

■ 論 文 ■

# 링크 통행시간 추정을 위한 데이터 퓨전 알고리즘의 개발

A Data Fusion Algorithm for Link Travel Time Estimation

최 기 주

(아주대학교 교통공학과 조교수)

정 연 식

(아주대학교 교통공학과 대학원)

## 목 차

- I. 서론
- II. 연구의 배경
  - 1. 교통정보 생성과정
  - 2. 자료의 가공시 발생하는 문제
  - 3. 정보의 추정에 필요한 사항
- III. 이론적 배경
  - 1. 퍼지논리(Fuzzy Logic)
  - 2. 퍼지회귀모형(Fuzzy regression model)
  - 3. 데이터 퓨전
  - 4. D-S 규칙을 이용한 데이터 퓨전
- IV. 통행 시간 추정을 위한 데이터 퓨전 알고리즘 개발
  - 1. 퍼지모형의 구축
  - 2. 퍼지선형회귀모형의 구축
  - 3. 각 그룹별 신뢰값의 퓨전
  - 4. 온라인 통행시간 추정을 위한 퓨전
- V. 모의자료의 적용
  - 1. 자료의 정리
  - 2. 실시간 자료의 그룹별 분류
  - 3. 소속함수를 통한 소속정도 산출
  - 4. 소속정도의 조합
  - 5. 온라인 통행시간 계산
  - 6. 동적 자료의 구축
- VI. 결론  
참고문헌

## 요 약

지능형교통체계(ITS: Intelligent Transport Systems)의 구현을 위한 가장 중요한 요소중의 하나는 교통정보의 생성이다. 교통정보의 생성은 루프 검지기, 폐쇄회로(CCTV), probe 차량, 경찰, 통신원 등을 통해 수집된 제보자료들을 분석 및 가공함으로써 이루어진다. 그러나 이들 수집원은 주어진 시간에 있어 모든 네트워크를 통해서 자료가 완전히 수집되어지는 것은 아니다. 즉, 특정 지역에 수집원이 몰려 있는 경우가 있는 반면, 전혀 수집되어지지 않는 지역이 발생할 수도 있다. 이러한 공간적인 불균형적 특성은 동시에 발생한 다량의 자료를 처리하는 기술과 자료가 수집되지 않은 지역에 대한 처리기술을 요하게 된다. 본 논문은 전술한 바와 같은 사항에 대하여 ITS의 진행 단계별로 드러날 수 있는 문제점을 검토하고, 자료통합에 대한 일반적인 개념을 우선 설명한다. 다음에 특정시간에 주어진 자료의 통합을 위해 퍼지선형회귀모형(fuzzy linear regression model)과 데이터 퓨전(data fusion)기법의 내용을 소개하고, 신뢰성있는 단일 교통정보 생성을 위한 데이터 퓨전 알고리즘을 제시한다. 또한 제시된 알고리즘을 토대로 가상의 자료를 이용하여 적용가능성을 타진해 보았다. 제시되어진 알고리즘은 향후 교통정보 수집환경이 어느 정도 형성된다고 볼 때, 예측치와 실측자료간의 자료검증을 통하여 신뢰도를 가질 경우 보다 광범위하게 사용되어질 수 있을 것으로 판단된다.

### I. 서론

ITS(지능형교통체계: Intelligent Transport Systems)는 혼잡의 완화, 용량증대의 효과, 안전성의 제고, 교통으로 야기되는 환경문제 해결 등의 목표 아래 기존의 교통체계에 첨단 컴퓨터, 통신, 실시간 자료 처리 기술 등을 접목시켜 효과적인 개선을 이룩하려는 새로운 시도이다.

ITS 기술의 적용은 혼잡상황의 예고와 대안경로의 안내 등의 정보를 도로 이용자에게 제공함으로써 도로에서 소비되는 총체적 비용을 최소화하는데 있으며, 이러한 ITS에는 세부분야(ATMS, ATIS, APTS, CVO, AVHS)<sup>1)</sup>가 있지만 특히 ATIS의 주목적은 적절한 통행 정보를 제공함으로써 O-D(Origin-Destination)간을 통행하는 차량의 통행시간을 최소화하는 것이다.

통행 정보를 제공하기 위해서는 우선 교통정보 수집체계로부터 자료를 충분히 수집해야 한다. 보다 정확하고 최신의 자료를 수집하기 위해서는 다양한 수집체계로부터 자료를 실시간으로 수집해야 하며, 수집된 자료는 제반 분석기법을 통해 교통정보로 변환되어 전달되게 된다.

한편, 분석되어 산출된 교통정보는 교통체계를 이용하는 이용자에게 만족되어야 한다. 즉, 이용자가 원하는 시기에 원하는 장소의 교통정보가 제공되어야 하는 것이다. 이러한 요구를 만족하기 위해 적어도 ATIS센터 또는 이와 유사한 기능을 부여 받은 단위는 교통 정보 수집체계의 확보와 더불어 교통 소통관련 자료가 중복적으로 수집된 구간 및 수집되지 않은 구간(missing link)에 대해서도 교통 상황을 추정할 수 있는 기능이 갖추어 져야 한다.

본 논문은 이러한 맥락에서 동일한 시공간적 환경하에 다양한 수집체계로부터 수집되는 교통 관련 정보를 효과적으로 통합하기 위한 방법으로 이른바 자료의 융합 또는 데이터 퓨전(이하 data fusion)이라 불리는 기법에 대한 소개와 이를 적용하여 보다 신뢰성 있는 교통정보를 추정하기 위한 알고리즘 제시를

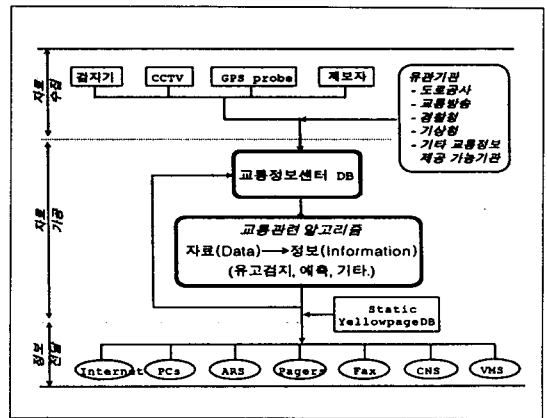
그 목적으로 한다.

### II. 연구의 배경

ITS환경 아래서 교통정보의 기초자료가 형성되어질 때 두가지의 큰 과업이 발생하게 된다. 첫째는 교통정보를 단위수집체계로부터 생성하는 것이고, 다른 하나는 복수의 수집체계로부터 생성된 정보를 통합하는 일이라고 볼 수 있다. 이러한 각각의 과업에 대해서 전체적으로 해법을 제공하기 위해서 이 논문은 작성되었는데 본 절에서는 교통정보 생성 및 통합과정에 있어서의 고려사항과 필요과업을 살펴보기로 한다.

#### 1. 교통정보 생성과정

일반적으로 교통정보 제공을 위한 진행과정은 교통관련 자료의 수집, 분석처리, 전달 시스템으로 나누어진다. 다음의 <그림 1>은 우리나라의 상황에서 연출되어 질 수 있는 교통정보 제공을 위한 과정의 한 예를 간략하게 보여주고 있다.



<그림 1> 교통정보 제공을 위한 과정

그림에 나타난 것처럼 일반적으로 교통관련 자료의 수집장치는 루프 검지기(loop detector), 폐쇄회로

1) ATMS : Advanced Traffic Management System(첨단교통관리체계), ATIS : Advanced Traveler Information System(첨단운전자 정보제공체계), APTS : Advance Public Transportation System(첨단대중교통체계), CVO : Commercial Vehicle Operations(첨단화물운송체계), AVHS : Advanced Vehicle & Highway System(첨단도로 및 첨단 차량 제어 체계)

텔레비전(CCTV), GPS probe 차량<sup>2)</sup>, 통신원의 제보자료, 기타 유관기관 등이 포함된다. 수집된 자료는 교통정보를 추출하기 위해 분석 처리되어 인터넷, PC통신, ARS(자동응답전화), 호출기, 팩스, VMS(Variable Message Sign:가변정보전광판), CNS(Car Navigation System:차량의 자동항법장치) 등을 통해 이용자에게 전달되게 된다.

이상의 일련 과정 중에서 자료의 수집과 전달 부분은 주로 하드웨어에 기반한 부분이라 할 수 있고, 자료의 가공 부분은 소프트웨어에 기인하는 부분이라 할 수 있다. 즉, 자료의 수집과 전달 부분은 많은 비용이 소요되는 인프라구축 부분과 관련이 있는 반면, 자료의 가공 부분은 주로 알고리즘의 개발과 같은 비교적 적은 비용으로 고품질의 교통 정보를 추출하는 부분에 관련이 있다고 볼 수 있다.

본 논문에서는 소프트웨어에 기반하는 부분 즉, 자료의 가공에 관한 부분을 주로 서술하게 될 것이다. 그러나, 교통에 관련된 자료에 대한 가공은 교통의 특수한 특성<sup>3)</sup> 때문에 일반적으로 다음과 같은 사항을 고려해야 한다.

## 2. 자료의 가공시 발생하는 문제

모든 정보체계가 그러하듯이 ITS에서도 정보(information)는 자료(data)를 가공해서 생성되는 것으로서 소위 원시자료(raw data)라고 불리는 자료가 정보의 질적 수준에 직접적인 영향을 미친다. 이러한 원시 자료는 검지기 결과 자료(교통량, 점유율), probe 차량 등에 의해 측정된 미 가공 통행시간, 혹은 운전자나 기타 통신원(경찰, 비상요원, 등등)에 의한 제보자료 등이 포함되며, 가공된 정보는 혼잡한 상황을 지도상(congestion map)에 표출하거나, 네트워크 상에서 주요지점 사이의 추정된 통행시간 또는 유고 등의 형태로 나타내어 전달된다. 이 과정에서 교통정보

를 계산 및 추정하기 위해서는 다음과 같은 문제를 해결해야 한다(Tarko & Roupail, 1997).

- ① 특정 위치에서 일정 시간동안 자료를 이용할 수 없거나 수집이 되지않은 경우(data sparsity)
- ② 특정 기간에 다양한 수집체계로부터 유사한 정보가 다량으로 수집된 경우(data redundancy)
- ③ 불규칙적으로 발생하는 교통량과 자료의 불완전성에서 발생하는 자료의 불확실성(data uncertainty)

위와 같은 자료처리에 관련된 문제는 ITS의 진행 상태별로 차이가 있을 수 있기 때문에 현재 ITS 환경 하에서의 자료 확보와 관련한 ITS의 성숙 정도가 어떤 단계인가를 구분하여 살펴봐야 할 필요가 있다. 따라서 본 고에서는 ITS 환경의 초기 및 진보 단계 각각에 대하여 살펴 보기로 한다.

