

## 제 II 부 교통존설정체계 구축

---

제1장 서 론

제2장 교통존 설정 개요

제3장 교통분석유형별 교통존 설정  
유형 분석 및 교통망과의 연동방안

제4장 지리정보시스템을 이용한 교통존  
설정 기법

제5장 결 론

# 목 차

## 제 II 부

<b>제1장 서론</b> .....	1
제1절 연구의 배경 및 목적 / 3	
제2절 연구의 구성 및 결과 / 3	
<b>제2장 교통존 설정 개요</b> .....	5
제1절 교통존의 정의, 존재이유 / 7	
제2절 교통존의 일반적인 구분원칙 / 7	
제3절 교통존 설정에 따른 결과의 상이성 / 8	
<b>제3장 교통분석유형별 교통존 설정 유형 분석및 교통망과의 연동방안</b> .....	11
제1절 거시적 국가교통계획에서의 교통존 설정 / 13	
제2절 미시적 교통영향평가에서의 교통존 설정 / 15	
제3절 교통망과 연동한 교통존설정: 이론과 실제 / 17	
<b>제4장 지리정보시스템을 이용한 교통존설정 기법</b> .....	21
제1절 개 요 / 23	
제2절 이론적 배경 / 25	
제3절 TAZ 구성 체계의 개발 / 31	
제4절 모형간 특성비교 / 41	
<b>제5장 결 론</b> .....	45
제1절 결 론 / 47	
제2절 향후과제 / 48	
<b>참고문헌</b> .....	49

## 표 목 차

<표 3- 1>	서울시 교통센서스에서의 소존 구분 현황 .....	13
<표 3- 2>	서울시 센서스에서의 대존 구분 현황 .....	14
<표 3- 3>	실시 설계에서의 교통존 구분내역 .....	16
<표 3- 4>	존과 네트워크의 관계 .....	20
<표 4- 1>	TAZ 구성 체계 내의 실행모듈 .....	32
<표 4- 2>	TAZ 구성 체계에서 이용하는 데이터 파일 .....	33
<표 4- 3>	응집성 강행 모형 .....	34
<표 4- 4>	행정구역 강행 모형 .....	35
<표 4- 5>	응집성 및 행정구역 강행 기준값 .....	40
<표 4- 6>	수정된 시뮬레이터드 애닐링 실행결과 .....	41
<표 4- 7>	수정된 Ward 알고리즘 실행결과 .....	42
<표 4- 9>	BSU 수 변화 허용비가 다를 때 알고리즘 수행결과 .....	43
<표 4-10>	수정된 알고리즘의 수행시간 비교 .....	43

## 그림목차

<그림 2-1> 콘의 centroid 위치 .....	9
<그림 2-2> 3콘체제에서의 교통량 .....	9
<그림 2-3> 5콘체제에서의 교통량 .....	9
<그림 2-4> 3콘체제시의 결과 .....	10
<그림 2-5> 5콘체제시의 결과 .....	10
<그림 3-1> Network for Regional Analysis .....	18
<그림 3-2> Network for Corridor Analysis .....	19
<그림 4-1> 계층적 군집형성을 위한 트리 .....	28
<그림 4-2> TAZ 구성 체계 개념도 .....	32
<그림 4-3> 수정된 Ward 알고리즘 수행결과 .....	44
<그림 4-4> 수정된 시뮬레이티드 애닐링과 k-means 알고리즘 수행결과 .....	44

## 제1장 서론

---

제1절 연구의 배경

제2절 연구의 구성 및 결과

## 제1절 연구의 배경 및 목적

최근의 교통개발연구원의 주관하에 전국교통 DB 사업이 활발히 수행되고 있으며 이에 따라서 교통수요예측에서의 존의 합리적 구성을 통한 보다 정확할 수 있는 현실적인 교통수요가 요구되고 있다고 할 수 있다. 특히 DB사업과 관련하여 수치지도의 연계사업등이 진행되고 있는만큼 수요로서의 OD와 공급으로서의 네트워크가 함께 진전을 이루고 있다고 할 수 있다.

본 연구에서는 TAZ의 일반적 원칙을 상기해 보고, 제반 교통사업에서의 존구성체계의 적용상황을 아울러 살펴보고 과연 문제점이 무엇인지 알아보고, 이러한 문제를 해결할 수 있는 방안으로서 지리정보시스템을 이용한 존구성체계의 프로토타입개발을 통해서 향후 교통 DB가 보다 원만히 교통분석에 활용될 수 있는 방안을 제시해 보는데 있어서 기초가 될 만한 예비연구를 해 보는데 그 목적이 있다.

## 제2절 연구의 구성 및 결과

연구의 구성은 존의 구획이유, 교과서적인 지식으로서 구획원칙을 review하고, 현재까지의 교통존분석체계를 검토하는 부분, 그리고 지리정보시스템을 이용하여 dynamic 하게 교통존의 형성을 도모할 수 있는 프로토타입의 제시로 연구가 구성되며 구체적인 연구의 결과는 지리정보시스템을 활용한 교통존구성체계의 제시가 일차적인 연구결과이고 부수적으로 활용방안을 제시하는 것이라고 할 수 있다.

## 제2장 교통존 설정 개요

---

제1절 교통존의 정의, 존재이유

제2절 교통존의 일반적인 구분원칙

제3절 교통존 설정에 따른 결과의 상이성

## 제1절 교통존의 정의, 존재이유

- 교통존이란 교통계획, 특히 수요의 예측부분에서 승객이나 화물이동과 흐름을 분석하고 추정하기 위해 (현상을 모형화 하기 위해) 같은 단위공간을 설정하는데 이를 교통존이라 한다. 교통존은 교통분석대상지역에 인위적으로 경계를 그어 각 교통지구의 사회경제적 특성 및 교통여건을 파악하고 이를 기본으로 하여 자료의 수집, 분석, 예측을 하기 위한 단위이다.
- 따라서 효과적인 계획의 작성 및 수요의 분석을 위해서는 과업대상지역을 교통존으로 적절히 분할하여야 한다.

## 제2절 교통존의 일반적인 구분원칙

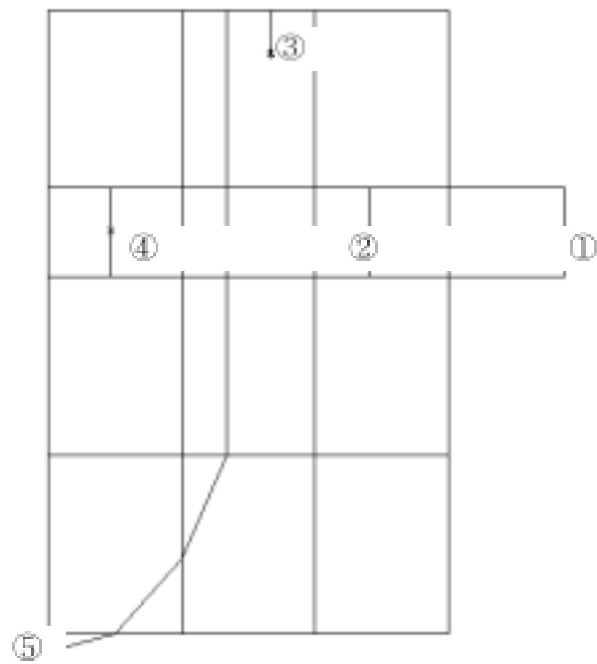
- 일반적으로 교통존 설정시 주의 사항은 아래와 같다. (Garber and Hoel, 1988)
  - ① 존규모에 대해서는 존간 통행 발생량이 비교적 균등하게 한다.
  - ② 도로용량과 존의 통행 발생량을 조선에 배정했을 때의 통행량과의 관계를 고려하여야 한다.
  - ③ 존구획선은 지형의 연속성을 고려하고 교통시설이 공통적이며 통행특성과 토지 이용이 균질 하도록 결정하여야 한다.
  - ④ 사회경제 자료의 이용을 용이하게 하기 위하여 센서스 조사체계를 상당부분 참조하여야 한다.
  - ⑤ 존의 규모와 수는 조사와 분석에 소요되는 신뢰도 및 비용, 시간을 고려하여 연구특성에 따라서 적정하게 결정하여야 한다.
- 특히 상기의 존 설정시 주의사항 하에서의 보다 세부적으로 요구되어지는 일반적인 교통존의 구획의 원칙은 아래와 같다.
  - ① 존은 모든 활동이 zone centroid에 집중되는 가정에 의해 발생하는 오차가 크지 않도록 결정되어야 한다.
  - ② 존의 형태는 크기는 일정하지 않더라도 zone centroid와 연결이 용이한 형태가 되어야 하며 가능한 상태가 되어야 하며 가능한 한 존에 통학권이 포함되도록 하여 zone centroid 접근 비용이 주요시설에 접근하는 비용을 대표할 수 있게 한다.
  - ③ 행정구역 및 센서스 조사존과 가급적 일치시킨다.
  - ④ 존내 토지이용과 인구수 및 인구구성이 동질해야 한다.



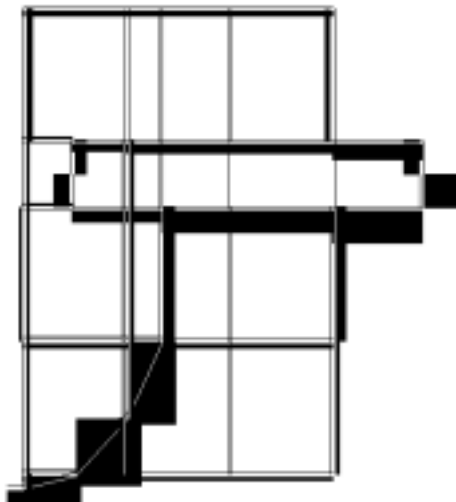
- ⑤ 도로가 가급적 존의 경계선과 일치하도록 하고 기존 존 체계의 코든, 스크린 라인에 적합해야한다.
- 존을 설정할 때 도시 전체를 다루는 도시종합교통계획이나 지하철로선로계획과 같은 경우에는 존의 규모가 커지게 마련이지만 가로망 계획 도로건설의 타당성 검토, 버스노선계획등과 같은 비교적 도시의 일부분을 대상으로 하는 교통계획에는 존을 작게 할 필요성이 있다. 또한 인구밀도가 낮은 교외지역은 존의 규모를 크게 해야 하고 도심지를 포함하는 도시 내부에서는 존을 작게 하는 것이 일반적이다.
- 도시종합교통계획에서는 존을 대중소로 나누어 분석 목적에 따라 적절히 활용하는데 이를 분류하는 방법은 다음과 같다.
  - ① 행정구역 혹은 도시 계획 구역에 의하던가 혹은 하천, 철도, 입야 등과 같이 지형적 조건에 따라 대존으로 나눈다.
  - ② 대존을 도시의 개발축과 도시기본계획 등에서 나누어진 생활권을 고려하여 지역특성이 유사한 것끼리 중존으로 나눈다.
  - ③ 중존은 다시 소존으로 나눈다.
- 조사 대상지역내의 교통존이 설정되면 조사대상지역 외의 도시외곽 지역을 존으로 나눈다. 이는 폐쇄선을 통과하는 통행을 조사하기 위한 것이다. 교통존이 설정되면 폐쇄선 내측에 대하여 가구를 표본 추출하여 가구방문 조사, 직장방문 조사, 차량방문조사등을 실시하게 된다.

### 제3절 교통존 설정에 따른 결과의 상이성

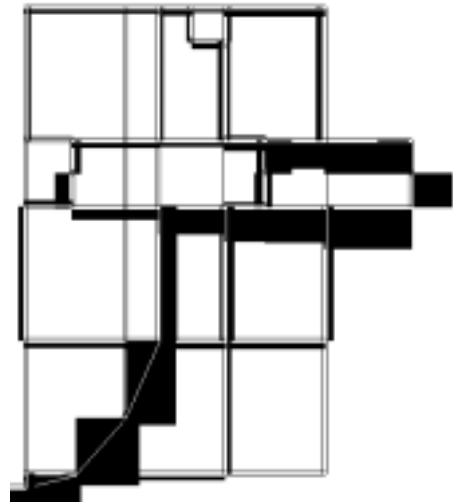
- 교통존의 설정에 따른 추정 가로교통량 (특히 미래의 시점에서)의 상이성이 존재하는 예를 보이기 위해서 다음의 UTOWN의 사례를 들어보기로 한다.
- UTOWN에는 총 5개의 존이 있다. <그림 2-1>은 UTOWN에서의 존의 centroid의 위치를 나타낸다. 즉, 원래의 5개의 존체계를 유지하는 경우와 3개의 존체계로 새로이 구성하였을 때 동일한 OD의 양을 가지고 그 변화의 정도를 알아보기로 한다.
- 새로운 3존체계의 OD는 TRANPLAN에서 \$MATRIX COMPRESS을 이용하여 1번과 2번, 3번과 4번 존을 합하여 새로운 존을 만들고 5번존과 합하여 3개의 존으로 만들었다. 이때 기존의 5존체제에서의 가로교통량과 3존체제에서의 그것을 비교하여 본다. 그림 2와 3은 교통량이 변화된 정도를 보기 쉽게 그림으로 나타내었다. 그림 4와 5는 하나의 교통축을 지나는 교통량의 변화를 보여준다.



