom Kommunikationsverkehr  $T_M$  in der Verkehrsmatrix unterscheiden sollte. Daher wird als Fehlermaß die quadratische Abweichung  $\Delta^2 T$  definiert:

$$\Delta^2 T = (T_D - T_M)^2 .$$

Außerdem wird die relative Abweichung  $\Delta_R T$  vom Originalwert definiert:

$$\Delta_R T = \frac{|T_D - T_M|}{T_M} .$$

Diese Abweichungen geben den Approximationsfehler für die gegebene Testfläche an der speziellen Position  $(x_j, y_j)$  an. Damit eine statistisch relevante Aussage über die Abweichung bei einer gegebenen Testfläche möglich ist, muß eine Reihe von Messungen an verschiedenen Positionen durchgeführt werden.

#### 4.1.2 Verwendung einer quadratischen Testfläche

Um den Einfluß der Form der Testfläche auf das Ergebnis zu untersuchen, wird zusätzlich mit einer quadratischen Testfläche geprüft. Da das Verfahren des rekursiven Partitionierens nur rechteckige Testflächen erzeugt, könnte sich dadurch ein Vorteil für dieses Verfahren ergeben.

Die Positionierung der quadratischen Testfläche wird analog zur kreisförmigen durchgeführt. Der Radius  $r$  entspricht bei der quadratischen Testfläche der halben Seitenlänge. Dadurch ist es möglich, die Positionierung und die Verfahren zur Berechnung der Meßkurven analog zur kreisförmigen Testfläche zu übernehmen. Der Zusammenhang zwischen einer quadratischen und einer kreisförmigen Testfläche mit gleichem Wert  $r$  ist in Abbildung 4.1.2 dargestellt.

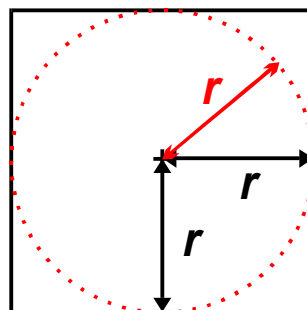


Abbildung 4.1.2: Kreisförmige und quadratische Testfläche

Für die Berechnung des Kommunikationsverkehrs innerhalb der quadratischen Testfläche in der Verkehrsmatrix werden folgende Formeln angewandt:

$$T_M^R = \sum_{j \in I_M^R(K_i)} T(x_j, y_j) ,$$

$$I_M^R(K_i) = \{ l \mid |x_i - x_l| \leq r \wedge |y_i - y_l| \leq r \} .$$

Der Kommunikationsverkehr der Demand-Nodes innerhalb der quadratischen Testfläche wird bestimmt mit:

$$T_D^R = \sum_{j \in I_D^R(K_i)} T_j ,$$

$$I_D^R(K_i) = \{ l \mid |x_i - x_l| \leq r \wedge |y_i - y_l| \leq r \} .$$

Die Fehlermaße werden analog zum Fall einer kreisförmigen Testfläche definiert.

### 4.1.3 Meßreihe einer gegebenen Testfläche

Für eine statistische Aussage ist das Ergebnis vieler Messungen an verschiedenen Positionen im Gebiet nötig. Dazu wird die oben beschriebene Messung  $M$  mal wiederholt und die einzelnen Fehler  $\Delta^2 T_i$  und  $\Delta_R T_i$  statistisch ausgewertet. Der Vorgang wird in Abbildung 4.1.3 gezeigt. Mit dem Index  $i$  ( $1 \leq i \leq M$ ) werden jeweils die Ergebnisse der  $i$ -ten Messung bezeichnet.

<i>Schleife:</i>	
	1. <i>Wähle zufällig die Position:</i> $x_i = \text{rnd}(r, X_{\max} - r) , y_i = \text{rnd}(r, Y_{\max} - r) .$
	2. <i>Messe den Kommunikationsverkehr <math>T_{Di}</math> und <math>T_{Mi}</math></i>
	3. <i>Berechne die Fehlermaße <math>\Delta^2 T_i</math> und <math>\Delta_R T_i</math></i>
<i>Über die Anzahl von Tests <math>M</math></i>	

Abbildung 4.1.3: Algorithmus der Berechnung einer Meßreihe

Als Ergebnis lassen sich folgende Größen der Verteilung der beiden Fehlermaße bestimmen:

- die maximalen Abweichungen  $\Delta^2 T_{\max}$  und  $\Delta_R T_{\max}$ :

$$\Delta^2 T_{\max} \geq \Delta^2 T_i, \forall (1 \leq i \leq M) ,$$



$$\Delta_R T_{max} \geq \Delta_R T_i, \forall (1 \leq i \leq M) .$$

- die minimalen Abweichungen  $\Delta^2 T_{min}$  und  $\Delta_R T_{min}$ :

$$\Delta^2 T_{min} \leq \Delta^2 T_i, \forall (1 \leq i \leq M) ,$$

$$\Delta_R T_{min} \leq \Delta_R T_i, \forall (1 \leq i \leq M) .$$

- die mittleren Abweichungen  $\Delta^2 T_{mean}$  und  $\Delta_R T_{mean}$ :

$$\Delta^2 T_{mean} = \frac{1}{M} \sum_{i=1}^M \Delta^2 T_i ,$$

$$\Delta_R T_{mean} = \frac{1}{M} \sum_{i=1}^M \Delta_R T_i .$$

- die Variationskoeffizienten  $\Delta^2 T_{varco}$  und  $\Delta_R T_{varco}$ :

$$\Delta^2 T_{varco} = \frac{\sqrt{\Delta^2 T_{var}}}{\Delta^2 T_{mean}} , \quad \Delta^2 T_{var} = \frac{1}{M-1} \sum_{i=1}^M (\Delta^2 T_i)^2 - (\Delta^2 T_{mean})^2 ,$$

$$\Delta_R T_{varco} = \frac{\sqrt{\Delta_R T_{var}}}{\Delta_R T_{mean}} , \quad \Delta_R T_{var} = \frac{1}{M-1} \sum_{i=1}^M (\Delta_R T_i)^2 - (\Delta_R T_{mean})^2 .$$

Generell werden alle Größen für jeden Testlauf gemessen, wobei im folgenden nur eine Auswahl der Ergebnisse gezeigt werden kann. Diese Werte erlauben eine quantitative Aussage über die Güte der verschiedenen Verfahren. Hierbei sind vor allem die maximalen und mittleren Abweichungen interessant.

Obige Meßwerte liefern Ergebnisse bei einer bestimmten Größe des Testkreises bzw. -quadrats, die durch den Radius  $r$  bestimmt wird. In der Praxis sind die Versorgungsgebiete der Sender unterschiedlich groß. Der Radius erstreckt sich von weniger als einem Kilometer bei Mikrozellen bis zu zehn oder mehr Kilometern in dünn besiedelten, ebenen Regionen. Daher ist es äußerst wichtig, die Approximationsgüte des Kommunikationsverkehrs abhängig von der Größe des Versorgungsgebietes, also hier des Kreisradius, zu untersuchen.

#### 4.1.4 Meßergebnisse abhängig vom Radius

Um die verschiedenen Verfahren zur Clusterung direkt vergleichen zu können wird die Approximationsgüte in Abhängigkeit von der Größe der Testgebiete dargestellt. Für die hier verwendeten Testkreise wird sie als Kurve über den Radius aufgetragen. Dazu erstellt man für einen bestimmten Radius die Meßreihen und bestimmt die Fehlergrößen, d.h. die oben genannten Abweichungen. Der Radius wird dabei systematisch verändert und jeweils die zugehörige Meßreihe bestimmt. Damit lassen sich die Fehlergrößen in Abhängigkeit vom Radius  $r$  errechnen. Für die quadratische Abweichung ergeben sich die Funktionen  $\Delta^2 T_{max}(r)$ ,  $\Delta^2 T_{min}(r)$ ,  $\Delta^2 T_{mean}(r)$  und  $\Delta^2 T_{varco}(r)$ . Für die relative Abweichung werden

entsprechende Funktionen erzeugt. Anhand der in Abhängigkeit vom Radius  $r$  gezeichneten Kurven können verschiedene Demand-Node Verteilungen verglichen werden. Dabei kann auch die Eignung bei unterschiedlicher Größe der Versorgungsgebiete eines Senders untersucht werden.

#### **4.1.5 Kriterien für geeignete Parameter der Meßkurven**

Für eine statistisch signifikante Aussage müssen genügend Einzelmessungen durchgeführt werden. Die Anzahl der nötigen Messungen hängt dabei vor allem vom Radius der Testkreise ab. Zum einem verringert sich bei zunehmendem Radius die Anzahl der möglichen Positionen der Testflächen, wie in Abbildung 4.1.1 dargestellt wird. Zum anderen nimmt die Varianz der Fehlermaße mit wachsendem Radius ab, d.h. bei größeren Radien sind weniger Messungen nötig, um die gleiche Aussagegenauigkeit zu erreichen. Die Anzahl der möglichen Positionen und die Varianz hängen linear von der Fläche der Testkreise, also quadratisch von deren Radius  $r$  ab.

Der mögliche Bereich für die Radien der Testkreise hängt von der Größe des Gebiets und der Granularität der Verkehrsmatrix ab. Der minimale Radius der Testkreise sollte deutlich größer sein als die Auflösung der Verkehrsmatrix, die in dieser Arbeit immer  $100\text{ m}$  beträgt. Da die Testkreise komplett im Gebiet liegen müssen, darf der maximale Durchmesser nicht größer als das Minimum von Höhe und Breite des Planungsgebietes sein. Um Randeffekte zu vermeiden, sollte der maximale Radius etwas kleiner als die halbe minimale Ausdehnung sein.

Bei der Berechnung einer kompletten Meßreihe sollten die Radien nicht mit konstanter Schrittweite erhöht werden. Es könnten dadurch Interferenzen der konstanten Schrittweite mit einem eventuell vorhandenen regelmäßigen Muster der Demand-Nodes entstehen. Statt dessen wird der Radius um einen zufälligen Wert vergrößert. Bei größeren Radien sind die Unterschiede der Fehlermaße zwischen einzelnen Meßreihen mit ähnlichem Radius gering. Aus diesem Grund kann die durchschnittliche Schrittweite mit wachsendem Radius erhöht werden, ohne die Genauigkeit der kompletten Meßkurve zu verringern.

#### **4.1.6 Parameter der Meßkurven**

Für den minimalen Radius wurde der Wert  $200\text{ m}$  in allen Tests genommen, wobei wegen der Auflösung der Verkehrsmatrix von  $100\text{ m}$  die Ergebnisse der Testkreise in der Nähe des minimalen Radius nicht sehr genau sind.

Durch die Ausdehnung des Würzburger, synthetischen und konstanten Gebiets von  $16.1\text{ km}$  auf  $15.1\text{ km}$  ist der maximale Radius in diesem Fall auf  $7.5\text{ km}$  beschränkt. Um Randeffekte zu vermeiden wird er auf  $5\text{ km}$  festgelegt. Das Dallas Gebiet hat eine Ausdehnung von  $192.4\text{ km}$  auf  $192.4\text{ km}$ , wobei hier der maxi-

male Radius auf 8.5 km beschränkt wird. Dieser Wert ist sinnvoll, da auch die Kapazität und Reichweite der Sender beschränkt sind.

Wie oben beschrieben, hängt die Anzahl  $M$  der nötigen Messungen vom Radius  $r$  ab. Für die Meßreihen wird folgende Beziehung zur Bestimmung von  $M(r)$  verwendet:

$$M(r) = c \cdot \left( \frac{2r}{r_{max} + r_{min}} \right)^{-c_{exp}} .$$

Durch die Konstanten  $c$  und  $c_{exp}$  wird die Anzahl der Tests festgelegt, wobei die Konstante  $c_{exp}$  die Nichtlinearität der Funktion bestimmt.

Zur Abschätzung der Anzahl der Messungen wird die Verkehrsmatrix betrachtet, die im Planungsgebiet von 16.1 km auf 15.1 km bei einem Raster von 100 m genau  $161 \cdot 151 = 24\,311$  Einträge enthält. Bei den Tests mit minimalem Radius wird im Schnitt nur ein Matrixelement überdeckt. Dabei können sich maximal so viele verschiedene Ergebnisse ergeben, wie Flächenelemente in der Verkehrsmatrix vorhanden sind. Deshalb kann der Parameter  $c$  so abgeschätzt werden, daß bei minimalem Radius nicht mehr Tests nötig sind, als die Verkehrsmatrix Einträge enthält.

Die Schrittweite bei der Erhöhung der Radien darf, wie oben erklärt, nicht konstant sein. Deshalb wird eine zufällige Schrittweite verwendet, die mit wachsendem Radius größer wird:

$$\Delta r = \text{rnd}(0, g) \cdot \left( \frac{2r}{r_{max} + r_{min}} \right)^{g_{exp}} .$$

Hierbei liefert die Funktion  $\text{rnd}(a, b)$  einen zufälligen Wert im Bereich  $[a, b]$ , und die Konstanten  $g$  und  $g_{exp}$  legen die Schrittweite des Radiuszuwachses fest.

	<b>Dallas Gebiet</b>	<b>übrige Gebiete</b>
<i>Minimaler Radius</i>	200 m	200 m
<i>Maximaler Radius</i>	8.500 m	5.000 m
<i>Konstante c</i>	1.800	2.500
<i>Konstante c<sub>exp</sub></i>	0.8	0.8
<i>Konstante g</i>	20 m	20 m
<i>Konstante g<sub>exp</sub></i>	0.8	0.8

*Tabelle 4.1.1: Parameter für die Meßkurven*

In Tabelle 4.1.1 werden die in den folgenden Meßkurven verwendeten Parameter aufgelistet. Dabei ergibt sich für das Dallas Gebiet eine Meßkurve mit 1300 Testreihen verschiedener Radien, wobei beim minimalen Radius ungefähr 21 000 und beim maximalen Radius etwa 1000 Einzeltests durchgeführt werden. Insgesamt werden beim Dallas Gebiet für eine Meßkurve 7.3 Mio. einzelne Testkreise gemessen. Für die restlichen Gebiete werden von 20 000 beim minimalen, bis 1500 Tests beim maximalen Radius durchgeführt. Die Meßkurve besteht dort aus etwa 700 Testreihen, wofür insgesamt 4.4 Mio. einzelne Testkreise geprüft werden. Für alle Positionen eines Testgebiets muß für jedes einzelne Flächenelement bestimmt werden, ob es innerhalb des Testgebiets liegt. Dazu müssen im Dallas Gebiet  $1924 \cdot 1924 \cdot 7\,300\,000 \approx 2.7 \cdot 10^{13}$  Prüfungen durchgeführt werden. In den anderen Gebieten sind es  $1.1 \cdot 10^{11}$ . Selbst auf modernen Prozessoren würde die Berechnung einer einzigen Meßkurve des Dallas Gebiets mehrere Wochen in Anspruch nehmen. Erst durch die starke Optimierung des Testprogramms kann die benötigte Rechenzeit auf ein vertretbares Maß gesenkt werden. Dazu werden die Demand-Nodes gemäß ihrer Lage sortiert, und intelligente Algorithmen zur Bestimmung der Flächenelemente innerhalb eines Testkreises verwendet. Die endgültige Version des Testprogramms benötigt auf dem Prozessor AMD K6<sup>4</sup> mit 233 MHz eine Stunde für die Berechnung des Dallas Gebiets. Bei den übrigen Gebieten beträgt die Rechenzeit etwa 20 Minuten.

## 4.2 Meßergebnisse des Verkehrsvergleichs

Im folgenden wird die Approximationsgüte des Kommunikationsverkehrs der beiden Clusterungsverfahren untersucht. Dazu werden für die Berechnung der Testkurven die Parameter aus Tabelle 4.1.1 verwendet. Bei der Erzeugung der Demand-Nodes mit dem Verfahren des Zusammenfassens von Knoten, im folgenden mit „Zusammenfassen“ bezeichnet, werden die Parameter aus Tabelle 4.2.1 verwendet. Für das Verfahren des rekursiven Partitionierens, im folgenden mit „Partitionieren“ abgekürzt, werden die Parameter aus Tabelle 4.2.2 genommen.

<b>Vergleichsfläche</b>	<b>Max. Verkehr</b>	<b>Max. Fläche</b>	<b>Optimierung</b>
<i>Würzburger Gebiet</i>	0.08	$0.75 \text{ km}^2$	<i>keine</i>
<i>Dallas Gebiet</i>	0.0002	$25 \text{ km}^2$	<i>keine</i>
<i>Synthetisches Gebiet</i>	0.08	$0.75 \text{ km}^2$	<i>keine</i>
<i>Konstantes Gebiet</i>	0.08	$0.75 \text{ km}^2$	<i>keine</i>

*Tabelle 4.2.1: Parameter des Zusammenfassens von Knoten*

<sup>4</sup> Vergleichbar mit einem Pentium® MMX gleicher Taktfrequenz

Das Verfahren des Zusammenfassens von Knoten wird zuerst ohne Optimierung untersucht, da diese Auswirkung auf die Güte der Verteilung der Demand-Nodes haben kann. Die Wirkung der Optimierung wird in Kapitel 4.2.4 untersucht. Als Testfläche wird die Kreisform gewählt, sofern nicht extra eine andere Form angegeben wird.

<b>Vergleichsfläche</b>	<b>Rekursionstiefe</b>
<i>Würzburger Gebiet</i>	10
<i>Dallas Gebiet</i>	15
<i>Synthetisches Gebiet</i>	10
<i>Konstantes Gebiet</i>	10

*Tabelle 4.2.2: Parameter des rekursiven Partitionierens*

#### **4.2.1 Auswahl der Meßkurven**

Zum Vergleich verschiedener Verteilungen der Demand-Nodes sind in Abbildung 4.2.1 die maximalen quadratischen Fehler der beiden Algorithmen aufgetragen. Der maximale Fehler nimmt mit wachsendem Radius  $r$  zu. Auffällig ist, daß der Fehler beim Verfahren des Zusammenfassens nur sehr langsam wächst, während er beim Verfahren des Partitionierens steil ansteigt. Der mittlere relative Fehler in Abbildung 4.2.2 nimmt mit steigendem Radius ab, was nach dem Gesetz der großen Zahlen auch zu erwarten ist.

Der maximale Fehler schwankt stark mit dem Radius  $r$ . Er hängt im wesentlichen von der zufälligen Position der Testfläche ab und wird für jeden Wert des Radius neu bestimmt. Ein hoher Wert für den maximalen Fehler tritt immer dann auf, wenn die Testfläche in einem Gebiet mit sehr geringer Verkehrsdichte liegt und dabei zusätzlich Demand-Nodes von benachbarten Regionen mit hohem Verkehrsaufkommen überdeckt. Dadurch täuschen die Demand-Nodes einen höheren Kommunikationsverkehr vor, als in der restlichen Testfläche mit niedriger Verkehrsdichte vorhanden ist. In der Praxis wirkt sich dieses Szenario kaum aus, weil an dieser Stelle kein Sender positioniert wäre. Dort könnte ein Sender nur sehr wenig Kommunikationsverkehr versorgen.

Bei der Abschätzung des Kommunikationsverkehrs ist es vor allem wichtig, den Kommunikationsverkehr in einem Gebiet nicht zu niedrig anzusetzen. Eine zu hohe Annahme über den Kommunikationsverkehr ist dabei wesentlich weniger nachteilig.

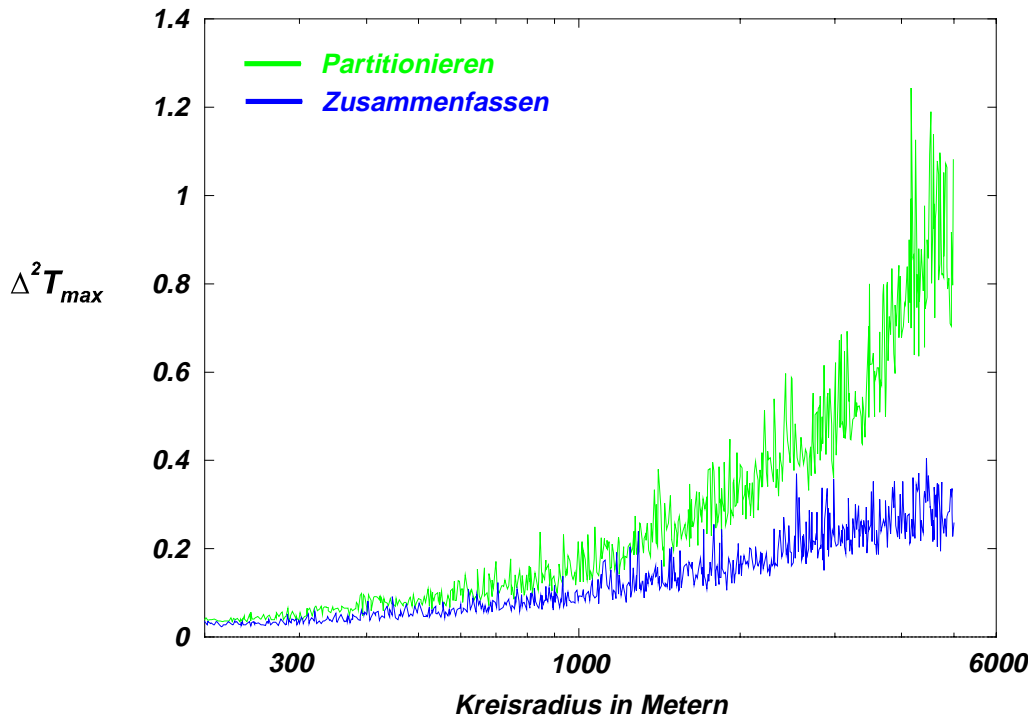


Abbildung 4.2.1: Maximaler quadratischer Fehler im Würzburger Gebiet

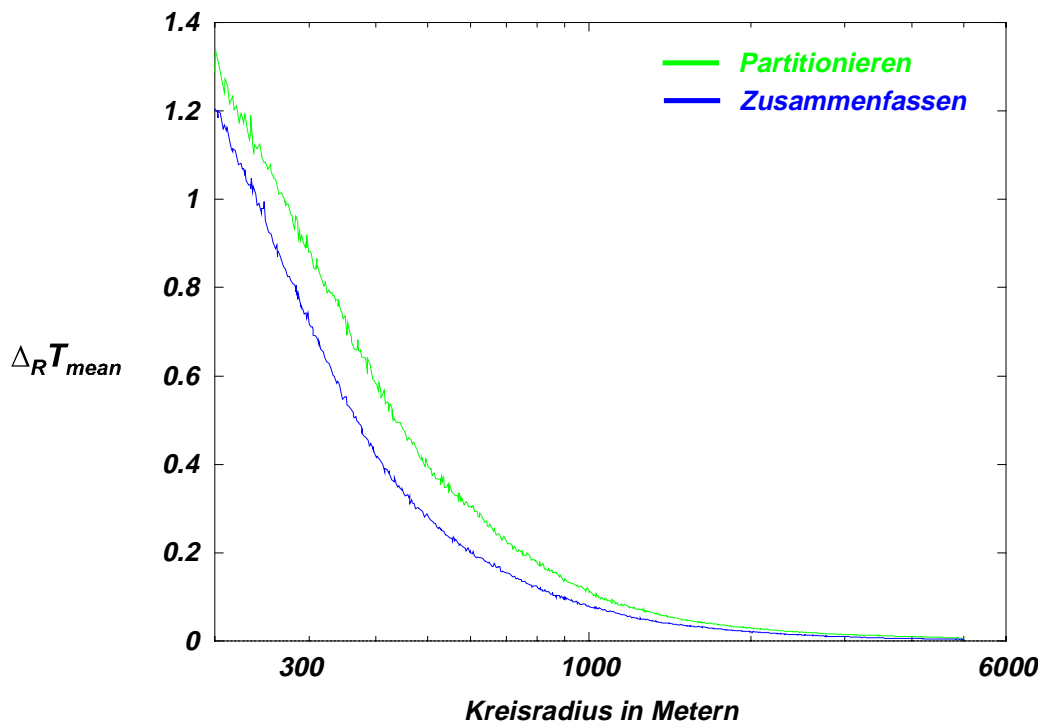


Abbildung 4.2.2: Mittlerer relativer Fehler im Würzburger Gebiet

In den folgenden Kurven wird meist der mittlere absolute Fehler aufgetragen. Bei ihm kann, im Gegensatz zum relativen Fehler, die Approximationsgüte der beiden Verfahren besser verglichen werden. Der relative Fehler kann, wenn die

wirkliche Verkehrsdichte in der Testfläche sehr gering ist, gemäß seiner Definition einen relativ hohen Wert annehmen. Dieser Wert ist aber nicht aussagekräftig, da für die Planung nur Versorgungsgebiete interessant sind, die wenigstens eine Mindesthöhe an Kommunikationsverkehr enthalten.