### 1) ITS 환경의 초기 단계

#### (1) 수집자료의 부족(Data sparsity)

ITS 진행의 초기단계에서 교통관련 자료에 대한 수집체계는 상대적으로 열악하여 주로 중요한 링크(link)에 국한되어 이루어지는 경우가 많다. 통행시간을 산출해 줄 수 있는 원천으로서의 probe 차량 및 검지기가 얼마 존재하지 않을 수 있으므로 대부분의 네트워크에 대한 실시간 자료를 기반으로 하는 통행시간 계산 방식의 접근은 불가능하다. 이와 같은 어려움은 실시간 자료가 수집되지 않을 때 누적된 교통 패턴 자료를 이용하여 부분적으로는 해결할 수 있다. 이 방법이 미국의 ADVANCE 프로젝트에서 사용된 접근 방식이다(Boyce et al. 1993). 또 다른 방식으로는 최근에 수집된 동일링크의 정보로부터 또는 인접 링크의 정보로부터 "missing" 링크의 교통 패턴을 추

2) GPS probe 차량 : GPS수신기를 탑재한 차량은 도로를 주행할 때 GPS 위성과의 통신으로 차량의 위치를 지속적으로 파악(x, y, z의 좌표 값과 수신시간 t의 자료를 수집할 수 있음)할 수 있다. 이 자료를 토대로 통행시간 정보를 수집하는 방식에 있어 실제 주행차량을 일컫는다. 통상 probe car 혹은 probe vehicle로 불리우나 본 논문에서는 probe 차량으로 통일하여 지칭할 것임.

3) 대표적인 특성으로는 교통 상황 변화의 무작위성(randomness), 교통 관련 자료의 막연함(vagueness), 불확실성(uncertainty) 등을 들 수 있다.

론하는 방식으로 이 경우 추론에 대한 메커니즘이 필요하게 된다.

## (2) 수집자료의 중복(Data redundancy)

비록 자료 수집체계가 네트워크에 희박하게 배치되어 있더라도 동시에 다양한 수집체계로부터 동일구간의 중복되는 정보를 수집할 수도 있다. 이러한 자료의 중복성으로 인해 분석되는 교통정보의 신뢰수준은 일반적으로 증가하게 된다. 한편, 다량의 정보를 이용하는 가장 간단한 방법은 동일한 시점에서 신뢰도가 가장 높은 수집체계에 대한 정보를 선택하는 것이다 (Sumner, 1991). 그러나, 보다 효율적인 방법은 다양한 수집체계로부터 수집된 정보를 조합하는 방안이다 (Tarko & Rouphail, 1993).

## (3) 수집자료의 불확실성(data uncertainty)

통행시간, 대기행렬, 교통량과 같은 자료를 분석 처리하는 데에는 많은 무작위성(randomness)이 관련하게 되는데, 특히 자료의 희박성으로 인해 발생하는 불확실성은 교통 특성을 추정하는데 많은 부담 요인이 된다. 이러한 경우가 적은 수의 probe 보고자료로부터 링크 평균 통행시간을 추정해야하는 ITS진행의 초기 단계에 발생하게 된다. 이러한 상황은 probe 보고 자료가 교통정보로 전환될 수 있기 때문에 검지기를 통한 링크 통행시간 추정 방법보다는 바람직한 방법이라 할 수 있다.

결론적으로 ITS 진행의 초기단계에서 자료와 관련하여 발생하는 주요 문제점은 완벽하고 양질의 교통 자료를 수집하기 어렵다는 데 있다. 따라서, 보다 완성된 정보를 추론하여 정보의 신뢰도를 개선하기 위한 '정보 통합 기법'이 요구된다. 이러한 맥락에서 데이터 퓨전은 ITS 진행의 초기 단계에서 정보의 통합 기법으로 매우 유용하게 이용될 수 있다고 본다.

## 2) 초기의 ITS 환경이 발전된 단계

ITS 환경이 어느 정도 발전된 단계에서는 자료수집 장치(예를 들면, 충분히 많은 probe 차량이 네트워크에 존재)가 어느 정도 존재한다고 가정해 볼 수 있으며, 또한, 새롭고 비교적 저렴한 검지기술이 보조간선

도로 이하의 도로에까지 확장되게 될 것이다(논의의 편의를 위해 그렇게 가정함). 위와 같은 환경 하에서는 초기 단계에 존재했던 자료의 희박성과 같은 문제는 어느정도 해결된다. 또한 정보의 불확실성은 원시 자료 규모가 풍부해지기 때문에 상대적으로 줄어들 것이다. 그러나, 이러한 환경에서도 다음의 두 가지의 경우를 고려해야 한다.

## (1) 대량의 자료 처리(Large amount of data)

비록 단일 수집체계로부터 수집된 많은 양의 자료가 교통정보의 정확성을 향상시키는 바람직한 방법일지라도, 추가적으로 통신장비, 저장장비, 계산장비가 필요하다. 이러한 상황에서 그리고 충분히 큰 네트워크에 대하여 수집되어진 데이터는 통신, 저장, 계산 시스템의 용량 초과에 대한 문제가 발생할 수 있다.

## (2) 수집자료의 중복(Data redundancy)

자료가 풍부하다는 것은 정보의 신뢰도를 제고하는 측면에서 바람직하며, 또한 풍부한 자료는 조사 시스템의 오류를 검지할 수 있는 보조기능을 제공할 수도 있다. 그러나, 정보의 질을 비교하지 않고서라도, 가장 낮은 신뢰도를 보이거나 높은 비용이 드는 수집체계는 사용하지 않는 것이 바람직하다.

다양한 수집체계로부터 자료를 수집할 때 특정 수집체계가 또 다른 수집체계보다 언제나 신뢰도가 높다고는 할 수 없을 것이다. 또한, 한 수집체계가 다양한 응용분야(유고, 용량, 차선 폐쇄에 대한 계획, 교통량, 교통밀도, 대기행렬, 현재 신호 제어, 등등)에 필요한 모든 교통 자료를 제공할 수는 없다. 때문에 다양한 수집체계에 대한 자료의 통합을 통해 단일의 정보를 추출하는 작업이 필요하게 되며 이것이 본 연구에서 성취하려는 목표이다.

결론적으로 ITS 진행이 발전된 단계에서 가장 중요한 문제는 정보의 과다량이다. 때문에 독립적인 수집체계로부터 수집된 원시자료의 량은 효율적인 자료 통합(데이터 퓨전)을 통하여 현실적으로 줄이는 것이 필요하다.

## 3. 정보의 추정에 필요한 사항

전술한 바와 같이 자료를 처리하거나 정보를 추정하

는데 있어서 통상 다음의 3가지 사항이 필요하게 된다. 첫째, 주변 링크에서 수집한 정보로부터 특정 링크에 대한 교통정보의 추정. 둘째, 불확실한 정보의 처리. 셋째, 다양한 수집체계로부터 정보의 조합 등이 그것이다.

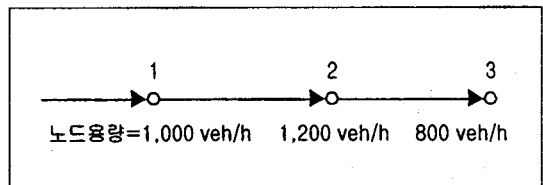
새로운 정보에 대한 추론은 관련된 제반 지식(knowledge)을 순차적으로 적용하여 구축될 수 있다(Tarko & Roupail, 1997). 즉, 지식을 기반(knowledge-based)으로 한 접근방식은 주변의 다양한 상호 보완적인 자료를 이용하여 특정 링크(missing link)에 대한 교통정보 추정에 이용된다. 이러한 지식 기반 접근 방식은 "missing" 링크 정보의 추정에 이용될 뿐 아니라 원시 자료의 해석에도 유용하다.

다음으로, 자료의 처리는 불확실한 상황을 지식으로 나타낼 수 있어야 하며, 이 자료는 추론 과정을 통해 최종 결론을 얻게 된다. 전통적인 논리에서 보면 명제는 참 혹은 거짓으로 나타내어지고, 불확실성은 계제될 수 없다. 그렇지만 베이지안 규칙(Bayesian rule)을 이용하여 전통적 논리를 확률적으로 확장하면 불확실성<sup>4)</sup>에 대한 표현이 어느 정도 가능해 진다(Genesereth & Nilsson, 1987). 이 접근 방식에서 적용된 불확실성은 무작위성에 대한 불확실성을 말한다. 그러나 교통에 관련된 불확실성을 처리할 때 특히 주의해야 할 사항은 대부분의 교통 특성이 정성적인 자료수집 방식(긴 대기행렬 발생과 같은 운전자들의 제보 자료)에 의해 이루어지고 있고, 교통류의 특성은 막연하게 정의("도로가 많이 혼잡되고 있다"와 같은 막연한 정의)되어 표현된다는 것이다.

무작위성과 퍼지성(fuzziness)은 본질적으로 차이가 있는 불확실성으로, 예를 들어 "내일 비가 올 것이다"라는 불확실성은 내일이 되기 전에 예측함으로써 생긴 것이고, 이 불확실성은 시간이 지나 내일이 되면 명백해진다. 그러나, "도로가 많이 혼잡되고 있다"와 같은 표현에서 나타나는 불확실성은 시간이 경과되거나 실험을 실시한다고 해서 명백해지는 것은 아니다. 이러한 의미가 퍼지성과 관련된 불확실성인데, 이것과 무작위성에 관련된 불확실성이 두 개념을 동

시에 처리하기 위해서는 특별한 방식이 요구된다. 퍼지연산논리(fuzzy operator logic:FOL)의 적용은 퍼지성과 무작위성이 포함된 불확실한 상황을 동시에 처리하는데 유용하게 적용된다.

FOL은 불확실성과 막연함을 0과 1사이의 한가지 값으로 표현할 수 있다. 이 값은 새로운 명제가 참이라는 것에 대한 정도의 신뢰도로 해석된다. 동일한 명제는 다량의 정보가 각기 다른 신뢰수준에서 하나 혹은 그 이상의 독립적인 수집체계로부터 정의되었을 때 발생하는 것이다. 그러므로 각기 다른 수집체계를 통해 수집된 다량의 정보에 대한 처리는 막연하고 애매한 표현(fuzziness)을 특정 값으로 표현하는 과정을 거쳐야 한다.



〈그림 2〉 신호가 설치된 간선도로의 예

교통류 이론을 이용하여 새로운 교통정보를 추가로 생성하는 간단한 예를 들어보기로 하자(Tarko & Roupail, 1997). 3개의 교통 신호가 설치되어 있고, 회전에 대한 교통량은 무시해도 좋은 간선도로를 생각해 보자. 〈그림 2〉에서처럼 신호에 대한 용량은 알고 있다고 가정하자.

첫째로 교통정보센터에 1번 노드와 2번 노드 사이(두 번째 링크)에 "혼잡"에 대한 자료가 보고되었다고 하자. 최근 통행시간 정보는 첫 번째와 세 번째 링크에서는 수집되지 않았고 다른 2개의 링크에 대해서는 가용한 온라인 정보가 없다고 하자. 이러한 사실을 기반으로 두 번째 링크의 평균 통행시간은 높게 추정될 수 있다.