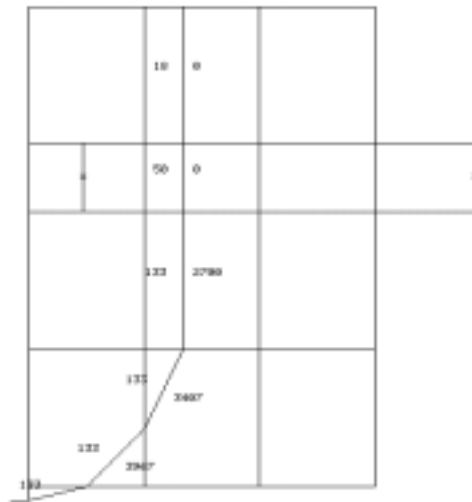
<그림 2-1> 존의 centroid 위치



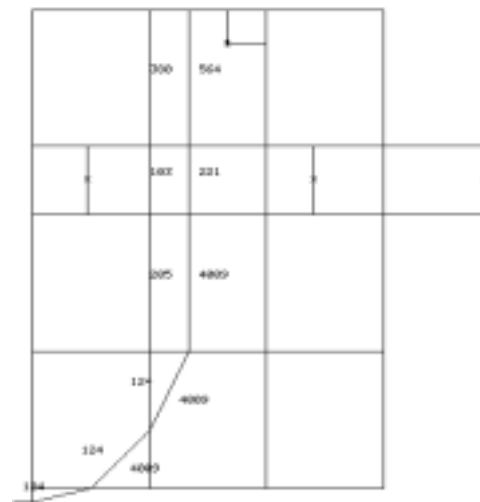
<그림 2-2> 3존체제에서의 교통량



<그림 2-3> 5존체제에서의 교통량



<그림 2-4> 3존체제시의 결과



<그림 2-5> 5존체제시의 결과

- 이상에서 보듯이 존체제의 설정은 가로교통량의 현격한 차이를 발생시킬 수 있음을 알수 있고 현실적인 예측치를 보장받기 위해서는 존체제의 설정에 있어 존과 네트워크의 조화성 (zone-network compatibility)를 고려한 신중한 접근이 요구된다고 할 수 있다.

## 제3장 교통분석유형별 교통존 설정 유형 분석및 교통망과의 연동방안

---

제1절 거시적 국가교통계획에서의 교통존 설정

제2절 미시적 교통영향평가에서의 교통존 설정

제3절 교통망과 연동한 교통존설정

## 제1절 거시적 국가교통계획에서의 교통존 설정

### 1. 서울시 교통센서스에서의 교통존 설정

- 서울시가 서울시정개발연구원과 1996년부터 추진한 『서울시 교통센서스 및 데이터 베이스 구축, 1997.12, 서울특별시』 교통존 설정 방식을 보면, 전국을 총 1,020개의 소존과 45개의 대존으로 나누었다. 소존은 사회경제지표에 관한 자료의 획득이 용이한 행정동(읍·면)단위로 설정하는 것을 원칙으로 하여 내부존으로 서울시의 경우 행정동 단위로 설정하되 1995년 이후 분동된 중랑구 신내1동과 신내 2동을 합하고, 강서구 등촌 1동과 등촌3동을 합하여 526존으로 설정하였으며, 인천시는 1995년 이후 분동된 남구 관교동과 문학동을 합하고, 연구수 동춘 2동과 청량동을 합하여 소존으로 설정하였고, 비조사지역은 행정동을 구(군)단위로 합하여 하나의 소존으로 설정하여 128개존으로 구분하였다. 또한 경기도의 경우도 인천시와 같은 규칙을 적용하여 가구 통행실태 조사지역에 해당하는 1995년 이후 분동지역을 합하여 분동이전 행정동 단위로 구분하고, 비조사지역의 행정동을 구(군)단위로 합하여 362개존으로 구분하였다.
- 한편, 외부존으로는 강원, 충남+충북+대전, 전남+전북+광주, 그리고 경북+경남+대구+부산을 구분하여 4개의 존을 설정하였다.
- 이와 같은 과정을 통하여 교통센서스에서 최종적으로 구축된 소존은 아래의 표와 같다.

<표 3-1> 서울시 교통센서스에서의 소존 구분 현황

지역		소존
내부존	서울시	526개
	인천시	128개
	경기도	362개
외부존		4개
총계		1,020개

- 소존을 묶어 대존을 설정하는데 서울의 경우는 서울시 25개구를 하나의 대존으로 설정하였고 경기도와 인천시는 단일 혹은 지리적, 사회경제적 특성이 유사한 여러개의 구(군)를 묶어서 하나의 대존으로 설정하였다.

- 이와 같은 과정을 통하여 설정된 대존은 내부존은 44개 외부존은 1개로 총 45개이고, 구체적인 대존의 구분현황은 아래 표와 같다.

<표 3-2> 서울시 센서스에서 대존 구분 현황

구분	일련번호	행정구역	구분	일련번호	행정구역
내부존	서울시	1 종로구	서울시	23 강남구	
		2 중구		24 송파구	
		3 용산구		25 강동구	
		4 성동구	인천시	26 중구,동구,남구	
		5 광진구		27 연수구,남동구	
		6 동대문구		28 부평구,계양구	
		7 중랑구		29 서구,강화군,옹진군	
		8 성북구	경기도	30 수원시	
		9 강북구		31 성남시	
		10 도봉구		32 의정부시,양주군	
		11 노원구		33 안양시,과천시,군포시,의왕시	
		12 은평구		34 부천시	
		13 서대문구		35 광명시	
		14 마포구		36 평택시,오산시,화성군	
		15 양천구		37 동두천시,포천군,연천군	
		16 강서구		38 시흥시,안산시	
		17 구로구		39 고양시,파주시,김포군	
		18 금천구		40 구리시,남양주시	
		19 영등포구		41 하남시	
		20 동작구		42 용인시,안성군	
		21 관악구		43 이천시,광주군,여주군	
		22 서초구		44 가평군,안평군	
			외부존	45	강원,충남·북,대전,전남·북,광주,경남·북,대구,부산

## 2. 기타 국가계획에서의 교통존 설정

- 철도청이 발주하고 교통개발연구원이 추가되어 기타 하도급 연구위탁기관이 제시한 「21세기 국가 철도망 구축기본계획수립, 1998.2, 교통개발연구원·철도기술연구원」에서는 광역철도가 교통망에 미치는 영향을 분석할 수 있도록 대도시 광역교통망 대상지역은 각 지역별로 별도로 세분화하였다. 전국을 대상으로 교통존을 설정하기 위해 설정기준은 다음과 같다.
  - 제주도를 제외한 전국지역의 교통영향권을 포함한다.
  - 각 교통존의 최소 단위는 시 또는 군을 원칙으로 하되 '95년 시군통합 개편지역을 감안하여 반영한다.

- 계획중인 타 교통시설물들의 교통영향을 고려할 수 있도록 타 교통계획 및 타당성 조사 등에서 설정한 교통지구체계를 참조한다.
- 위의 설정기준에 따른 교통존 설정결과는 다음과 같다.
  - 소존구분 : 노선대가 통과하거나 그 영향이 예상되는 지역을 세분화하여 구분하되, 특별·광역시에는 각각 1개의 지구로 설정하고 도별로 지역생활권을 감안하여 전체 교통지구는 모두 132개로 설정하였다.
  - 대존구분: 교통수요분석을 위해 생활권 단위의 도별단위의 지구로 조정하고 기타 특별, 광역시를 1개지구로 총 15개의 교통지구를 설정하였다.

### 3. 자치구 교통개선사업에서의 교통존 설정

- 용마에서 실시한 「도봉구 교통개선5개년계획, 1997.2, (주)용마부설교통기술연구원」에서는 직접적인 영향을 받는 1차 영향권역인 도봉구 내부를 주요 간선도로와 지역적인 통행분포특성을 고려하여 20개 지구단위를 우선적으로 「존」으로 설정한다. 도봉구를 제외한 주변의 노원구, 강북구, 의정부시 등 주요 간선도로로 연결되거나 지형상(산지, 하천 등) 및 도시계획 경계선으로 분리될 수밖에 없고 간접적인 영향을 받을 것으로 판단되는 서울시 및 경기도 지역까지를 고려한다. 그러므로 도봉구 내부를 총 20개 「존」으로 구분하고, 그 밖의 서울시 전역과 경기도 전역을 27개 「존」으로 구분하여 총 47개 「존」로 설정하였다.

## 제2절 미시적 교통영향평가에서의 교통존 설정

### 1. 대규모 공공 시설의 영향평가에서의 교통존 설정

- 『인천도시철도 3호선 건설에 따른 교통영향평가, 1998.2, 인천광역시지하철건설본부』에서는 1995년 1월부터 기존의 인천시가 인천광역시로 시역확장이 이루어진 바, 이 영향평가에서는 인천광역시의 행정구역 개편을 반영하여 인천시 내부 126개 권역, 인천시 외부 38개 권역으로 구분, 총 164개 권역으로 교통권역을 설정하였다. 인천시의 행정동 및 인천시 외곽지역에 대해서는 시, 군, 구를 교통권역으로 구분하여 설정하였다. 한편 인천시의 경계지역인 부천시, 김포군, 시흥시에 대해서는 동 및 면 단위로 세분한 지역도 있다. 그리하여 164개의 소존과 이러한 소존을 합하여 17개의 대존으로 구분하였다.

- 한편, 『인천국제공항 전용철도 1단계구간 교통영향평가, 2000.12, 철도청』 교통지구 구분은 자료수집이 용이하도록 적절한 지구크기를 구성하였고, 가능한 한 사회·경제적 특성이 유사한 지역을 하나로 묶었으며, 교통망과 토지이용형태를 고려하여 충분히 반영되도록 구분하였다. 교통지구의 구분은 전국을 대상으로 하되, 인천국제공항 전용철도의 노선을 따라 영향권역을 세분하고 그 밖의 지역은 크게 하나로 구분하였다. 이와 같은 과정을 통해 인천국제공항 전용철도의 직접영향권으로 서울·인천·경기지역을 설정하였으며, 이에 따라 서울은 구단위로 25개존으로 구분하였고, 인천도구를 단위로 하여 10개존으로 구분하였으며, 경기지역은 영향정도에 따라 6개존으로 구분하여 총 41개존으로 설정하고 나머지 기타 권역은 외부존으로 설정하였다.

## 2. 실시 설계에서의 교통존 설정

- 『영동지구~상원리간 도로개설공사 실시 설계 종합보고서, 1999.11, 수원시』에서는 자료의 이용가능성, 지역조건 등을 고려하여 교통권역을 설정하였다. 설정된 본 과업의 교통권역체계는 자료의 이용 가능한 최소단위라 할 수 있는 동단위의 교통권역을 기초단위로 하였고, 동단위에서 본 과업노선 주변지역은 세밀하게 존을 구성하였으며 그 외 지역은 이 지역의 특성에 따라 존을 grouping하여 사용하였다. 그 결과 내부존은 10개 외부존은 8개, 총 18개존으로 교통존을 설정하여 교통수요예측작업을 하였다.