## 4.2.2 Vergleich beider Verfahren

In Abbildung 4.2.3 werden die mittleren quadratischen Fehler beider Verfahren dargestellt. Deutlich ist der starke Anstieg beim Partitionieren zu sehen. Die Fehlerkurve des Zusammenfassens liegt im ganzen Bereich unterhalb der des Partitionierens und steigt auch nur geringfügig an.

Der große Unterschied in den Fehlerkurven wird aber nicht durch die Flächenbegrenzung beim Zusammenfassen hervorgerufen, sondern liegt in der Arbeitsweise der beiden Algorithmen begründet. Um dies zu verifizieren, wird auch eine Kurve des Zusammenfassens mit abgeschalteter Flächenbegrenzung gezeigt. Diese ist nahezu identisch zur Kurve mit auf  $0.75 \text{ km}^2$  begrenzter Fläche der Demand-Nodes. Da die Darstellung des Verfahrens ohne Flächenbegrenzung kein signifikant anderes Ergebnis liefert, wird in den nächsten Kurven auf seine Abbildung verzichtet.

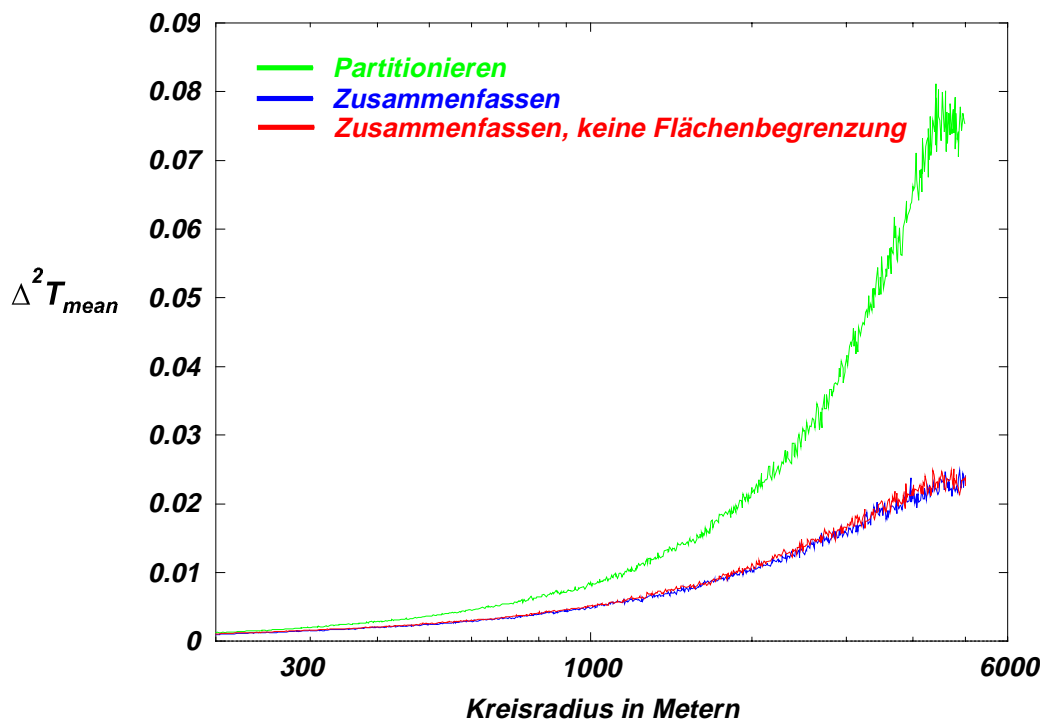


Abbildung 4.2.3: Vergleich beider Verfahren im Würzburger Gebiet

In der Abbildung 4.2.4 sind die Ergebnisse für das Dallas Gebiet aufgetragen, wobei hier das Anwachsen des Fehlers beim Partitionieren noch wesentlich

ausgeprägter auftritt. Bei einem Radius von 9 km für den Testkreis ist der Fehler des Partitionierens um den Faktor acht größer.

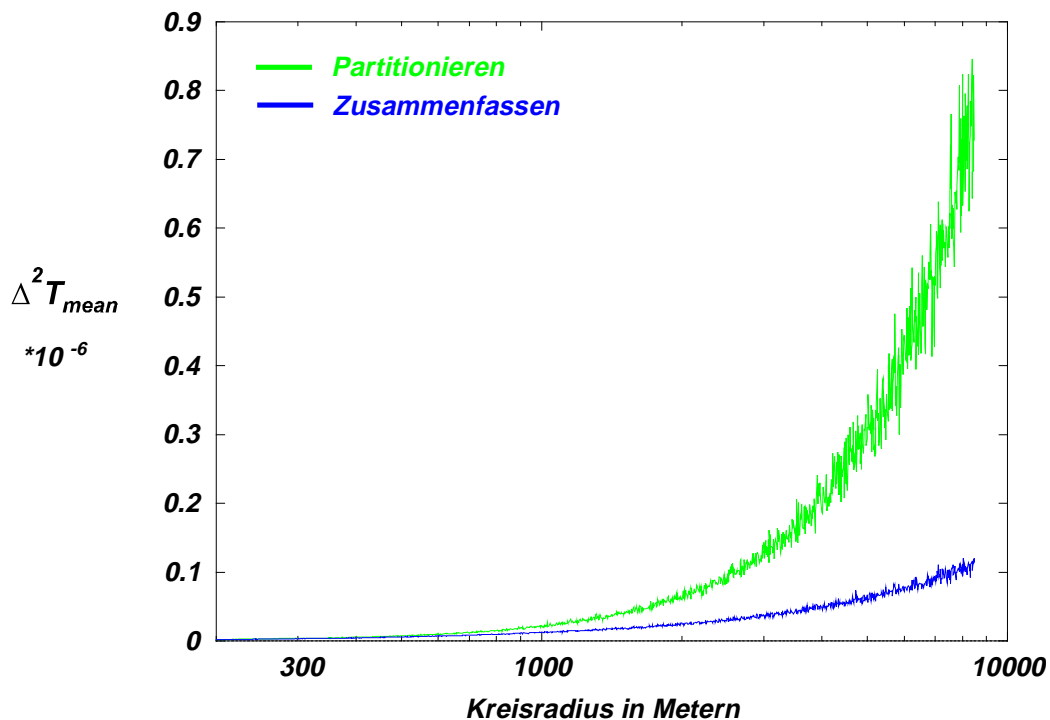


Abbildung 4.2.4: Vergleich beider Verfahren im Dallas Gebiet

Das konstante Gebiet in Abbildung 4.2.5 liefert neue Erkenntnisse. Die beiden Kurven liegen ungefähr auf gleicher Höhe, wobei die Kurve des Partitionierens stark oszilliert. Die Wellenlänge  $\nu$  der Schwingung ergibt sich durch den konstanten Abstand  $d$  der einzelnen Demand-Nodes in Abbildung 3.1.9 nach der Formel  $\nu = 2d$ .

Die Ausbildung regelmäßiger Muster der Demand-Nodes durch das Partitionieren stellt ein großes Problem dar. Die Genauigkeit der Approximation hängt in diesem Fall wesentlich von der Form des Versorgungsgebietes ab. Vor allem kann eine minimale Änderung der Größe des Versorgungsgebietes eine große Änderung der Höhe des innerhalb des Gebietes vorhandenen Kommunikationsverkehrs bewirken. Der Kommunikationsverkehr steigt nicht stetig mit Vergrößerung der Fläche des Versorgungsgebietes an. Vielmehr springt er bei der Ausdehnung des Versorgungsgebietes schlagartig auf einen wesentlich höheren Wert. Die Demand-Nodes repräsentieren dadurch innerhalb des Versorgungsgebietes teilweise einen wesentlich zu niedrigen Kommunikationsverkehr. Dies ist der Fall, wenn der Radius der Testfläche gerade etwas kleiner als ein Vielfaches des Abstandes  $d$  der Demand-Nodes ist. Wenn der Radius der Testfläche etwas über diesen Wert erhöht wird, liegen dann gleich mehrere Demand-Nodes zusätzlich innerhalb der Testfläche und repräsentieren einen zu hohen Kommunikationsverkehr.



Dieser Effekt macht eine sinnvolle Planung nahezu unmöglich. Durch die Erzeugung der Verkehrsmatrix bedingt, sind Gebiete mit konstanter Verkehrsdichte kein Sonderfall, sondern nach Kapitel 2.2.4 vielmehr der Regelfall.

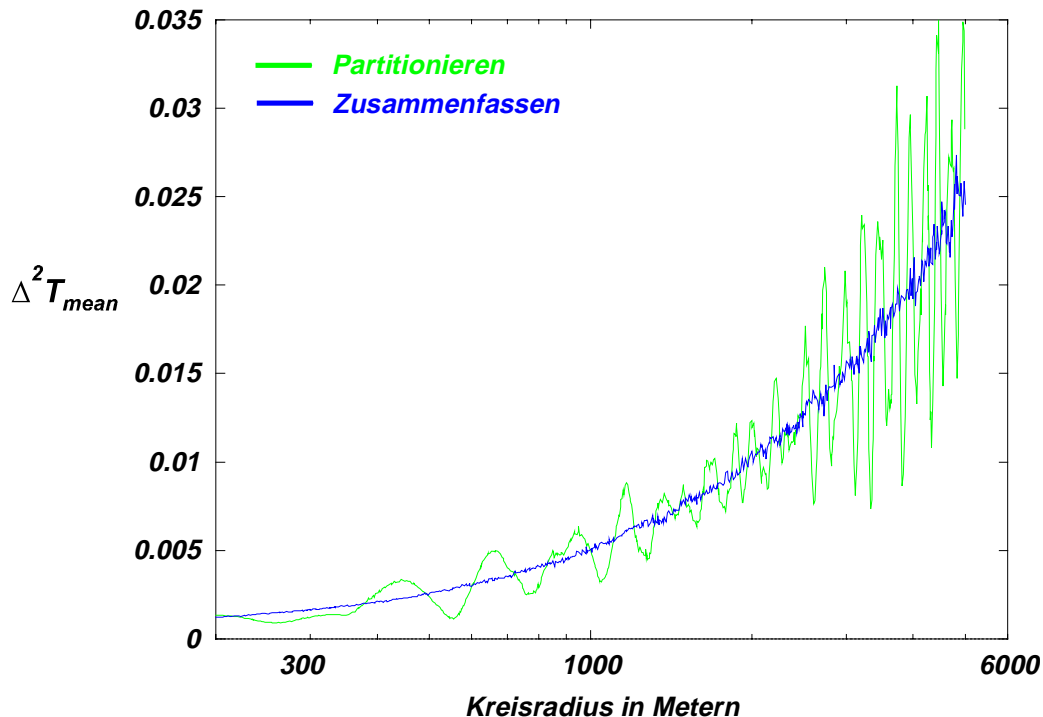


Abbildung 4.2.5 Vergleich beider Verfahren im konstanten Gebiet

### 4.2.3 Ergebnisse mit quadratischer Testfläche

Bei der Untersuchung mit der kreisförmigen Testfläche könnte das Verfahren des Partitionierens benachteiligt sein, da es nur rechteckige Demand-Node Flächen erzeugen kann. Deshalb werden sämtliche Meßreihen auch für eine quadratische Testfläche durchgeführt, wobei zusätzlich die Abhängigkeit der Meßergebnisse von der Form der Testfläche bewertet werden kann.

Die Meßergebnisse des Würzburger und Dallas Gebiets unterscheiden sich dabei kaum vom Fall einer kreisförmigen Testfläche und werden deshalb hier nicht dargestellt.

Ein interessantes Ergebnis bietet sich im synthetischen Gebiet, das in Abbildung 4.2.6 gezeigt wird. Der quadratische Fehler ist allgemein höher als beim Test mit der kreisförmigen Fläche. Vor allem steigt er beim Partitionieren kontinuierlich stark an, wohingegen der Fehler des Zusammenfassens wie im kreisförmigen Fall nur moderat wächst. Auffällig ist noch eine wesentlich stärkere Schwankung bei beiden Kurven.

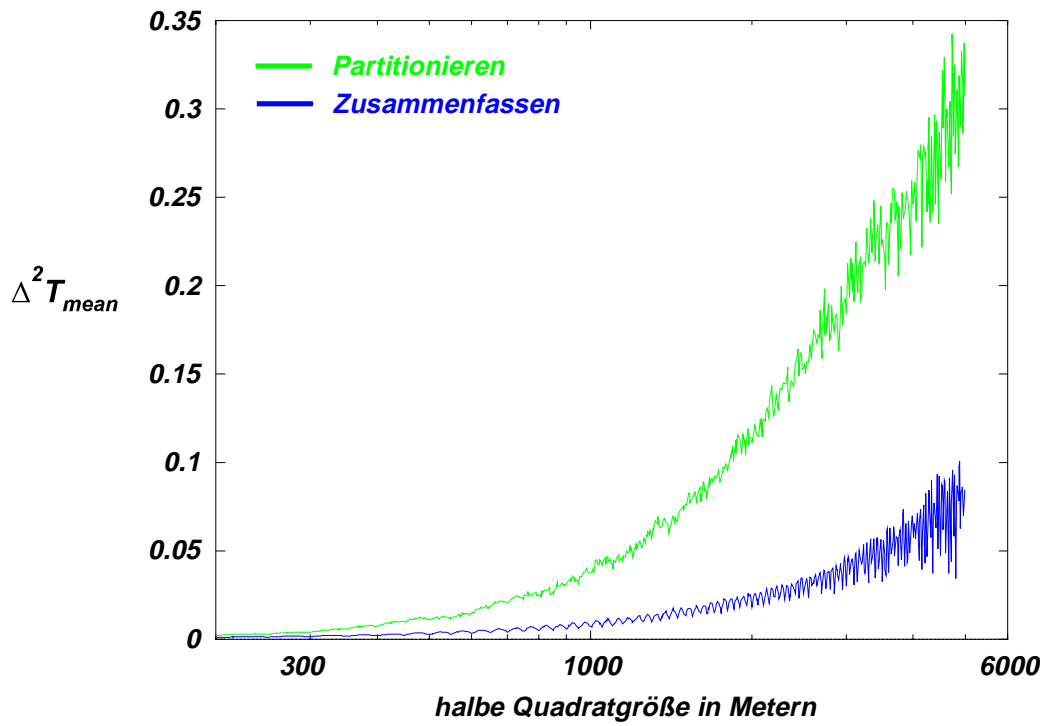


Abbildung 4.2.6: Vergleich mit quadratischer Testfläche im synthetischen Gebiet

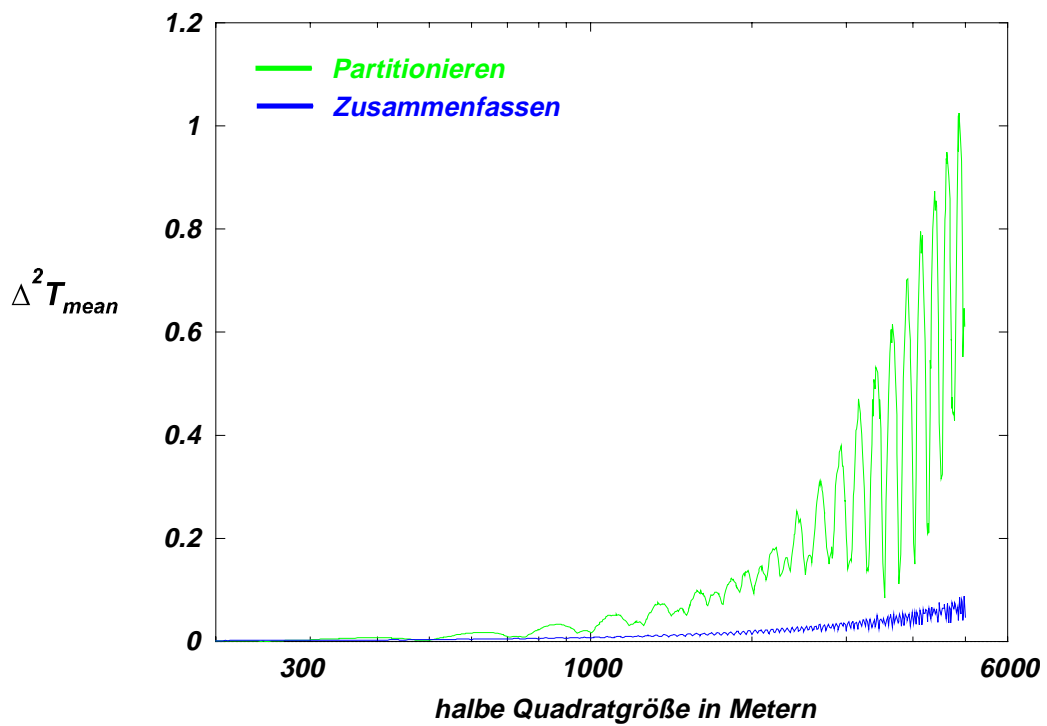


Abbildung 4.2.7: Vergleich mit quadratischer Testfläche im konstanten Gebiet

Im konstanten Gebiet, welches in Abbildung 4.2.7 dargestellt wird, oszilliert die Kurve des Partitionierens wie beim kreisförmigen Fall, wobei die Amplitude aber stark ansteigt. Der quadratische Fehler ist beim Verfahren des Partitionierens bei einer Seitenlänge des Quadrats von 10 km um das 20fache größer als beim Zusammenfassen. Die Kurve des Partitionierens liegt bei größeren Radien weit oberhalb der des Zusammenfassens, was im Fall der kreisförmigen Testfläche nicht zutrifft.

Insgesamt zeigt sich, daß der Test mit einer quadratischen Fläche dem Verfahren des Partitionierens nicht zugute kommt, sondern im Gegenteil die Einschränkungen dieses Verfahrens noch deutlicher in Erscheinung treten. Die Ergebnisse zeigen eine wesentlich schlechtere Approximationsgüte dieses Verfahrens, die durch Interferenz der quadratischen Testfläche mit den rechteckigen Demand-Node Gebieten hervorgerufen wird.

#### 4.2.4 Untersuchung der Optimierung

Die bis jetzt betrachteten Ergebnisse des Verfahrens des Zusammenfassens wurden ohne Optimierung erzeugt, um ihren Einfluß beim Vergleich der beiden Verfahren zu umgehen. Die Verwendung der Optimierung ist beim Verfahren des Zusammenfassens von Knoten zur Verbesserung der Verkehrsverteilung aber nahezu immer nötig. Im folgenden wird daher die Auswirkung der Optimierung auf die Approximationsgüte der Demand-Nodes untersucht.

<b>Kurve</b>	<b>Max. Verkehr</b>	<b>Max. Fläche</b>	<b>Anzahl N der Demand-Nodes</b>
<i>Keine Optimierung, <math>N'</math></i>	0.08	$0.75 \text{ km}^2$	1275
<i>Optimierung II, <math>N &lt; N'</math></i>	0.08	$0.75 \text{ km}^2$	1130
<i>Verbesserte Optimierung, <math>N &lt; N'</math></i>	0.08	$0.75 \text{ km}^2$	1100
<i>Verbesserte Optimierung, <math>N = N'</math></i>	0.07	$0.75 \text{ km}^2$	1269
<i>Verbesserte Optimierung, <math>N &gt; N'</math></i>	0.06	$0.75 \text{ km}^2$	1444

*Tabelle 4.2.3: Parameter der verschiedenen Vergleichskurven*

Dazu werden verschiedene Parametersätze verwendet, die in Tabelle 4.2.3 aufgelistet sind. Die dabei angegebene Anzahl der Demand-Nodes entspricht jeweils einem Durchlauf des Algorithmus im Würzburger Gebiet. Bei den anderen

Gebieten weichen sie leicht ab, wobei die Ergebnisse qualitativ gleich bleiben. Hierbei soll noch einmal in Erinnerung gerufen werden, daß die folgenden Angaben über die Anzahl der Demand-Nodes beim Verfahren des Zusammenfassens exakt nur für den dargestellten Testlauf gelten.

Im Zuge der Optimierung wird die Anzahl der Demand-Nodes verringert, was von sich aus schon die Genauigkeit der Repräsentation des Kommunikationsverkehrs verringert. Um die Demand-Node Verteilungen aussagekräftig vergleichen zu können, muß die Anzahl der Demand-Nodes ungefähr gleich sein. Ansonsten ergibt sich bei den Testläufen mit höherer Anzahl Demand-Nodes schon alleine dadurch eine bessere Approximationsgüte.

Insgesamt werden für die Untersuchung fünf Parametersätze verwendet: Eine Kurve ohne Optimierung, eine mit Optimierung II und eine mit verbesserter Optimierung, wobei die Grundparameter bei allen drei identisch gewählt sind. Zum Vergleich wird eine Kurve mit verbesserter Optimierung bestimmt, bei der die Anzahl der Demand-Nodes ungefähr der des Verfahrens ohne Optimierung entspricht. Schließlich wird noch eine Kurve mit verbesserter Optimierung mit einer erhöhten Anzahl der Demand-Nodes ermittelt. Für die Testläufe mit Optimierung II werden die in Tabelle 4.2.4 gezeigten Parameter verwendet.

<b>Parameter</b>	<b>Wert</b>
$T_{select\_min}$	0.7
$T_{select\_max}$	0.9
$T_{min}$	1.1
$R_{Tmax}$	0.6

**Tabelle 4.2.4: Parameter der Optimierung II**

Zur Untersuchung werden die fünf Kurven für alle Vergleichsgebiete berechnet, wobei in Abbildung 4.2.8 das Würzburger Gebiet und in Abbildung 4.2.9 das Dallas Gebiet gezeigt werden. In Abbildung 4.2.10 wird das synthetische Gebiet und in Abbildung 4.2.11 das konstante Gebiet dargestellt.

Aus allen vier Gebieten lassen sich die selben Erkenntnisse gewinnen. Die Ergebnisse mit Optimierung II und mit verbesserter Optimierung sind schlechter als die ohne Optimierung, da sich die Anzahl der Demand-Nodes bei der Optimierung verringert. Wie oben erläutert, führt dies alleine schon zu einer schlechteren Approximationsgüte. Die beiden Kurven mit unterschiedlicher Optimierung sind ungefähr auf gleicher Höhe, wobei die Ergebnisse mit Optimierung II tendenziell etwas schlechter sind als die mit verbesserter Optimierung.