그러나, 네트워크를 보다 자세히 살펴보면 두 번째 링크의 교통량은 첫 번째 신호의 용량(1,000대/시)을 초과할 수 없다. 물론, 두 번째 노드의 용량은 1,200대

4) 수학적 이론에서는 퍼지성(fuzziness)이라는 것을 불확실성(uncertainty)의 한 일면으로 이해한다. 그러나, 퍼지이론이 등장하기 이전에는 불확실성에 대한 취급은 확률적 이론을 통해 이루어졌다. 확률의 불확실성은 일반적으로 랜덤성의 개념으로 상징되는 현상의 발생과 관계가 있다.

/시 라는 것은 알고 있다. 일반적인 교통상황에서는 두 번째 링크에 혼잡상황은 발생하지 않기 때문에 혼잡이 발생했을 경우 두 번째 노드 용량은 감소하게 된다(Tarko & Rouphail, 1997).

용량이 감소하는 원인은 다음의 2가지 경우에 가능하다.

첫째, 세 번째 노드(교통신호)에 발생한 대기 행렬이 두 번째 노드까지 스페일백(spill back)이 발생한 경우.

둘째, 유고(incident)로 인해 두 번째 노드의 용량이 감소한 경우.

대기행렬의 스페일백은 3번째 신호의 용량에 민감하므로(용량이 가장 적다) 쉽게 발생할 수 있다. 그러므로 우리는 3번째 링크가 대기 차량으로 꽉 차여있다는 어느 정도의 신뢰성 있는 결론을 내릴 수 있다. 이때 링크에 대한 통행시간은 3번째 링크에 누적된 총 차량의 숫자, 3번째 신호의 유출율(용량), 3번째 링크의 길이 등의 자료를 이용하여 추정할 수 있다(Tarko & Rouphail, 1997).

Probe에 의한 보고 자료가 2번째나 3번째 링크에서 수집되었다고 가정하자. 만일 3번째 링크에 심한 혼잡이 발생했다는 가정이 확실하다면, 이때 두 개의 자료 수집체계는(probe자료와 네트워크 속성자료) 단일 수집체계로부터 수집된 것보다 혼잡에 관한 높은 신뢰율을 산출하게 된다. 또한 이는 링크통행시간의 추정을 개선하게 될 것이다. 그러나, 만일 probe 보고자료가 비혼잡을 나타냈다면(즉, 3번째 링크에 대한 probe 자료가 “짧은 통행시간”을 나타낼 때) 이때, 유고로 인한 두 번째 노드의 용량감소에 관한 우리의 가설은 보다 신뢰성을 띄게 된다. 그리고, probe 보고는 2번째 링크의 상황과는 무관하게 3번째 링크의 평균링크 통행시간 추정에 단독으로 사용되게 된다.

앞의 예에서는 교통류 이론을 이용하여 통행시간 정보를 정확하게 해석하는 과정을 기술하고 있지만, 이와 같은 과정은 데이터 퓨전에 있어서 다음의 세 가지 사항을 이해하는데 도움을 줄 것이다.

첫째, 온라인 자료가 수집되지 않은 링크의

현재 통행시간 추정

둘째, 풍부한 정보의 통합에 의한 통행시간 추정

셋째, 신호화된 간선도로의 유고 감지.

위의 예에서와 같이 특정 지역에서는 소수의 probe 보고자료, 노드의 용량 자료만으로 정보를 추정한 것과 마찬가지로, 적은 수의 자료만으로도 혹은 인접 주변의 자료를 이용하여 교통정보를 추정할 수 있다.

각 링크 통행시간은 운전자가 좋아하는 속도, 다른 차의 방해, 교통제어의 영향, 날씨, 특정시간대별(Time of Day:TOD)과 같이 정량화하기 힘든 많은 요소에 의해 영향을 받는다. 또한 교통 상황에 대한 막연한 정의의 예를 들면, “혼잡”과 같은 막연한 정의는 혼잡과 비혼잡 사이의 통행시간에 대한 정확한 경계가 없기 때문에 정량적인 한 값으로의 표현이 요구된다. 다음에 소개되는 퍼지 논리는 이러한 애매한 개념을 정확한 지식으로 표현할 수 있게 도와준다는데 매력이 있다.

이상에서 우리는 자료의 통합시 발생하는 전반적인 사항에 대해서 알아보았다. 본 논문에서는 “missing” 링크에 대한 추정보다는 불확실한 자료의 처리와 다양한 수집체계로부터 수집된 정보의 조합 및 추론에 대해서 논하게 될 것이다.

### III. 이론적 배경

#### 1. 퍼지논리 (Fuzzy Logic)

퍼지 집합 이론은 1965년 Zadeh에 의해 처음 소개되었으며, 이후 Zadeh는 퍼지논리 체계를 발전시켰고, 이 이론은 인공지능, 기상예측, 경제, 조류학, 운영과학 등에 적용되고 있다.

전통적인 집합 이론을 기초로 하고 있는 이항 논리(binary logic)는 특정 원소가 단지 2개의 가능성만을 가지고 있다. 즉, 집합에 속하거나 속하지 않는 2가지의 경우만을 다루게 된다; 소속함수는 0과 1로 구성된다(일반적으로 크리스프 집합(crisp set)이라 한다). 퍼지 집합은 각 원소가 소속정도(degree of membership)에 결합되어 일반적인 집합의 개념을 보다 확장시켰

다. 소속정도는 0과 1사이의 값을 가질 수 있다. 집합  $X = \{x_1, x_2, \dots, x_n\}$  라하고 A를 X의 부분집합( $X \supset A$ )이라 하자. 이때 소속함수(membership function)  $\mu_A(x)$ 는 다음과 같이 표현할 수 있다 :

$$\mu_A : (x) \rightarrow [0, 1], \forall x \in A \quad (1)$$

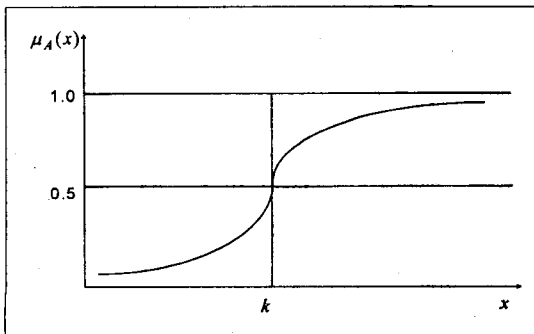
집합에서 소속정도가 0.5인 원소는 이 집합에 50% 속해 있다는 것을 의미한다. 또한 소속정도가 0인 원소는 이 집합에서 제외되지만 소속정도가 1인 원소는 이 집합에 100%속해 있다는 것을 나타낸다. 그러므로 위의 함수에서  $\mu_A$ 는 퍼지 집합의 소속정도를 나타내는 것이다.

또한 집합 A는 X의 퍼지 부분 집합이라 하고  $[x, \mu_A(x)]$ 에 의해 구축된다. 즉,

$$A = \{ \frac{\mu_A(x)}{x} \mid x \in X \} \quad (2)$$

각각의  $\mu_A$  값은 과거의 경험이나 전문가적인 지식과 판단에 근거하여 결정된 퍼지연산함수(fuzzy operator function)에 의해 결정된다. 일반적인 퍼지 연산값은 0, 0.5, 1과 같이 단순한 함수에 의해 간단하게 결정된다. 이와 같은 결정 방법은 적절한 분석 모형을 선정하고 모형을 이해하는데 용이하다는 이점이 있다. 예를 들어, 연산 함수는 과거의 자료나 수집된 자료를 토대로 다음과 같이 정의될 수 있다:

$$\mu_A(x) = \begin{cases} 0 & , x \leq 0 \\ \frac{1}{1 + \frac{k^2}{x^2}} & , x > 0 \end{cases} \quad (3)$$



<그림 3> 퍼지 연산 함수의 예

여기에서 k는 일반적인 상수를 나타낸다. 식 (3)으로부터 연산 함수값  $\mu_A(x)$ 와 속성 원소 x 사이의 관계는 아래 <그림 3>에 나타난 것과 같다.

## 2. 퍼지회귀모형(Fuzzy regression model)

퍼지회귀모형은 기존의 전통적인 회귀식에서 설정한 변수간의 크리스프 관계가 비현실적이라는 것을 인식하고, 실제 자료들간에 존재하는 퍼지성(fuzziness)을 모형에 반영하기 위해서 등장했다.

1982년 Tanaka, Uejima, Asai는 최초로 퍼지 선형 회귀 모형을 개발하여 주택 가격 결정 메커니즘을 구축하였다. 일반적인 회귀모형에서 관측 값과 추정 값의 편차는 관측 에러(error)에 의해 발생한다고 가정한다.

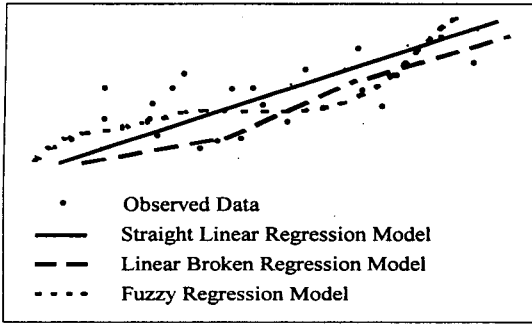
그러나, 퍼지 선형 회귀 모형에서는 이들 편차가 구축 모형의 불확실성에 의해 발생한다고 가정한다 (Tanaka et al., 1982). 즉, 이들 편차를 모형내 파라미터의 퍼지성으로 간주하여 퍼지 선형 회귀 모형을 구축한 것이다.

선형회귀형식의 모형(straight linear model, broken linear model)은 현재의 자료를 이용하여 미래의 예측이 쉽다는 이점이 있으나, 때때로 모형 내에 예러가 현저하게 증가하기도 한다. 많은 불확실성과 퍼지 현상이 있는 상황에서 의사결정을 할 때, 퍼지 선형 회귀 모형은 전통적인 회귀모형 보다 유용하게 접근할 수 있다.

퍼지 선형 회귀 모형은 일반적인 선형 회귀 모형과 비교해 볼 때, 다음과 같은 이점이 있다.

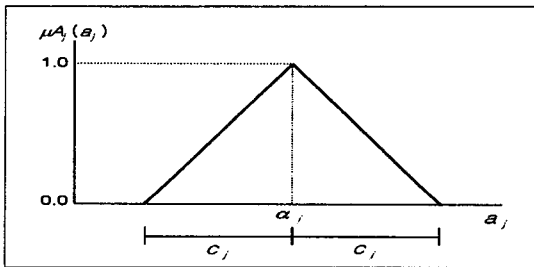
- ① 보다 높은 정도의 결과를 얻는다.  
(Higher production accuracy)
- ② 모형의 유연성을 증가시킨다.  
(Increase the modeling flexibility)
- ③ 비선형 관계를 포착할 수 있다.  
(Non-linear relationship can be captured)

다음의 <그림 4>는 3가지 회귀 모형간의 변동에 대한 예를 나타내고 있다. 이 그림은 전술된 퍼지 회귀 모형의 이점을 잘 나타내고 있다.