<표 3-3> 실시 설계에서의 교통존 구분내역

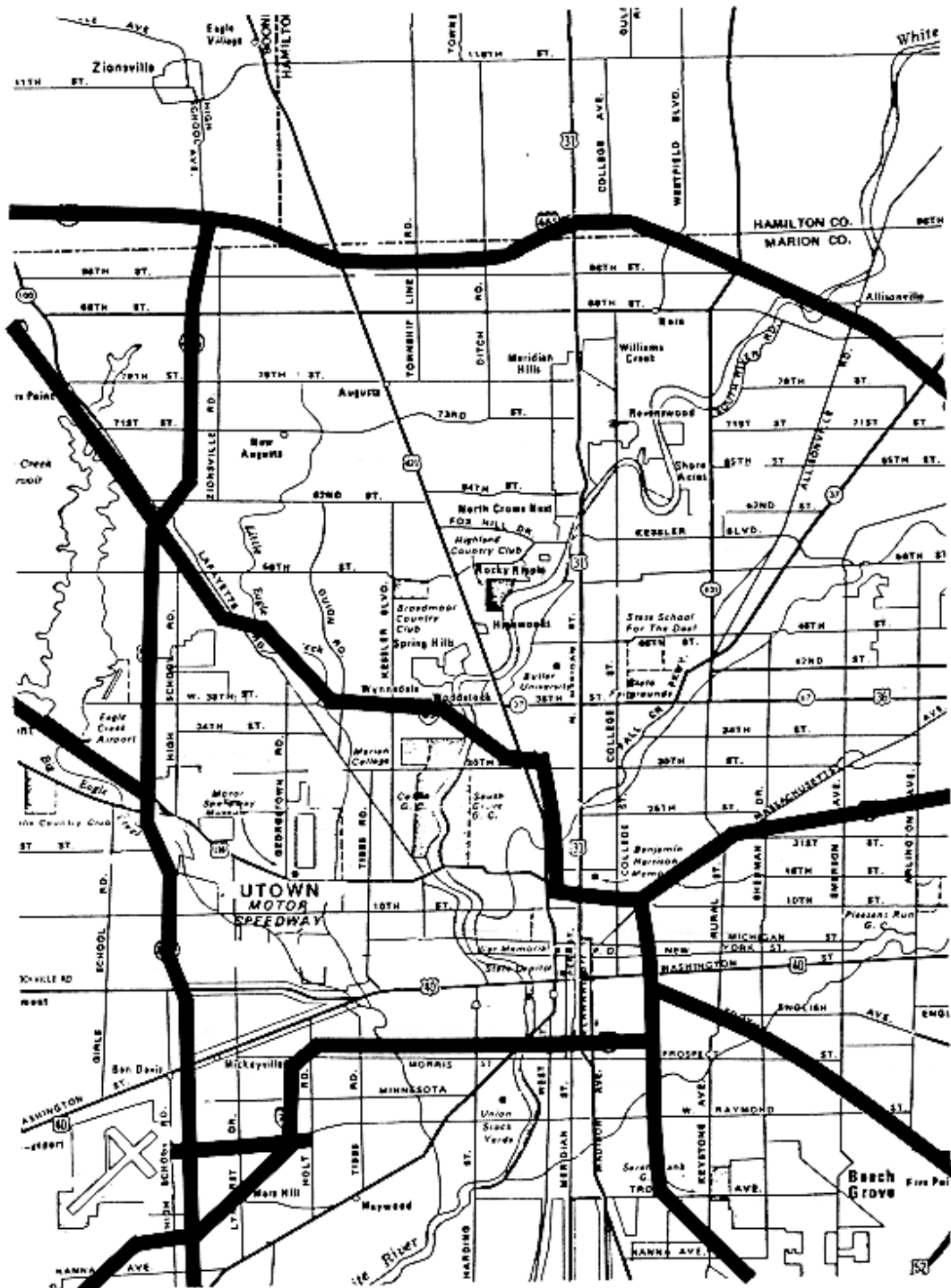
구분	존번호	내역	구분	존번호	내역
수원시	1	정지2동, 화서1동, 화서2동, 신안동, 팔달동	외곽존	11	국도42호선(만산방면)
	2	연무동, 조원동, 영화동, 남창동, 지동, 우만1동		12	국도43호선(성남, 하남방면)
	3	미의동		13	국도1호선(의왕방면)
	4	우만2동, 원천동		14	국도42호선(군포방면)
	5	인계동, 매탄1동, 매탄2동, 매탄3동, 매탄4동		15	국도1호선(오산방면)
	6	세류2동, 국선동, 권선동		16	국도42호선(용인방면)
	7	행동		17	국도43호선(화성방면)
	8	입북동, 구운동		18	경부고속도로(수원I/C)
	9	운천동, 정자1동, 파장동			
	10	서둔동			



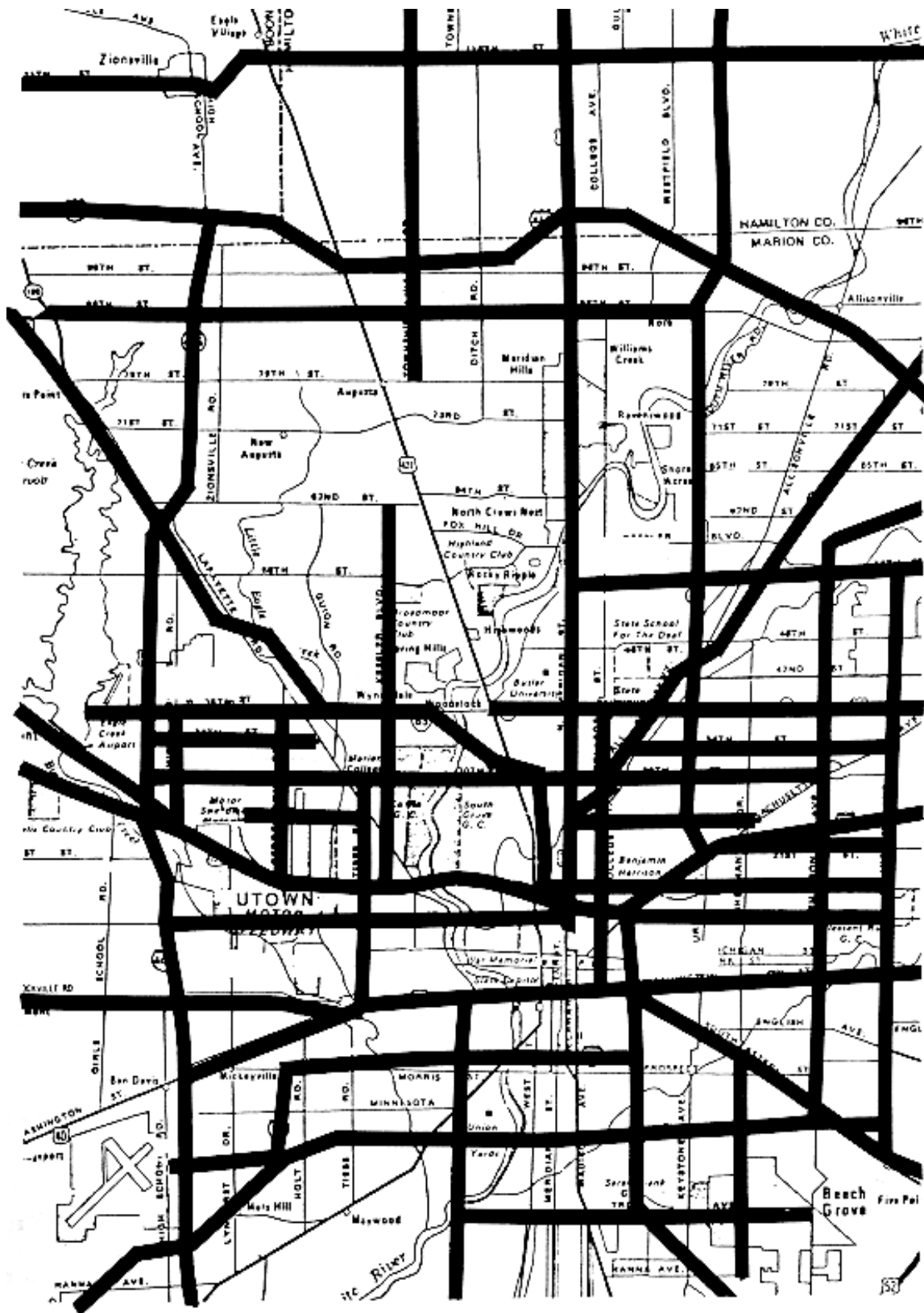
### 제3절 교통망과 연동한 교통존설정: 이론과 실제

#### 1. 교통존과 네트워크의 이용상의 양립성

- 더욱 효과적인 분석을 하기 위해 배정가로망 (assignment network)의 세분화 수준을 적절히 하여야 한다. 예를 들면 지역 네트워크는 corridor 또는 sub-area 분석처럼 세밀할 필요가 없다. CBD 지역 연구와 비슷한 소지역 연구에서는 시설의 모든 수준에서 배정을 해야 한다. 일반적으로 network와 zone system은 동행배정의 목적에 맞는 정도를 확보하여야 한다. 이 개념은 그림 6, 그림7을 통해 알 수 있다. 그림 6 은 지역간의 교통수요분석에, 그림7은 도시내부의 교통분석에 사용될 수 있다.
- 존 규모는 직접적으로 존내부의 동행수 (intrazonal trips)와 존외부 동행수 (interzonal trips)의 관계에 영향을 준다. 존이 커질수록 존내부 동행의 비율이 커진다. 그러므로 존내부 동행은 network에 배정되지 않기 때문에 총 교통량은 존이 작은 경우와 다를 수밖에 없다. 따라서 규모가 결정되었을 지라도 존의 크기나 링크의 수는 이러한 이유에서 다소 호환적일 필요가 있다.
- 존 네트워크의 조화성 (zone-network compatability)는 바로 이러한 이유에서 강조가 되는 동행배정단계에서의 유의성인데 assignment 결과가 분석목적과 주어진 상세수준 (level of detail)이 확보되도록 돕는다. 또한 이것의 결정시에는 사회경제적 지표와 그것의 활용도 역시 고려되어야 하는바, 계획하고 분석하는 자의 경험과 지식이 상당히 중요하다고 볼 수 있다.



<그림 3-1> Network for Regional Analysis



<그림 3-2> Network for Corridor Analysis

## 2. 교통존과 네트워크의 조화부재시의 결과

- 상기에서 지적한대로 존과 네트워크는 이용에 있어서 조화가 이루어져야한다. 즉, 존이 dense 하게 잘려서 구분되어져 있으며, 네트워크 역시, 미세한 가로까지 포함하여하고, 반대로 존이 coarse하게 (듬성듬성하게) 구분되어져 있는 경우, 상기의 그림6에서 보듯이 regional 수준에서의 대표적 간선도로만이 분석에 이용되는 것이 통례이다.
- 존과 네트워크가 조화가 되지 않는 경우, 즉 양자간의 수준이 조화롭지 못한 경우에 있어서는 가로교통량의 과대추정 (overestimate)과 과소추정(underestimate)문제가 발생하게 된다. 과대추정은 존은 dense 하게 나누어져 있으나, 이용가로망은 coarse한 경우이고, 과소추정은 그 반대이다.
- 이것을 표로서 나타내면 다음과 같다.