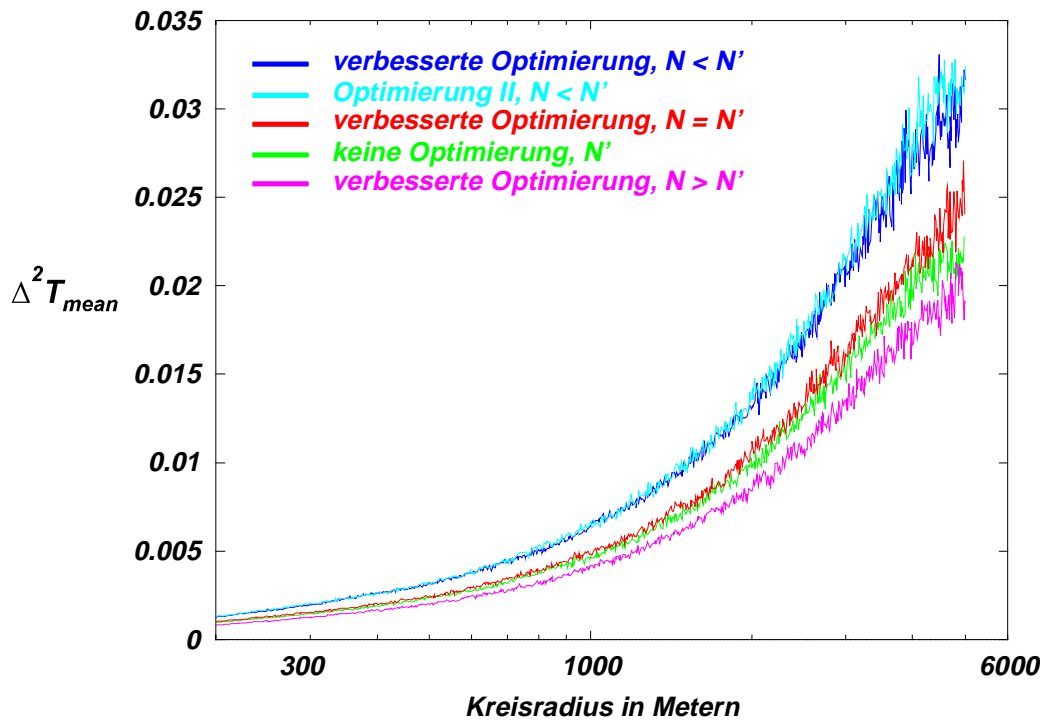


Abbildung 4.2.8: Vergleich verschiedener Parameter im Würzburger Gebiet

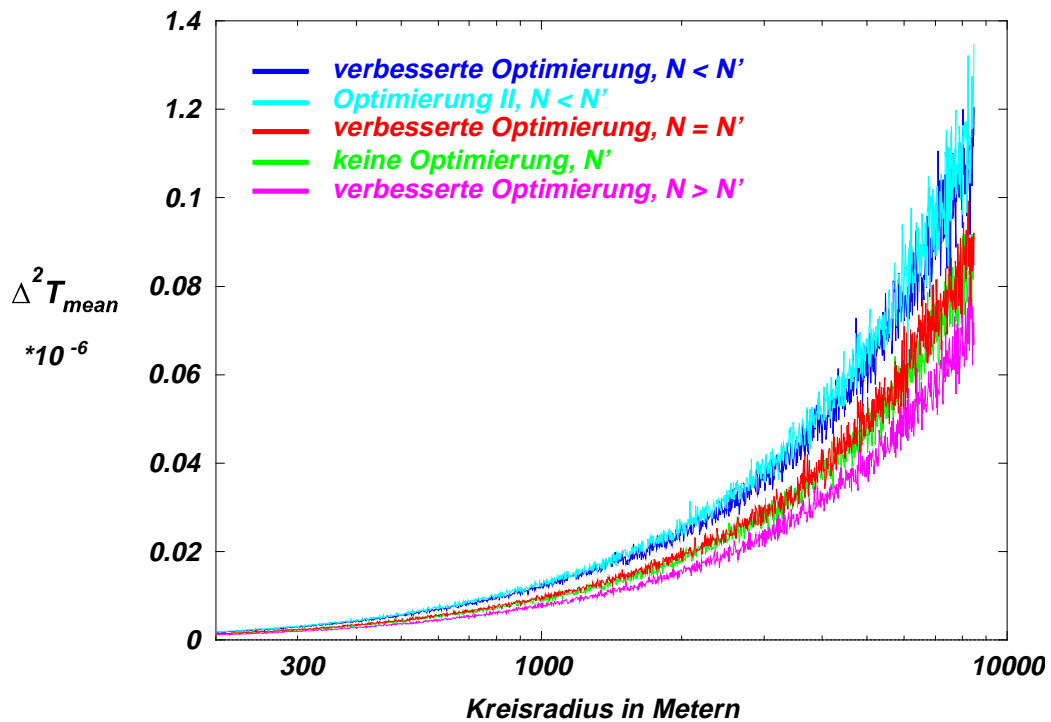


Abbildung 4.2.9: Vergleich verschiedener Parameter im Dallas Gebiet

Die Kurve mit verbesserter Optimierung, bei der die Anzahl der Demand-Nodes mit der ohne Optimierung übereinstimmt, ergibt ungefähr den gleichen Fehler wie die Kurve ohne Optimierung. Bei identischem Planungsaufwand, d.h. über-

einstimmender Anzahl der Demand-Nodes, ist der Fehler beider Verfahren gleich, wobei beim Verfahren mit Optimierung die Verkehrsverteilung der Demand-Nodes optimal ist.

Um mit Optimierung die gleiche Approximationsgüte zu erreichen wie ohne, reicht es aus, wenn die Anzahl der Demand-Nodes ungefähr übereinstimmt. Dazu verringert man den maximalen Kommunikationsverkehr  $T_{max}$  leicht, was die Anzahl der Demand-Nodes vor der Optimierung erhöht. Durch den Optimierungsschritt ergibt sich dann am Ende ungefähr die gleiche Anzahl an Demand-Nodes wie ohne Optimierung. Ein kleinerer Wert von  $T_{max}$  ist auch bei der Planung hilfreich, da die Auflösung der Demand-Nodes steigt, wenn der von ihnen repräsentierte Kommunikationsverkehr geringer ist.

Insgesamt zeigt sich, daß durch die Optimierung die Approximationsgüte des gesamten Verfahrens nicht abnimmt, obwohl die Verkehrsverteilung der Demand-Nodes wesentlich verbessert wird. Dadurch wird auch die Korrektheit und Güte der Optimierung verifiziert. Die Parameter müssen dabei so gewählt werden, daß die Anzahl der Demand-Nodes ungefähr mit der Anzahl ohne Optimierung übereinstimmt.

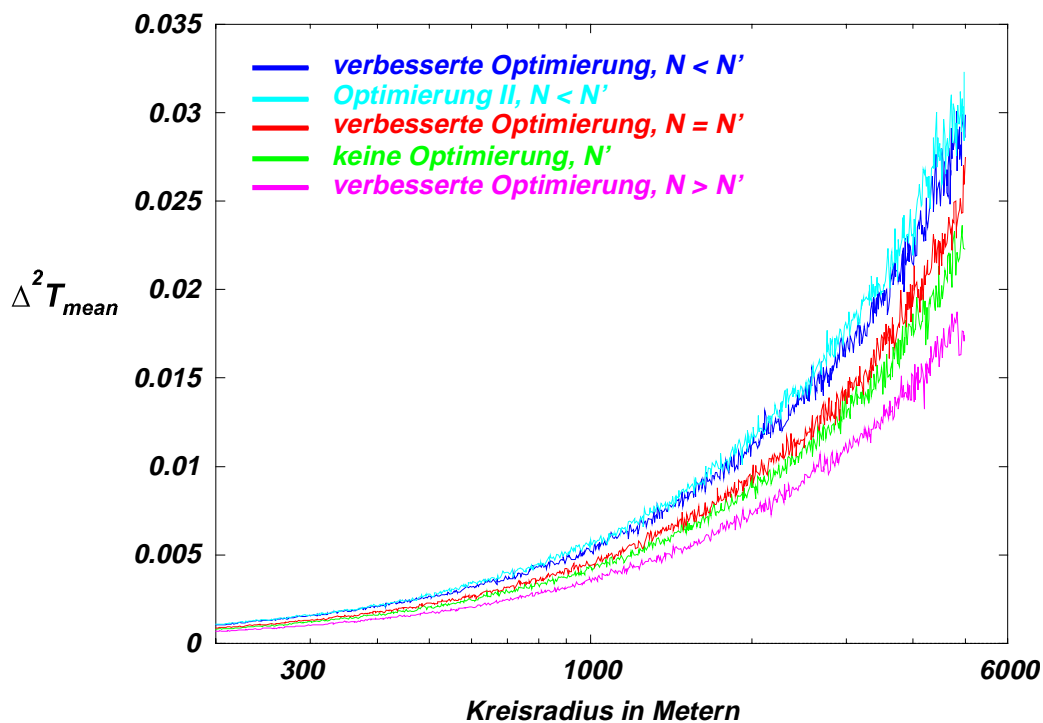


Abbildung 4.2.10: Vergleich verschiedener Parameter im synthetischen Gebiet

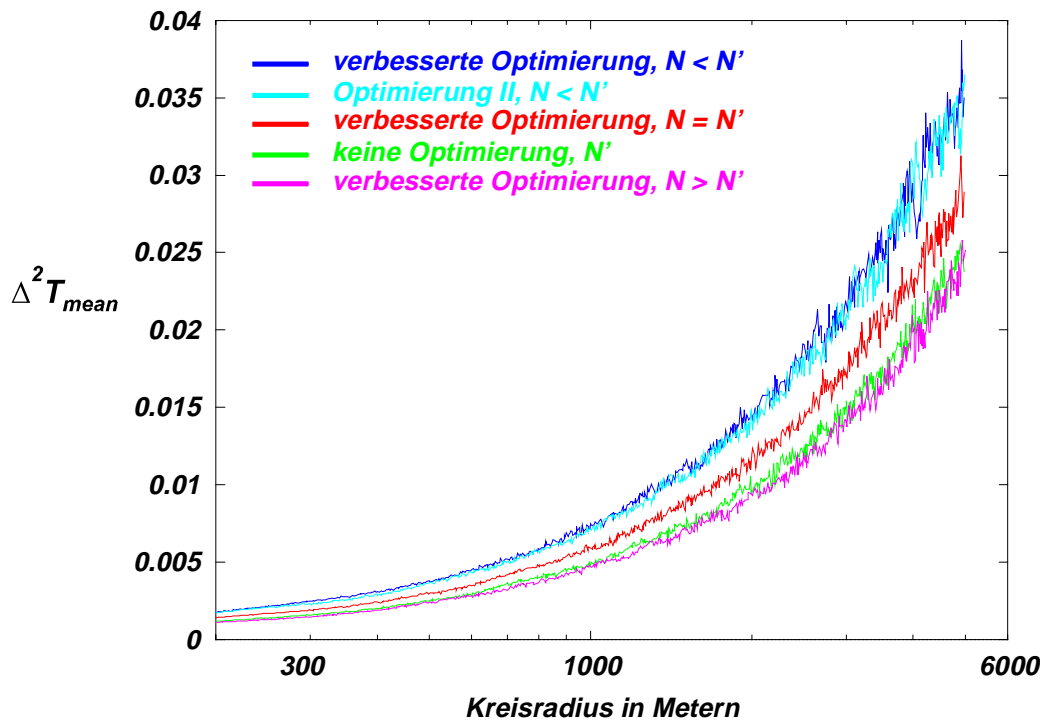


Abbildung 4.2.11: Vergleich verschiedener Parameter im konstanten Gebiet

#### 4.2.5 Vergleich verschiedener Anzahlen an Demand-Nodes

Die Anzahl der Demand-Nodes ist das entscheidende Kriterium für die Approximationsgüte der Verkehrsverteilung. Beim Verfahren des Zusammenfassens ist sie frei bestimmbar durch die Wahl des maximalen Kommunikationsverkehrs und der maximalen Fläche eines Demand-Nodes. In Abbildung 4.2.12 werden die Fehlerkurven für verschiedene Anzahlen von Demand-Nodes aufgetragen. Wie zu erwarten, nimmt die Approximationsgüte mit der Zahl der Demand-Nodes zu, wobei der Zuwachs der Genauigkeit im gesamten Bereich der Testradien gleich ist.

Um eine bestimmte Approximationsgüte der Demand-Node Verteilung zu bekommen, kann eine entsprechende Anzahl von Demand-Nodes erzeugt werden, wobei die Zahl der benötigten Demand-Nodes vom Planungsgebiet abhängt. Maximal können so viele Demand-Nodes erstellt werden, wie die Verkehrsmatrix Flächenelemente enthält. Dann hat das hier verwendete Fehlermaß konstant den Wert 0.

Beim Verfahren des rekursiven Partitionierens sollte der Fehler in ähnlicher Weise mit der Anzahl der Demand-Nodes abnehmen, wobei die Anzahl von Demand-Nodes hier immer eine Zweierpotenz sein muß. In Abbildung 4.2.13 sind die Meßergebnisse beim Partitionieren aufgetragen.

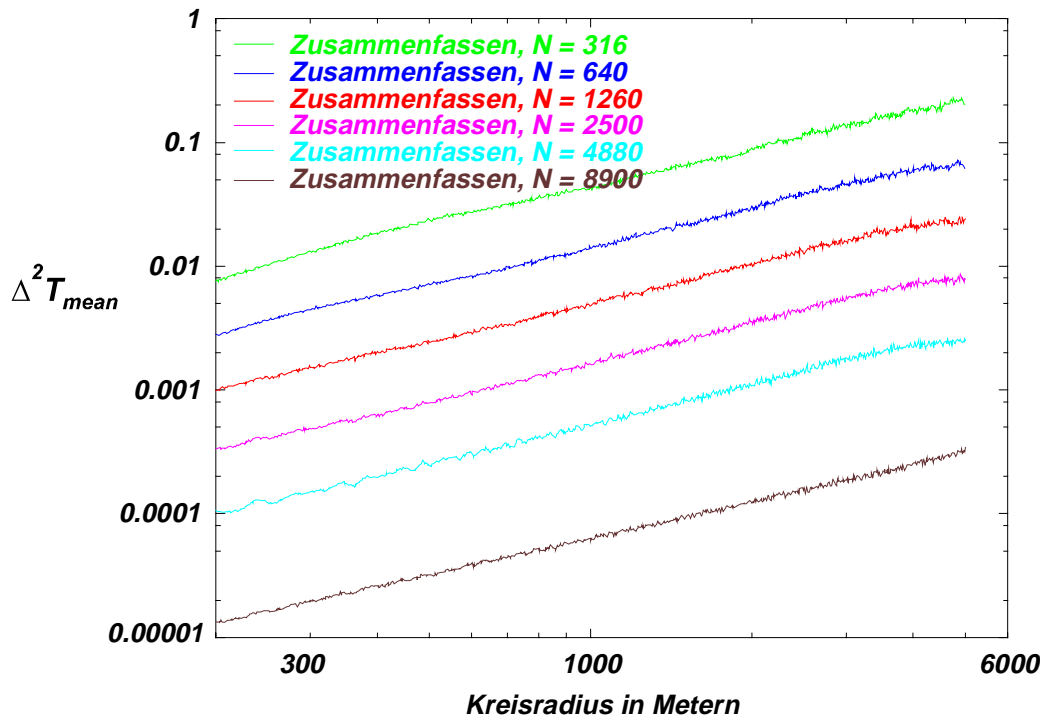


Abbildung 4.2.12: Vergleich der Demand-Node Anzahl beim Zusammenfassen

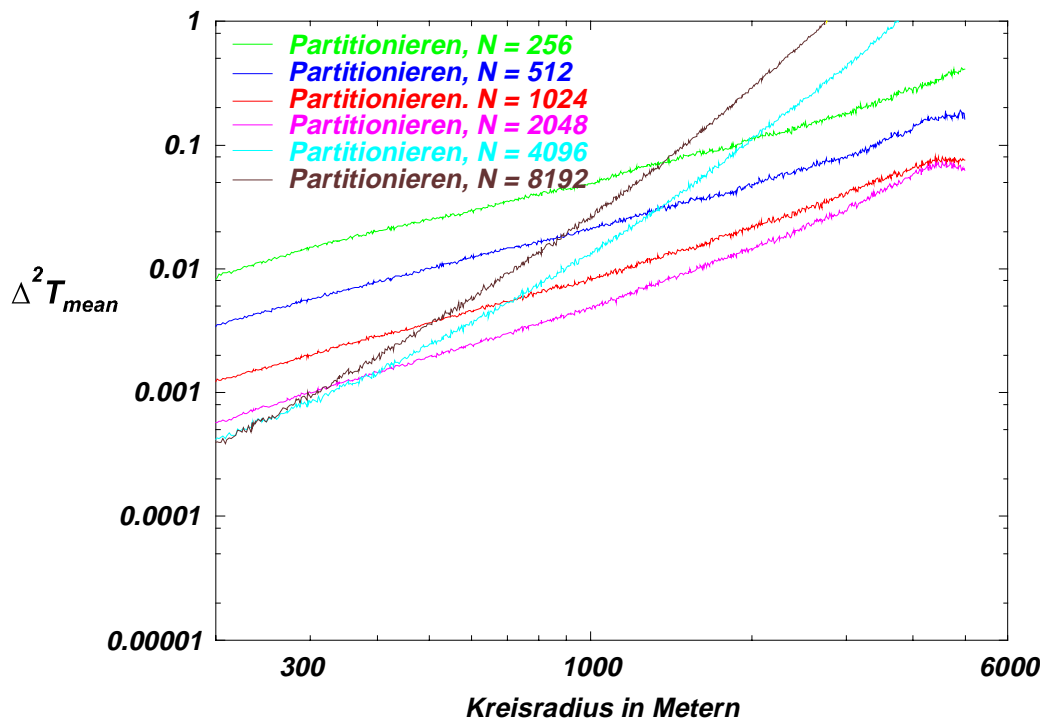


Abbildung 4.2.13: Vergleich der Demand-Node Anzahl beim Partitionieren

Hierbei ergibt sich ein unerwartetes Ergebnis. Bis zur Zahl von 1024 Demand-Nodes nimmt der Fehler mit wachsender Anzahl ab. Bei der Kurve mit 2048 Demand-Nodes ist der Fehler nur noch wenig kleiner und nähert sich bei höheren



Radien der Kurve mit 1024 Demand-Nodes an. Die Fehlerkurven mit 4096 und 8192 Demand-Nodes zeigen ein vollkommen anomales Verhalten, denn sie steigen mit wachsendem Radius steil an. Der absolute quadratische Fehler wächst um einen Faktor von mehr als  $10^4$  an und übertrifft bei höheren Radien die Fehlerkurven aller anderen Demand-Node Verteilungen.

Der Grund für diesen extremen Anstieg des Fehlers liegt im Algorithmus selbst begründet. Durch das anteilige Aufteilen von Matrixelementen an benachbarte Demand-Nodes erzeugt er einen identischen Kommunikationsverkehr für alle Demand-Nodes. Bei einer hohen Anzahl der Demand-Nodes kann der von einem Demand-Node repräsentierte Kommunikationsverkehr kleiner sein als der Verkehr eines Flächenelements. Einem Demand-Node müßte in diesem Fall nur ein Teil eines atomaren Flächenelements zugeordnet werden. Dies ist nicht möglich, da keine höhere Auflösung als die der Flächenelemente zur Verfügung steht. Der Algorithmus erzeugt dann mehrere Demand-Nodes an derselben Stelle, die insgesamt einen höheren Kommunikationsverkehr darstellen als wirklich vorhanden ist. In der Dokumentation zum Algorithmus [3] wird deshalb auch gefordert, daß ein Demand-Node aus mindestens vier Flächenelementen besteht. Diese Bedingung wird hier nur bis zu einer Anzahl von 2048 Demand-Nodes erfüllt.

In Abbildung 4.2.14 ist der relative Fehler aufgetragen, der bei den beiden Kurven auf ein bestimmtes Minimum fällt und mit wachsendem Radius nicht weiter abnimmt, was auf den gerade beschriebenen systematischen Fehler zurückzuführen ist.

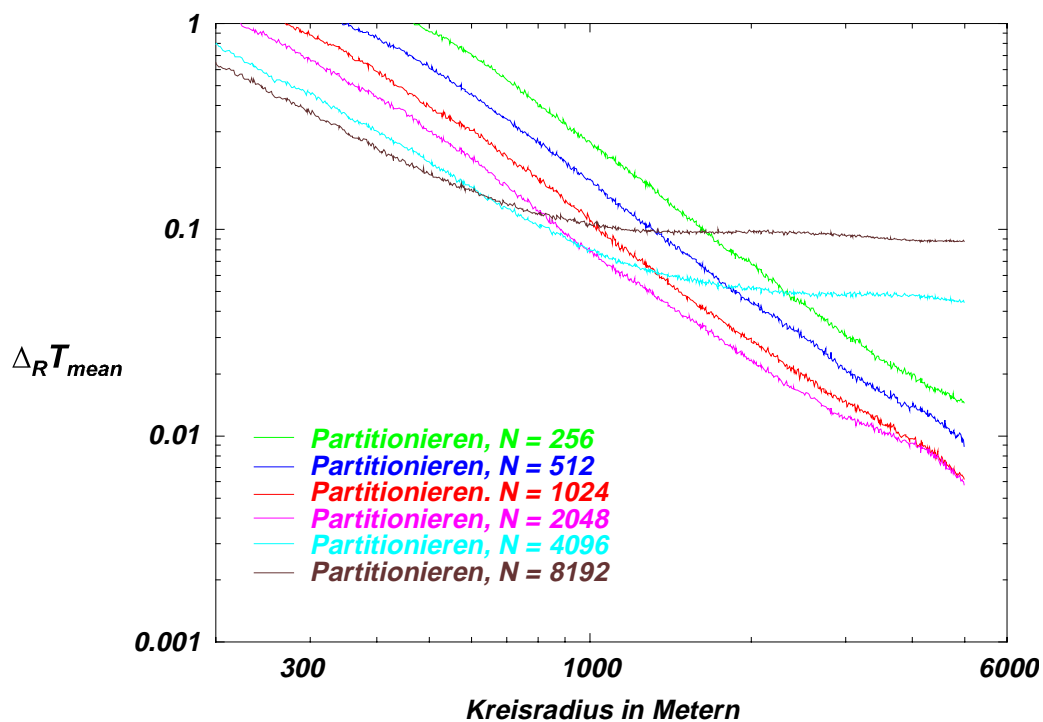


Abbildung 4.2.14: Vergleich der Demand-Node Anzahl beim Partitionieren

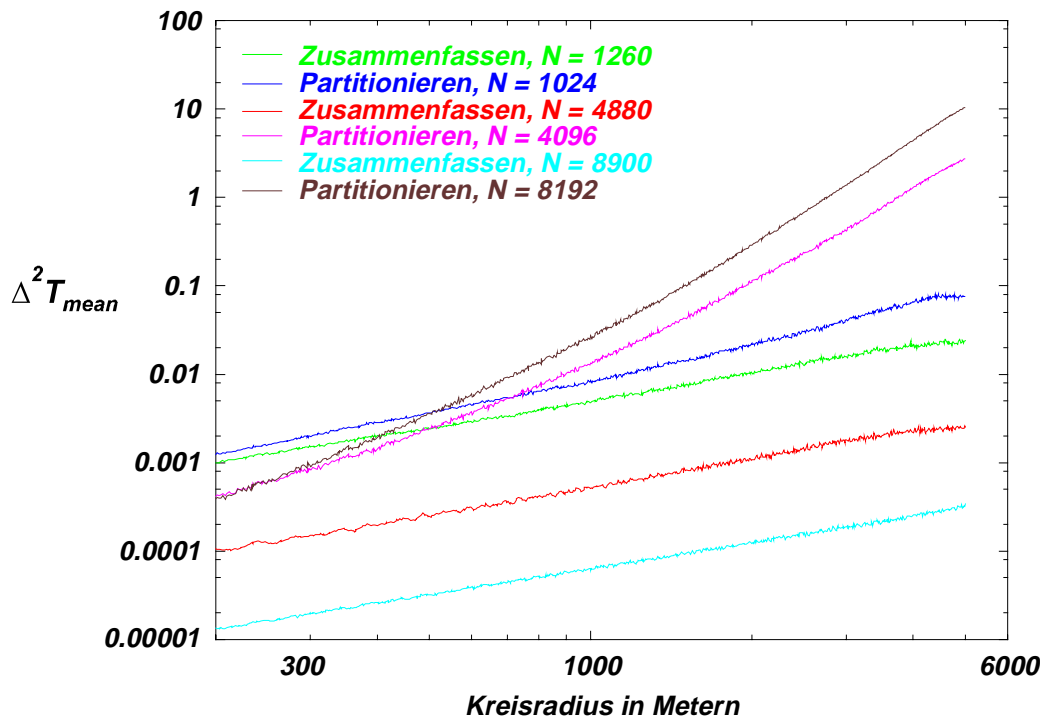


Abbildung 4.2.15: Vergleich der Demand-Node Anzahl beider Verfahren

Insgesamt zeigt sich, daß die Bedingung von mindestens vier Flächenelementen pro Demand-Node die Approximationsgüte des Partitionierens beschränkt. Die Approximationsgüte erreicht dann ein Maximum, welches auch durch eine Erhöhung der Zahl der Demand-Nodes nicht überschritten werden kann. Da zusätzlich nur die Anzahl der Demand-Nodes eine Zweierpotenz sein muß, ergibt sich eine starke Restriktion dieses Verfahrens. Im Falle des Würzburger Gebiets z.B., kann beim rekursiven Partitionieren keine höhere Approximationsgüte erreicht werden, als die der Verteilung mit 2048 Demand-Nodes.