〈그림 4〉 각 회귀 모형간의 변동

다음의 〈그림 5〉는 퍼지 회귀 모형에서 일반적으로 적용되는 퍼지 파라미터인 “삼각” 소속함수를 나타낸다. 〈그림 5〉에서 퍼지 집합은 식 (4)와 같이 표현될 수 있다.



〈그림 5〉 퍼지회귀 모형의 파라미터 모드

$$\mu_{A_j}(a_j) = \begin{cases} 1 - \frac{|a_j - a_j|}{c_j}, & a_j - c_j \leq a_j \leq a_j + c_j \\ 0, & c_j \geq 0 \end{cases} \quad (4)$$

여기에서  $A_j$ 는  $(a_k, c_k)$ 에 의해 구성된 퍼지 파라미터.  $a_k$ 는  $k$ 번째 회귀 계수의 중앙값,  $c_k$ 는  $k$ 번째 파라미터 계수의 폭을 각각 나타낸다.

실수  $\lambda$  ( $0 \leq \lambda \leq 1$ )를 신뢰도로 정의하고,  $x$ 를 소속도 집합  $A$ 에 대응(mapping)시켰을 때,  $x$ 의 원소가  $\lambda$ 보다 크면 원소  $x$ 는  $A$ 에 속한다고 한다. 이 관계를 다음과 같이 표현한다.

$$A\lambda = \{ x \mid \mu_A(x) \geq \lambda \} \quad (5)$$

단,  $\lambda \leq \mu_A(x_0) \leq 1, x_0 \in A\lambda$

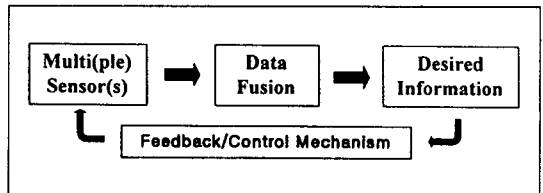
본 논문에서는  $\lambda$ 값은 각기 다른 자료의 수집체계를

조합하여 적합한 자료를 선정하는데 적용되기 때문에 자료의 통합과정에서 매우 중요한 인자로 작용하게 된다.

### 3. 데이터 퓨전

#### 1) 데이터 퓨전기법의 개요

데이터 퓨전은 다양한 검지기(수집체계)로부터 수집한 자료들을 어떻게 조합할 것인가를 다루는 기술을 말한다(Hall, 1992). 즉, 각 검지기(수집체계)에 대한 물리적 이벤트(event), 활동(activity), 또는 상황(situation)에 관한 추론을 하기 위해 다양한 수집자료들을 적절히 활용하고 이들의 영향도를 평가하여 최적의 결과값을 산출하는 것이다. 〈그림 6〉은 데이터 퓨전의 일반적인 수행과정을 나타낸다. 이러한 데이터 퓨전 기술은 초창기에는 전장 감시(battlefield surveillance), 전술적 상황평가(tactical situation assessment) 등의 목적으로 군(軍)에서 주로 적용되었으나, 근래에는 로봇공학(robotics), 자동생산(automated manufacturing), 원격탐사(remote sensing) 등의 일반 과학 및 공학 분야에도 널리 확산되어 적용되고 있다. 최근에는 데이터 퓨전이 군에서 적용되는 분야와 이론적 발전을 연구하는 일반 분야의 2개 분야로 나뉘어 발전되고 있다.



〈그림 6〉 Data fusion의 일반적인 수행과정

#### 2) 데이터 퓨전의 기대효과

데이터 퓨전의 기대 효과는 정량적으로 정확한 수치를 계산할 수는 없지만 정성적인 측면으로 운영적 수행도 향상, 검지 공간의 확장, 신뢰도 향상, 추론의 애매성 감소, 검지율 향상 등의 효과가 있다(Hall, 1992).

일반적으로 단일 검지 체계로부터 수집된 자료를



통해 추론된 정보보다는 둘 이상의 검지 체계로부터 추론된 정보의 기대 효과가 더 크다. 그렇다면, 검지 체계의 수와 기대효과는 비례하는 것일까? 이 문제를 설명하기 위해 Nahie과 Pokoski(1980)는 다음의 4가지 항목으로 구분하여 설명하였다.

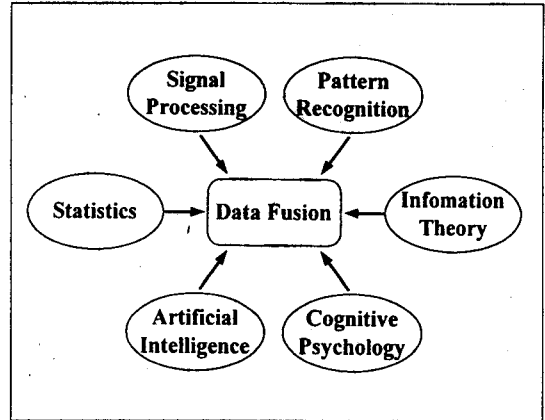
- 첫째. 정확하지 않은 검지 체계(추론 확률 0.5이하)로부터 수집한 자료를 조합할 때 데이터 퓨전의 장점이 현저하게 드러나는 것은 아니다.
- 둘째. 보다 정확한 검지 체계(추론 확률 0.95 이상)로부터 수집된 자료를 조합할 때, 정확도가 현저하게 증가하지는 않는다.
- 셋째. 검지 체계의 수가 많이 존재할 때(예를 들면 8~10이상의 검지 체계), 검지기 수를 추가하더라도 추론의 정도가 현저하게 증가하는 것은 아니다. 그러나, 검지체계 수의 추가는 측정 가능한 분야의 수가 증가시키기 때문에 추론 능력에 중요한 영향을 미칠 수 있다.
- 넷째. 개개의 검지 체계가 적절한 검지 확률을 가지고 있을 때, 1~7개의 검지 체계가 가장 적절하다.

### 3) 데이터 퓨전 기법의 분류

데이터 퓨전의 연구는 적용 폭이 넓고 적용 기술이 다양하기 때문에 복잡하다. 그리고 근래의 데이터 퓨전 기술은 통계, 인공지능(Artificial Intelligence), 운영과학(Operations Research), 디지털 신호 처리, 의사결정 분야의 수학적 기법과 경험적 기법이 얽혀있는 양상을 보이고 있다. 그래서, 일반 연구원이나, 엔지니어들이 특정 문제를 해결하기 위해 어떠한 기법을 적용시킬 것인가에 대해 많은 어려움을 겪고 있다. 다음의 <그림 7>은 일반적으로 데이터 퓨전에 적용되는 이론을 나타내고 있다.

데이터 퓨전의 기능은 기능적 모형에 따라 3가지<sup>5)</sup> 혹은 4가지 단계로 나뉘어 설명하지만 본 논문에서는 1단계에서 주로 사용되는 기법을 이용하여 교통관련

정보의 추정에 이용하게 될 것이다. 여기에서 1단계에 적용되는 기법에 대해 알아보기로 하겠다.



<그림 7> Data fusion의 이론적 분류

일반적으로 1단계에서 적용되는 데이터 퓨전이론은 크게 다음의 3가지로 분류한다(Hall, 1992).

- ① 보다 정확하게 물체의 위치와 속도를 추정하는 위한 기법(positional fusion)
- ② 물체를 확인하기 위해 여러 개의 센서로부터 자료를 조합하는 기법(identity fusion)
- ③ 1단계 과정을 보조하기 위해 필요한 부수적인 기법(ancillary support algorithms)

본 논문에서는 identity 퓨전 기법을 이용하여 알고리즘을 개발하게 될 것이다. 그리고, <표 1>은 identity 퓨전에 적용되는 분석 기법에 대해 간략하게 소개하고 있다.

<표 1>에서 소개한 기법 중에서 Dempster-Shafer(이하 D-S)규칙은 통계를 기반으로한 데이터 퓨전 접근방식으로 개개의 검지 체계가 100%의 확률로 의사결정을 수행할 수 없을 경우에 유용하게 적용된다(Klein, 1993). 각 검지 체계가 특정 상황에 대한 관찰을 목적으로 하고 있다면, 각 검지 체계는 자기 다른 검지 수준으로 정보의 추론에 기여한다. 이때 각 검지 체계는 0과 1사이의 확률을 가지게 되며, 의사결

5) Level 1 : Fused position and identity estimates  
 Level 2 : Hostile or friendly military situation assessments  
 Level 3 : Hostile force threat assessments

정을 위해 이들 확률은 조합의 과정을 거치게 된다. 실시간 자료에 대한 조합에서 D-S 규칙은 조합에 대한 정보를 지속적으로 개선하고 이를 바탕으로 의사 결정 또한 개선되게 된다.

소개된 D-S 규칙은 교통관련 자료의 불확실성과 실시간 수집, 지역별, 시간대별 수집체계의 신뢰도 변화 등을 고려할 때, 교통정보 추론을 위한 데이터 퓨전 알고리즘으로 D-S 규칙이 적절하다고 판단된다.

#### 4. D-S 규칙을 이용한 데이터 퓨전

1960년대에 Dempster는 두 공간 사이에 불확실성의 상한과 하한을 측정하는 불확실성의 대응(mapping)에 대한 수학적 근거를 마련하였다. 이 결과 확신(belief) 혹은 베이스(Bayes) 이론과 같은 확률 질량 함수를 이용하는 조합의 규칙이 만들어졌다. Dempster의 제자인 Shafer는 확신 함수의 개념을 발전시켜 D-S 이론을 개발하였다. D-S 이론은 불확실성에 대한 정확한 확률적 개념 대신에 개인적인 확신으로 통합하는 접근 방식이다(Shafer, 1976). 이 이론을 좀더 자세히 알아보기 위해 다음과 같은 예를 들어보기로 하겠다.