<표 3-4> 존과 네트워크의 관계

존 \ 네트워크	coarse	dense
coarse	OK	과소추정
dense	과대추정	OK

## 제4장   지리정보시스템을 이용한           교통존설정 기법

---

제1절   개   요

제2절   이론적 배경

## 제1절 개 요

- 이상에서 살펴본 바와 같이 교통분석 존의 구성은 교통계획 및 기타 공간 또는 비공간(예, 선거계획)계획 절차에 있어 매우 중요한 위치를 차지하고 있다. 이는 존 구성 이후에 이루어지는 제반 분석의 결과가 존 구성문제에 완전히 독립적이지 않다는 데 기인한다(Openshaw, 1978). 교통 분석에서도 이러한 현상은 예외가 아니며, 특히 교통 수요예측에 있어서는 수요추정의 최종 산물인 가로 교통량 결과가 교통존의 구성에 밀접한 연관을 가지게 된다는 것은 주지의 사실이다.
- 그러나 1960년대 이후, 몇몇 군집형성 알고리즘(clustering algorithm)의 변형을 통하여 공간계획 또는 교통계획의 존 구성 문제를 발전시켜보려는 연구가 이루어져 왔으나 교통존의 구성시 요구되는 일반원칙, 데이터의 수집 및 관리상의 어려움, 혹은 존 구성 방법들에 대한 표준화된 원칙 등의 제시된 바가 적어 교통 부분에서는 활용이 미미한 실정이다.
- 특히, 존 구성에 이용되는 각종 사회-경제적 자료의 체계적인 수집 및 관리가 어려운 현재의 상황으로 인해 존 구성 알고리즘의 구현이 일회적이고 특정 자료에 국한되어, 설사 그러한 표준화된 원칙 또는 알고리즘이 있다고 해도 실제 응용시에 많은 어려움이 존재한다. 따라서, 현재 각 지자체별 혹은 국가적으로 지리 정보 체계 구축 사업이 진행되고 있음을 감안하면, 존 구성 알고리즘을 지리정보체계와 연계하여 자료의 준비 및 존 구성의 노력에 소요되는 중복투자를 피하고 자료의 활용도를 증가시키려는 노력을 하게 되었다.
- 또한, 개발된 존 구성체계는 교통외 분야에서의 모든 공간계획모형에 응용이 가능하리라고 판단된다. 이러한 분야의 대표적인 예로서 선거구 및 행정구역 분할 등을 들 수 있으며, 기타 도시 및 지역계획에서의 존 구성을 필요로 하는 분야에서 이용할 수 있으리라고 판단된다.
- 이러한 맥락에서, 최소공간단위로부터 그보다 큰 존을 이끌어내어 그룹화(grouping)할 수 있는 일반적 기법을 파악하고 이를 토대로 기존의 군집형성알고리즘, 군집분석 및 지리정보체계(Geographic Information System; 이하 GIS라 칭함)를 이용하여 교통 분석 존(Traffic Analysis Zone; 이하 TAZ라 칭함)을 구성시키는 제반 알고리즘의 비교검토, 수정 및 GIS기반의 프로토타입 시스템의 구현을 본 절에서는 제시함으로써 향후 연구의 토대를 마련하고자 한다.
- TAZ 구성의 일반화된 경우라 할 수 있는 존 구성 혹은 군집 분석(clustering analysis)에 대한 연구는 오래전부터 이루어져 왔으며, 이러한 연구들은 크게 군집

형성을 위한 알고리즘에 관한 것(Ward, 1963, cliff and Hagget, 1970, and Keane, 1975)과 존 구성을 위한 지표 및 기준에 관한 것(Wright, 1990)으로 나눌 수 있다. 주된 내용은 이러한 알고리즘들을 검토하여 적절한 기준에 따라 분류함은 물론, 주요 알고리즘을 선정하여 TAZ 구성에 적절하도록 알고리즘의 수정을 통해 이를 프로그램화 한 후, 알고리즘의 상호간의 비교 분석을 시행함을 우선적인 보고서 내용으로 삼고 있다. 부차적으로 이러한 비교를 통해 각각의 군집형성 알고리즘이 TAZ 구성에 사용되기 위하여 어떠한 수정이 필요한가를 파악하고, 이러한 수정으로 인해 발생하거나 군집형성 알고리즘 자체가 가지고 있는 장단점들을 분석하여, 교통계획시의 TAZ 구성에 있어서 본질적 4단계 수요추정모형의 단점에 피드백 (feedback) 기능을 보장함으로서 존 구성에 있어서 유연성과 효율성을 제고하고자 한다.<sup>1)</sup> 이는 통행배정의 결과를 바탕으로 새롭게 모형을 적용함에 있어서 여러 단계로부터 출발할 수 있으나 통행발생, 배분, 수단선택의 계단제로 회귀하거나 존체계 구성으로 다시 회귀한 후 다시 시작하여 수요를 추정할 때 원래의 결과와 회귀한 후의 결과와의 차이가 발생할 수 있다는 측면에서도 존체계의 구성체계모형개발 및 구현은 연구의 가치가 있다고 보여진다.

- 내용으로는 알고리즘의 비교분석 및 알고리즘 도출을 위해 선결되어야 할 사항으로서 기존의 알고리즘을 면밀히 분석하고 이를 수용하기 위한 각 알고리즘의 실행 모듈들과 GIS 패키지 (PC ARC/INFO 3.4D Plus를 사용하였음)를 연결하는 인터페이스를 개발함은 물론 알고리즘의 수정이 이루어졌으며 이러한 제반비교분석이 함께 이루어졌다.
- 이러한 수정 알고리즘, GIS기반의 인터페이스의 개발을 통해 장기적으로는 존 설정기능을 지닌 모듈을 수요추정과정에 자동내재화 시킴으로서 계획작성 및 수요추정과정에 유연성과 효율성을 도모할 수 있기를 기대하며, 한편, 공간적인 테스트 지역으로는 서울시 및 관련데이터 베이스를 중심으로 분석하였다.

1) 실제 수요추정과정을 교통계획의 주요 단계로 볼때, 수요추정의 산출로서 링크 교통량의 민감도가 교통존 구성여부가 가장 직접적으로 영향을 주고 있는바 최적의 교통존 체계는 매우 중요하며 또한 이를 구현함으로서 결과의 유효성을 증대시키고, 존 구성의 용이함 및 수요추정과정에서의 분석시간의 단축이 효율성의 증진으로 귀착될 수도 있음을 시사함.

## 제2절 이론적 배경

### 1. 존 구성 문제의 정의

- 존 구성 (zone design) 문제는 군집 형성(clustering 또는 agglomeration) 문제의 특수한 형태로 볼 수 있다. 군집형성은 일반적으로 주어진 자료들을 성격이 유사한 그룹으로 구분하는 것으로 정의할 수 있으며, M개의 개체(entity)를 N개의 군집(cluster)으로 구성하는 군집형성 문제의 결과는  $M \times N$  행렬  $X$ 로 표현할 수 있다. 즉  $i = 1, \dots, M$  이고  $j = 1, \dots, N$ 이라면 (단,  $M > N$ ) 행렬  $X$ 의 원소(element)인  $x_{ij}$ 는 아래와 같은 값을 가진다.

$$\begin{cases} x_{ij} = 1, \text{ 개체 } i \text{가 군집 } j \text{에 속할 때} \\ x_{ij} = 0, \text{ 개체 } i \text{가 군집 } j \text{에 속하지 않을 때} \end{cases}$$

- 이와 같은 군집형성 문제를 이용하여 존 구성 문제를 표현하기 위해서는 최소단위 개체에 해당하는 기본 공간 단위(basic spatial unit: 이하 BSU라 칭함)의 개념을 도입하게된다. BSU는 특정 연구에서는 사용가능한 자료를 얻을 수 있는 공간적인 최소단위이다. 따라서, 전체 연구영역은 BSU라는 최소단위들의 집합으로 표현될 수 있으며, 각각의 BSU들은 더이상 작은 단위로 나누어질 수 없고, 서로 겹치지 않으며, 연구영역 전체를 빈틈없이 채워야 한다. Keane (1975)은 집합  $E$ 를 몇몇의 독립변수들의 함수로 파악하여 존 구성 문제를 다음과 같이 표현하였다.

$$E = f(S, R, A, F) \dots\dots\dots <식 1>$$

여기서,

$E$  : 존 구성의 결과 벡터집합

$S$  : 유한개이고 겹치지 않는 BSU의 집합

$R$  : 존 구성 기준의 집합

$A$  : 존 구성 프로세스(process)

$F$  : 가능한 해(파티션; partition)의 집합

- 이상에서 볼 때, BSU의 집합 내에서 적절한 존 구성 알고리즘을 통해 가능한 파티션 중 주어진 존 구성 기준을 만족하는 파티션을 찾아내는 것이 존 구성 문제의 본질이라 할 수 있다. 여기에서 사용되는 존 구성 기준 및 군집 형성 알고리즘은 아래에 언급한다.



## 2. 존 구성을 위한 기준

- 존 구성의 원칙은 해당 존 구성의 목적에 따라 달라질 수 있으나 앞절에서 소개한 바와 같다.
  - 존내부의 사회, 경제적 특성이 동일해야 한다.
  - 존 내부의 통행이 적어야 한다
  - 가능한한 지형적이거나 행정적인 경계선을 사용해야 한다.
  - 존 안에 다른 존이 포함되지 않아야 한다.
  - 각 존의 가구수, 인구, 면적 및 통행량이 비슷한 것이 좋다.
  - 가능한한 센서스 단위와 일치해야 한다.
- 한편, 존 구성 과정에서 위와 같은 원칙들을 만족시키기 위해서 제안된 기본적인 지리적 및 속성관점의 지표들은 아래와 같은 지표로서 대변될 수 있다.

### 가. 유사성(Similarity)

- 존 구성에서 가장 기본적으로 고려해야 할 사항으로서 동일 영역을 구성하는 BSU들간의 동질성을 의미한다. 예를 들어, n개의 BSU를 m개의 존으로 구성할 때, 이러한 동질성의 달성은 아래와 같은 제곱의 합(sum of square: Z)의 최소화로 표현할 수 있다.

$$Z = \sum_{k=1}^m \sum_{i=1}^n \sum_{j=1}^n d_{ij}^2 \lambda_{ki} \lambda_{kj} \dots\dots\dots <식 2>$$

여기서

$d_{ij}$  : 두 BSU i, j의 속성값의 차

$\lambda_{ki}$  : BSU i가 존 k에 속해 있을 때 1, 그렇지 않을 때 0 (  $\lambda_{ki}$ 도 동일함)

### 나. 연속성(Contiguity)

- 형성된 영역은 연속적으로 배치되어 있어야 한다. 연속성은 해당 영역을 떠나지 않고 영역내의 어떤 점에서 출발하여 다른 점에 도달할 수 있음을 의미한다. 이것은 행정구역 편성, 학군조정, 그리고 기타 존 구성 모델에서도 중요한 문제로 다루어지고 있으며, 특히 TAZ 구성 문제에서는 필수적인 기준이 된다. 기존의 수학적, 통계적 모델에서 명시적으로 연속성을 모델 화하고 이를 강행하는 것은 매우 어렵다. 따라서, 연속성 제약은 TAZ 구성에 있어서 일반적인 군집형성 알고리즘의 활용을 어렵게 해왔으며, 기타 기준들의 경우도 마찬가지로 악재로 작용하게 된다.

#### 다. 응집성(Compactness)

- 응집성은 형태를 기술하기 위해 사용되는 상대적인 기준이다. 응집성은 종종 행정구역편성 모델에서 자주 인용되는 게리맨더링(gerrymandering)에 상반되는 단위로서 사용되기도 한다. (Wright, 1990) 응집성을 측정하기 위한 표준 지표는 존재하지 않지만 현재까지 제안된 몇 가지 기준은 다음과 생성된 존의 면적과 구해진 응집성 값이 서로 독립적인 것을 선택하여야 한다. 즉, 같은 형상을 가지지만 다른 면적을 가진 두 형태는 같은 응집성 값을 주어야 한다. 일례로, Horton (1932)에 의해 제안된  $D^2/A$ 는 면적  $A$ 와 최대직경  $D$ 의 함수이다. 따라서, 이 비율에 어떤 상한값  $\lambda$ 를 책정하여 존의 응집성에 대해 어느 정도의 요구조건을 부여하는 것이 가능하여 이 한계값이 더 작으면, 상대적으로 더 응집된 형태를 얻는다.

### 3. 군집형성 알고리즘

- 군집형성은 일반적으로 주어진 자료들을 성격이 유사한 그룹으로 구분하는 것이라고 표현할 수 있으므로, 존 구성 알고리즘이 요구하는 기준 중 가장 기본이 되는 유사성을 만족한다. 따라서, 존 구성 알고리즘은 기본적으로 일반적인 군집형성 문제를 위한 알고리즘을 근간으로 하고, 이에 존구성의 특징을 반영하도록 적절하게 수정하여야 구현할 수 있다.
- 군집형성 알고리즘은 다양한 접근과 변형이 존재하나, 접근 방식에 따라 크게 계층적 방법과 비계층적 방법으로 구분할 수 있으며, 이 두 방법의 기본 개념 및 절차를 요약하면 아래와 같다.