In Abbildung 4.2.15 werden die Fehlerkurven beider Verfahren verglichen. Beim Partitionieren kann keine höhere Genauigkeit erreicht werden als die der gerade beschriebenen Kurve. Beim Zusammenfassen von Knoten kann die Approximationsgüte der Verkehrsverteilung durch Erhöhen der Anzahl der Demand-Nodes bis zur Übereinstimmung mit der Verkehrsmatrix erhöht werden.

#### 4.2.6 Untersuchung der Auswirkung der Flächenbegrenzung

Eine nur in dem Verfahren des Zusammenfassens vorhandene Funktion ist die Möglichkeit zur Begrenzung der Fläche, die ein Demand-Node repräsentiert. Die Flächenbegrenzung bewirkt verschiedene Effekte bei der Verteilung der Demand-Nodes. Die Demand-Nodes liegen durch die Begrenzung ihrer Fläche in Bereichen mit niedrigem Kommunikationsverkehr dichter. Weiterhin enthält jede Region, deren Fläche größer ist als der Wert der Flächenbegrenzung, mindestens einen Demand-Node.

Der Kommunikationsverkehr dieser in der Fläche begrenzten Demand-Nodes ist offensichtlich niedriger als der maximal mögliche Kommunikationsverkehr  $T_{max}$ . Dadurch wird zum einen die Gesamtzahl der Demand-Nodes im Planungsgebiet erhöht. Zum anderen wird ihre Verkehrsverteilung negativ beeinflusst, da weniger Demand-Nodes den maximalen Kommunikationsverkehr erreichen.

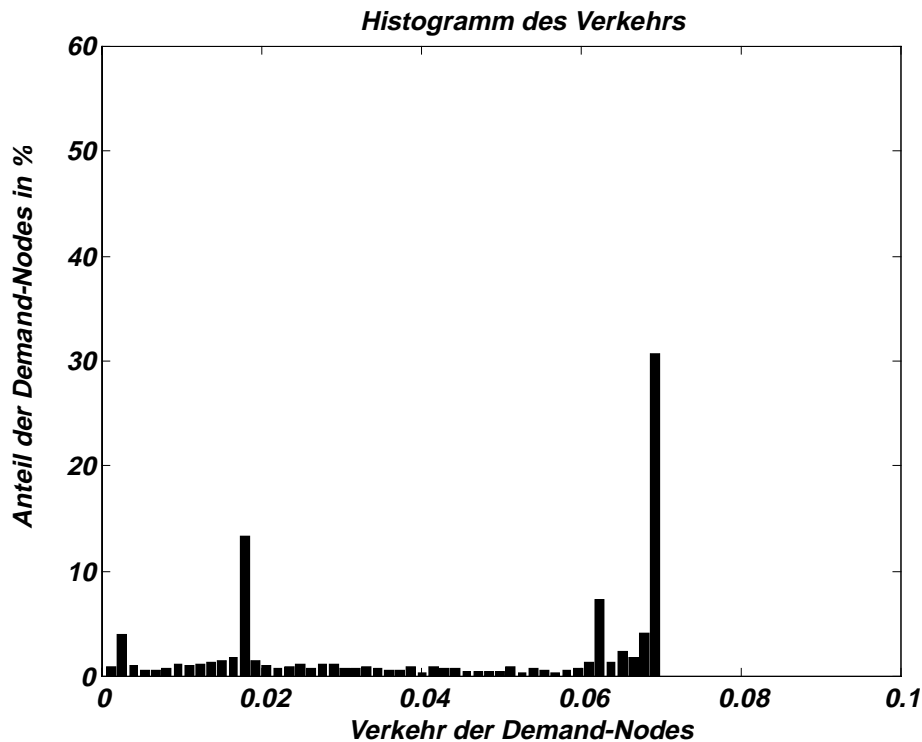
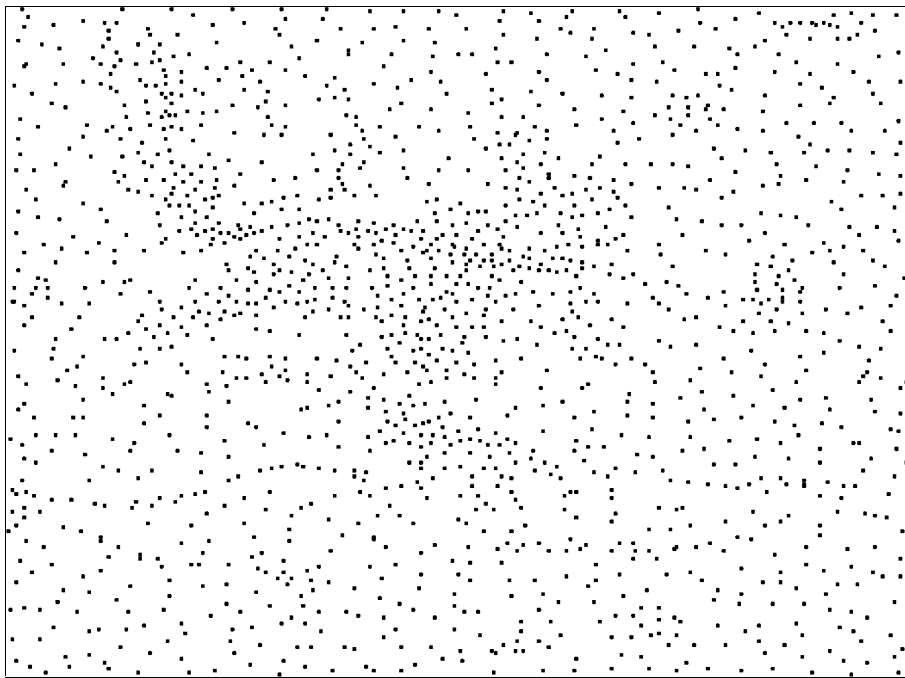


Abbildung 4.2.16: Würzburger Gebiet, verbesserte Optimierung mit Flächenbegrenzung der Demand-Nodes auf  $0.2 \text{ km}^2$

Bei einem sehr kleinen Wert von z.B.  $0.2 \text{ km}^2$  für die Flächenbegrenzung ergibt sich die in Abbildung 4.2.16 gezeigte Verkehrsverteilung. Dort sind, zusätzlich zu der am maximalen Kommunikationsverkehr, mehrere Spitzen von Demand-Nodes mit gleichem Kommunikationsverkehr zu erkennen. Da der Kommunikationsverkehr in großen Bereichen des Gebiets konstant ist, wie in Kapitel 3.2 beschrieben, stellt jede Spitze Demand-Nodes mit maximaler Fläche in einem Bereich des Planungsgebietes mit einem bestimmten Kommunikationsverkehr dar. Bei einem sinnvollem Wert für die Flächenbegrenzung von z.B.  $0.75 \text{ km}^2$  ergibt sich eine Verteilung wie in Abbildung 3.3.13, bei der ansonsten die gleichen Parameter wie hier verwendet werden.

Die Anzahl der Demand-Nodes erhöht sich bei dieser extremen Wahl der Flächenbegrenzung stark, da keine Fläche größer als  $0.2 \text{ km}^2$  im Planungsgebiet ohne einen Demand-Node ist. Die Verteilung der Demand-Nodes ist in Abbildung 4.2.17 dargestellt, welche direkt mit der Verteilung mit einer Flächenbegrenzung von  $0.75 \text{ km}^2$  aus Abbildung 3.3.16 verglichen werden kann. Deutlich ist zu sehen, daß auch in den Bereichen mit sehr niedrigem Kommuni-

kationsverkehr die Demand-Nodes dicht liegen, und keine größere Fläche ohne einen Demand-Node ist. Der Kommunikationsverkehr der einzelnen Demand-Nodes ist aber hier sehr unterschiedlich. Die Demand-Nodes in der Stadt haben z.B. einen wesentlich höheren Kommunikationsverkehr als die in den Waldgebieten, deren Kommunikationsverkehr durch die Flächenbegrenzung begrenzt wird. Die Dichte der Demand-Nodes in einem Bereich erlaubt dort keine direkte Aussage mehr über die Höhe des Kommunikationsverkehrs.



*Abbildung 4.2.17: Demand-Nodes des Würzburger Gebiets mit Flächenbegrenzung der Demand-Nodes auf  $0.2 \text{ km}^2$*

Zum Vergleich der Approximationsgüte verschiedener Werte der Flächenbegrenzung werden vier Parametersätze verwendet, die in Tabelle 4.2.5 aufgelistet werden. Damit die Flächenbegrenzung auch wirksam ist, setzt man zur Berechnung die verbesserte Optimierung mit Standardwerten ein. Ohne Optimierung erreichen viele Demand-Nodes nicht den maximalen Verkehr, obwohl sie noch nicht die maximale Fläche erreicht haben. Wie in den Kapiteln 3.2 und 3.3 beschrieben wird, repräsentieren nach der Optimierung nahezu nur die Demand-Nodes mit begrenzter Fläche einen niedrigeren Verkehr.

Als Vergleich dient dabei die Kurve ohne Flächenbegrenzung. Als erste Vergleichskurve wird die maximale Fläche eines Demand-Nodes auf  $1.00 \text{ km}^2$  beschränkt. Dies zeigt kaum eine Auswirkung, da fast kein Demand-Node eine größere Fläche repräsentieren würde. Bei der dritten Kurve wird der Wert  $0.3 \text{ km}^2$  für die maximale Fläche verwendet. Außerdem erhöht man den maximalen Kommunikationsverkehr eines Demand-Nodes so, daß die Anzahl der erzeugten Demand-Nodes ungefähr der ohne Flächenbegrenzung entspricht. Dadurch

wird die Vergleichbarkeit der beiden Kurven sichergestellt. Bei der vierten Kurve wird die Flächenbegrenzung auf  $0.2 \text{ km}^2$  sehr restriktiv gesetzt, wodurch die meisten Demand-Nodes in ihrer Fläche begrenzt werden. Dabei erhöht sich auch die Anzahl der Demand-Nodes stark.

<b>Kurve</b>	<b>Max. Verkehr</b>	<b>Max. Fläche</b>	<b>Anzahl N der Demand-Nodes</b>
Keine Flächenbegrenzung, $N \neq N'$	0.07	unbeschränkt	1226
Flächenbegrenzung, $N \approx N'$	0.07	$1.0 \text{ km}^2$	1238
Flächenbegrenzung, $N = N'$	0.088	$0.3 \text{ km}^2$	1226
Flächenbegrenzung, $N > N'$	0.07	$0.2 \text{ km}^2$	1672

Tabelle 4.2.5: Parametersätze für die Untersuchung der Flächenbegrenzung

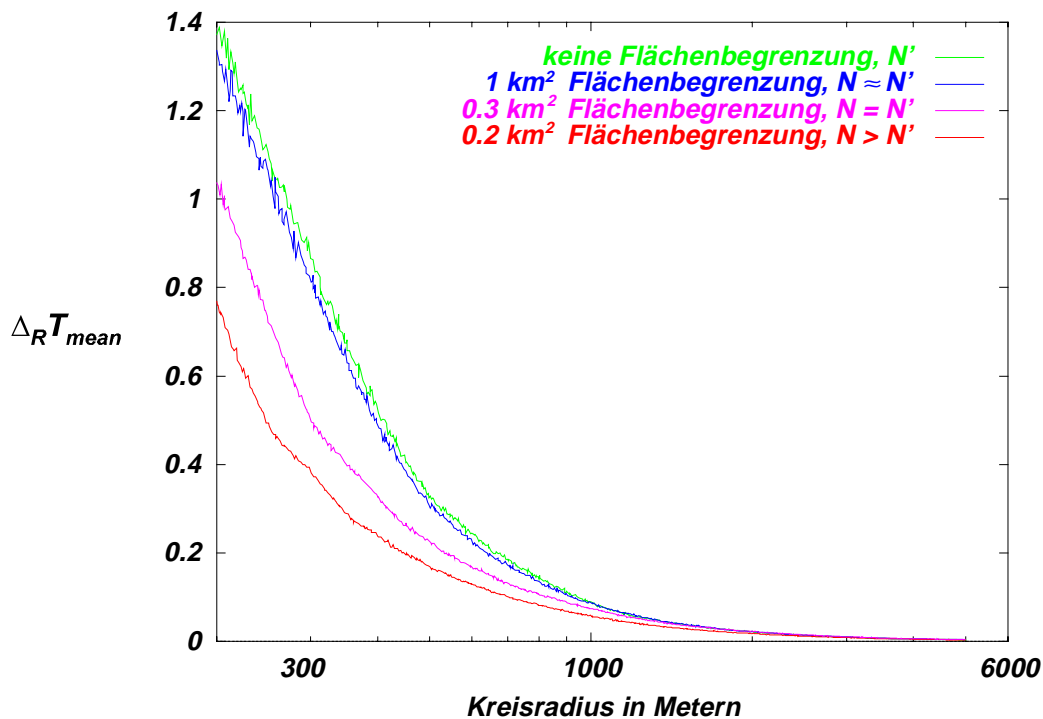


Abbildung 4.2.18: Untersuchung der Flächenbegrenzung bei kleiner Testfläche

In Abbildung 4.2.18 sind die Ergebnisse für die vier Parametersätze dargestellt. Die Kurve mit einer Begrenzung der Fläche auf  $1.00 \text{ km}^2$  unterscheidet sich kaum von der ohne Begrenzung. Die Kurve mit extremer Begrenzung der Fläche auf  $0.2 \text{ km}^2$  zeigt vor allem im Bereich kleiner Testflächen einen deutlich niedrigeren Fehler. Dies wird zum Teil aber durch die hohe Anzahl an Demand-Nodes

hervorgerufen. Interessanter ist die Kurve mit  $0.3 \text{ km}^2$  Flächenbegrenzung, da sie ungefähr die gleiche Anzahl Demand-Nodes besitzt wie die Kurve ohne Flächenbegrenzung. Auch hier ist im Bereich von kleinen Testflächen ein signifikant niedrigerer Fehler beobachtbar.

Der Bereich größerer Radien der Testflächen ist in Abbildung 4.2.19 besser zu erkennen. Die Flächenbegrenzung zeigt ihre Auswirkung vor allem im Bereich kleiner Testflächen. Bei großen Radien ist die Kurve mit  $0.3 \text{ km}^2$  Begrenzung sogar schlechter als die Kurve ohne Begrenzung der Fläche. Dies lässt sich leicht durch den höheren maximalen Kommunikationsverkehr eines Demand-Nodes bei diesem Parametersatz erklären. Der höhere Kommunikationsverkehr ist nötig, damit trotz Flächenbegrenzung ungefähr die gleiche Anzahl Demand-Nodes erzeugt wird wie im Fall ohne Begrenzung der Fläche. Die Kurve mit der Beschränkung auf  $0.2 \text{ km}^2$  ist nicht besonders aussagekräftig, da bei ihr alleine durch die hohe Anzahl der Demand-Nodes der Fehler im ganzen Bereich kleiner ist.

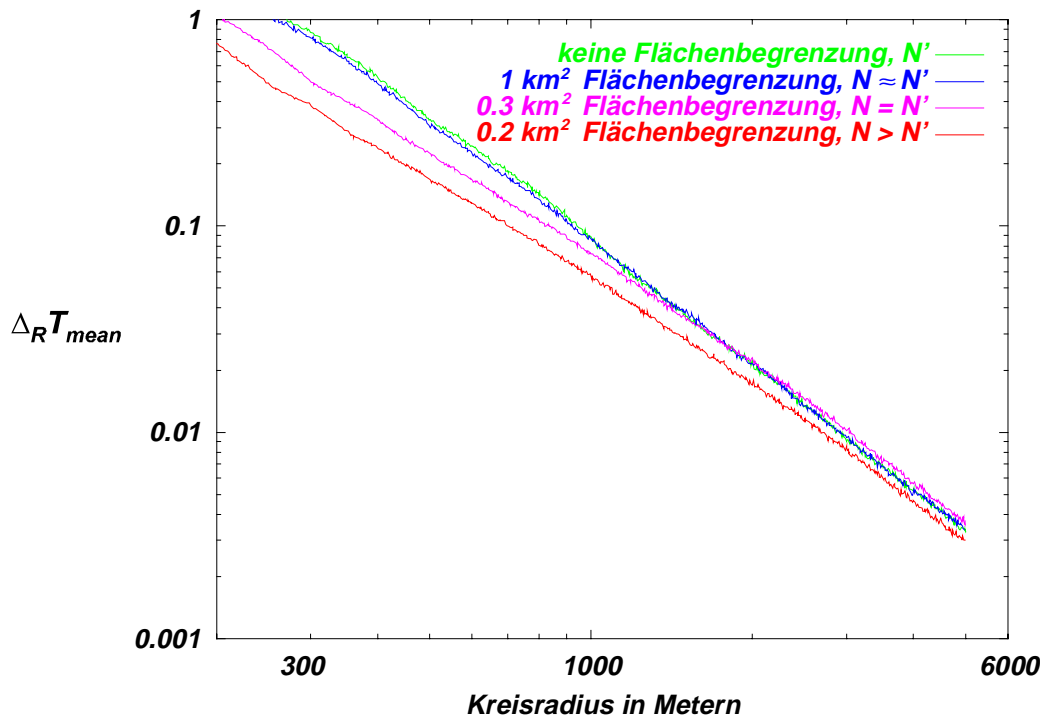


Abbildung 4.2.19: Untersuchung der Flächenbegrenzung bei großer Testfläche

Der Hauptvorteil der Flächenbegrenzung liegt darin, daß selbst in Gebieten mit sehr niedriger Verkehrsdichte kein Bereich größer als die Flächenbegrenzung ohne einen Demand-Node bleibt. Dadurch kann nicht nur in diesem Bereich eine Versorgung des Kommunikationsverkehrs sichergestellt werden, sondern auch eine Versorgung der Fläche. Dieser Vorteil kann durch das statistische Meßverfahren nicht effektiv überprüft werden. Durch die Mittelung des Fehler-

maßes über das gesamte Planungsgebiet wirkt sich die bessere Approximationsgüte in Bereichen mit niedrigem Verkehr kaum in der Meßkurve aus.

### 4.3 Repräsentation der Landnutzungsklassen

Bei der Bewertung der verschiedenen Verfahren zur Clusterung sind zum einen die Abweichungen der Verkehrsverteilung der Demand-Nodes von der Verkehrsmatrix interessant. Zum anderen ist auch die Genauigkeit der Repräsentation der Landnutzungsklassen durch die Demand-Nodes wichtig, da sie für die Wellenausbreitungsmodelle nötig sind. Diese Landnutzungsklassen legen fest, welche Modelle und Parameter zur Bestimmung der Signalstärke, die über die Versorgung eines Teilnehmers entscheidet, zugrunde gelegt werden. Aus diesem Grund müssen die Landnutzungsklassen der Demand-Nodes die wirkliche Landnutzung möglichst genau wiedergeben. Die Daten über die Landnutzungsklasse an einer Position werden in der Landnutzungsmatrix gespeichert, deren Erstellung in Kapitel 2.2.4 erläutert wird.

#### 4.3.1 Meßverfahren

Zur Bewertung der Güte der Repräsentation der Landnutzungsklassen im Planungsgebiet werden die selben Tests wie in Kapitel 4.1 verwendet. Nun wird aber die Verteilung der Landnutzungsklassen innerhalb der Testfläche bestimmt und nicht der Kommunikationsverkehr summiert.

Zum einen mißt man diese Verteilung innerhalb der Landnutzungsmatrix, wie in Abbildung 4.3.1 gezeigt. Dabei wird für jede Landnutzungsklasse  $i$  die Anzahl der Flächenelemente  $N_M(i)$  innerhalb der Testfläche bestimmt. Daraus läßt sich die Verteilung  $D_M(i)$  der einzelnen Landnutzungsklassen in der Matrix berechnen, wobei die Landklassen von 1 bis  $L$  durchnummeriert werden:

$$D_M(i) = \frac{N_M(i)}{\sum_{j=1}^L N_M(j)} .$$

Zum anderen wird die Verteilung der Landnutzungsklassen der Demand-Nodes innerhalb der Testfläche berechnet, wie in Abbildung 4.3.2 dargestellt wird:

$$D_D(i) = \frac{N_D(i)}{\sum_{j=1}^L N_D(j)} .$$

Hierbei wird mit  $N_D(i)$  die Anzahl der Demand-Nodes mit der Landnutzungsklasse  $i$  innerhalb der Testfläche bezeichnet.

Wichtig dabei ist, daß diese Verteilung nur dann bestimmt werden kann, wenn innerhalb der Testfläche mindestens ein Demand-Node liegt. Andernfalls kann keine Aussage über die Landnutzungsklasse durch die Demand-Nodes gemacht werden. In der Praxis tritt dieser Fall nicht auf, da kein Sender eine Versor-

gungsfläche hat, die keinen Kommunikationsverkehr versorgt, d.h. keinen einzigen Demand-Node überdeckt.