사실(fact)과 0 사실(0 fact), 즉, 2개의 대립가설을 고려해보자. 단일 지식기반 i로부터 추론된 결론은 사실 가설에 참이라는 신뢰값  $\lambda_i$ 를 배정하고, 0사실 가설에

는 신뢰값  $1-\lambda_i$ 를 배정할 수 있다. D-S규칙에 따라 만일 자료가 두 개의 독립된 수집체계(probe 차량과 루프 검지기)로부터 나왔다면, 주어진 가설에 대한 조합된 신뢰도는 개개의 수집체계들로부터 신뢰값  $\lambda_f = \lambda_{\lambda_1} \lambda_{\lambda_2}$  인 비례적인 결과를 도출한다. 스케일 팩터 k는 가설의 조합이 실행 불가능한 경우에는 직교곱(orthogonal product)이 모두 배정되지 않기 때문에 필요하다. 그래서 사건에 대한 가설을 지지하는 2개의 독립된 수집체계(즉, probe 차량과 루프 검지기)로부터 도출된 두 신뢰값  $\lambda_1, \lambda_2$ 의 퓨전값  $\lambda_f$ 는 다음과 같이 계산된다.

$$\lambda_f = \frac{\lambda_1 \lambda_2}{\lambda_1 \lambda_2 + (1-\lambda_1)(1-\lambda_2)} \quad (6)$$

각 독립된 수집체계에 대하여, 식(6)은 모든 신뢰값이 융합될 때까지 계속해서 반복 적용된다. 반복 작업시 적용되는 순서에는 관계없이 동일한 결과값이 산출된다.

한편, 앞에서 설명한 정보의 조합 방법은 정보 수집체계의 신뢰도가 동일할 경우에만 해당된다. 즉, 정보 수집체계의 신뢰도가 동일하지 않은 경우에는 다른 방법으로 접근되어야 하는데, 예를 들어 누적된(historical) 혼잡의 패턴과 최근 수집된 probe자료로부터 혼잡에 관한 결론이 추론되었다면, 최근 probe 자료에 의한 결론은 누적된 혼잡 패턴보다는 높은 신뢰성을 가지게 된다. 이러한 경우 다음의 2가지의 상황

<표 1> Identity 퓨전 분석 기법의 분류

방법	핵심 처리과정(Kernel process)	필요한 입력자료의 특징	결과 자료의 특징
전통적 추론	가설에 대하여 관측된 확률	경험적 확률, 모집단의 분포	가설에 대한 예외의 확률
베이지안 추론	주어진 새로운 자료에 대한 신념을 개선하는 확률	경험적/주관적 확률	사건의 발생 가능성의 개선
Dempster /Shafer	불확실성에 대한 많은 증거가 있는 가설의 확률과 진실된 가설의 확률	경험적/주관적 확률	사건의 발생 가능성과 무시 가능한 수준을 개선
퍼지 집합이론	소속함수를 가지는 대수 집합 원소	주관적 소속함수	집합 원소와 소속함수의 윤곽
군집분석	관측자료를 유사 척도를 기반으로 한 기본 그룹으로 구분	파라메트릭 혹은 주관적 자료	군집 원소와 유사성 척도
추정 이론	관측치에 대한 최적의 추정	양적 관측 상태/관측 모형	상태 벡터
엔트로피	정보 항목의 척도 계산	경험적 혹은 주관적 확률	최적 확률
FOM (figure of merit)	파라미터와 본체사이의 유사성의 정도 계산	센서 관측 자료로부터의 속성 벡터	비교 대상의 유사성에 대한 수치 자료
전문가 시스템	인간의 추론 과정을 모방한 컴퓨터 프로그램	추론을 입증하는 관측 자료	논리 조건에 대한 최적의 추론
템플릿 (Templates)	패턴 매칭 기법	관측된 자료의 기록	템플릿을 입증하는 자료의 선언
뉴럴네트워	관측 공간으로부터 인공공간까지의 비선형 변환	관측된 파라메트릭 자료	기 분류된 카테고리 입력자료의 매치

주) Waltz & Liras, 1990.

을 고려할 수 있다.

첫째, 결과가 일치하는 경우. 이들 자료는 독립적으로 수집되었기 때문에 최종 신뢰도는 개개의 신뢰값보다 높다.

둘째, 결과가 상호 반대인 경우. 최종 신뢰값은 2개의 반대결과 사이의 값을 가진다. 물론, 이 경우에 조합된 신뢰값은 보다 높은 신뢰도를 가지고 있는 수집체계로부터 추론된 결과에 더욱 근접하게 된다.

식 (6)은 D-S 퓨전규칙의 초기 버전으로 정보의 단일 수집체계가 보다 높은 신뢰성을 가질 수 없음을 나타낸다.

수집체계별로 신뢰도의 차별화를 고려하기 위해 Tarko와 Roupail(1997)은 다음과 같은 방식을 적용하였다. 이들은 첫 번째 수집체계( $\lambda_1$ )를 우선화시키기 위해 다음과 같은 조합방식을 가정했다.

$$\lambda_f = \lambda_1 \oplus (\lambda_2^* \cdot \lambda_2) \quad (7)$$

여기에서 연산자  $\oplus$ 는 식 (6)의 조합 규칙을 따르는 퓨전 연산을 나타내며,  $\cdot$ 은 아래의 식 (8)과 같이 수행된다.

$$\lambda_2^* \cdot \lambda_2 = (1 - \lambda_2^*)(1 - \lambda_2) + \lambda_2^* \lambda_2 \quad (8)$$

식 (8)에서 새로 도입된  $\lambda_2^*$ 는 퍼지연산자를 나타내며, 만일  $\lambda_1$ 과  $\lambda_2$  값이 서로 상반된 결과가 나타났을 때, 수집체계 2에 대한 신뢰도를 축소시키기 위해 도입된 것이다. 이 값은 다음 식으로 계산된다.

$$\lambda_2^* = 1 - \frac{1}{2} |\lambda_1 - \lambda_2| \quad (9)$$

퓨전된 신뢰값  $\lambda_f$ 는  $\lambda_1$ 과  $\lambda_2$  사이의 차이가 최대 값 1에 근접할 때  $\lambda_1$ 에 접근한다. 일반적으로 조합 규칙은  $\lambda_1 = \lambda_2$  일 때 D-S이론에 근접하게 된다.

이상과 같이 알아본 데이터 퓨전 기법의 이론적 개

념을 바탕으로 본 논문에서는 통행 시간 추정을 위한 데이터 퓨전 알고리즘을 제시한다.

#### IV. 통행 시간 추정을 위한 데이터 퓨전 알고리즘 개발

전술된 바와 같이 교통 관련 정보의 수집체계는 각각 그 신뢰도가 다르게 나타날 것이다. 예를 들어, CCTV, probe 차량, 루프 검지기 등을 통해 교통 정보를 수집한다고 할 때 각 검지 체계별로 CCTV=0.7, probe 차량=0.9, 루프 검지기=0.6과 같은 임의의 신뢰율을 적용할 수는 없다. 또한 이러한 신뢰율이 모든 구간에 걸쳐 언제나 동일하게 나타나는 것은 결코 아니다.

이러한 동기에서 본 논문에서는 퍼지 선형 회귀 모형을 이용하여 수집된 자료에 대한 1차 변환을 하였다. 퍼지 선형 회귀 모형은 측정 자료와 추정자료간의 에러(error)를 퍼지성으로 인식하고 이 값을 퍼지 파라미터(parameter)에 적용시킨다(Tanaka et al., 1982). 퍼지 선형 회귀는 이상치 자료(직관적으로 판단할 때 발생 불가능한 자료로 예를 들면 지점속도가 200Km/h로 수집된 경우)에 대한 필터링(filtering)이 용이(낮은 소속정도를 부여받게 됨)할 것으로 판단되며, 이러한 특성은 교통 수집체계에서 드러나는 문제점을 1차적으로 필터링하여 각 수집체계별로 각기 다른 신뢰율을 부여하게 된다.

각기 다른 신뢰율이 배정된 이들 수집체계는 조합을 위해 D-S 규칙과 베이지안 이론을 적용하여 최종적으로 보다 신뢰성있는 교통 정보를 추정하게 되는 것이다.

##### 1. 퍼지 모형의 구축

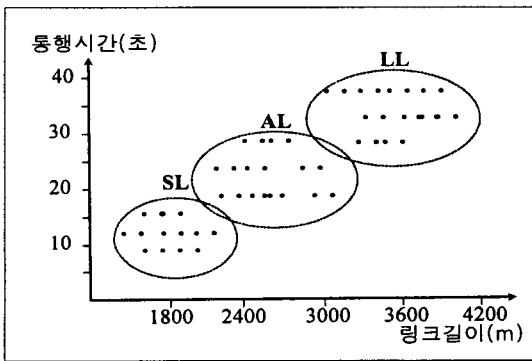
만일 동일한 교통 조건을 고려한다면, 일반적으로 길이가 긴 링크는 보다 많은 통행 시간이 요구된다. 링크 길이와 통행 시간 사이의 상관 관계를 알아보기 위해 산점(scatter point plot)을 표시하고 겹쳐지는 부위가 많은 지역에 대해 그룹(group)을 지을 수 있다<sup>(i)</sup>.

(i) 본 논문에서 제시하고 있는 링크의 그룹핑은 알고리즘의 전개 과정상 직관적으로 상관관계가 있는 링크의 길이와 통행시간과의 상관 관계만을 고려하였으나, 향후 본 논문에서 제시된 알고리즘을 통해 많은 인자를 고려한 실용적인 우수한 모형의 개발이 요구된다.

예를 들어, <그림 8>은 산점이 개략적으로 3개의 그룹으로 나뉘었음을 나타내고 있고, 다음과 같이 3개의 그룹으로 정의할 수 있다.

- ① 긴 링크와 긴 통행 시간(LL)
- ② 중간 링크와 중간 통행 시간(AL)
- ③ 짧은 링크와 짧은 통행 시간(SL)

<그림 8>에 나타난 것처럼, 산점에 대한 각 그룹은 각각의 소속함수(memberhip function)와 관련된 퍼지 파라미터를 가짐을 나타낸다. 여기에서 우리는 퍼지 선형 회귀방정식을 다음과 같은 유형으로 설정할 수 있다(Tanaka & Lee, 1997).



<그림 8> 링크 통행시간에 적용된 퍼지집합

$$Y = A_0 + A_1x_1 + \dots + A_nx_n = AX \tag{10}$$

여기에서,  $A = [A_0, A_1, \dots, A_n]$ ,  $X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$

식 (10)에서 파라미터  $A_i$ 를 대칭 퍼지수(symmetric fuzzy number)로 가정한다면,  $A_i = (a_i, c_i)_L$ 로 표현되는  $A_i$ 를 다음과 같이 정의할 수 있다.

$$\mu_{A_i}(x) = L\left(\frac{x_i - a_i}{c_i}\right), \quad c_i > 0 \tag{11}$$

식 (11)에서  $a_i$ 는 중심을 나타내고,  $c_i$ 는 폭을 나타낸다.  $L(x)$ 는 퍼지수에 대한 함수 유형을 나타낸다. 대

칭 삼각 퍼지수(symmetric triangular fuzzy number)  $A_i$ 에 대하여,  $L(x) = \max(0, 1 - |x|)$ , 그리고,  $A_i$ 의 소속함수는 다음과 같이 표시된다.

$$\mu_{A_i}(x) = 1 - \frac{|x - a_i|}{c_i} \tag{12}$$

식 (10)에서 회귀 계수  $A_i$ 는 퍼지수이기 때문에 추정된 결과  $Y$ 도 퍼지수가 된다. 그러므로, 이 개념을 보다 확장하여 식 (10)을 다음과 같이 나타낼 수 있다.

$$Y(x_j) = (a_0, c_0)_L + (a_1, c_1)_L x_{j1} + \dots + (a_n, c_n)_L x_{jn} = (a^T x_j, c^T |x_j|)_L \tag{13}$$

여기에서  $a^T x_j$ 와  $c^T |x_j|$ 는 퍼지 결과  $Y(x_j)$  각각에 대한 중심과 폭을 나타내며, 다음과 같다.

$$a^T x_j = a_0 + a_1 x_{j1} + \dots + a_n x_{jn} \tag{14}$$

$$c^T |x_j| = c_0 + c_1 |x_{j1}| + \dots + c_n |x_{jn}| \tag{15}$$

우리는 여기에서 식 (13)에 대한 소속함수를 다음과 같이 나타낼 수 있다.

$$\mu_Y(y) = \begin{cases} 1 - \frac{|y - a^T x|}{c^T |x|}, & x \neq 0 \\ 1, & x = 0, y = 0 \\ 0, & x = 0, y \neq 0 \end{cases} \tag{16}$$

## 2. 퍼지선형회귀모형의 구축

일반적으로 회귀 문제를 다룰 때는 다음의 2가지 기본적인 특징을 처리하게 된다.