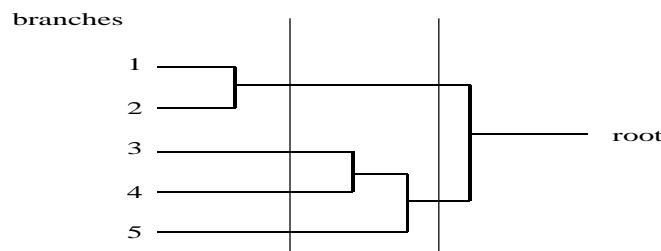
#### 가. 계층적 방법

##### 1) 기본개념

- 계층적 방법은 <그림 4-1>과 같이 개체(기본 단위 혹은 군집)간의 관계를 나타내는 트리(tree)로서 표현할 수 있다. 루트(root)가 완전하게 하나로 결합된 개체들의 모임을 표현하는데 반해, 왼쪽의 각각의 가지(branch)들은 하나의 개체를 표현한다. 트리가 가지로부터 루트로 이동함에 따라 기본 단위에서 군집으로의 병합이 이루어진다. 가지에서 루트로의 계층적 군집형성 방법은 응집식 방법(agglomerative method)이라 불리고, 그 반대로의 방법은 분할식 방법(diverse method)이라 불리며, 일반적으로 응집식 방법이 널리 사용된다(Anderberg, 1973).

- 즉,  $n$ 개의 개체를 위한 트리가 구성되면, 트리에서 원하는 군집의 수를 가진 위치를 절단하여 해를 얻을 수 있는 바, 예를 들어, <그림 4-1>에서 ㉠ 위치에서의 절단은 3개의 군집을, ㉡ 위치에서의 절단은 2개의 군집을 얻게 된다.
- 응집식 방법의 관점에서는 두개의 개체가 하나로 합병되고, 분할식 방법에서는 하나의 개체가 2개로 나누어진다. 계층적 방법에서 이러한 결합 또는 분할은 영구적이며, 일단 수행되면 차후에는 고려의 대상에서 제외된다. 이것은 계층적 방법의 약점이자 단점이다. 즉, 영구적인 결합 및 분할로 인해 탐색해야 할 해(solution)의 범위가 급격하게 줄어드는 반면, 전체적인 군집 형성과정에서 최적해의 획득을 보장하지 못하게 되는 단점이 있다.

## 2) Ward의 계층적 알고리즘



<그림 4-1> 계층적 군집형성을 위한 트리

- 계층적 응집식 (hierarchical agglomerative) 알고리즘의 일종으로서, 지정된 속성값의 유사성에 기반을 두고, 중첩되지 않는 계층적 그룹을 형성한다. 이 알고리즘은 초기에 주어진  $k$  개의 개체를, 각각 하나의 개체만을 가지는  $k$  개의 그룹으로 본다. 여기에서  $k(k-1)/2$  개의 조합가능한 모든 짝을 고려하고, 최적의 목적함수값을 갖는 짝을 선택하여 전체를  $(k-1)$  개의 그룹으로 만든다. 이와 같은 과정의 반복을 통해 최종적으로는 전체 개체를 모두 포함하는 하나의 큰 그룹을 만들어 낼 수 있으며, 실제 사용에 있어서는 원하는 수의 그룹이 남았을 때 알고리즘을 종료시켜서 해를 찾을 수 있다.
- 목적함수는 형성된 그룹의 오차 제곱의 합(SSE: Square Sum Error)을 사용하며 아래와 같이 정의된다(Ward, 1963).

$$SSE = \sum_{i=1}^n x_i^2 - \frac{1}{n} \left( \sum_{i=1}^n x_i \right)^2 \quad \text{<식 3>}$$

- 여기서,  $x_i$ 는 각 개체들의 속성값이다. 따라서, 이 알고리즘은 목적함수인 SSE를 최소화하는 그룹의 짝을 찾아 결합해 나가게 된다. 덧붙여서,  $n$ 개의 개체를  $m$ 개의 그룹으로 만드는 Ward 알고리즘은 하수처리 플랜트의 가파른 하수처리에서 오는 폐다스택에서의 그룹수이다.

<단계 1>  $k \leftarrow n$

<단계 2> 활성화(active)된 모든 그룹  $i$ 와  $j$ 의 짝에서 결합을 가정할 때 최소의 SSE를 가지는  $i$ 와  $j$  값을 찾는다.

<단계 3> <단계 2>에서 결정된 그룹의 짝  $i$ 와  $j$ 에서  $i$ 번째 그룹에는  $j$ 번째 그룹이 속했음을 표시하고,  $j$ 번째 개체는 비활성화한다.

<단계 4>  $k \leftarrow k - 1$

<단계 5>  $k$ 값이 2이거나  $m$ 이면 종료, 그렇지 않으면 <단계 2>로 간다.

## 나. 비계층적 방법

### 1) 기본개념

- 비계층적 방법의 대부분은 반복적 분할 방법(iterative partitioning method)으로 볼 수 있다. 이들 방법의 기본 아이디어는 어떤 초기 파티션을 선택하고 이에서 더 나은 파티션을 얻기 위해 기본단위의 소속 군집을 바꾸어나가는 것이다(Anderberg, 1973). 대부분의 반복적 분할 방법은 아래와 같이 동작한다(Aldenderfer and Blashfield, 1984).

<단계 1> 기본 단위를 지정된 수의 군집으로 나누는 초기 분할을 실시하고, 이들 군집들의 중심점(centroid)를 계산한다.

<단계 2> 각각의 기본 단위를 가장 가까운 중심점을 가진 군집에 할당한다.

<단계 3> 군집들의 새로운 중심점을 계산한다.

<단계 4> 군집을 바꿀 기본단위가 없으면 종료 그렇지 않으면 <단계 2>로 감.

- 반복적 분할 방법은 기본 단위를 군집에 할당하는 방식에 따라 여러 변형이 존재한다.

### 2) 시뮬레이티드어닐링(Simulated Annealing)

- 최적화 문제에서 빈번히 발생하는 해의 지역 극소 문제를 해결하기 위한 비계층적 방법의 하나로서, 고온의 물체를 서서히 식힘으로써 최종적으로 가장 낮은 에너지 상태를 얻을 수 있다는 어닐링 방법을 시뮬레이션한 것이다. 최초의 가상온도를 고온으로

하고, 최적화가 진행됨에 따라 온도를 서서히 낮추어 간다. 이 때, 온도에 따라 클러스터의 재구성이 확률적으로 진행되는 방식이다.

- 주어진 온도에서, 시스템은 특정 에너지 레벨을 가지게 된다. 주어진 에너지 레벨에서 시스템의 확률은  $\exp(-E/T)$ 의 함수로 나타내진다. 여기서  $E$ 는 에너지 레벨이고,  $T$ 는 시스템의 온도이다. 따라서, 온도가 낮아지면 시스템은 항상 더 높은 에너지 상태로 변경될 기회를 가지게 된다. 이를 계속해서, 시스템의 온도가 0도에 접근하면, 원자는 최소 에너지 상태를 가지게 된다. 간단히 말해, 전통적인 최적화 알고리즘들은 급속한 냉각 또는 담금질과 유사하게 해를 향해 돌진하므로 지역 극소에 도달하는데 비해, 시뮬레이티드 어닐링은 서서히 온도를 낮추어 가며 최소 에너지 상태, 즉 전역 최적 해를 얻을 수 있다는 것이다(Macmillan and Pierce, 1994). 알고리즘을 약술하면 아래와 같다.

<단계 1> 초기 파티션 설정 및 초기 온도( $T$ ) 설정

<단계 2> 임의로 선택한 한 개체를 다른 클러스터로 이동

<단계 3> 목적(에너지) 함수값의 변화폭( $\Delta E$ )을 구한다

$$\Delta E = \text{개체 이동 후의 시스템 에너지} - \text{이동전의 시스템 에너지}$$

<단계 4>  $\Delta E$ 가 음수이면, 이 이동을 수락한다.

- 이값이 양수이면, 이 이동을 확률적으로 수락한다. 확률함수는 아래와 같다.

$$P(\Delta E) = \exp\left(-\frac{\Delta E}{T}\right) \dots\dots\dots \text{<식 4>}$$

<단계 5> 시스템의 온도( $T$ )를 아래와 같이 감소시킨다.

$$T = h \cdot T \dots\dots\dots \text{<식 5>}$$

<단계 6>  $T$ 가 0도에 충분히 접근하면 중단, 그렇지 않으면 <단계 2>로 간다.

### 3) k-means 알고리즘

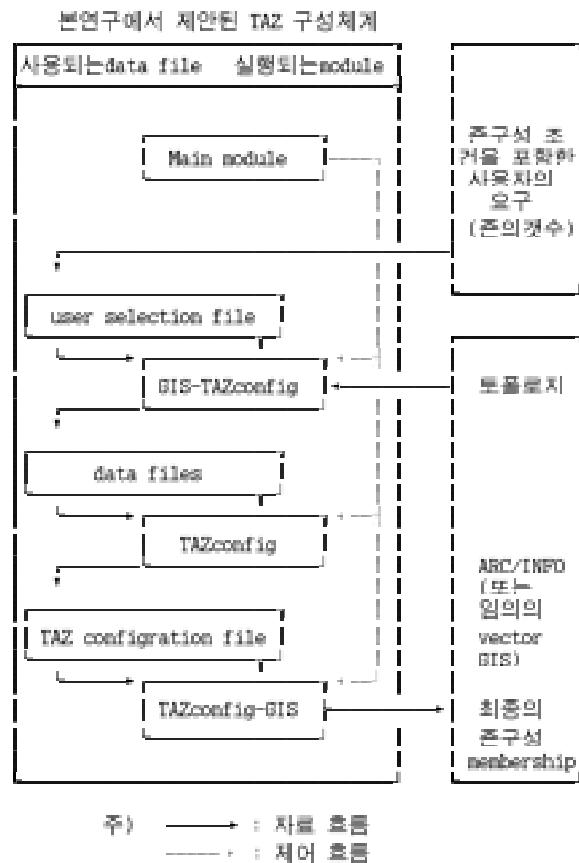
- MacQueen(1967)은 각각의 자료단위를 가장 가까운 중심점(nearest centroid)을 가지는 군집( $k$  개의 군집중 하나)에 할당하는 과정을 표기하기 위해 "k-means"라는 용어를 사용하였다. 이 과정의 핵심은 중심점의 계산이 현재의 군집 구성(cluster membership)에 기반하여 계산된다는 것이다. 이에 반해 다른 교환(swapping)에 의한 방법에서는 중심점의 재계산(update)이 한 사이클(cycle)이 끝난 후에 이루어진다.  $m$  개의 자료 단위를  $k$ 개의 클러스터로 나누는 MacQueen의 알고리즘은 아래와 같다.

- <단계 1> 자료들 중에서 선두  $k$  개를 뽑아, 각각 하나의 멤버만을 가지는 군집으로 만든다.
  - <단계 2> 남아있는  $(m-k)$ 개의 자료 단위들 각각을, 가장 가까운 중심점을 가진 군집에 할당한다. 각각의 할당이 이루어질 때마다 멤버를 얻은 군집의 중심점을 다시 계산한다.
  - <단계 3> 모든 자료가 <단계 2>에서 할당된 후, 기존의 중심점을 고정점(fixed seed point)으로 결정하고, 모든자료단위가 가장 가까운 고정점을 가진 군집에 할당되게 한다.
- 이 방법은 수렴이 달성될 때까지 과정이 계속되지 않고, 자료 단위의 최초할당을 그 기본해로 수락하므로 비용이 적게 든다는 장점이 있다. 군집 구성의 가장 큰 변화는 첫번째 재할당에서 일어나며, 이어지는 재할당 과정에서의 변화는 상대적으로 적으므로, 이 방법은 유용한 결과를 제시할 수 있다.

### 제3절 TAZ 구성 체계의 개발

#### 1. 존구성 체계의 개요

- 구현된 TAZ 구성 체계는 <그림 4-2>에서 제시된바와 같은 전반적인 구성을 가지며, 실제로 세부존의 최종 산출물로서의 생성TAZ 소속 (membership)을 결정하는 모듈 등 대부분은 C언어로 작성되었다. 각 모듈의 역할과 자료 화일의 용도는 <표 4-1>와 <표 4-2>에 제시하였다.



<그림 4-2> TAZ 구성 체계 개념도

<표 4-1> TAZ 구성 체계 내의 실행모듈

모듈명	기능	수행내용
Main module	주 모듈	본 TAZ 구성 체계(이하 본 체계라 칭함)를 운영/제어 한다. 즉, 각 구성요소들을 사용자 지시 및 상황에 따라 적절히 수행시킨다.
GIS-TAZconfig	GIS-TAZ 구성 모듈 인터페이스	ARC/INFO 또는 이와 상응하는 벡터형 GIS 데이터베이스에서 TAZ 구성에 필요한 자료를 뽑아 적절한 형식으로 변환한다. 필요한 자료는 위상(topology) 자료 및 각 존에 대한 사회-경제적 자료들이다.
TAZconfig	TAZ 구성 모듈	실질적인 TAZ 구성을 준거 기준에 의해서 시행하여 각 세부 존의 대존으로의 변환에 있어서의 membership을 제공한다.
TAZconfig-GIS	TAZ 구성 모듈-GIS 인터페이스	TAZ 구성 모듈에서 만들어진 존 구성 결과를 적절한 형식으로 변환하여 ARC/INFO로 돌려준다.