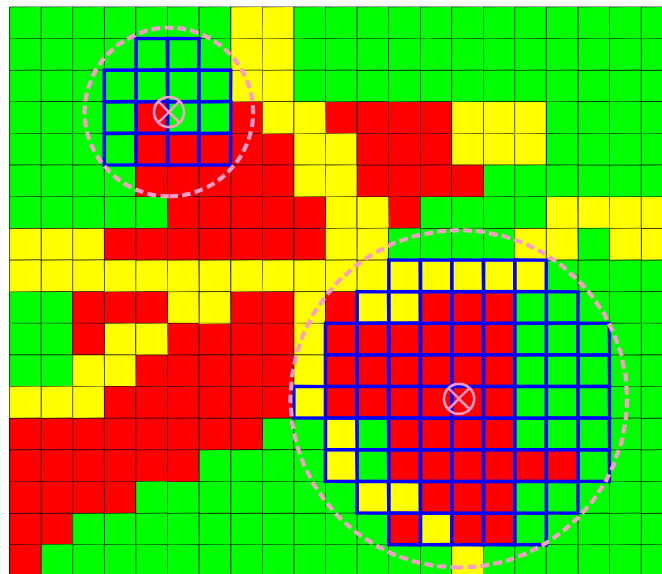


Abbildung 4.3.1: Landnutzungsklassen in der Matrix

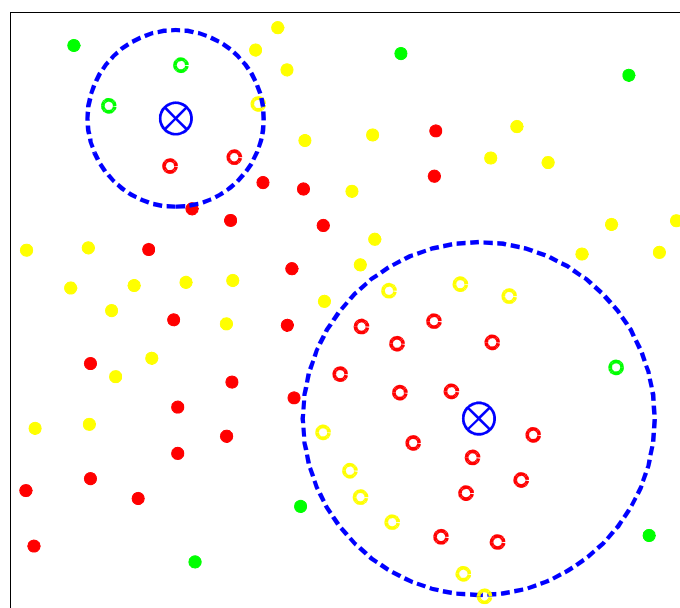


Abbildung 4.3.2: Landnutzungsklassen der Demand-Nodes

Die beiden Verteilungen der Landnutzungsklassen sollten im Idealfall identisch sein. Zum Vergleich verschiedener Demand-Node Verteilungen wird folgendes Fehlermaß definiert:



$$\Delta_R L = \frac{1}{2} \sum_{j=1}^L |D_M(j) - D_D(j)| .$$

Die Definition von  $\Delta_R L$  ist dabei so angelegt, daß sich der Wert 0 ergibt, falls der prozentuale Anteil jeder Landnutzungsklasse sowohl in der Matrix, als auch bei den Demand-Nodes übereinstimmt. Falls alle Demand-Nodes eine andere Landnutzungsklasse besitzen als die entsprechenden Flächenelemente in der Matrix, so liefert  $\Delta_R L$  den maximalen Wert von 1.0. Der Wert des Fehlermaßes  $\Delta_R L$  ermöglicht eine genaue Aussage über die Approximationsgüte der Landnutzungsklassen in der Testfläche.

Zur Erläuterung wird als Beispiel der Wert von  $\Delta_R L$  für die kleinere der beiden Testflächen in Abbildung 4.3.1 berechnet.

Für die Matrix ergeben sich die in Tabelle 4.3.1 aufgelisteten Werte.

<b>Landnutzungsklasse</b>	<b><math>N_M(i)</math></b>	<b><math>D_M(i)</math></b>
<i>Rot</i>	4	$\frac{4}{14}$
<i>Grün</i>	10	$\frac{10}{14}$
<i>Gelb</i>	0	$\frac{0}{14}$

*Tabelle 4.3.1: Landnutzung der Testfläche in der Matrix*

Für die Demand-Nodes ergeben sich die Werte aus Tabelle 4.3.2.

<b>Landnutzungsklasse</b>	<b><math>N_D(i)</math></b>	<b><math>D_D(i)</math></b>
<i>Rot</i>	2	0.4
<i>Grün</i>	2	0.4
<i>Gelb</i>	1	0.2

*Tabelle 4.3.2: Landnutzung der Testfläche der Demand-Nodes*

Das Fehlermaß  $\Delta_R L$  läßt sich nun berechnen zu:

$$\Delta_R L = \frac{1}{2} \sum_{j=1}^L |D_M(j) - D_D(j)| = \frac{1}{2} \left( \left| \frac{4}{14} - 0,4 \right| + \left| \frac{10}{14} - 0,4 \right| + \left| \frac{0}{14} - 0,2 \right| \right) = 0,31$$

Für ein statistisch aussagekräftiges Ergebnis wird, analog zu den in den Kapiteln 4.1.3 bis 4.1.4 beschriebenen Verfahren, die mittlere Abweichung  $\Delta_{RL_{mean}}$  in Abhängigkeit vom Radius  $r$  bestimmt. Die Funktion  $\Delta_{RL_{mean}}(r)$  wird als Gütemaß für die Approximationsgüte der Landnutzung herangezogen.

### 4.3.2 Meßergebnisse

In der Abbildung 4.3.3 ist das Gütemaß  $\Delta_{RL_{mean}}(r)$  für beide Verfahren dargestellt. In der Grafik fällt der wannenförmige Verlauf der Kurve des Zusammenfassens auf, die bei einem Radius von etwa 1000 m ein Minimum erreicht.

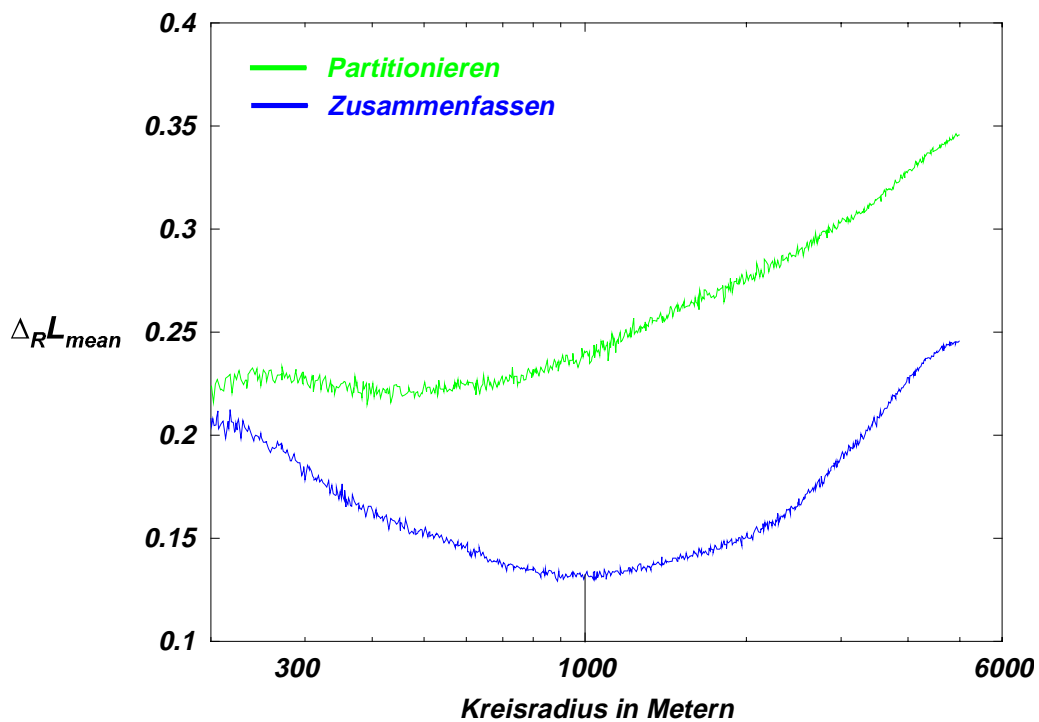


Abbildung 4.3.3: Vergleich der Landnutzungsklassen beider Verfahren

Der höhere Fehler bei sehr kleinen Radien ist damit zu erklären, daß oft nur ein oder einige Demand-Nodes in der Testfläche liegen, und damit die Genauigkeit der Darstellung geringer ist. Der Anstieg des Fehlers mit wachsendem Radius kommt daher, daß die prozentuale Verteilung der Landnutzungsklassen in der Matrix nie genau der Verteilung der Demand-Nodes entsprechen kann. Der Grund ist der Unterschied in der Granularität der Daten über die Verkehrsverteilung zwischen der Verkehrsmatrix und den Demand-Nodes.

In der Grafik ist deutlich der Vorteil des Verfahrens des Zusammenfassens von Knoten zu sehen, dessen Fehler in allen Bereichen niedriger ausfällt. Vor allem im wichtigen Bereich der Versorgungsgebiete mit einigen Kilometern Durchmesser ist dieses Verfahren überlegen, was an dem besseren Algorithmus zur Festlegung der Landnutzungsklasse eines Demand-Nodes liegt. Der beim Zu-

sammenfassen verwendete Algorithmus zur Festlegung der Landnutzungsklasse wird in Kapitel 3.2.1.4 und der beim Verfahren des Partitionierens eingesetzte in Kapitel 3.1.1 beschrieben.

In Abbildung 4.3.4 ist der Fehler bei unterschiedlicher Anzahl von Demand-Nodes dargestellt. Vor allem der Fehler bei kleinen Radien nimmt mit höherer Anzahl der Demand-Nodes ab, wobei das Minimum des Fehlers früher erreicht wird. Bei größeren Radien ergibt sich kaum ein Unterschied zwischen den einzelnen Kurven. Beim Verfahren des Partitionierens läßt sich der Fehler durch Erhöhung der Anzahl der Demand-Nodes nur wenig verringern.

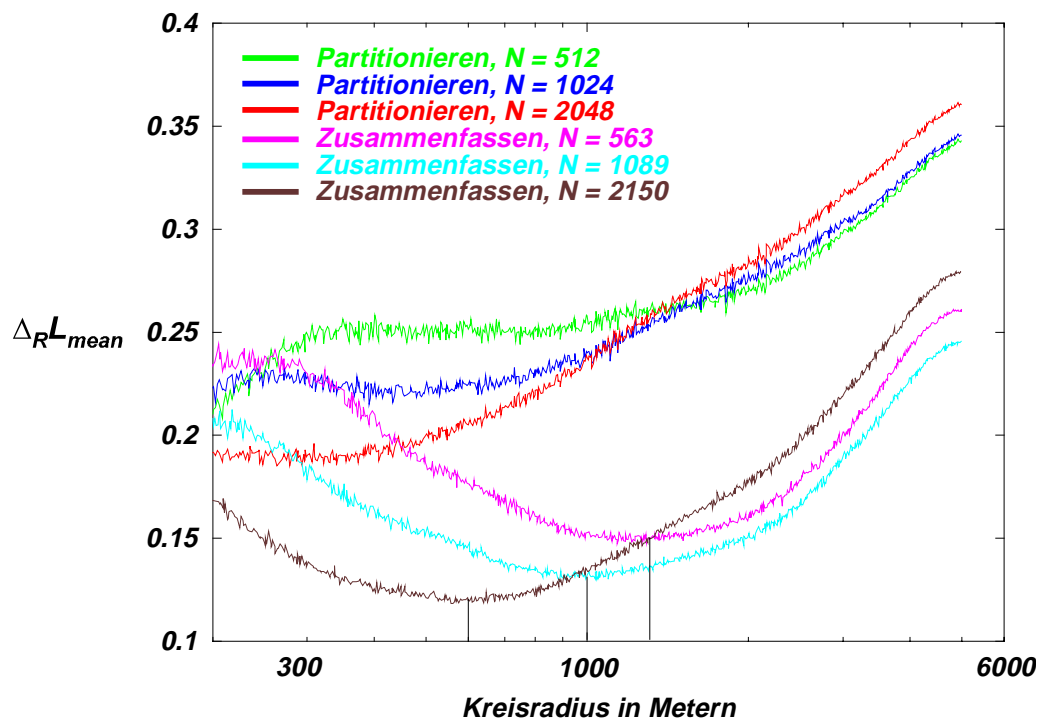


Abbildung 4.3.4: Vergleich bei verschiedener Anzahl der Demand-Nodes

## 5 Anwendungsbeispiele der Clusterung

Im folgenden werden verschiedene Bereiche vorgestellt, bei denen die Clusterung eine Rolle spielt. Außerdem werden Anwendungsbeispiele für die durch die Clusterungsverfahren erzeugten Punktmuster aufgezeigt. Die meisten Beispiele sind aus dem Bereich Mobilfunk, obwohl das Konzept der Demand-Nodes universell zur Planung von Versorgungsstandorten dienen kann.

### 5.1 Einfluß der Clusterung der Teilnehmer auf die subjektive Dienstgüte

In zellularen mobilen Kommunikationssystemen ist die Verteilung der Teilnehmer im Versorgungsgebiet häufig geclustert, d.h. die Teilnehmerkonzentration schwankt stark abhängig von der Position. Diesen Effekt zu berücksichtigen, der die subjektive Dienstgüte der einzelnen Teilnehmer beeinflusst, ist ein wichtiger Aspekt bei der Netzwerkplanung. Die subjektive Dienstgüte, im Englischen als „quality of service“ (QoS) bezeichnet, hängt in entscheidendem Maß vom Charakter der Clusterung ab, wobei vor allem die Intensität der Clusterung ausschlaggebend ist. In [26] werden dazu verschiedene zweidimensionale stochastische Prozesse untersucht, die zur Modellierung der Clusterung der Teilnehmer herangezogen werden können. Aufgrund dieser Prozesse wird die subjektive Dienstgüte berechnet, die ein Testteilnehmer dort erfahren würde. Diese Prozesse werden in der entsprechenden Literatur [2] und [21] ausführlich beschrieben. Für die korrekte Wahl eines Modells der stochastischen Prozesse ist die wirkliche Struktur der Clusterung der Teilnehmer entscheidend. Diese Untersuchung, basierend auf den Daten des Würzburger und des Dallas Gebiets, ist eine Validierung und Erweiterung mit Meßwerten der Ergebnisse von [26].

#### 5.1.1 Modellierung der subjektiven Dienstgüte

Der Einfluß der Struktur der Clusterung auf die subjektive Dienstgüte eines Teilnehmers wird im folgenden untersucht. Dazu wird ein Testteilnehmer im Netz betrachtet. Falls dieser in einer Umgebung mit Clusterung liegt, und die Teilnehmerkonzentration durch die Clusterung nicht bei der Netzwerkplanung berücksichtigt wurde, wird die subjektive Dienstgüte dieses Testteilnehmers verringert. Um die Senkung der Dienstgüte wegen Konzentrationseffekten der Clusterung quantitativ zu beschreiben, wird angenommen, daß bei der Planung der Versorgungsgebiete die Clusterung vernachlässigt wurde. Dies entspricht einem Planungsansatz, bei dem nur die geographische Fläche versorgt wird und nicht der Kommunikationsverkehr. Diesen klassischen, analytischen Planungsansatz, der in [6], [12] und [20] beschrieben wird, verwenden viele bekannte Planungsverfahren und -werkzeuge, wie z.B. *PEGASOS* [11] oder *PLANET* [18].

## 5.1.2 Beschreibung des verwendeten Modells

Im folgenden wird ein Testteilnehmer in einem mobilen Netz mit Clusterung der Benutzer untersucht. Die Struktur der Clusterung wird durch einen zweidimensionalen Punktprozeß beschrieben. Das Versorgungsgebiet eines Senders, auch als Zelle bezeichnet, bedient hier eine Anzahl  $X$  von Teilnehmern. Die Zelle bedeckt dabei die Fläche  $F$  und verfügt über  $K$  Kanäle. Durch die Clusterung ist die Anzahl der zu versorgenden Teilnehmer bei einer frei gewählten Zelle nicht konstant. Deshalb wird die Anzahl als Zufallsvariable  $X$  definiert. In Abbildung 5.1.1 wird das Modellsystem dargestellt. Für die Zufallsvariable  $X$  können verschiedene Verteilungen angenommen werden, die sich durch die räumliche Verteilung der Teilnehmer im Gebiet ergeben. Für diese räumliche Verteilung können als Modell verschiedene zweidimensionale stochastische Punktprozesse angenommen werden. Zusätzlich kann die wirkliche Verteilung im Planungsgebiet, die durch die Demand-Nodes repräsentiert wird, gemessen werden.

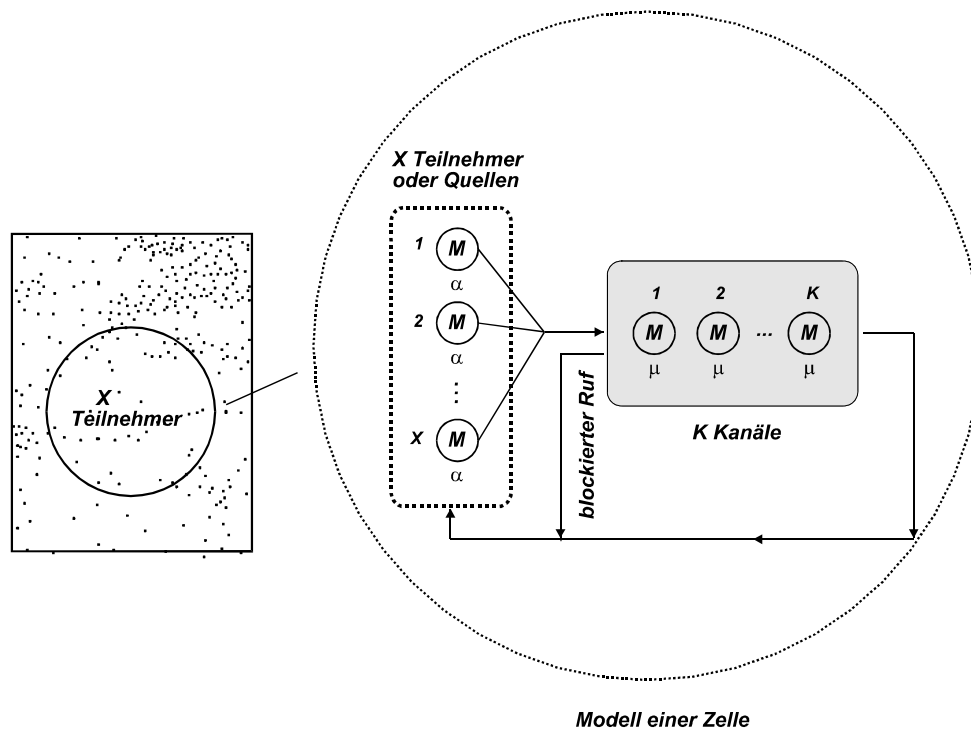


Abbildung 5.1.1: Modell einer Zelle mit Clusterung der Teilnehmer

Der Verkehrsprozeß der Teilnehmer ist in das Modell eines Verlustsystems mit endlicher Quellenzahl eingebettet. Ein Teilnehmer kann dabei zwei Zustände annehmen: *Passiv* oder *Aktiv*. Die Länge der Zeitintervalle, in denen der Teilnehmer in einer der beiden Zustände verharret, wird als exponentiell verteilt angenommen. Die Zufallsvariable  $I$  mit einem Mittelwert von  $\frac{1}{\alpha}$  beschreibt die Aufenthaltszeit im Zustand *Passiv*. Nach dem Ende eines Gesprächs bleibt ein

Teilnehmer in diesem Zustand, bis er den nächsten Anruf startet. Falls der Versuch eines Anrufs zurückgewiesen wird, da kein Kanal frei ist, verbleibt der Teilnehmer im Zustand *Passiv*. Das Zeitintervall im Zustand *Aktiv* wird mit der Zufallsvariable  $B$  mit einem Mittelwert von  $\frac{1}{\mu}$  beschrieben, wobei dieser Zustand die Nutzung eines Kanals darstellt. Solange der Teilnehmer verbunden ist, verbleibt er in diesem Zustand und geht nach dem Ende der Verbindung wieder in den Zustand *Passiv* über.

Die Verkehrsaktivität, die ein Teilnehmer bewirkt, d.h. die Wahrscheinlichkeit, daß ein Teilnehmer aktiv ist, falls eine eventuelle Blockierung vernachlässigt wird, ist:

$$\rho_M = \frac{\frac{\alpha}{\mu}}{1 + \frac{\alpha}{\mu}} = \frac{a^*}{1 + a^*} .$$

Das Verkehrsangebot, welches in dieser Arbeit zur Planung herangezogen wird, ergibt sich zu:

$$a^* = \frac{\alpha}{\mu} .$$

Hierbei beschreibt  $a^*$  auch die Verkehrsintensität, die durch den Kommunikationsverkehr eines Teilnehmers erzeugt wird. Dieser Wert wird häufig für die Netzwerkplanung verwendet und in der Pseudoeinheit „Erlang“ [*Erl*] angegeben. Er liegt im Bereich von 0.015 *Erl* bis 0.05 *Erl*. Die Verkehrsintensität ist der Wert, der in der Verkehrsmatrix abgelegt wird und als Grundlage zur Erstellung der Demand-Nodes dient. Um die Verkehrsintensität zu berechnen, die durch mehrere Teilnehmer erzeugt wird, werden die einzelnen Verkehrsangebote  $a^*$  summiert.

Das Modell einer Zelle stellt ein Standard-Verlustsystem mit  $K$  Bedieneinheiten und einer endlichen Anzahl  $X$  von Quellen dar, wie es z.B. in [10] beschrieben wird. Die Zufallsvariable  $X$  bezeichnet die zufällige Anzahl der Teilnehmer, die sich bei einer beliebigen Positionierung einer Zelle ergibt, ohne dabei die Clustering zu berücksichtigen.

### 5.1.3 Definition der subjektiven Dienstgüte

Die subjektive Dienstgüte wird definiert als die Blockierungswahrscheinlichkeit, die ein beliebiger Testteilnehmer im untersuchten Gebiet sieht.

Mit  $x(i) = P(X = i)$  ergibt sich die Wahrscheinlichkeit  $x^*(i)$ , daß der Testteilnehmer sich in einer Zelle mit  $X = i$  Teilnehmern aufhält, nach [23] zu:

$$x^*(i) = i \cdot x(i) \cdot \frac{1}{E[X]} .$$

Die Wahrscheinlichkeit  $x^*(i)$  ist proportional zur Anzahl der Teilnehmer  $i$  in der Zelle und zur Auftrittswahrscheinlichkeit  $x(i) = P(X = i)$  einer Zelle mit  $i$  Teilnehmern. Weiterhin muß  $c \cdot \sum_{i=1}^{\infty} i \cdot x(i) = 1$  gelten, woraus  $c = \frac{1}{E[X]}$  resultiert.

Insgesamt ergibt sich daraus die obige Formel für  $x^*(i)$ .

Mit der Verwendung der aus [10] bekannten Engset-Formel bekommt man die bedingte Blockierungswahrscheinlichkeit des Testteilnehmers, falls er in einer Zelle mit  $X = i$  Teilnehmern ist:

$$p_B(i) = P(\{\text{"Testteilnehmer zurückgewiesen"} \mid X = i\})$$

*Gleichung 5.1.1*

$$= \frac{\binom{i-1}{K} a^{*K}}{\sum_{k=0}^K \binom{i-1}{k} a^{*k}} .$$

Beide Gleichungen zusammen geben die Blockierungswahrscheinlichkeit eines beliebigen Testteilnehmers an, die der oben beschriebenen subjektiven Dienstgüte entspricht:

$$p_B = \sum_{i=K+1}^{\infty} p_B(i) x^*(i) .$$

#### 5.1.4 Modelle der Verteilung der Teilnehmer

Für die Zufallsvariable  $X$  können folgende Verteilungen herangezogen werden:

- Deterministische Verteilung: Die Anzahl der Teilnehmer in einer Zelle ist konstant:

$$P(X = i) = \begin{cases} 1 & \text{für } i = d \\ 0 & \text{sonst} \end{cases} ,$$

$$\text{mit } E[X] = d, c_X = 0 .$$

- Poisson Verteilung:

$$P(X = i) = \frac{y^i}{i!} e^{-y} ,$$

$$\text{mit } E[X] = y, c_X = 1 .$$

- Negativ-binomiale Verteilung:

$$P(X = i) = \binom{y+i-1}{i} \cdot (1-q)^y q^i ,$$

$$\text{mit } E[X] = \frac{yq}{1-q} , c_x = \frac{1}{\sqrt{yq}} .$$

wobei sich die Parameter  $y$  und  $q$  auch aus gegebenen  $E[X]$  und  $c_x$  bestimmen lassen:

$$q = \frac{1}{E[X] \cdot c_x^2} , y = \frac{E[X]}{E[X] \cdot c_x - 1} .$$

- Messung der Verteilung mit Hilfe von Testzellen, die beliebig über das Gebiet gelegt werden.