첫째, 최적 수리모형이 무엇인가?

둘째, 어떻게 관련자료들에 대한 최적 모형을 수립할 것인가?

전통적인 선형 회귀 모형을  $y_i = \beta^T x$ 라 하고, 추정값을  $y_i^* = \beta^*{}^T x$ 라 하면 관측값과 추정값 사이의 편차는

7) 일반적으로 다음과 같이 표기한다.  $i.i.d. e_i \sim N(0, \sigma^2)$

정규분포를 갖는 확률변수<sup>7)</sup>로 가정하게 된다.

$$y_i - y_i^* = \epsilon_i, i = 1, \dots, N \quad (17)$$

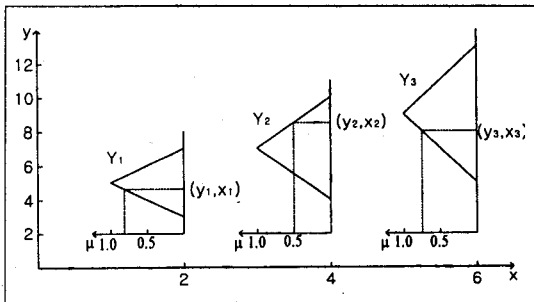
그러나 우리는 편차가 모형내 파라메터의 퍼지성에 관련이 있다고 가정하여 다음의 퍼지 선형 함수를 구축한다(식 (10)과 동일한 식임).

$$Y = A_0 + A_1 x_1 + \dots + A_n x_n = AX \quad (18)$$

우리가 비퍼지 자료(non-fuzzy data)를 가지고 있을 때, 퍼지 선형 회귀 모형의 문제는 퍼지 파라메터  $A^*$ 를 결정하는 것이다. 즉, 퍼지 결과 집합  $Y_i^* = A^* x_i$ 는 소속정도  $h$  이상을 만족하는  $y_i$ 를 포함하게 된다(〈그림 9〉참고).

아래 〈표 2〉와 같이 비 퍼지 입력-결과 자료에 대하여, 퍼지 결과 자료를  $Y_i = (y_i, e_i)$ 라 표현하면, 입력-결과 자료는 〈그림 10〉와 〈표 3〉과 같이 표현된다. 이때 소속함수  $Y$ 는 다음과 같이 나타내어진다.

$$\mu_{Y_i}(y) = 1 - \frac{|y_i - y|}{e_i} \quad (19)$$



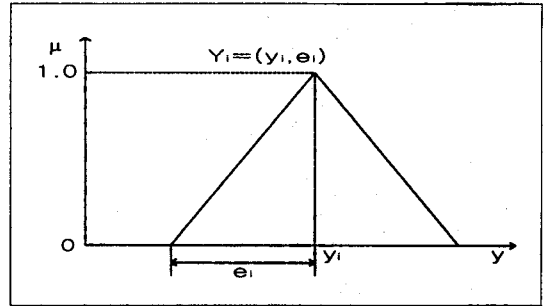
〈그림 9〉 퍼지선형회귀모형  $Y_i = A_0 + A_1 x_i$ 에 대한 설명  
 $A_0 = (3, 1)$   $A_1 = (1, 0.5)$   $h = (0.5)$

〈표 2〉 입력-결과 자료

샘플 수	결과자료(y)	입력자료(x)
1	$y_1$	$x_{11}, \dots, x_{1n}$
⋮	⋮	⋮
N	$y_N$	$x_{N1}, \dots, x_{Nn}$

〈표 3〉 비퍼지 입력자료에 대한 퍼지결과자료

샘플 수	결과자료(y)	입력자료(x)
1	$y_1(y_1, e_1)$	$x_{11}, \dots, x_{1n}$
⋮	⋮	⋮
N	$y_N(y_N, e_N)$	$x_{N1}, \dots, x_{Nn}$



〈그림 10〉 퍼지 결과

퍼지 선형 회귀 모형을 구축하기 위해서는 다음과 같은 가정이 필요하다(Tanaka et al., 1982).

① 자료는 퍼지선형모형으로 표현될 수 있다.

$$Y_i = A_0 + A_1 x_{i1} + \dots + A_n x_{in} \triangleq A^* x_i \quad (20)$$

여기에서 퍼지 파라메터  $A_i$ 의 유형은 식 (12)와 동일하며,  $x_i, Y_i$ 는 식 (16)을 이용하여 다음과 같이 구할 수 있다.

$$\mu_{Y_i^*}(y) = 1 - \frac{|y_i - x_i' a|}{c |x_i|} \quad (21)$$

② 퍼지 결과 자료  $Y_i = (y_i, e_i)$ 에 대하여 추정된 퍼지선형모형  $Y_i^* = A^* x_i$ 의 적합도(degree of fitting)는 제약 조건  $Y_i^h \subset Y_i^{*h}$ 에 대하여  $h$ 를 최대화하는  $\bar{h}$ 로 측정된다.

$$Y_i^h = \{ y \mid \mu_{Y_i}(y) \geq h \} \\ Y_i^{*h} = \{ y \mid \mu_{Y_i^*}(y) \geq h \} \quad (22)$$

여기에서  $\bar{h}$ 는 〈그림 11〉를 통해 증명되며, 모든 자료  $Y_1, \dots, Y_n$ 에 대한 퍼지 모형 모형의 적합도는  $\min_i \{\bar{h}_i\}$ 로 정의된다.

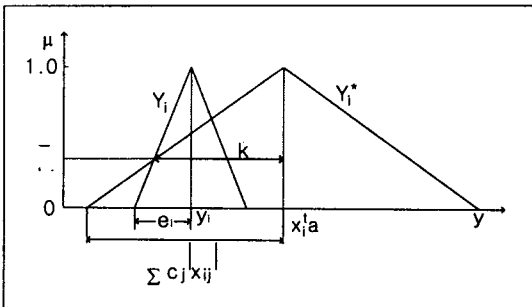
③ 퍼지 선형 모형의 애매성(vagueness)은 다음과 같이 정의된다.

$$J = \sum_j c_j |x_{ij}| \quad (23)$$

퍼지 선형 회귀 모형에 대한 문제는 모든  $i$ 에 대하여 제약조건  $\bar{h}_i \geq H$ 하에서  $J$ 를 최소화하는 퍼지 파라미터  $A_i^*$ 를 구하는 것으로 요약 설명할 수 있다. 여기에서  $H$ 는 의사결정자가 정한 모형의 적합도이며,  $\bar{h}_i$ 는 다음 식으로 구할 수 있다.

$$\bar{h}_i = 1 - \frac{|y_i - x_i^t a|}{\sum_j c_j |x_{ij} - e_i} \quad (24)$$

이 식은 <그림 11>를 참고한다면 다음과 같이 증명된다.



<그림 11> 퍼지자료  $Y_i$ 에 대한  $Y_i^*$ 의 적합도

$$1 : (1 - \bar{h}_i) = (\sum_j c_j |x_{ij}|) : k \quad (25)$$

$$k = |y_i - x_i^t a| + e_i (1 - \bar{h}_i) \quad (26)$$

우리의 주목적인  $A_i^* = (a_i, c_i)$ 는 식 (23)과 식 (24)를 이용하여 LP(Linear Programming)문제의 해를 통해서 구하게 된다. 즉,

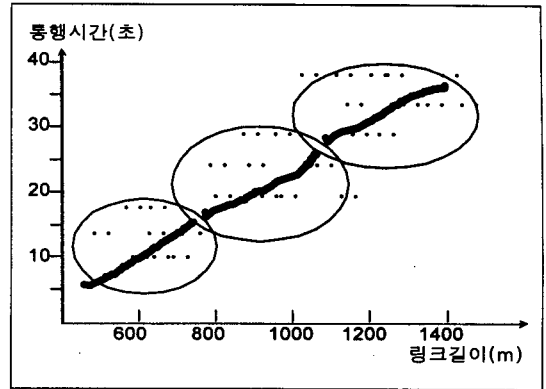
$$\min_{a,c} J = \sum_j c_j |x_{ij}| \quad (27)$$

$$s.t. \quad a x_i + (1-H) \sum_j c_j |x_{ij}| \geq y_i + (1-H)e_i$$

$$-a x_i + (1-H) \sum_j c_j |x_{ij}| \geq -y_i + (1-H)e_i$$

$$c_j \geq 0 \quad i=1, \dots, N \quad j=0, \dots, n$$

산점도(<그림 8>참고)를 통해 각 자료에 대한 그룹핑을 한 후 위의 절차에 따라 각 그룹에 대한 퍼지 선형 회귀 모형과 소속함수를 구축하게 된다. 아래 <그림 12>는 링크 길이에 대한 통행시간의 그룹핑 결과와 각 그룹에 대한 퍼지 회귀를 나타낸다.



<그림 12> 각 그룹에 대한 퍼지 회귀

### 3. 각 그룹별 신뢰값의 퓨전

전술한 바와 같이 교통 관련 자료에 대한 퓨전은 수집체계별 각각 신뢰도의 차별화가 이루어 져야 한다. 퓨전에 대한 기본적인 규칙은 D-S 규칙(식 (6)참고)을 적용하게 될 것이다. 수집된 자료는 <그림 12>과 같이 그룹을 이루어 분석하게 될 것이기 때문에 각 그룹에 대한 신뢰도의 차별화가 필요하게 된다.

$$\lambda_i = \frac{\lambda_1 \lambda_2}{\lambda_1 \lambda_2 + (1 - \lambda_1)(1 - \lambda_2)}$$

퍼지 선형 회귀 분석을 통해 구해진 각 자료에 대한 신뢰도(소속정도)는 그룹별로 각각 다른 신뢰도(소속정도)를 부여하기 위해 식 (24)에서 구한 자료의 각 그룹별 소속정도  $\bar{h}_i$ 를 이용하게 된다.