## 2. 수정된 알고리즘

### 가. 기본 수정 사항

#### 1) 가상 인접 관계의 생성

<표 4-2> TAZ 구성 체계에서 이용하는 데이터 파일

파일명	기능	저장내용
User selection file	사용자선택사항 저장 파일	사용자가 원하는 프로그램 실행 옵션들을 기록한다. 이 파일은 직접적으로는 존 구성에 필요한 데이터 파일을 만들기 위해 GIS-TAZ 구성 모듈 인터페이스(GIS-TAZconfig)에 의해 사용되며, 일부 옵션들은 주 모듈(Main module)이 프로그램의 실행 시퀀스를 제어하는데 이용된다.
data file	존 구성을 위한 자료파일	존 구성에 필요한 자료를 저장하고 있는 파일들로서 GIS-TAZ 구성 모듈 인터페이스(GIS-TAZconfig)에 의해 생성된다. 이 파일은 TAZ 구성 모듈(TAZconfig)에 의해 사용된다.
TAZ configuration file	존 구성 결과 파일	존 구성 결과를 담은 파일이다. 역시 텍스트 파일이며 TAZ 구성 모듈-GIS 인터페이스(TAZconfig-GIS)에 의해 ARC/INFO에서 사용가능한 파일 형태로 변환된다. 파일에는 BSU 번호와 해당 BSU가 속하는 TAZ 번호가 나열된다.

- 실제 연구영역에서는 다른 BSU들과 전혀 인접관계가 없는 고립된 BSU가 존재할 수 있다. 이는 연구영역이 도서(islands)지역을 포함할 때 발생할 수 있으며, 이러한 경우 아무런 부가적인 처리를 하지 않은 상태에서 TAZ 구성을 실시하면 해당 BSU는 독립적인 하나의 TAZ로 존재하게 된다. 따라서 본 연구에서는 BSU들간의 인접관계를 생성하는 과정에서 가상적인 인접관계를 생성하여 이러한 문제를 해결하도록 하였다. 이 처리는 GIS-TAZ 구성 모듈 인터페이스(GIS-TAZconfig) 부분에서 이루어지는데, 그 과정은 아래와 같다.

<단계 1> 직접적인 인접관계에 있는 BSU들의 리스트를 작성함.

이것은 ARC/INFO로부터 얻어진다.

<단계 2> 사용자가 지시한 가상 인접관계 리스트를 덧붙인다. 이는 존 간의 거리에 의해 생성되는 가상인접관계가 적절치 않을 경우를 고려한 것임.

<단계 3> 현재까지 만들어진 리스트를 조사하여, 이에 포함되어 있지 않는 고립된 BSU들의 리스트를 작성함.

<단계 4> 고립된 각각의 BSU에 대해서 중심점간의 거리가 가장 가까운 것을 찾아 짝지움.

<단계 5> <단계 4>에서 얻어진 BSU들의 짝을 인접관계 리스트에 덧붙임.



## 2) 존내 BSU 갯수의 불균형 조정

- 연속성 제약외에 다른 제약을 강행하지 않고 존 구성을 실시할 경우, 그 속성값이 주변의 값들과 큰 차이를 보이는 BSU는 존 구성이 계속 진행됨에도 불구하고 다른 존에 잘 편입되지 않는 경우가 발생한다. 이것은 결과적으로 다른 존에 비해 월등히 많거나 적은 BSU로 이루어진 존을 생성하게 하므로 존내 BSU 갯수의 불균형을 조정하는 별도의 고려가 필요하다. 따라서, 본 체계에서는 존내 BSU 갯수의 변화 허용폭을 지정하는 옵션을 두어 이를 처리하고 있으며, 이 변수는 0부터 1사이의 값을 가진다. 즉, 이 변수값이 0.2이고 132 개의 존을 8개로 나눈다면 존당 평균 BSU 갯수는 16.5가 된다. 또,  $16.5 * 0.2 = 3.3$  이므로 결과에서 존당 BSU의 갯수  $x$ 는  $16.5 - 3.3 \leq x \leq 16.5 + 3.3$ 가 된다. 이러한 기준의 실제적인 강행은 각 알고리즘마다 다소의 차이가 있으므로, 이에 대해서는 각각의 수정된 알고리즘 설명에 기술한다.

## 3) 응집성 및 행정구역 고려를 위한 모델

- 존 구성시에 응집성 혹은 현 행정구역의 고려는 실제 구현된 예가 적으며, 그 강행(enforcement)<sup>2)</sup>을 위한 모델도 표준화된 것이 존재하지 않는다. 따라서, 본 연구에서는 응집성 및 행정구역 강행을 위하여 각각 2개의 모델을 선정하고 이를 각 알고리즘에 적용하여 비교해 보았다. 먼저, 각 존의 응집성 지표는 <표 4-3>과 같이 두가지의 모형을 이용하였다.
- 첫번째 모형은 위에서 제시된 Horton의 식을 그대로 따른 것이다. 즉, 대상 존내에서 가장 먼 두 점의 거리를 최대 직경으로 하고, 이의 제곱을 면적으로 나눈 것이다. 두번째 모형은 경계선 길이(perimeter)에 기반을 둔 모형으로

<표 4-3> 응집성 강행 모형

응집성 강행 모형	식
1	최대직경 <sup>2</sup> / 면적
2	경계선길이 <sup>2</sup> / 면적

2) 강행(enforcement or enforcing)의 의미는 연속성, 응집성, 행정구역순응들을 확보하기위해 어느 정도의 제어(control)를 함을 의미함.

- 대상 존의 경계선 길이의 제곱을 면적으로 나눈 것이다. 경계선길이에 제곱을 한 것은 대상 존의 크기에 관계없이 같은 형태라면 동일한 지표값을 얻기 위한 것이다. 따라서, 가장 응집된 형태로 간주되고 있는 원의 경우  $4\pi 2r^2 / \pi r^2 = 4\pi$ 를 얻을 수 있고, 이 값은 원의 크기에 관계없이 일정하다.
- 행정구역 강행을 위한 모형도 역시 두 가지를 사용하였는데, 이는 <표 4-4>와 같다.
- 첫 번째 모형은 존구성 단계에서 결합할 존간(계층적 알고리즘의 경우) 혹은 이동하려는 BSU와 대상 존간(비계층적 알고리즘의 경우)의 행정구역이 동일한가의 여부에 중점을 둔 것이다. 두 번째 모형은 결합 또는 이동후 해당 존의 상태에 중점을 둔 것으로 0과 1사이의 값을 가진다. 여기서 근거 행정구역이란 해당 존이 걸쳐 있는 행정구역 중 가장 많은 BSU가 속한 행정구역을 뜻한다. 따라서, 근거 행정구역에 속하지 않은 BSU를 해당 존의 전체 BSU로 나누어 원하는 지표값을 얻을 수 있다.

&lt;표 4-4&gt; 행정구역 강행 모형

행정구역 강행 모형	식
1	1 : 결합 또는 이동하려는 존간에 행정구역 경계선이 존재하는 경우 0 : 그외의 경우 0
2	근거 행정구역에 속하지 않은 BSU 수 / 해당 존의 BSU 수

#### 나. Ward의 계층적 알고리즘

##### 1) 목적함수의 계산

- 전술한 바와 같이 존 구성 문제의 목적함수는 연속성 및 기타 제약으로 인해 수학적, 통계적 모델화가 어려우므로 경험적인(heuristic) 득점함수를 이용하여, 존의 결합 순위를 결정한다. 이 득점함수는 데이터 값에서 얻어지는 지표(index)외에도, 앞서 제시한 응집성 혹은 현 행정구역을 고려해야 한다. 따라서, 최종적인 선택은 알고리즘을 통해 계산된 점수(score) S를 이용하여, 이 점수가 가장 낮은 조합을 하나의 존으로 구성한다. 이 S는 아래와 같이 계산된다.

$$S = Z + C \times W_c + F \times W_f \dots\dots\dots <식 6>$$

여기서

S : 선택 점수

Z : 데이터 값에서 얻어진 지표

C : 응집성 지표

Wc : 응집성 강행정도

P : 현 행정구역 점수

Wp : 행정구역 강행정도

- 위 식에서 S를 제외한 나머지 변수들은 0과 1 사이의 값을 갖는다. 응집성 지표 및 행정구역 점수는 상술했던 식에서 얻어진 값들로서, 정규화된 후 주어진 강행도를 곱하여 선택점수에 합산된다.
- Z score는 Ward의 기본 알고리즘에서의 목적함수인 <식 3>을 사용한다. 이를 위해 최초 데이터에 전처리 과정이 요구된다. 먼저, 반영할 데이터들 각각을 0과 1 사이의 값으로 정규화한다. 각 자료의 정규화를 위한 식은 아래와 같다.

$$D_n = \frac{D_o - Min}{Max - Min} \dots\dots\dots <식 7>$$

여기서,

Dn : 정규화된 데이터

Do : 최초 데이터

Max : 데이터 중 최대값

Min : 데이터 중 최소값

- 따라서, 최종적으로 얻어지는 대상 존의 SSE값은 아래와 같이 정의된다.

$$SSE_{zone} = \sum_i (SSE_i \times W_i) \dots\dots\dots <식 8>$$

여기서,

SSEzone : 대상 존의 SSE값

SSEi : 대상 존에 속한 BSU들의 i번째 필드의 데이터 값들의 SSE

Wi : i번째 필드의 가중치

- 따라서, 본 알고리즘에서 최종적인 득점함수값의 산출은 아래와 같은 과정을 통해 이루어진다.

- <단계 1> 각 필드별로 주어진 데이터를 정규화한다.
- <단계 2> 각 필드별 SSE를 구한다.
- <단계 3> 각 필드별 SSE들의 가중합을 구하고 이를 Z라 놓는다.
- <단계 4> 응집성 지표와 행정구역 점수를 구하고, 각각 정규화한다.
- <단계 5> Z, 응집성 지표, 행정구역점수의 가중합을 구한다. 이 값이 최종적인 득점합수값이다.

## 2) 존내 BSU 갯수의 불균형 조정

- BSU 갯수의 불균형을 막기 위해 존 구성 프로시저 내에서 특정 존에 응집이 편중되지 않도록 매 스텝마다 한 존내의 최대 BSU 갯수와 최소 BSU 갯수를 비교하여 이것이 주어진 비율을 벗어날 경우 존내 BSU의 수가 최소인 존을 먼저 고려하도록 하였다. 이 때 판단의 기준이 되는 비율은 사용자에게 의해서 주어지며, 이것은 실제로는 최종 결과에서의 BSU 갯수의 변화량을 지시하는 값이다. 그러나, 실제 경우에서 연속성 및 기타 제약, 그리고 존간의 결합을 통해 군집을 이루는 계층적 알고리즘 자체의 특성으로 인해 이 비율이 항상 보장되지는 않는다.
- 본 알고리즘의 경우, 수행과정에서 BSU의 갯수가 최대인 존을 완성된 존으로 가정하고 이 때의  $BSU\text{갯수} * (1 - BSU\text{변화율})$ 에 미치지 못하는 존이 있을 때, 다음 단계에서 이를 우선적으로 고려하여 존 구성의 불균형을 조정하도록 하였다. 이 과정을 정리하면 아래와 같다.

<단계 1> 존당 최대 BSU 수와 최소 BSU 수를 구한다.

<단계 2> 만약  $(\text{최대BSU수} * (1 - BSU\text{변화율}) > \text{최소BSU수})$  이면 최소 BSU 수를 가지는 존에 인접한 존 리스트 작성 그렇지 않으면 활성화(active)된 모든 인접한 존의 리스트를 만든다.

<단계 3> 작성된 리스트내에 존재하는 짝들의 득점합수값을 계산

<단계 4> 작성된 리스트에서 최소의 득점합수값을 가지는 짝을 찾는다.

## 3) 수정된 알고리즘

- 이상의 내용을 반영하여, 수정된 최종 알고리즘은 아래와 같다. 수정은 <단계 2>에서만 이루어졌으며, 사용된 변수의 의미는 상술된 Ward 알고리즘에서와 같다.