Mit  $E[X]$  wird hierbei der Mittelwert der Zufallsvariable  $X$  und mit  $c_x$  ihr Variationskoeffizient bezeichnet.

Im folgenden Kapitel wird das Messen der Verteilung der Zufallsvariable  $X$  im Würzburger und im Dallas Gebiet beschrieben.

### 5.1.5 Messung der Verteilung der Teilnehmer

Für die Messung der Zufallsvariable  $X$  wird ausgenutzt, daß jeder Demand-Node die Teilnehmer in seiner Fläche zusammenfaßt und ihren Kommunikationsverkehr summiert. Dazu wird eine Menge von Testkreisen mit konstanter Fläche  $F$  zufällig in das Gebiet gelegt, und der Kommunikationsverkehr innerhalb jedes Testkreises summiert, wie in Abbildung 5.1.2 gezeigt. Daraus ergeben sich für  $N$  Tests die Verkehrswerte  $T_j$  ( $1 \leq j \leq N$ ), bei konstantem Radius  $r = \sqrt{\frac{F}{\pi}}$  der Testkreise.

Das mittlere Verkehrsangebot pro Teilnehmer kann aus der mittleren Aktivität eines Teilnehmers  $\rho_M$  berechnet werden:

$$a_{mean}^* = \frac{\rho_M}{1 - \rho_M} .$$

Mit Hilfe des mittleren Verkehrsangebots  $a_{mean}^*$  und den Verkehrswerten  $T_j$  werden die Werte für die Zufallsvariable  $X$  berechnet, die nicht ganzzahlig sein müssen:

$$X_j = \frac{T_j}{a_{mean}^*} , \forall (1 \leq j \leq N) .$$



Die Wahrscheinlichkeit  $P(X = i)$  ergibt sich dann zu:

$$P(X = i) = \frac{|M(i)|}{N} .$$

Mit  $M(i)$  wird die Menge der Indizes  $j$  der Zufallsvariablen  $X$  bezeichnet, bei der  $X_j$  gerundet eine Anzahl von  $i$  Teilnehmern darstellt:

$$M(i) = \{1 \leq j \leq N \mid i-1 < X_j \leq i\} ,$$

$|M(i)|$  gibt die Anzahl der Testkreise an, die  $i$  Teilnehmer enthalten, wobei

$$\sum_{i=0}^{\infty} |M(i)| = N \text{ ist.}$$

Daraus läßt sich der Mittelwert  $E[X]$  von  $X$  berechnen:

$$E[X] = \sum_{j=1}^N i \cdot P(X = i) .$$

Der Mittelwert  $E[X]$  hängt dabei von der Größe der Testkreise, d.h. von deren Radius  $r$  ab. Zu einem vorgegebenem  $E[X]$  kann für einen bestimmten Wert von  $\rho_M$  ein geeigneter Radius der Testkreise gefunden werden. Die subjektive Dienstgüte ergibt sich dann aus der Verteilung der mit diesem Radius gemessenen Zufallsvariablen  $X$ .

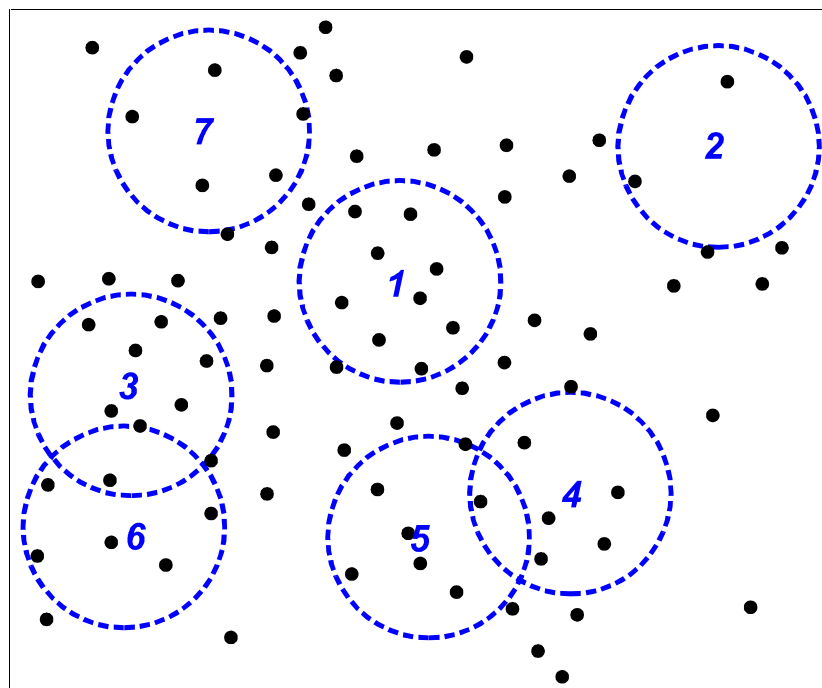


Abbildung 5.1.2: Beispiel für zufällige Testkreise

## 5.1.6 Ergebnisse der subjektiven Dienstgüte

Bei den beiden folgenden Kurven wird die subjektive Dienstgüte für eine Zelle mit  $K = 7$  Kanälen gemessen, was einer typischen Anzahl von Kanälen einer Zelle in GSM-Systemen entspricht. Dargestellt wird die subjektive Dienstgüte für die Fälle einer deterministischen, einer Poisson und verschiedener negativ-binomialer Verteilungen. Bei der negativ-binomialen Verteilung wird die Dienstgüte für einen Variationskoeffizienten  $c$  von 0.5, 0.7, 1.0 und 2.0 aufgetragen. In Abbildung 5.1.3 sind zusätzlich die Meßwerte des Würzburger Gebiets und in Abbildung 5.1.4 die Meßwerte des Dallas Gebiets eingetragen.

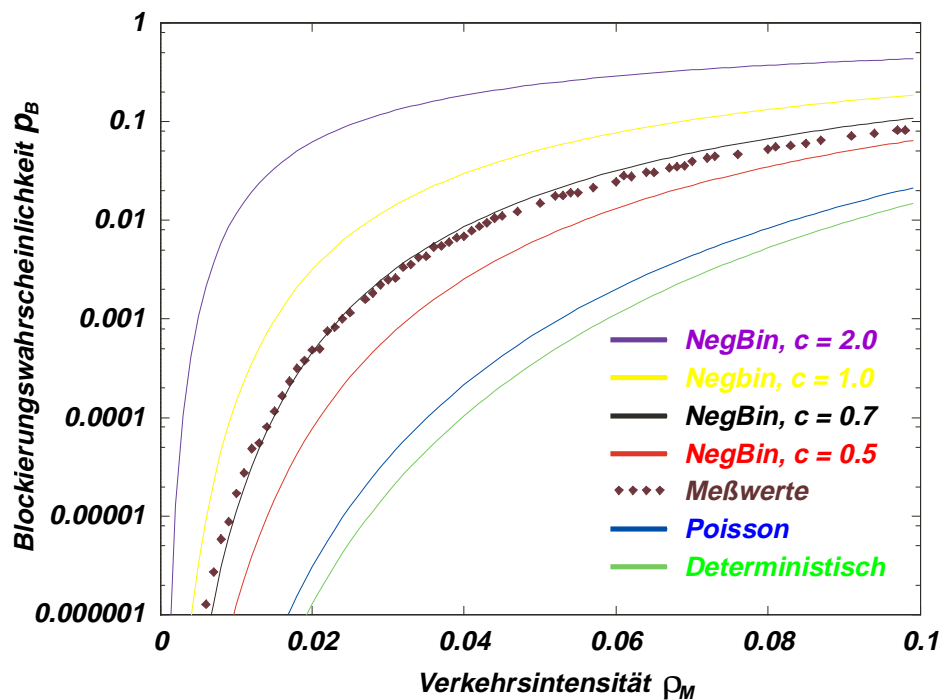


Abbildung 5.1.3: Subjektive Dienstgüte bei  $E[X] = 30$ ,  $K = 7$ , Würzburger Gebiet

Das generelle Verhalten für alle Arten der Teilnehmerverteilung sieht ähnlich aus, wobei der Unterschied zwischen einer nach Poisson verteilten und einer festen Anzahl von Teilnehmern pro Zelle klein ist. Die Dienstgüte nimmt aber stark ab, falls die Verteilung der Teilnehmer einer negativ-binomialen Verteilung folgt. Bei der negativ-binomialen Verteilung mit  $c = 1$  ist die Dienstgüte um den Faktor 10 bis 100 geringer als bei der Poisson Verteilung mit  $c = 1$ , was auf einen starken Einfluß der höheren Momente der Verteilung hinweist.

Im Falle einer gleichmäßigen Verteilung der Teilnehmer im Gebiet ergibt sich offensichtlich die beste Dienstgüte. Falls die Planung der Zellen ohne Berücksichtigung der Clusterung der Teilnehmer durchgeführt wurde, wird die Dienstgüte in Gebieten mit stärkerer Clusterung, d.h. größeren Unterschieden der Teilnehmerdichte, stark abnehmen.

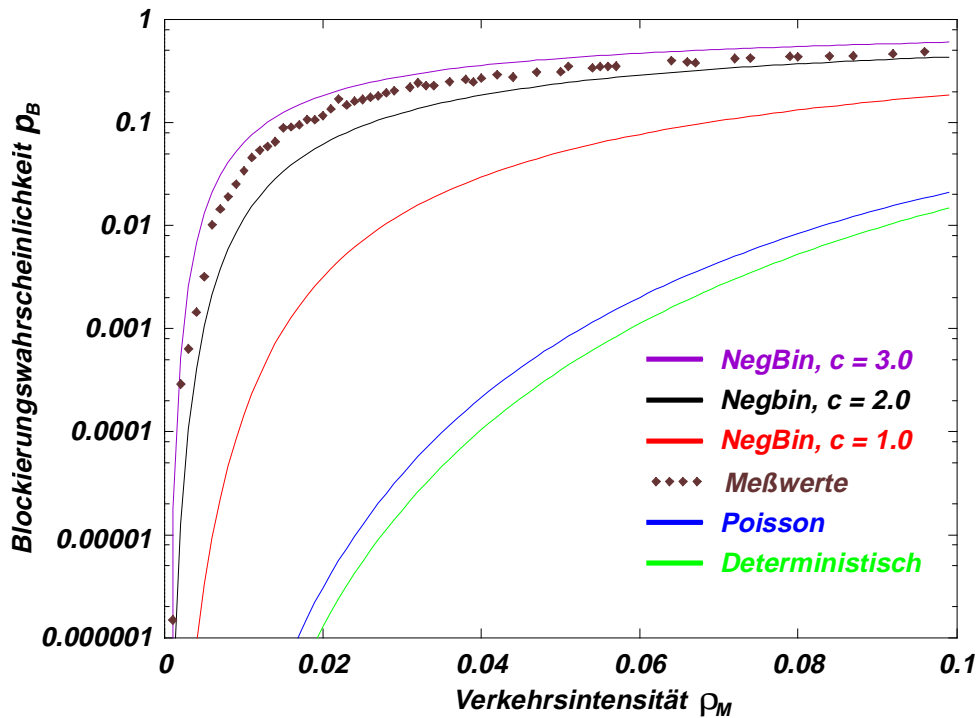


Abbildung 5.1.4: Subjektive Dienstgüte bei  $E[X] = 30$ ,  $K = 7$ , Dallas Gebiet

### 5.1.7 Modellierung mit Wiederholungsversuchen

Das Modell einer Zelle, welches in den vorhergehenden Kapiteln untersucht wurde, wird im folgenden modifiziert. Das Teilnehmermodell aus zwei Zuständen wird ersetzt durch ein Modell mit Wiederholungsversuchen, welches in Abbildung 5.1.5 dargestellt wird.

Dieses Modell wird ausführlich in [22] und [25] untersucht. Ein Teilnehmer kann sich dort in drei Zuständen befinden: *Aktiv*, *Passiv* oder *Wartend*. Der Zustand *Wartend* stellt die Zeit zwischen Wiederholungsversuchen dar, die kürzer ist als die zwischen neuen Gesprächsanforderungen.

Die Zufallsvariable  $R$  mit einem Mittelwert von  $\frac{1}{\beta}$  beschreibt die Aufenthaltszeit im Zustand *Wartend*. Wenn die Gesprächsanforderung abgewiesen wird, geht der Teilnehmer mit der Wahrscheinlichkeit  $\theta$  in den Zustand *Wartend* über. Mit der Wahrscheinlichkeit  $1 - \theta$  verharrt er im Zustand *Passiv*, bis er die nächste Gesprächsanforderung startet. Falls diese fehlschlägt, bleibt er mit der Wahrscheinlichkeit  $\theta$  weiterhin im Zustand *Wartend*, und mit der Wahrscheinlichkeit  $1 - \theta$  geht er in den Zustand *Passiv* über.

Eine genauere Analyse der subjektiven Dienstgüte ist in [25] enthalten, wobei dort die Anzahl der Teilnehmer in einer Zelle konstant ist. Für ein Gebiet mit Clusterung muß die Wahrscheinlichkeit  $p_B$  aus der Gleichung 5.1.1 durch die numerisch gewonnene Wahrscheinlichkeit aus [22] ersetzt werden.

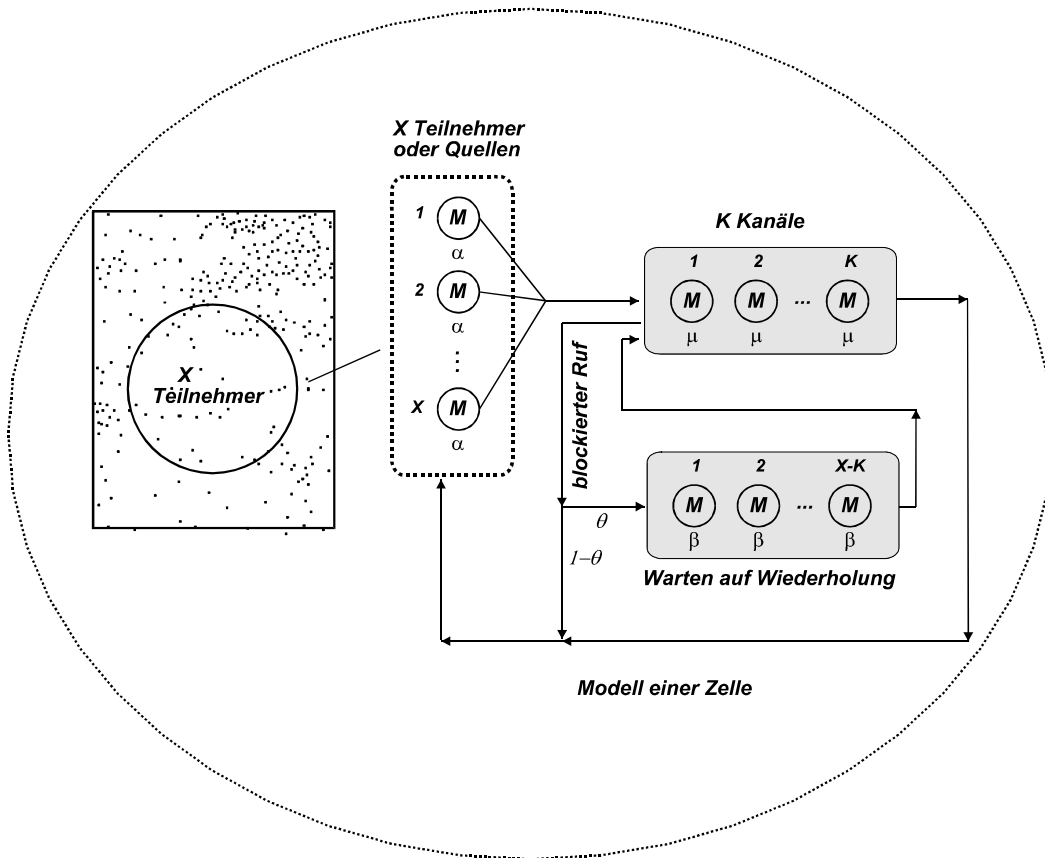


Abbildung 5.1.5: Modell einer Zelle mit Wiederholungsversuch

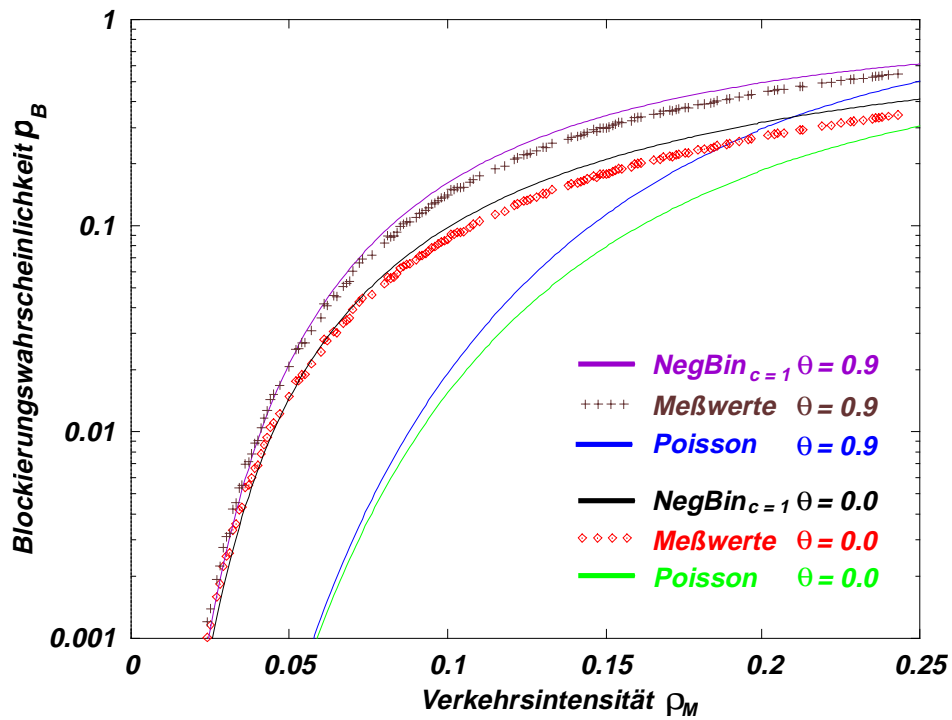


Abbildung 5.1.6: Subjektive Dienstgüte mit Wiederholungsversuch

In Abbildung 5.1.6 wird das Ergebnis für die Wiederholungswahrscheinlichkeiten  $\theta = 0.0$  und  $\theta = 0.9$  gezeigt, wobei die Teilnehmer jeweils gemäß einer Poisson und einer negativ-binomialen Verteilung verteilt sind. Zusätzlich ist die im Würzburger Gebiet gemessene Verteilung für beide Wiederholungswahrscheinlichkeiten aufgetragen. Die übrigen Parameter entsprechen dem Fall ohne Wiederholungsversuch.

Das Aussehen der Kurven entspricht in etwa dem Fall ohne Wiederholungsversuch, wobei die Wiederholungsversuche die Dienstgüte in höherem Maße beeinflussen. Bei den Wiederholungsversuchen tritt ein Selbstverstärkungseffekt auf, da eine höhere Systemauslastung zu mehr Wiederholungen führt, die wiederum die Systemlast weiter erhöhen.

### 5.1.8 Bewertung der Ergebnisse

Die Meßwerte zeigen, daß in beiden Testgebieten eine starke Clusterung vorliegt, die im Dallas Gebiet wesentlich größer ausfällt. Die häufig gemachte Annahme von nach Poisson verteilten Teilnehmern ist als vollkommen falsch abzulehnen, da die Blockierwahrscheinlichkeit im Vergleich zur wirklichen Verteilung bis zum Faktor  $10^5$  zu niedrig angenommen würde. Die subjektive Dienstgüte eines Teilnehmers wäre in Wirklichkeit um diesen Faktor schlechter. Bei der Planung von Mobilfunknetzen ist es aus diesem Grund essentiell, die Clusterung der Teilnehmer- oder Verkehrsverteilung zu berücksichtigen, um korrekte Planungsergebnisse zu erzielen.

Interessant ist auch, daß für die Berechnung der subjektiven Dienstgüte die wirkliche Verteilung der Teilnehmer sehr genau durch die zweiparametrische negativ-binomiale Verteilung beschrieben werden kann. Verteilungen mit einem Parameter, wie z.B. Poisson, können die große Dynamik der Verteilung durch die Clusterung nicht beschreiben.

### 5.1.9 Stochastische Punktprozesse als Planungsmodell

In Bezug auf die subjektive Dienstgüte würde es ausreichen, eine Verteilung der Teilnehmer zu erzeugen, die eine negativ-binomiale Verteilung der oben definierten Zufallsvariable  $X$  liefert. Man müßte dann nicht aus oft unzulänglichen Daten über den Kommunikationsverkehr die Demand-Nodes berechnen. Aus den obigen Ergebnissen zeigt sich aber auch, daß ein einfacher zweidimensionaler Prozeß, wie die homogene räumliche Poisson Verteilung, als Modell nicht ausreicht.

Um eine komplexere Verteilung der Zufallsvariable  $X$ , z.B. eine negativ-binomiale Verteilung, durch einen zweidimensionalen Punktprozeß zu erzeugen, fehlen geeignete Verfahren. Bis jetzt ist kein Verfahren bekannt, das eine zweidimensionale Punktverteilung erzeugen kann, bei der die Zufallsvariable  $X$  einer negativ-binomialen Verteilung mit gegebenen Parametern gehorcht.

## 5.2 Integriertes Planungstool *ICEPT*

Eine effiziente Planung von mobilen Kommunikationsnetzen wird immer wichtiger, da der Bedarf an drahtloser Kommunikation stark gestiegen ist und weiter ansteigt. Die Ressourcen der Übertragungskanäle, d.h. der verfügbare Frequenzbereich, sind aber stark begrenzt. Indes müssen für neue Dienste zusätzliche Systeme aufgebaut oder vorhandene optimiert werden. Dazu wurde das integrierte Planungstool für Kommunikationsnetze *ICEPT* (Integrated network Cellular Planning Tool) am Lehrstuhl für verteilte Systeme (Informatik III) der Universität Würzburg entwickelt. Dieses System zur automatischen Erstellung und Optimierung von mobilen Kommunikationssystemen wird in [27] und [29] beschrieben. Weitere Informationen über das Planungstool befinden sich in [13].