식 (24)에 의해, 자료 수집체계는 각 그룹별  $\bar{h}_i$ 자료를 생성하고, 각 그룹의 샘플 크기(N)를 이용하여 각 자료의 신뢰도 평균값  $\bar{h}_i$ 를 계산할 수 있다. 즉, 예를 들어 수집체계 1에 대한 k그룹의 평균값  $\bar{h}_1$ 은 다음과 같이 구하게 된다.

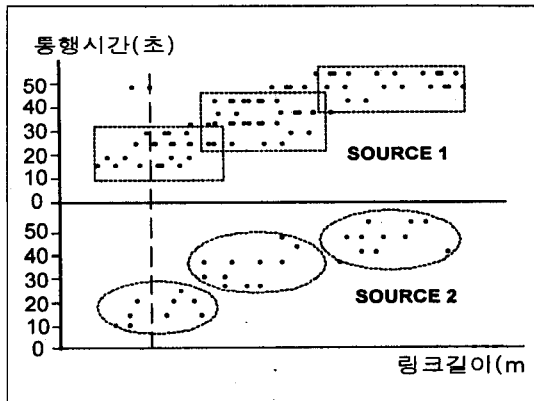
$$\bar{h}_1 = \frac{1}{N} \sum_{i=1}^k \bar{h}_i(n_i) \quad (28)$$

여기에서, 샘플 크기는  $N = n_1 + \dots + n_k$ 를 나타낸다.

그래서, 각 수집체계에 대한 퓨전된 신뢰값  $\lambda_i$ 는 식 (6)과 식 (28)을 적용하여 구하게 된다.

#### 4. 온라인 통행시간 추정을 위한 퓨전

서로 다른 2가지 이상의 교통 정보 수집체계를 고려해보자(〈그림 13〉 참고). 고정된 링크의 길이에 대하여 각 정보 수집체계는 각기 다른 통행 시간정보를 제공하게 될 것이다(물론 동일할 수 있겠지만 흔하지는 않을 것이며, 동일한 경우에도 관계없다). 이때 각 정보 수집체계를 통해 제공되는 정보는 각기 다른 신뢰도를 가지고 정보를 제공하게 된다(5.4절 참고).



〈그림 13〉 서로 다른 수집원으로부터 수집된 자료

그래서, 동일 시점에 특정 링크에 대하여 다량의 정보가 수집되는 경우에는 수집체계간의 퓨전 과정이 필요하게 된다. 본 논문에서는 베이지안 이론을 기초로 하여 접근하였다.

$$a_i = \frac{e_i^2}{\sum_{k=1}^n e_k^2} \quad (29)$$

여기에서  $e_i$ 는 퍼지 선형 회귀 모형을 통한 추정자료와 측정자료간에 발생한 에러를 나타낸다.

링크별 동적 통행 시간에 대한 추정을 위해서는 식

(29)에서 구한 각 수집 체계별 가중치를 고려한 퓨전이 최종적으로 이루어지게 되며 이를 수식화하면 다음과 같다.

$$OTT = \sum_{i=1}^n a_i T_i \quad (30)$$

여기에서  $OTT$ 는 온라인 통행시간,  $T_i$ 는 자료 수집 체계  $i$ 에 대한 예상 통행시간,  $n$ 은 자료 수집체계의 수를 각각 나타낸다.

#### V. 모의자료의 적용

본 알고리즘은 아직 수집체계가 확립된 단계가 아니기 때문에 실제로 데이터 퓨전에 적용할 상황은 아니다. 이러한 이유에서 본 연구에서는 링크 30개에 대하여 가상의 자료를 설정하고 이에 대한 퓨전 알고리즘을 적용하여 그 결과를 통해 수집된 자료에 대한 가중치 부여 방법을 적용해 자료의 우선 순위를 적용하고 통행 시간을 추정한다. 물론 가상의 자료는 실제 상황과 유사한 상황을 연출하고자 4가지의 시나리오<sup>8)</sup>를 설정하고 적절하게 자료를 구축하였다.

예제에 사용된 자료는 다음의 3가지 온라인 자료<sup>9)</sup>를 적용하였다.

- ① CCTV 자료
- ② 검지기 자료
- ③ Probe 자료

적용 과정은 총 6단계로 나뉘며 각 단계별 주요 내용은 자료의 정리, 실시간 자료의 그룹별 분류, 소속 함수를 통한 소속정도 산출, 소속 정도의 조합, 온라인 통행 시간 계산, 동적 자료의 구축이다.

8) 실시간으로 수집되는 자료는 CCTV자료, probe 자료, 루프 검지기 자료로 가정하고, 수집되는 자료는 최대한 현실을 반영하기 위해 다음과 같은 시나리오를 설정하였다.

- 시나리오 I : 3가지 실시간 자료가 모두 수집된 링크
- 시나리오 II : 2가지 실시간 자료가 모두 수집된 링크
- 시나리오 III : 1가지 실시간 자료가 모두 수집된 링크
- 시나리오 IV : 실시간 자료가 수집되지 않은 링크

9) 사실상 CCTV와 루프 검지기를 통해 통행시간을 측정하는 것은 어려운 일이나, 향후 구축될 교통정보 수집체계를 대변하기 위해 현실적으로 교통정보 수집에 이용되고 있는 이들 수집체계를 예로 들어 설명하기로 한다.

〈표 4〉 3개의 수집원에 대한 링크 통행시간

Link Number	CCTV 자료(초)	검지기 자료(초)	Probe 자료(초)	링크 길이(m)
1	45	N/A	50	2200
2	N/A	N/A	N/A	2000
3	60	63	58	2800
4	20	17	N/A	800
5	N/A	48	N/A	2100
6	N/A	N/A	44	2000
7	54	N/A	N/A	2500
8	N/A	41	37	1900
9	40	36	N/A	1800
10	N/A	48	45	2200
11	60	55	58	2800
12	20	22	N/A	1000
13	N/A	31	N/A	1500
14	72	75	N/A	3400
15	N/A	61	53	2800
16	59	66	67	3000
17	N/A	N/A	N/A	2300
18	25	N/A	27	1200
19	N/A	59	N/A	2700
20	72	65	74	3300
21	N/A	N/A	44	2000
22	34	36	N/A	1700
23	N/A	51	N/A	2400
24	65	69	70	3100
25	31	N/A	29	1400
26	48	44	N/A	2200
27	43	40	N/A	2000
28	41	N/A	36	1900
29	N/A	50	47	2300
30	62	58	55	2469

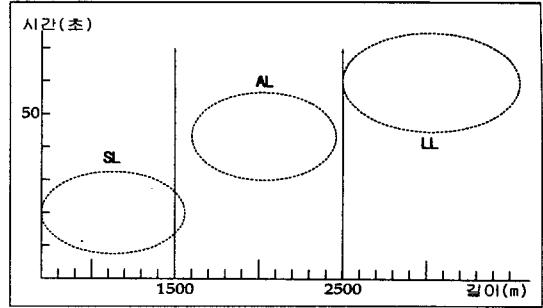
주) N/A는 수집되지 않음을 나타냄

### 1. 자료의 정리

전술한 바와 같이 가상의 자료는 30개의 링크에 대하여, 특정 시간 간격(15초 혹은 20초) 내에서 3가지 수집체계로부터 수집된 모든 자료를 적용하였다. 〈표 4〉는 특정 시간 간격 동안 차량이 진입하여 소요된 각 링크의 통행 시간을 나타낸다.

### 2. 실시간 자료의 그룹별 분류

검지기로부터 수집한 자료는 3가지 그룹으로 분류한다<sup>10)</sup>. 즉, 〈그림 14〉와 같이 링크의 길이에 따라 SL(1500 m 이하), AL(1500~2500 m), LL(2500 m 이상)로 분류한다.



〈그림 15〉 그룹화된 자료

### 3. 소속함수를 통한 소속정도 산출

〈그림 14〉에 나타난 것처럼 그룹 지어진 각각의 자료는 검지기 자료, CCTV자료, probe 자료로 각각 나누어 수집된 자료와 예측된 자료에 대한 오차의 소속함수를 구축하게 된다. 이때 소속함수에 의해 각 그룹별 자료에 대한 소속정도(신뢰도)가 구축된다. 계산과정에 대한 설명을 위해 SL 링크 그룹중에서 검지기 자료에 대한 소속정도를 구하는 과정을 설명하기로 한다.

〈표 5〉 SL 링크에 대한 검지기 자료 분석

링크길이(m)	관측자료( $t_i$ )	예측자료( $\hat{t}_i$ )	잔차제곱( $\epsilon^2$ )
800	17	18.2179	1.4833
1000	22	20.2949	2.9073
1500	25	25.4872	0.2374

〈그림 11〉과 식 (24)에 의해  $\sum_j c_j |x_{ij}|$ 를 다음과 같이 구할 수 있다.

$$\begin{aligned} \sum_j c_j |x_{ij}| &= \sum |t_i - \hat{t}_i| \\ &= |(17-18.2179)| + |(22-20.2949)| + |(25-25.4872)| \\ &= 3.4102 \end{aligned}$$

실시간으로 수집된 자료는 퍼지성이 없기 때문에 식 (24)에서  $\epsilon_i$  값은 0이 된다. 여기에서 각 수집 자료  $i$ 에 대한 소속정도는 다음과 같이 구하게 된다.

$$\bar{h}_i = 1 - \frac{|t_i - \hat{t}_i|}{3.4102}$$

10) 분류는 산점도를 통해 오차가 특히 심한 자료를 제외하고 유사 자료들을 모아 3개의 그룹으로 분류하였다.



짧은 링크그룹 SL에 대한 최소 소속정도는 다음과 같다.

$$\bar{h}_{SL} = 1 - \frac{122 - 20.29491}{3.4102} = 0.5$$

여기에서 소속정도가 낮은(0.5) 이유는 그룹화된 자료의 수(3개)가 적기 때문에 발생된 것이다.

나머지 AL과 LL그룹에 대한 소속정도와 CCTV 자료, probe 자료에 대한 그룹별 소속정도도 동일한 계산과정을 통해 구할 수 있으며, 그 결과 값은 <표 6>과 같다.

<표 6> 그룹별 최소 소속정도

수집체계	최소 소속정도		
	SL	AL	LL
검지기	0.50	0.76	0.74
CCTV	0.65	0.75	0.69
probe 차량	*	0.80	0.63

\* : 해당 자료가 2개이므로 분석에 의미가 없는 것으로 판단되어 AL에 포함시켜 분석함.