<단계 1>  $k \leftarrow n$

<단계 2> 최소의 득점합수값을 갖는 i와 j를 찾는다. 이 때, 현저히 적은 BSU를 가지는 존이 있으면 이를 먼저 고려한다.

<단계 3> <단계 2>에서 결정된 그룹의 짝  $i$ 와  $j$ 에서  $i$  번째 그룹에는  $j$  번째 그룹이 속했음을 표시하고,  $j$  번째 개체는 비활성화(inactive) 시킨다.

<단계 4>  $k \leftarrow k - 1$

<단계 5>  $k$ 값이 2이거나  $m$ 이면 종료, 그렇지 않으면 <단계 2>로 간다.

다. 시뮬레이티드 애닐링과  $k$ -means 알고리즘

- 시뮬레이티드 애닐링과  $k$ -means 알고리즘은 모두 비계층적 알고리즘으로서, 동일한 수정을 가하였다. 수정된 사항들은 아래와 같다.

#### 1) 존내 BSU 갯수의 불균형 조정

- 계층적 알고리즘과는 달리 각각의 BSU의 존간 이동을 통해 동작하는 비계층적 방법에서는 존당 BSU의 수를 보다 엄격하게 제어할 수 있다. 이는 상술된 시뮬레이티드 애닐링 알고리즘이 <단계 2>에서 하나의 개체를 선택하는 과정에 의해 이동이 수락하기 전 선별을 거침으로서 이루어진다. 즉, 존당 BSU 갯수의 수락가능한 간격(interval)을 계산한 후, 매 이동시마다 그 이동으로 인하여 그 간격을 벗어나는 존이 생겨나면 이 이동을 사전에 막는다. 그 과정은 다음과 같다.

<단계 1> 존당 평균 BSU 갯수  $A$ 를 구한다.

$$A = \text{연구 영역의 전체 BSU 수} / \text{생성될 존 수}$$

<단계 2> 사용자에게 의해 주어진 BSU 갯수의 변화 허용폭을 이 값에 곱하여 실제 BSU의 변동폭  $V$ 를 얻는다.

<단계 3> 존당 BSU의 수락가능한 간격을 얻는다.

$$A - V \leq \text{존당 BSU 갯수} \leq A + V$$

<단계 4> 이 이동을 수락했을 경우 BSU를 획득/손실하는 두 존의 BSU수가 <단계 3>에서 구한 범위에 속하는지 검사하여 이 범위를 벗어나면 이 이동을 취소한다.

#### 2) 연속성 검사

- 계층적 알고리즘에서는 인접한 두 존을 묶어 하나의 존을 만들도록 알고리즘을 수정하였으므로 별도의 연속성 검사를 시행할 필요가 없지만, BSU 교환에 의존하는 비계층적 알고리즘에서는 매번 그 교환으로 인해 연속성이 상실되는지를 검사하여야 한다. 그 알고리즘은 아래와 같다.

<단계 1> 대상 존에 속하는 BSU의 리스트를 만든다.

<단계 2> 리스트에서 하나의 BSU를 임의로 선택하여 태그(tag)한다.

<단계 3> 리스트의 모든 BSU를 검사하여 태그된 BSU와 공간적으로 인접해있는 BSU에 역시 태그한다.

<단계 4> 모든 BSU가 태그되었으면 연속성이 유지되고 있는 것이므로 종료한다. 그렇지 않으면 <단계 2>를 (리스트내의 BSU 수 - 1)만큼 반복한다.

<단계 5> 정해진 수대로 반복하였음에도 태그되지 않은 BSU가 있다면 연속성이 깨진 것으로 보고 종료한다.

- 연속성 검사의 다른 한 부분은 BSU 이동후의 존이 도우넛 형태가 되는 것을 막는 것이다. 이는 대상 존의 외각 BSU, 즉 다른 존과 인접해있는 BSU들의 리스트를 만들고 이 BSU들이 인접해 있는 존이 모두 하나라면 대상존이 다른 존에 둘러싸여 있다고 판단한다.

### 3) 응집성 및 행정구역의 강행

- Ward 알고리즘에서는 모든 결합 가능한 존의 짝에 대하여 각 지표를 구하고 이것을 각각 정규화시켜 하나의 지표를 생성하였으나, 시뮬레이티드 애닐링에서는 하나의 BSU를 택하고 이 이동을 수락할 것인지의 여부를 결정하여야 하므로 달리 비교의 대상이 존재하지 않는다. 따라서, 비계층적 방법에서는 먼저 주어진 데이터 값에 의하여 존 구성을 실시하고, 응집성 및 행정구역 강행을 위한 후처리 과정을 별도로 수행하는 방식을 택하였다. 따라서, 수정된 알고리즘은 그 원형의 <단계 2>에서 임의의 한 개체를 선택하는 과정만 달라진다. 첫째, 임의의 BSU를 모두 선택가능하게 하는 것이 아니라 존의 외곽에 위치한 BSU들 중 하나를 선택한다. 이는 존 내부에 있는 BSU는 다른 존으로 이동할 수 없기 때문이다. 두번째, BSU와 이것을 이동할 존을 택한 후에, 이 이동에 대하여 존당 BSU 수와 연속성 유지여부를 검사하는 과정이 추가된다. 응집성 강행을 위한 후처리 과정의 알고리즘은 아래와 같다.

<단계 1> 시뮬레이티드 애닐링에 의해 얻어진 결과를 초기 파티션으로 정의한다.

<단계 2> 존 외곽의 BSU 중 가장 낮은 번호를 가진 것을 선택한다.

<단계 3> 선택된 BSU와 인접한 모든 존에 대하여, BSU 이동 전후의 응집성

Cbefore, Cafter를 구한다.

Cbefore = BSU를 잃는 존의 이동전 응집성 + BSU를 얻는 존의 이동전 응집성

Cafter = BSU를 잃는 존의 이동후 응집성 + BSU를 얻는 존의 이동후 응집성

- 단, 선택된 BSU가 속한 존과 대상 존의 응집성이 이미 기준값보다 작은 경우 이 존은 고려에서 제외된다.

<단계 4> Cbefore > Cafter이고 가장 작은 Cafter값을 만드는 존을 선정한다.  
그러한 존이 없으면 <단계 6>로 간다.

<단계 5> 선정된 BSU를 선정된 존으로 옮긴다.

<단계 6> 존 외곽에 위치한 모든 BSU가 선택되었으면 다음단계로, 그렇지 않으면 다음 BSU를 선택하고 <단계 3>으로 간다.

<단계 7> 이번 반복(iteration)에서 BSU의 이동 수락이 있었으면 <단계 2>로 간다.  
그렇지 않으면 종료한다

<표 4-5> 응집성 및 행정구역 강행 기준값

구분	식
응집성	모형 1 : $4 / (\pi \times \text{응집성강행도})$ 모형 2 : $4\pi / \text{응집성 강행도}$
행정구역	$1 - \text{행정구역 강행도}$

- 행정구역 강행 알고리즘의 경우도 위와 동일하다. 다만, <단계 3>에서 언급된 응집성의 기준값은 주어진 응집성 강행도를 반영하기 위한 것으로서, 응집성 및 행정구역 강행의 기준값은 <표 4-5>와 같다.
- 즉, 강행모형 2에서 응집성 강행 지표값은  $4\pi$ 보다 큰 값이므로 이를 응집성 강행도로 나누어 주면 응집성 강행도 1일 때  $4\pi$ , 그리고 0일 때  $\infty$ 의 값을 얻을 수 있다. 따라서, 강행도가 1일 때에는 대상 존이 완전한 원형이 되기 전까지는 모든 이동을 고려하게 되며, 0 일 때에는 어떠한 값이 나오더라도 고려 대상에서 제외된다. 또한, 행정구역의 경우에 그 지표값이 (근거 행정구역에 속하지 않은 BSU 수 / 해당 존의 BSU 수)이므로 이값은 0~1 의 범위에 속하게 된다. 따라서, 제시된 기준값 식을 통해 강행도 1일 때 기준값 0을, 0일 때 1값을 얻어 이 값보다 작은 값을 가진 대상 존은 고려대상에서 제외시킴으로서 주어진 강행도를 반영할 수 있게 된다.

## 제4절 모형간 특성비교

### 1. 수정알고리즘의 구현 및 적용

- 이장과 같이 수정된 알고리즘의 시험을 위해 서울시를 연구 영역으로 하여 알고리즘을 수행해 보았다.

<표 4-6> 수정된 시뮬레이티드 애닐링 실행결과

응집성 초기치	행정 구역 초기치	모형	유사성	응집성		행정구역	
				모형 1	모형 2	모형 1	모형 2
0.5	0.5	1	2.338089	82.102539	2082.952637	102	9.918253
		2	2.762145	67.081032	1288.583008	73	3.887302
1.0	1.0	1	2.721489	57.678795	1659.335815	87	6.387007
		2	2.718163	64.397774	1292.845215	73	4.476002

- 전술한 바와 같이 알고리즘 구현은 마이크로 소프트웨어 C 및 C++로, 그리고 GIS부분의 구현은 PC ARC/INFO 및 SML(Simple macro language)를 사용 하였으며, 사용된 자료는 1993년 서울시의 것으로 494개의 동과 22개 구로 구성되어 있다. 존구성을 위해 이용한 자료는 과세 연상면적으로서, 주거, 상업, 사무실, 공장 면적이다. 속성값에 대한 가중치는 동일하게 부여하였으며, 존당 BSU수의 허용비는 Ward 알고리즘의 경우 0.2, 비계층적 알고리즘의 경우 0.5로 하여, 50개 존을 구성하였다.
- Ward 알고리즘의 경우 수정된 알고리즘은 4종이 있으며, 이는 응집성 및 행정구역 강행 모형이 각각 2종씩 있으므로 이의 조합으로 이루어진 것이다. 즉, 아래 <표 4-6>에 제시된 실행결과에서, 모형 1-2는 응집성 모형 1, 행정구역 강행 모형 2를 사용한 알고리즘이다. 또한 제시된 수치는 모두 각 존의 지표값을 구한 뒤 전체 연구 영역에 존재하는 모든 존의 지표값을 합산한 것이다. “응집성”과 “행정구역”은 각각의 강행도를 뜻한다.
- 수정된 시뮬레이티드 애닐링과 k-means 역시 Ward 알고리즘에서와 동일한 조건으로 실행하였으며, 이 때의 결과는 <표 4-7>과 <표 4-8>에 제시하였다. 다만, BSU 교환에 의한 알고리즘에서는 행정구역 강행 모형 1번이 의미가 없어지게 된다. 이는 하나의 행정구역 내에서의 이동이 모두 행정구역 지표값 0으로 처리되므로 생기는 문제이다. 따라서, 여기서는 행정구역 강행 모형은 모형 2 하나로 했다.



- 각각의 평가지표는 강행 모형식과 같은 것을 사용한다. 즉, 주어진 식으로 응집성 및 행정구역을 강행하였을 때, 그 강행이 어느정도 잘 수행되는가를 평가한 것이라 할 수 있다. 또한, 표에서 제시한 값은 각 존의 해당 지표값을 모두 더하여 얻은 값이므로, 그 값이 적은 것이 해당 기준에 더 잘 부합되는 것이라 할 수 있다. 다만, 행정구역 평가 모형 1은 강행 모형과는 다르게 하나의 존이 걸쳐 있는 행정 구역의 수로 하였다. 이는 이 모형이 조합 혹은 이동시에 대상 존간 혹은 대상 BSU간에 행정구역의 존재여부에 따라 값이 주어지는 것이므로, 모든 과정이 완료된 후의 평가지표로는 적절치 않기 때문이다.

## 2. 각 알고리즘기반의 모형간 비교분석

- 이상의 실행결과에서 다음과 같은 사실을 확인할 수 있다.
- 첫째, <표 4-7>의 실행결과에서와 같이 계층적 알고리즘의 경우, 응집성 강행도는 존의 외곽선을 사용한 모형 2가, 그리고 행정구역 강행도는 조합된 존 간의 경계선의 존재 유무를 기반으로 하는 모형 1이 상대적으로 우수한 강행결과를 생성해 내었다.