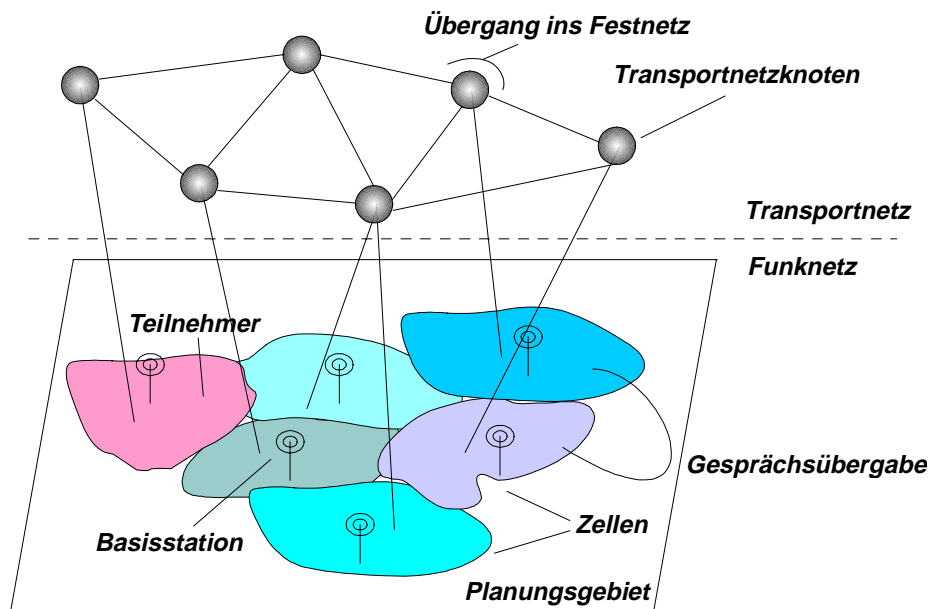


Abbildung 5.2.1: Schematischer Aufbau eines mobilen Kommunikationsnetzes

### 5.2.1 Struktur heutiger mobiler Kommunikationsnetze

In modernen mobilen Kommunikationsnetzen ist das sogenannte zellulare Konzept verwirklicht, d.h. die gesamte Versorgungsfläche wird dicht von Zellen überdeckt. Jede Zelle stellt dabei das Versorgungsgebiet eines Senders da, wie in Abbildung 5.2.1 dargestellt wird. Die Zellgröße schwankt von weniger als einem Kilometer Durchmesser bei Mikrozellen bis zu über 10 km in flachen Gebieten mit wenig Kommunikationsverkehr. Die Form einer Zelle ist in der Praxis nicht kreisförmig, sondern hängt von der Wellenausbreitung ab.

Das wichtigste Problem bei der Planung eines mobilen Kommunikationsnetzes ist die Positionierung von Sendern und die Bestimmung ihrer Kapazität. Die

effiziente Ausnutzung des zur Verfügung stehenden begrenzten Frequenzbereichs zwingt zu einer optimalen Planung der Kapazität eines Senders.

## 5.2.2 Planung mobiler Kommunikationssysteme mit *ICEPT*

In *ICEPT* wird ein neuer Ansatz zur Funknetzplanung und Optimierung verwendet. Grundlage hierzu ist die Verteilung des Verkehrsaufkommens im Planungsgebiet. Aus geographischen Daten wird gemäß dem in 2.2.4 dargelegten Konzept eine Matrix über den Bedarf an Kommunikation erstellt. Auf dieser Grundlage werden die Demand-Nodes mit Hilfe eines der beiden in dieser Arbeit vorgestellten Verfahrens erzeugt. Durch das Konzept der Demand-Nodes wird das kontinuierliche Problem der Versorgung des Verkehrsbedarfs im Planungsgebiet ersetzt durch ein diskretes Problem der Überdeckung von Punkten. Das Ziel der Planung ist nun die Versorgung von möglichst vielen Demand-Nodes durch eine möglichst kleine Zahl von Zellen. Ein Demand-Node wird hierbei als versorgt angesehen, falls die Signalstärke nach einem Wellenausbreitungsmodelle, wie Hata [8] oder COST231 [19], bei Berücksichtigung etwaiger Störungen infolge von Interferenz ausreichend hoch ist.

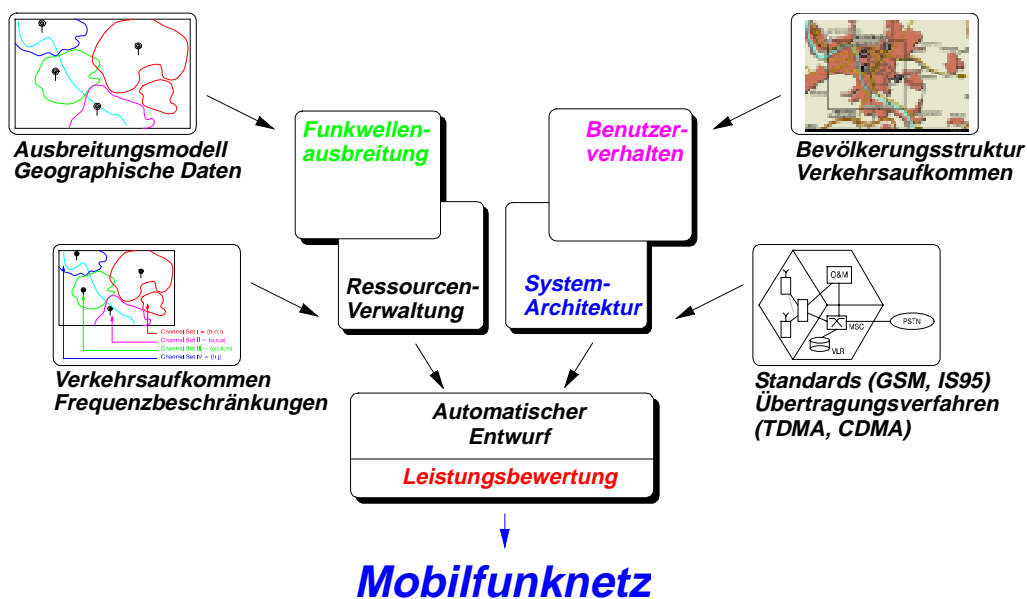


Abbildung 5.2.2: Hauptentwurfskriterien mobiler Kommunikationssysteme

Zur Bestimmung der Zellen, die das Planungsgebiet versorgen sollen, wird in *ICEPT* das Konzept der Überdeckung mit Mengen verwendet. Für jeden Sender wird die Menge der von ihm versorgten Demand-Nodes bestimmt. Zur Planung werden die Sender gemäß vorgegebener Parameter im Planungsgebiet positioniert, wie in [15] beschrieben. Zur Auswahl einer endgültigen Menge von Senderstandorten werden die Versorgungsgebiete der Sender kombiniert, wobei auf Basis der Mengen von versorgten Demand-Nodes gearbeitet wird. Die Planung

des Netzes wird somit auf ein kombinatorisches Problem zurückgeführt, für das bekannte Verfahren existieren. Zur Wahl der optimalen Kombination wird eine Heuristik zur Gewinnung einer maximalen Überdeckung nach [32] verwendet. Das Ergebnis ist eine optimierte Menge von Senderpositionen und der damit versorgten Demand-Nodes.

Aufgrund dieser Versorgungsgebiete und der Systemparameter des verwendeten Kommunikationssystems wird zusätzlich die erforderliche Kapazität der Sender, d.h. die Anzahl der benötigten Kanäle, berechnet. Weiterhin ist die Belegung von Frequenzen für die Kanäle bei Berücksichtigung der physikalischen Einschränkungen, wie z.B. des Wiederholungsabstands, integriert. In Abbildung 5.2.2 werden die von ICEPT berücksichtigten Entwurfskriterien dargestellt.

Zur Visualisierung der Ergebnisse enthält das Tool eine graphische Oberfläche unter X-Window, die in Abbildung 5.2.3 dargestellt wird. Das ganze Programm läuft unter verschiedenen Unix Varianten, wie Solaris oder Linux. Bei Verzicht auf die graphische Oberfläche kann das Planungsmodul auch unter dem Betriebssystem Windows 95/NT eingesetzt werden.



Abbildung 5.2.3: Bildschirmdarstellung von ICEPT



### 5.2.3 Planungsergebnisse von *ICEPT*

In Abbildung 5.2.4 ist ein Planungsergebnis von *ICEPT* für das Würzburger Gebiet dargestellt. Dazu wurde die Konfiguration von sieben Sendern mit *ICEPT* bestimmt. In der Abbildung sind zusätzlich die Versorgungsgebiete der jeweiligen Sender eingezeichnet. Die Überlappung der Versorgungsgebiete ist nötig, damit ein sich bewegender Teilnehmer an der Grenze zweier Gebiete von einem Sender zum anderen übergeben werden kann, ohne dabei die Verbindung zu verlieren. Dieser Vorgang wird als „Handover“ bezeichnet.

Die Demand-Nodes sind in der Farbe ihres versorgenden Senders eingezeichnet, wobei nicht versorgte Demand-Nodes grau gefärbt sind. Bei den Versorgungsgebieten ist auffällig, daß ihre Form sehr unterschiedlich ist, was durch die Geographie verursacht wird. Die in der Verkehrsplanung für die Form der Versorgungsgebiete verwendeten Kreise oder Hexagone treffen in der Praxis kaum zu. Die wirkliche Form hängt von der Wellenausbreitung zwischen der Funkstation und den Teilnehmern ab. Bei vielen Planungstools werden keine realitätsnahen Versorgungsgebiete bei der Planung verwendet. Durch diese Maßnahme wird die Komplexität der Planung wesentlich verringert. Die wirkliche Versorgungsgüte der Benutzer kann damit aber nur sehr eingeschränkt bestimmt werden.

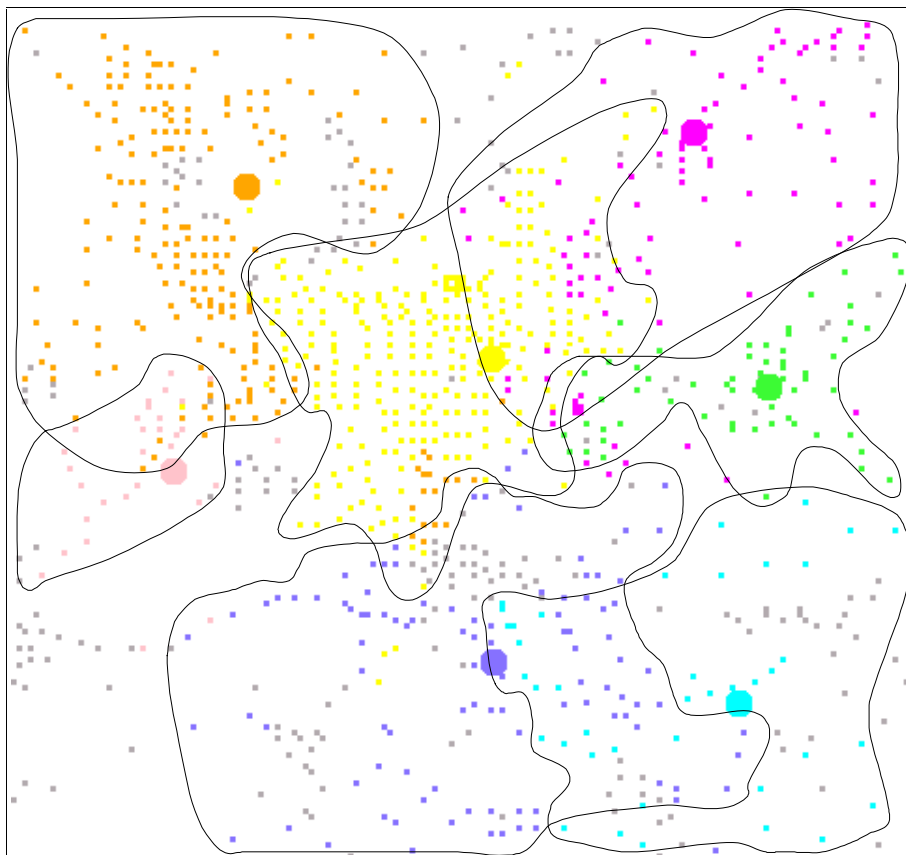


Abbildung 5.2.4: Versorgungsgebiete des Würzburger Gebiets, [29]

### 5.3 Bedarfsabschätzungswerkzeug *CUTE*

Das auf Grundlage dieser Diplomarbeit für die Betriebssystemfamilie Windows 95/NT entwickelte Programm *CUTE* (*CU*stomer *T*raffic *E*stimation *T*ool) soll die Planung von Kommunikationsnetzen erleichtern. Es dient dazu, aus geographischen Daten Informationen über den Kommunikationsbedarf in einem Planungsgebiet zu erzeugen. Dieses Werkzeug ist dabei nicht auf mobile Kommunikationssysteme beschränkt, sondern für beliebige Probleme gedacht, bei denen ein bestimmter Bedarf modelliert werden soll. In [28] wird die Anwendung dieses Konzept in mobilen Kommunikationssystemen beschrieben.

Die Arbeit von *CUTE* ist in vier getrennte, in sich abgeschlossene Abschnitte, aufgeteilt, die in Abbildung 5.3.1 gezeigt werden. Dadurch ist es möglich, einzelne Teile zu ersetzen, um das System den unterschiedlichen Bedürfnissen der Kunden anzupassen.

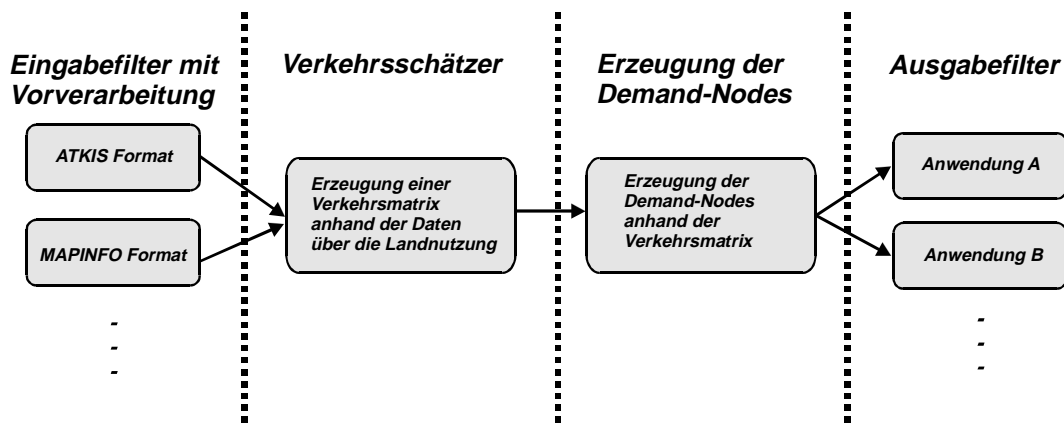


Abbildung 5.3.1: Aufbau von *CUTE*

Als Eingabedaten werden geographische Informationen verschiedener Formate, wie z.B. *ATKIS* [1] oder *MAPINFO* [17], genommen. Diese Daten werden im ersten Schritt vorverarbeitet, wobei verschiedene Inkonsistenzen entfernt werden, wie in [14] beschrieben wird. Als Ergebnis wird eine Matrix mit Informationen über die Landnutzung im Planungsgebiet erstellt. Genauere Angaben dazu finden sich auch in Kapitel 2.2.4.

Im zweiten Schritt arbeitet der Verkehrsschätzer und erzeugt aus den Daten über die Landnutzung eine Verkehrsmatrix, die den Kommunikationsverkehr an jeder Stelle im Planungsgebiet angibt. Dabei werden die Einstellungen des Anwenders bei der Berechnung des Kommunikationsverkehrs aus den Landnutzungsdaten verwendet.

Im dritten Schritt werden die Demand-Nodes gemäß der Verkehrsmatrix nach dem Verfahren des Zusammenfassens von Knoten erzeugt.

Die Daten der Demand-Nodes werden im vierten Schritt durch einen Ausgabe-filter an das vom Benutzer benötigte Format angepaßt, damit sie direkt weiterverarbeitet werden können.

In der Abbildung 5.3.2 ist die Oberfläche von *CUTE* zu sehen, wobei die Daten des Würzburger Gebiets dargestellt werden. In dieser Abbildung können die bereits beschriebenen Schritte zur Erzeugung der Demand-Nodes nochmals nachvollzogen werden. Aus den links oben abgebildeten geographischen Daten, die in einem Vektorformat vorliegen, wird die rechts oben abgebildete Matrix über die Landnutzungsklassen erzeugt. Der Verkehrsschätzer berechnet daraus die rechts unten dargestellte Verkehrsmatrix. Durch das Verfahren des Zusammenfassens von Knoten werden daraus die links unten gezeichneten Demand-Nodes erstellt. Diese werden auch im Fenster der geographischen Daten angezeigt, um die Repräsentation dieser Daten durch die Demand-Nodes verifizieren zu können.

Das Planungstool *CUTE* wird kontinuierlich weiterentwickelt und mit zusätzlichen Funktionen ausgestattet. In dieser Arbeit kann nur eine sehr frühe Version dieses Tools beschrieben werden. Der interessierte Leser möge sich aktuellere Informationen von [9] erwerben.

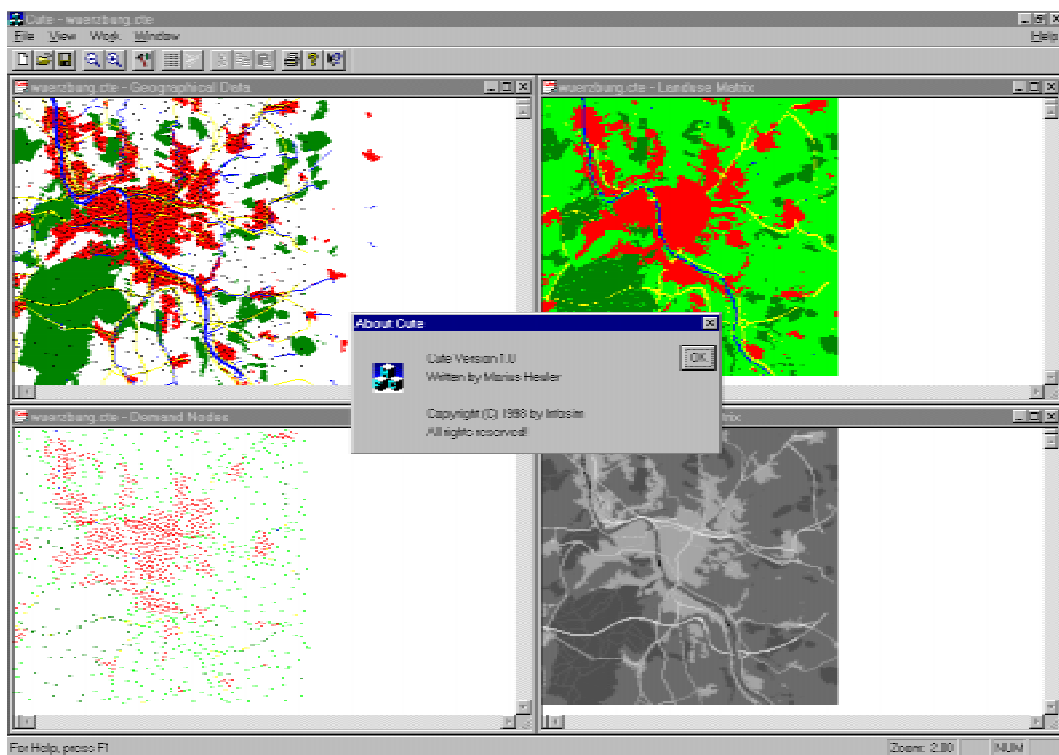


Abbildung 5.3.2: Bildschirmdarstellung von *CUTE*

## 5.4 Senderpositionierung anhand der Demand-Nodes

Eine Möglichkeit zur Positionierung von Sendern ist es, die Positionen von Demand-Nodes direkt als Standorte für Sender zu verwenden. Dazu kann das Verfahren des Zusammenfassens von Knoten herangezogen werden. Um die Positionen der Demand-Nodes direkt als Senderstandorte verwenden zu können, werden genauso viele Demand-Nodes erzeugt, wie Sender geplant sind. Dazu müssen die Parameter des Algorithmus entsprechend gesetzt werden. Die maximale Kapazität eines Senders muß dabei berücksichtigt werden, d.h. der maximale Kommunikationsverkehr der Demand-Nodes darf nicht höher eingestellt werden als die größtmögliche Kapazität eines Senders. Durch die Flächenbeschränkung beim Erzeugen der Demand-Nodes kann die Beschränkung der Reichweite eines Senders simuliert werden. Die Höhe des Kommunikationsverkehrs eines Demand-Nodes gibt dann direkt den zu versorgenden Kommunikationsverkehr dieses Senders an. Die Fläche der Demand-Nodes entspricht dabei der Versorgungsfläche des Senders.

Bei diesem Verfahren wird die Wellenausbreitung zwischen Sender und Teilnehmer überhaupt nicht berücksichtigt. Der wirkliche Versorgungsgrad der Teilnehmer kann aus diesem Grund nicht bestimmt werden. Oft reichen die Daten bei der Planung auch nicht für eine korrekte Modellierung der Wellenausbreitung aus.

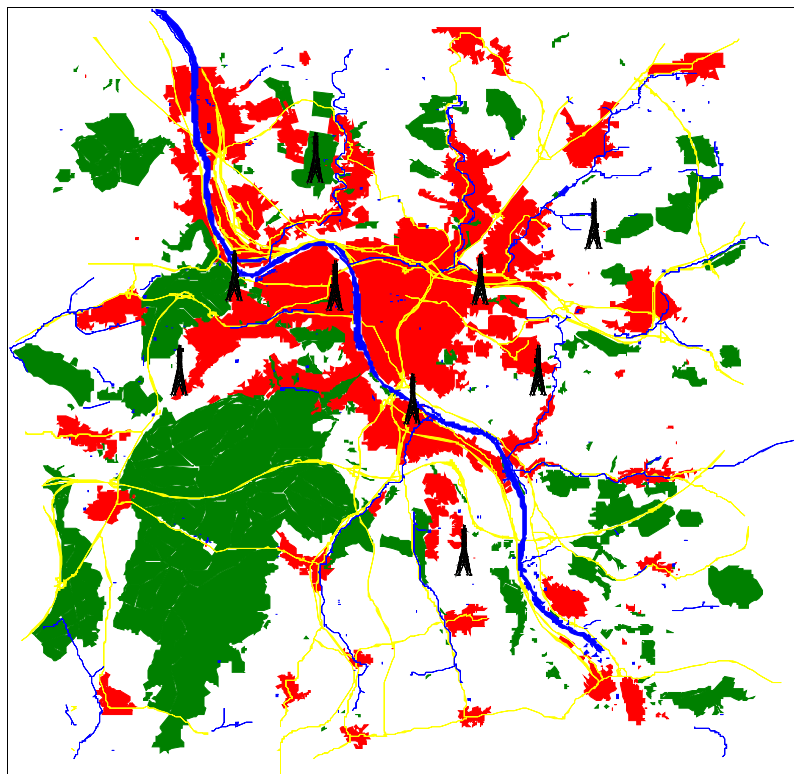


Abbildung 5.4.1: Sender auf Positionen von Demand-Nodes im Würzburger Gebiet

In Abbildung 5.4.1 wird für das Würzburger Gebiet die Lage von Sendern gezeigt, die sich an der Position von neun Demand-Nodes befinden. Die Flächenbegrenzung war dabei abgeschaltet. Der Kommunikationsverkehr wird ziemlich gut versorgt, wobei wegen der fehlenden Flächenbegrenzung die Waldgebiete kaum überdeckt werden. Mit zusätzlichen Sendern und bei Verwendung der Flächenbegrenzung könnten auch diese Gebiete versorgt werden. Die Positionen der Sender sind im Bereich der Stadt Würzburg sinnvoll gewählt: ein Sender in der Innenstadt wird von fünf Sendern am Rande des Stadtgebiets flankiert, die auch den Kommunikationsverkehr der großen stadtnahen Gemeinden versorgen.