#### 4. 소속정도의 조합

식 (28)에 따라 각 수집체계로부터 수집한 자료에 대한 소속정도의 조합은 각 그룹에 해당되는 자료의 수에 따라 결정된다. 그래서 최종 소속정도를 다음과 같이 구할 수 있다.

$$\bar{h}_{\text{검지기}} = \frac{1}{21}(3 \times 0.50 + 9 \times 0.76 + 9 \times 0.74) = 0.71$$

$$\bar{h}_{\text{CCTV}} = \frac{1}{18}(4 \times 0.65 + 6 \times 0.75 + 8 \times 0.69) = 0.70$$

$$\bar{h}_{\text{probe}} = \frac{1}{16}(9 \times 0.80 + 7 \times 0.63) = 0.73$$

각 검지 체계별 최종 소속정도를 토대로 D-S 규칙 (식 (6) 참고)을 적용하여 검지기 자료( $\bar{h}_{\text{검지기}}$ )와 CCTV 자료( $\bar{h}_{\text{probe}}$ )로부터 조합한 1차 신뢰값은 다음과 같이 구한다.

$$\lambda_f^1 = \frac{0.70 \times 0.71}{0.70 \times 0.71 + (1 - 0.70)(1 - 0.71)} = 0.85$$

위의 1차 계산값과 probe 자료( $\bar{h}_{\text{probe}}$ )의 조합에 의해 최종 신뢰값을 다음과 같이 구하게 된다.

$$\lambda_f = \frac{0.85 \times 0.73}{0.85 \times 0.73 + (1 - 0.85)(1 - 0.73)} = 0.94$$

실시간으로 각각 관측된 자료에 대한 신뢰 값은 최종 신뢰값(0.94)을 토대로 90%의 신뢰값을 설정한다.

#### 5. 온라인 통행시간 계산

기 다른 수집체계로부터 수집된 자료를 통해 링크별 단일 온라인 통행 시간을 구하기 위해 식 (29)를 적용하고 각각의 수집체계에 대하여 가중치를 부여한다. 각 수집체계에 대한 잔차제곱의 평균은 다음과 같다.

$$e^2_{\text{검지기}} = 5.0200$$

$$e^2_{\text{CCTV}} = 3.8121$$

$$e^2_{\text{probe}} = 5.2999$$

이 자료를 토대로 각 수집체계에 대한 가중인자 (weighting factor)를 구하면 다음과 같다.

$$a_{\text{검지기}} = \frac{e^{-2}_{\text{검지기}}}{e^{-2}_{\text{검지기}} + e^{-2}_{\text{CCTV}} + e^{-2}_{\text{probe}}} = 0.3$$

$$a_{\text{CCTV}} = \frac{e^{-2}_{\text{CCTV}}}{e^{-2}_{\text{검지기}} + e^{-2}_{\text{CCTV}} + e^{-2}_{\text{probe}}} = 0.4$$

$$a_{\text{probe}} = \frac{e^{-2}_{\text{probe}}}{e^{-2}_{\text{검지기}} + e^{-2}_{\text{CCTV}} + e^{-2}_{\text{probe}}} = 0.3$$

#### 6. 동적 자료의 구축

각각의 자료 수집체계에 대하여 계산된 가중 인자는 각 링크에 대한 실시간 통행시간 계산에 적용된다. 예를 들어 링크 3에 대하여 CCTV 자료가 60초, 검지기 자료가 63초, probe 자료가 58초라면, 이때 실시간 동적 자료는 다음과 같이 계산한다.

$$t_3 = 0.4(60) + 0.3(63) + 0.3(58) = 60\text{초}$$

<표 7>은 위의 계산과정을 통해 구축된 30개 링크별 동적 통행시간 자료이다.

**VI. 결론**

교통량 검지시스템은 ATMS/ATIS에 대하여 핵심적인 인프라라 할 수 있다. 실시간 교통관리와 여행자 정보체계는 ITS 기능을 위해 지속적인 시·공간적 교통정보를 필요로 한다. 데이터 퓨전은 다양한 수집체계로부터 수집된 교통 상황 정보 중 가장 유용한 정보를 추출하는 효율적 방법으로 ITS의 초기 단계와 발전된 단계에서 모두 필요한 방법이라 여겨진다.

<표 7> 데이터 퓨전을 통해 구축된 실시간 링크 통행시간

Link Number	CCTV 자료(초)	검지기 자료(초)	Probe 자료(초)	동적 자료(초)
1	45	N/A	50	47
2	N/A	N/A	N/A	N/A
3	60	63	58	60
4	20	17	N/A	19
5	N/A	48	N/A	48
6	N/A	N/A	44	44
7	54	N/A	N/A	54
8	N/A	41	37	39
9	40	36	N/A	38
10	N/A	48	45	47
11	60	55	58	52
12	20	22	N/A	21
13	N/A	31	N/A	31
14	72	75	N/A	73
15	N/A	61	53	57
16	59	66	67	68
17	N/A	N/A	N/A	N/A
18	25	N/A	27	26
19	N/A	59	N/A	59
20	72	65	74	70
21	N/A	N/A	44	44
22	34	36	N/A	35
23	N/A	51	N/A	51
24	65	69	70	68
25	31	N/A	29	30
26	48	44	N/A	46
27	43	40	N/A	42
28	41	N/A	36	39
29	N/A	50	47	48
30	62	58	55	59

교통 상황 정보의 퓨전을 위해 특히 요구되는 사항은 링크의 속성별, 수집체계의 유형별 각각 그 신뢰도가 다르게 나타날 수 있다는 것이다. 이러한 동기에서 본 논문에서는 퍼지 선형 회귀 모형을 이용하여 수집된 자료에 대한 1차 변환을 실시하였다. 퍼지 선형 회귀 모형은 모형에 발생하는 에러를 퍼지성으로 인식하고 이 값을 퍼지 파라미터로 적용한다. 이 결과 퍼지 선형 회귀 모형은 교통 상황 정보의 수집 중에 발생하는 문제점을 1차적으로 필터링하여 각기 다른 신뢰도를 부여하여 분석하게 됨으로서 보다 현실적으로 정보를 처리할 수 있게 된다.

본 연구에서는 자료 수집의 어려움으로 인해 구체적인 사례를 들어 설명하지는 않았고 단지 알고리즘의 일반적인 설명과 모의 자료의 적용을 통해 알고리즘을 수행하였지만, 향후 본 논문을 기초로 실제 자료의 적용을 통한 구체적인 알고리즘의 구축이 요구된다. 또한 알고리즘의 적용과정에서 링크의 길이와 통행시간과의 관계를 통해 그룹핑을 하였지만 차선수, 차선폭, 도로의 유형, 기타 도로의 용량분석시 고려되는 인자들을 고려한 알고리즘의 개발이 이루어져야 할 것이다<sup>11)</sup>. 또한 이들 인자를 고려한 링크의 그룹핑에 대한 일반적인 기준의 설정이 함께 연구되어야 할 것이다.

데이터 퓨전 이론은 교통이론, 네트워크 분석, 전문가적 지식 등의 배경 지식을 이용하여 혼잡/유고관리, 통행시간 추정, 노선계획에 매우 유용하게 이용될 수 있는 바 이러한 분야에 퓨전 이론의 적용은 앞으로 연구의 가치가 많을 것으로 판단된다.

**참고문헌**

1. William R. McShane, Roger P. Roess, "Traffic Engineering", Prentice Hall, 1990.
2. Andrzej P. Tarko and Nagui M. Roupail,

11) 실제로 본 알고리즘에 적용된 링크길이와 통행시간 관계는 직관적으로 선형 비례관계를 나타내며 네트워크 관련 기본 속성자료로 구축되어 있을 것이라 판단되어 적용한 것이며, 차선수, 도로의 유형, 차선폭, 도로의 선형 등과 같은 속성 자료는 직관적 판단의 어려움으로 본 알고리즘에서는 고려하지 않았다. 그러나, 이들 링크 속성자료가 구축되어 통계적인 분석이 이루어진다면 본 알고리즘에서 제시한 퍼지 회귀 모형의 적용에는 어려움이 없을 것으로 판단된다.

- "Intelligent Traffic Data Processing for ITS Applications", *Journal of Transportation Engineering*, pp.298~307, July/August, 1997.
3. David L. Hall, "Mathematical Techniques in Multisensor Data Fusion", Artech House, 1992.
  4. "Introduction to Multisensor Data Fusion : Multimedia Software", TECH REACH Inc. 1996.
  5. Edward Waltz and James Llinas, "Multisensor Data Fusion", Artech House, 1990.
  6. Lawrence A. Klein, "Sensor and data fusion concepts and applications", SPIE press, 1993.
  7. Stanislaw Berka, Xin Tian & Andrew Tarko, "Data Fusion Algorithm for ADVANCE Release 2.0", ADVANCE Working Paper Series # 48, May, 1995.
  8. Andrezej Tarko & Nagui M. Roupail, "Data Fusion for ADVANCE Release 1.5 Data Fusion", ADVANCE Working Paper Series # 33, February, 1994.
  9. Sumner R., "Data Fusion in Pathfinder and TravTek", *Vehicle Navigation and Information Systems Conference Proceedings*, P-253, pp.71-75, 1991.
  10. Nahie, P.J. and J.L. Pokoski, "NCTR Plus Sensor Fusion Equals IFFN", *Proceedings of the IEEE Transactions on Aerospace Electronic Systems*, Vol. AES-16, pp.320~327, May 1980.
  11. B. Hofmann-Wellenhof, H. Lichtenegger, and J. Collins, "Global Positioning System Theory and Practice", Springer-Verlag Wien New York, 1994.
  12. Michael R. Genesereth & Nils J. Nilsson, "Logical Foundations of Artificial Intelligence", Morgan Kaufmann, 1987.
  13. Hideo Tanaka, HaeKwan Lee, "Fuzzy Regression Combining Central Tendency and Possibilistic Properties", *Fuzzy-IEEE*, pp.63~68, 1997.
  14. Hideo Tanaka, Satoru Uejima, Kiyoji Asai, "Linear regression Analysis with Fuzzy Model", *IEEE trans. on systems, man, and cybernetics*, pp 903-907, vol 12, No. 6, 1982.
  15. H. Tanaka, H. Ishibuchi, S. Yoshikawa, "Exponential possibility regression analysis", *Fuzzy Sets and Systems* 69, pp.305~318, 1995.
  16. George J. Klir and Bo Yuan, "Fuzzy sets and fuzzy logic: Theory and application", Prentice Hall, 1995.
  17. HaKyun Kim, "Applying fuzzy logic to managerial decision making : estimating price-earnings ratios using fuzzy linear regression analysis" New York, New York, June 1994.
  18. T. Terano, K. Asai, and M. Sugeno, "Fuzzy systems theory and its application", Academic Press, 1992.
  19. 이광형, 오길록, "퍼지이론 및 응용 I 권", 홍릉과학출판사, 1991. 6.
  20. 유동선, 이교원 공저, "알기 쉬운 퍼지입문", 교우사, 1996. 5.
  21. "첨단교통체계 연구기획보고서", 국토개발연구원, 1996. 3.
  22. 채석, 오영석, "퍼지이론과 제어", 청문각, 1997.