<표 4-7> 수정된 Ward 알고리즘 실행결과

응집성 초기치	행정구역 초기치	모형	유사성	응집성		행정구역	
				모형 1	모형 2	모형 1	모형 2
0.5	0.5	1-1	2.852478	68.191284	1292.138184	52	0.733333
		1-2	2.824088	68.725296	1307.129272	53	0.739560
		2-1	3.020945	64.305588	1208.095337	52	0.375000
		2-2	3.027459	63.278999	1200.289795	55	1.200000
1.0	1.0	1-1	2.909623	60.553429	1236.379028	53	0.800000
		1-2	2.899560	59.588608	1249.800293	54	0.850672
		2-1	2.967692	67.239838	1206.386597	51	0.200000
		2-2	2.964513	66.753365	1184.333862	54	0.626136

<표 4-8> 수정된 k-means 실행결과

응집성 초기치	행정 구역 초기치	모형	유사성	응집성		행정구역	
				모형 1	모형 2	모형 1	모형 2
0.5	0.5	1	2.616133	78.151497	1755.130005	79	4.731805
		2	2.707003	66.018349	1305.923096	66	1.981974
1.0	1.0	1	2.711228	61.513149	1540.994385	74	2.835198
		2	2.674829	67.292603	1264.949097	65	2.113331

- 둘째, 비계층적 알고리즘에서는 외곽선에 기반한 응집성 강행 모형 2가 우수한 강행도를 보여주었다. 따라서, Ward 알고리즘의 예와 연관지어 생각해 볼 때, 최대적정에 기반한 응집성 모형 1 보다는 외곽선에 기반한 것이 우수한 강행도를 보임을 알 수 있다.
- 셋째, 이상과 같이 수정된 알고리즘들을 응집성 및 행정구역 강행면에서 보면, 수정된 Ward 알고리즘이 우세한 것으로 나타난다. 유사성면에서 볼 때는 다른 비계층적 알고리즘이 우세하다고 볼 수 있으나, 유사성 이외의 기준들이 강행될 때 유사성에 약재로 작용한다는 점을 생각하면 당연한 결과라 할 수 있다. 이것은 비계층적 알고리즘이 개체의 그룹간 이동에 의해서 동작하는데서 기인하는데, 존 구성시 BSU의 이동에는 단순한 군집형
- 성 과정과 달리 각종 공간적인 제약이 가해지므로, 그 이동이 원활하게 일어나지 못하기 때문이다. 따라서, 존 구성 알고리즘으로 수정된 비계층적 알고리즘은, 존의 크기가 작고 존당 BSU 수 변화비가 클 때, 좋은 결과를 보이게 된다.

<표 4-9> BSU 수 변화 허용비가 다를 때 알고리즘 수행결과

기준 허용비	유사성	응집성		행정구역	
		모형 1	모형 2	모형 1	모형 2
0.2	2.641967	72.930870	1775.113525	89	8.477381
0.5	2.674829	67.292603	1264.949097	65	2.113331

- 실행 예에서 비계층적 알고리즘의 BSU 수의 변화비를 크게 한 것도 이같은 이유에서이며, 실제로 그 값을 낮추었을 때는 <표 4-9>와 같이 결과에 현격한 차이를 보인다. 이것은 시뮬레이티드 애닐링에서 BSU 수의 변화 허용비가 0.2와 0.5일 때 그 수행결과를 비교한 것으로 응집성 강행도 1.0, 행정구역 강행도 1.0으로 실행하였으며, 응집성 모형 2를 사용하였다.

<표 4-10> 수정된 알고리즘의 수행시간 비교

알고리즘	Ward	Simulated Annealing	k-means
실행시간(초)	236.04	651.59	440.05

- 마지막으로, 수정된 각 알고리즘의 실행시간은 <표 4-10>에 나타난 바와 같다. 실행시간 측정은 위해 사용한 시스템은 IBM-PC 호환기종으로서 CPU 클럭 66MHz인 486 DX-2이다. 또한 각 알고리즘에 사용한 응집성 및 행정구역 강행 모형은 각각에서 가장 강행도가 좋은 모형을 선택하였다. 따라서, Ward 알고리즘은 응집성 모형 2와 행정구역 모형 1을, 비계층적 알고리즘은 응집성 모형 2를 사용한 것이다. 수행결과를 비교해 볼 때, 실행시간면에서도 수정된 Ward 알고리즘이 우수하게 나타난 것을 확인할 수 있다.
- 이상과 같이 각 알고리즘을 실행한 결과는 <그림 4-3>과 <그림 4-4>에 나타난 바와 같다. 여기에서는 응집성 및 행정구역 강행도가 1일 때의 결과를 제시하였으며, 예시를 위해 Ward 알고리즘의 경우에만 그 강행도가 모두 0일 때의 결과를 제시하였다.



(a) 응집성, 행정구역 강행도 0.0



(b) 응집성, 행정구역 강행도 1.0

&lt;그림 4-3&gt; 수정된 Ward 알고리즘 수행결과



(a) 수정된 시뮬레이티드 애닐링



(b) 수정된 k-means

&lt;그림 4-4&gt; 수정된 시뮬레이티드 애닐링과 k-means 알고리즘 수행결과

## 제5장 결 론

---

제1절 결 론

제2절 향후과제

## 제1절 결 론

- 본 연구에서는 TAZ의 일반적 원칙, 제반 교통사업에서의 존구성체계의 적용상황을 살펴보고, 최근의 교통 DB 사업의 진척 및 이에 따른 수치지도의 연계사업등을 고려하여 향후의 존구성체계의 개발을 목적으로 프로토타입의 사례연구를 수행하였다.
- 존구성원칙아래 일반적인 존의 구획기법으로 대표적인 군집형성 알고리즘을 수정하여 존 구성 알고리즘제안하고 이를 바탕으로 응집성과 행정구역의 강행여부에 따라 여러 모형을 상정한후, GIS를 기반으로 존 구성체계를 구축하였다. 이 과정에서 3가지의 큰 알고리즘의 수정과정을 통해서 응집성 및 행정구역 강행에 따른 모형을 제시하고 수행결과를 유사성과 응집성 및 수행시간 측면에서 비교하였다.
- 실행결과를 토대로하여 응집성 및 행정 구역 강행면에서 볼 때, 응집성 강행 모형은 생성된 존의 외각선의 길이에 기반한 모형이 최대직경에 기반한 모형보다 우수하였다. 또한, 행정구역 강행 모형은 대상 존간 행정구역 경계선의 유무에 기반한 모형이 우수하였다.
- 구현된 알고리즘을 비교해 볼 때, 첫째, 구현면에서는 Ward 알고리즘이 개념이 단순하며, 비계층적 알고리즘과 같이 별도의 연속성 검사를 필요로 하지 않으므로 구현이 용이하였다. 둘째, 유사성은 알고리즘별 변화는 두드러지지 않았으며, 응집성과 행정 구역 강행도에 반비례하는 경향이 더욱 뚜렷하였다. 셋째, 응집성 및 행정구역 강행은 비계층적 알고리즘에 비해 Ward 알고리즘이 우수하였다. 넷째, 수행시간은 Ward 알고리즘이 가장 적었고, k-means, 시뮬레이티드 애닐링의 순서를 보였다.
- 이상의 결과를 볼 때, 계층적 알고리즘이 전반적으로 존 구성 알고리즘 구축에 유리함을 알 수 있다. 물론, 현재의 수정된 비계층적 알고리즘은 유사성과 기타 기준을 동시에 강행하지 못하고 별도의 후처리를 행한다는데서 문제점을 찾을 수 있으나, 공간적인 여러제약으로 인해 BSU의 교환이 원활히 일어나지 못하는 것이 존 구성에 비계층적 방법을 적용할 때의 문제점이라 할 수 있다. 따라서, 비계층적 알고리즘을 존 구성에 본격적으로 활용하기 위해서는 이같은 공간적인 제약을 극복할 수 있는 다른 접근이 요구된다.

## 제2절 향후과제

- 한편, 본고에서 제시된 알고리즘과 모형을 토대로 수행한 결과는 어디까지나 특정알고리즘과 모형에 기반한 유사성, 응집성을 수치로 비교할 때 이것이 작은 것을 우수한 모형으로 보았음을 밝히며, 사실 어떠한 알고리즘 기반의 어떤 모형이 실제 목전에 둔 교통계획과정에서 우세한지를 판단하기는 어렵다고 본다. 왜냐하면, 유사성과 응집성을 기반으로 구축된 모형기반의 교통계획과정이 (최종산출물을 가로망에서의 1일 교통수요라고 보면) 결과에 대한 정확도를 보장하지 못하기 때문이다. 다시 말하면 수정된 Ward알고리즘 기반의 어떤모형이 k-means알고리즘 기반의 어떤모형보다 유사성과 응집성면에서는 우세하더라도 수요의 정확도를 다음과 같은 예측과 실측치의 차이로서 나타낼때는 (예를들면, RMS-root mean square) 반전되는 결과를 나타낼 수도 있기 때문이다.

$$RMS = \sqrt{\sum_i (O_i - E_i)^2}$$

(여기서 여기서 i는 링크를, O는 실제 교통량을, E는 모형에 의한 교통량을 나타냄.)

- 따라서 모형과 알고리즘의 선택은 계획가에 따라서 현상을 잘 모사하는 것을 선택할 필요가 있다고 보여진다.
- 끝으로 교통계획의 효율성의 증진이란 관점에서 보면, 이러한 체계 구축 및 실행과정을 통해, GIS와 TAZ 구성 모듈을 연계하여 자동화된 TAZ 구성체계를 구축함으로써 일련의 교통계획과정에 유연성을 부여할 수 있는 일은 의미가 있으며 특히 제반 자료가 최근 GIS화 하는 과정중임을 생각하면 상당한 장래의 효율성 증진이 기대된다.
- 향후 연구로서는 이러한 존 구성체계를 실제 적용하여 존-네트워크(zone-network) 적합성(compatability)을 각 단계의 교통계획별로 검증하고 존 구성단계의 본고에서 제시된 모형간의 초기치 부여등 어느정도의 관련 적절한 표준값 도시, 지역 및 국가 차원에서 제시하는 작업이 필요할 것이며 존 구성이 가로망에 미치는 영향의 민감도에 대한 규명도 꼭 필요한 작업으로서 일반적인 경향을 파악하는 작업도 부가적으로 필요할 것이다. 아울러 상세히 존을 나누지 않고도 비교적 정확한 가로교통량을 얻어내는 제반기법 (Daganzo등이 1980년대에 시도한 연구등)에 대한연구등이 필요할 것이다.

## 참 고 문 헌

- 최기주, "GIS-T의 필요성과 활용방안", 우리나라 지리정보 시스템의 기본구상과 교통부문 활용방안, 교통개발연구원, 1995
- A. D. Cliff and P. Haggett. On the efficiency of alternative in region-building problems. *Environment and Planning A*, 2:285-294, 1970.
- Allen J. Scott. *Combinatorial Programming, Spatial Analysis and Planning*. Methuen and Co. Ltd., London, UK, 1971.
- ESRI(Environmental Systems Research Institute). *PC ARC/INFO STARTER KIT Users Guide*. Environmental Systems Research Institute Inc., Redland, CA, 1990.
- J. H. Ward. Hierarchical grouping to optimize an objective function. *Journal of the American statistical association*, 58:236-244, 1963.
- Jeff R. Wright. ISIS: Toward an Integrated Spatial Information System. In T.J. Kim, L.L. Wiggings and J.R. Wright, editors, *Expert Systems: Applications to Urban Planning*, chapter 3, pages 43-65, Springer-Verlag, New York, NY, 1990.
- Keechoo Choi. *The Implementation of an intergrated transportation planning model with GIS and expert systems for interactive transportation planning*. PhD thesis, University of Illinois, 1993.
- M. Keane. The size of the region-building problem. *Environment and Planning A*, 7:575-577, 1975.
- Mark S. Aldenderfer and Roger K. Blashfield. *Cluster analysis*. Sage Publications Inc., 1984.
- Michael R. Anderberg. *Cluster analysis for applications*. Academic press, 1973.
- N.J. Garber and L.A. Hoel. *Traffic and Highway Engineering*. West Publishing Co., New York, NY, 1988.
- S. Openshaw. Optimal zoning system for spatial interation modes. *Environment and Planning A*, 9:169-184, 1977.
- William Macmillan and Todd Pierce. Optimization modeling in GIS framework: the problem of political redistricting. In Stewart Fortheringham and Peter Rogerson, editors, *Spatial Analysis and GIS*, chapter 11, Taylor & Francis Ltd, 1994.