Eine andere Möglichkeit ist es, wesentlich mehr mögliche Positionen von Sendern zu bestimmen, als für die endgültige Versorgung des Gebiets vorgesehen sind. Dabei können entweder die Positionen der Demand-Nodes als potentielle Standorte herangezogen werden, oder man positioniert die Sender unabhängig von den Demand-Nodes z.B. gemäß eines Rasters. Verschiedene Möglichkeiten dazu sind in [15] zu finden. Als Ergebnis wird eine Auswahl der Sender geliefert, die aufgrund gegebener Bedingungen optimal ist. Das Problem der Auswahl entspricht dabei genau der Problemstellung, eine maximale Überdeckung von Punkten mit Hilfe gegebener Flächen vordefinierter Form zu finden. Dies ist eine bekannte Fragestellung aus der Graphentheorie, die in der allgemeinen Form NP-vollständig ist. Deshalb werden meist Heuristiken, wie z.B. die in dem in Kapitel 5.2 beschriebenen Tool *ICEPT*, zur Planung herangezogen. Zusätzlich müssen die physikalischen Bedingungen der Wellenausbreitung, wie z.B. Interferenz, berücksichtigt werden.

Bei dem neuen Zugriffsverfahren *CDMA*, welches in [19] und [31] beschrieben wird, gibt es kaum Interferenz zwischen den einzelnen Kanälen. Das gesamte verfügbare Frequenzspektrum wird dort von allen Teilnehmern und Sendern gemeinsam genutzt. Durch die Verwendung von orthogonalen Codes bei der Modulation ist es trotzdem möglich, aus den über dem gesamten Spektrum verteilten Signalen die Daten jedes einzelnen Teilnehmers zu separieren. Dazu ist es aber nötig, daß alle Teilnehmer vom Sender mit ungefähr gleicher Feldstärke empfangen werden. Wegen der gemeinsamen Nutzung des gesamten Spektrums ist auch kein Wiederholungsabstand zwischen Sendern gleicher Frequenz, wie beim *GSM*-System, mehr nötig. Insgesamt ist bei *CDMA* die Planung in wesentlichen Punkten einfacher, wenngleich zusätzliche Aspekte berücksichtigt werden müssen. Das Positionieren von Sendern anhand des Kommunikationsverkehrs der Demand-Nodes erscheint gerade deshalb bei *CDMA* sehr erfolgversprechend. Die Wichtigkeit dieses Zugriffsverfahrens zeigt sich auch durch seine Verwendung im zukünftigen Mobilfunksystem der dritten Generation *UMTS* (*Universal Mobile Telecommunications System*). Bereits im Einsatz ist es im amerikanischen Mobilfunkstandard *IS-95*.

## 6 Zusammenfassung und Ausblick

In dieser Arbeit hat sich gezeigt, daß in vielen Bereichen Objekte nicht gleichmäßig verteilt auftreten, sondern meist Anhäufungen von Objekten mit ähnlichen Eigenschaften vorliegen. Diese Clusterung trifft nach [16] für die Bevölkerungsverteilung, nach [33] für den Datenstrom in Netzwerken und schließlich auch für den Verkehrsbedarf in Kommunikationsnetzen zu, wie in dieser Arbeit gezeigt wird. Die bei Telefongesprächen gültige Annahme einer Poisson Verteilung der Gesprächszeitpunkte ist in oben genannten Bereichen unzutreffend, da dort starke Abhängigkeiten der Objekte untereinander vorliegen. Falls diese Clusterung nicht berücksichtigt wird, können vollkommen falsche Ergebnisse erzielt werden, wie in Kapitel 5.1 gezeigt wird. Um die Clusterung leicht zu modellieren, ist es immens wichtig, die Eigenschaften der Clusterung auf einfache Weise zu beschreiben. Dazu ist das in Kapitel 2.2 eingeführte Konzept der Demand-Nodes sehr hilfreich.

Mit Hilfe dieser Demand-Nodes kann die Clusterung von Objekten direkt durch eine Verteilung von Punkten dargestellt werden. Obwohl der Begriff Demand-Node aus dem Bereich Mobilfunk stammt, ist das Konzept universell für die meisten Problemstellungen der Standortplanung einsetzbar. In allen Fällen, in denen ein bestimmter ortsabhängiger Bedarf durch eine Anzahl Versorger bedient werden muß, ist es direkt verwendbar.

Zur Erzeugung der Demand-Nodes sind spezielle Algorithmen nötig, die maßgeblich über die Approximationsgüte der Verkehrsverteilung durch die erstellten Demand-Nodes entscheiden. Dazu wurden zwei Verfahren, das „Rekursive Partitionieren“ und das „Zusammenfassen von Knoten“, vorgestellt und detailliert analysiert. Das rekursive Partitionieren weist entscheidende Mängel auf, die seine Einsatzmöglichkeiten restringieren. Das im Rahmen dieser Arbeit entwickelte Verfahren des Zusammenfassens von Knoten beseitigt diese Einschränkungen und erschließt neue Anwendungsgebiete für das Konzept der Demand-Nodes. In Kapitel 5.4 wird ein Konzept zur Planung mobiler Kommunikationssysteme auf Basis von Demand-Nodes gezeigt. In dem in Kapitel 5.2 beschriebenen Planungstool *ICEPT* wird dieses Konzept erfolgreich angewandt. Eine allgemeinere Anwendung der Demand-Nodes ist die Abschätzung des Bedarfs, die in dem in Kapitel 5.3 beschriebenen Werkzeug *CUTE* eingesetzt wird.

Zur Bewertung der Güte der Demand-Nodes, d.h. der Genauigkeit der Repräsentation der Clusterung, ist ein quantitatives Gütemaß hilfreich. Dazu wurde in der Arbeit ein aus der räumlichen Statistik bekanntes Verfahren eingeführt, welches mit Hilfe von zufällig positionierten Testflächen ein Vergleichsmaß liefert. Mit diesem Fehlermaß ist ein quantitativer Vergleich differenter Ausprägungen von Demand-Nodes möglich. Der Vergleich beider Verfahren anhand unterschiedlicher Kriterien weist auch quantitativ essentielle Vorteile des Algorithmus des Zusammenfassens von Knoten nach. Die Approximationsgüte der Clusterung durch die Demand-Nodes ist bei diesem Verfahren wesentlich höher.

Insgesamt ist dieses neu entwickelte Clusterungsverfahren für diverse Problemstellungen besser geeignet als das rekursive Partitionieren.

Für die Mobilfunkplanung können die Demand-Nodes nutzbringend zur Modellierung des Benutzerverkehrs eingesetzt werden, wie in den Kapiteln 5.2 und 5.4 gezeigt wird. Die Planung von Senderstandorten anhand des durch die Demand-Nodes repräsentierten Kommunikationsverkehrs wird dort erfolgreich durchgeführt. Gerade auch für die Planung von mobilen Kommunikationssystemen der dritten Generation wie *UMTS*, die das *CDMA*-Zugriffsverfahren benutzen, ist die Verwendung von Demand-Nodes zur Verkehrsrepräsentation ein sinnvolles Mittel.

In [9] und Kapitel 5.3 wird das Programm *CUTE* zur Abschätzung des Verkehrsbedarfs in einem Planungsgebiet vorgestellt. Dieses Tool verwendet den Algorithmus des Zusammenfassens von Knoten, um Demand-Nodes aufgrund geographischer Daten zu erzeugen. Durch den allgemeinen Ansatz mit Adaptionmöglichkeiten an unterschiedliche Anforderungen, kann es zur Planung des Bedarfs bei mannigfaltigen Problemstellungen verwendet werden.

Zukünftig soll das Konzept der Demand-Nodes auch für die Planung von IP-Netzwerken dienen, wobei hier der Bedarf an Bandbreite eines Übertragungsmediums durch die Demand-Nodes dargestellt wird. Bei der Planung eines neuen Kommunikationsnetzes kann damit die benötigte Kapazität einzelner Leitungen dimensioniert werden. Noch wichtiger ist wohl der Fall, bei einem bestehenden Netz die Auslastung oder verschiedene Dienstqualitäten in Abhängigkeit von unterschiedlichen Annahmen über den Bedarf der Teilnehmer zu berechnen. Durch das Konzept der Demand-Nodes ist es leicht möglich, z.B. einen höheren Bedarf an Kommunikationsverkehr der Teilnehmer anzunehmen und für diesen Fall obige Größen des Kommunikationsnetzes zu bestimmen.

Zusammengefaßt ist das Konzept der Demand-Nodes in vielen Bereichen auch außerhalb des Mobilfunks einsetzbar, um den Bedarf von Teilnehmern zu modellieren, wobei die Clusterung des Bedarfs berücksichtigt wird. Die Clusterung, die hier vor allem die Ungleichmäßigkeit der Verteilung des Bedarfs beschreibt, hat in vielen Bereichen einen starken Effekt, der nicht vernachlässigt werden darf. Das neu entwickelte Clusterungsverfahren des Zusammenfassens von Knoten erweitert das Einsatzgebiet der damit erstellten Demand-Nodes erheblich. Neben den bereits erwähnten Einsatzmöglichkeiten wird dieses Verfahren in der Zukunft weitere neue Anwendungsgebiete für das Konzept der Demand-Nodes erschließen.

## 7 Literaturverzeichnis

- [1] ATKIS (1991): Amtliches Topographisches Kartographisches Informationssystem. Bayerisches Landesvermessungsamt. München.
- [2] Cressie, N. A. (1991): *Statistics for spatial data*. New York: Wiley.
- [3] Dubes, R. und A. K. Jain (1988): *Algorithms for Clustering Data*. Prentice Hall advanced reference series. New Jersey: Prentice Hall
- [4] Erlang, A. K. (1948): Solution of some problems in the theory of probabilities of significance in automatic telephone exchanges. *A.T.S. Transactions of the Danish Academy of Technical Sciences Vol. II*, Seite 138-155. Copenhagen.
- [5] Fritsch, Th., K. Tutschku und K. Leibnitz (1995): Field strength prediction by ray-tracing for adaptive base station positioning in mobile communication networks. *Proceedings of the 2nd ITG Conference on Mobile Communication '95*. Neu-Ulm.
- [6] Gamst, A., E.-G. Zinn, R. Beck und R. Simon (1986): Cellular radio network planning. *IEEE Aerospace and Electronic Systems Magazine 1*, Seite 8-11.
- [7] Ghosh, A. und S. L. McLafferty (1987): *Location strategies for retail and service firms*. Massachusetts: Lexington Books.
- [8] Hata, M. (1980): Empirical formula for propagation loss in land mobile radio services. *IEEE Transactions on Vehicular Technology 29(3)*, Seite 317-325.
- [9] InfoSim (1998): Customer Traffic Estimation Tool CUTE. Internetseite: <http://www.infosim-usa.com>. Irvine: InfoSim USA Inc.
- [10] Kleinrock, L. (1975): *Queueing Systems, Vol. 1: Theory*. New York: Wiley.
- [11] Krüger, M. und R. Beck (1990): GRAND – Ein Programmsystem zur Funknetzplanung. *Philips: Technische Mitteilung 2*, Seite 7-12.
- [12] Lambrecht, F. und A. Baier (1994): Methodik der Funknetzplanung für zellulare Mobilfunksysteme. 2. *ITG-Fachtagung Mobile Kommunikation*, Seite 517-524.
- [13] Leibnitz, K. und K. Tutschku (1998): Internetseite des Planungssystems ICEPT (Integrated network CELLular Planning Tool): <http://www-info3.informatik.uni-wuerzburg.de/icept>.



- [14] Leskien, T. (1997): Erfassung und Charakterisierung geographischer Daten für die Planung von Mobilfunksystemen. Diplomarbeit am Lehrstuhl für Informatik III. Universität Würzburg, Institut für Informatik.
- [15] Liebler, P. (1998): Optimization algorithms for the design of cellular communications systems. Diplomarbeit am Lehrstuhl für Informatik III. Universität Würzburg, Institut für Informatik.
- [16] Mannarsalo, P., A. Koski und I. Norros (1998): Telecommunication networks and multifractal analysis of human population distribution. *COST257*.
- [17] MAPINFO: Internetseite: <http://www.mapinfo.com>. NY: MapInfo Corporation.
- [18] Mobile Systems International Ltd (1996): *PLANET Technical reference guide*. MSI.
- [19] Rappaport, T. S. (1996): *Wireless Communications – Principles & Practice*. New Jersey: Prentice Hall.
- [20] Simon, R. et al. (1985): Rechnergestützter integrierter Funknetzentwurf für zellulare Funkssysteme. *NTG Fachtagung Bewegliche Funkdienste*, Seite 90-95. München.
- [21] Stoyan, D., W. Kendall und J. Mecke (1992): *Stochastic Geometry and Its Applications*. Chichester: Wiley.
- [22] Tran-Gia, P. (1982): *Überlastprobleme in rechnergesteuerten Fernsprechvermittlungssystemen. Modellbildung und Analyse*. Doktorarbeit, Universität-Gesamthochschule-Siegen.
- [23] Tran-Gia, P. (1988): *Zeitdiskrete Analyse verkehrstheoretischer Modelle in Rechner- und Kommunikationssystemen*. Universität Stuttgart, Institut für Nachrichtenvermittlung und Datenverarbeitung.
- [24] Tran-Gia, P. (1996): *Analytische Leistungsbewertung verteilter Systeme*. Berlin: Springer.
- [25] Tran-Gia, P. und M. Mandjes (1996): Modeling of customer retrial phenomenon in cellular mobile networks. Forschungsbericht. Preprint-Reihe Nr. 142. Universität Würzburg, Institut für Informatik.
- [26] Tran-Gia, P. und N. Gerlich (1996): Impact of customer clustering on mobile network performance. Forschungsbericht. Preprint-Reihe Nr. 143. Universität Würzburg, Institut für Informatik.

- [27] Tutschku, K. (1998): Demand-based radio network planning of cellular mobile communication systems. *IEEE Infocom 1998*. San Francisco.
- [28] Tutschku, K. und P. Tran-Gia (1998): Spatial traffic estimation and characterization for mobile communication network design. *IEEE Journal of Selected Areas in Communications – Special issue on advances in computational aspects of teletraffic models*. Erscheint Herbst 1998.
- [29] Tutschku, K., K. Leibnitz und P. Tran-Gia (1997): *ICEPT* – An integrated cellular network planning tool. *IEEE Conference on Vehicular Technology 1997*. Phoenix.
- [30] Tutschku, K., N. Gerlich und P. Tran-Gia (1996): An integrated approach to cellular network planning. *Proceedings of the 7th International Network Planning Symposium (Networks 96)*. Sydney.
- [31] Viterbi, A. J. (1995): *CDMA – Principles for spread spectrum communications*. MA: Addison Wesley
- [32] Vohra, R. V. und N. G. Hall (1993): A probabilistic analysis of the maximal covering location problem. *Discrete Applied Mathematics* 43, Seite 175-183.
- [33] Willinger, W., W. E. Leland, M. S. Taqqu und D. V. Wilson (1993): On the self-similar nature of ethernet traffic. *ACM SIGComm* 93. San Francisco.

# 8 Abbildungsverzeichnis

Abbildung 2.1.1: Phasen der Planung eines mobilen Kommunikationssystems.....	8
Abbildung 2.2.1: Diskretisierung einer kontinuierlichen Funktion zu einer Matrix.....	14
Abbildung 2.2.2: Landnutzungsklassen des Würzburger Gebiets.....	15
Abbildung 2.2.3: Verkehrsdichte des Würzburger Gebiets .....	17
Abbildung 2.3.1: Verkehrsdichte des Dallas Gebiets .....	19
Abbildung 2.3.2: Verkehrsdichte des synthetischen Gebiets .....	20
Abbildung 3.1.1: 1. Schritt des Algorithmus.....	24
Abbildung 3.1.2: 2. Schritt des Algorithmus.....	25
Abbildung 3.1.3: Ergebnis des Algorithmus für $T = 4$ Teilungen .....	26
Abbildung 3.1.4: Position des Demand-Nodes.....	27
Abbildung 3.1.5: Bestimmung der Landklasse .....	27
Abbildung 3.1.6: Demand-Nodes des Würzburger Gebiets .....	29
Abbildung 3.1.7: Demand-Nodes des Dallas Gebiets .....	29
Abbildung 3.1.8: Demand-Nodes des synthetisches Gebiets .....	30
Abbildung 3.1.9: Demand-Nodes des konstanten Gebiets .....	30
Abbildung 3.2.1: Verkehrsmatrix.....	34
Abbildung 3.2.2: Algorithmus des Zusammenfassens von Knoten .....	34
Abbildung 3.2.3: Zufällige Auswahl eines Knotens.....	35
Abbildung 3.2.4: Nachbarn eines Knoten .....	36
Abbildung 3.2.5: Suchfolge.....	36
Abbildung 3.2.6: Zum Vereinigen ausgewählter Knoten $N$ .....	37
Abbildung 3.2.7: Aus der Vereinigung hervorgegangener neuer Knoten $Z$ .....	38
Abbildung 3.2.8: Position des Demand-Nodes.....	39
Abbildung 3.2.9: Wahl der Landklasse .....	40
Abbildung 3.2.10: Demand-Nodes des Würzburger Gebiets .....	41
Abbildung 3.2.11: Demand-Nodes des Dallas Gebiets .....	41
Abbildung 3.2.12: Demand-Nodes des synthetischen Gebiets .....	42
Abbildung 3.2.13: Demand-Nodes des konstanten Gebiets .....	42
Abbildung 3.2.14: Verkehrshistogramm des Würzburger Gebiets .....	44
Abbildung 3.2.15: Verkehrshistogramm des Dallas Gebiets .....	45
Abbildung 3.3.1: Algorithmus mit Optimierung .....	46
Abbildung 3.3.2: Algorithmus des Löschens von verkehrsarmen Knoten.....	47
Abbildung 3.3.3: Algorithmus für zu optimierenden Knoten .....	48
Abbildung 3.3.4: Beispiel für zwei zu löschende Knoten.....	49
Abbildung 3.3.5: Gebiet nach der Entfernung der zwei Knoten .....	49
Abbildung 3.3.6: Würzburger Gebiet, mit $T_{min} = 0.6 \cdot T_{max}$ , $R_{Tmax} = 1.1$ .....	51
Abbildung 3.3.7: Dallas Gebiet, mit $T_{min} = 0.6 \cdot T_{max}$ , $R_{Tmax} = 1.1$ .....	51
Abbildung 3.3.8: Algorithmus der Optimierung von Flächenelementen .....	53
Abbildung 3.3.9: Beispiel zur Feldauswahl der Optimierung.....	55
Abbildung 3.3.10: Felder nach der Optimierung.....	55
Abbildung 3.3.11: Würzburger Gebiet, mit Parametern gemäß Tabelle 3.3.1 .....	56
Abbildung 3.3.12: Dallas Gebiet, mit Parametern gemäß Tabelle 3.3.1 .....	57
Abbildung 3.3.13: Würzburger Gebiet, verbesserte Optimierung.....	59
Abbildung 3.3.14: Dallas Gebiet, verbesserte Optimierung .....	59
Abbildung 3.3.15: Konstantes Gebiet.....	60
Abbildung 3.3.16: Demand-Nodes des Würzburger Gebiets .....	63

Abbildung 3.3.17: Demand-Nodes des Dallas Gebiets .....	63
Abbildung 4.1: Kommunikationsverkehr eines Testkreises auf der Matrix.....	66
Abbildung 4.2: Kommunikationsverkehr der Demand-Nodes eines Testkreises .....	66
Abbildung 4.1.1: Mögliche Positionen von Testkreisen.....	68
Abbildung 4.1.2: Kreisförmige und quadratische Testfläche .....	69
Abbildung 4.1.3: Algorithmus der Berechnung einer Meßreihe.....	70
Abbildung 4.2.1: Maximaler quadratischer Fehler im Würzburger Gebiet .....	76
Abbildung 4.2.2: Mittlerer relativer Fehler im Würzburger Gebiet .....	76
Abbildung 4.2.3: Vergleich beider Verfahren im Würzburger Gebiet .....	77
Abbildung 4.2.4: Vergleich beider Verfahren im Dallas Gebiet .....	78
Abbildung 4.2.5 Vergleich beider Verfahren im konstanten Gebiet.....	79
Abbildung 4.2.6: Vergleich mit quadratischer Testfläche im synthetischen Gebiet.....	80
Abbildung 4.2.7: Vergleich mit quadratischer Testfläche im konstanten Gebiet.....	80
Abbildung 4.2.8: Vergleich verschiedener Parameter im Würzburger Gebiet.....	83
Abbildung 4.2.9: Vergleich verschiedener Parameter im Dallas Gebiet.....	83
Abbildung 4.2.10: Vergleich verschiedener Parameter im synthetischen Gebiet .....	84
Abbildung 4.2.11: Vergleich verschiedener Parameter im konstanten Gebiet.....	85
Abbildung 4.2.12: Vergleich der Demand-Node Anzahl beim Zusammenfassen .....	86
Abbildung 4.2.13: Vergleich der Demand-Node Anzahl beim Partitionieren.....	86
Abbildung 4.2.14: Vergleich der Demand-Node Anzahl beim Partitionieren.....	87
Abbildung 4.2.15: Vergleich der Demand-Node Anzahl beider Verfahren.....	88
Abbildung 4.2.16: Würzburger Gebiet, verbesserte Optimierung mit Flächenbegrenzung der Demand-Nodes auf $0.2 \text{ km}^2$ .....	89
Abbildung 4.2.17: Demand-Nodes des Würzburger Gebiets mit Flächenbegrenzung der Demand-Nodes auf $0.2 \text{ km}^2$ .....	90
Abbildung 4.2.18: Untersuchung der Flächenbegrenzung bei kleiner Testfläche.....	91
Abbildung 4.2.19: Untersuchung der Flächenbegrenzung bei großer Testfläche.....	92
Abbildung 4.3.1: Landnutzungsklassen in der Matrix.....	94
Abbildung 4.3.2: Landnutzungsklassen der Demand-Nodes .....	94
Abbildung 4.3.3: Vergleich der Landnutzungsklassen beider Verfahren .....	96
Abbildung 4.3.4: Vergleich bei verschiedener Anzahl der Demand-Nodes .....	97
Abbildung 5.1.1: Modell einer Zelle mit Clusterung der Teilnehmer.....	99
Abbildung 5.1.2: Beispiel für zufällige Testkreise .....	103
Abbildung 5.1.3: Subjektive Dienstgüte bei $E[X] = 30$ , $K = 7$ , Würzburger Gebiet.....	104
Abbildung 5.1.4: Subjektive Dienstgüte bei $E[X] = 30$ , $K = 7$ , Dallas Gebiet.....	105
Abbildung 5.1.5: Modell einer Zelle mit Wiederholungsversuch.....	106
Abbildung 5.1.6: Subjektive Dienstgüte mit Wiederholungsversuch .....	106
Abbildung 5.2.1: Schematischer Aufbau eines mobilen Kommunikationsnetzes.....	108
Abbildung 5.2.2: Hauptentwurfskriterien mobiler Kommunikationssysteme .....	109
Abbildung 5.2.3: Bildschirmdarstellung von ICEPT.....	110
Abbildung 5.2.4: Versorgungsgebiete des Würzburger Gebiets, [29].....	111
Abbildung 5.3.1: Aufbau von CUTE.....	112
Abbildung 5.3.2: Bildschirmdarstellung von CUTE.....	113
Abbildung 5.4.1: Sender auf Positionen von Demand-Nodes im Würzburger Gebiet.....	114