

GPS 데이터를 이용한 링크통행속도 예측에 관한 연구

최진우, 홍남관, 양영규
경원대학교 전자계산학과

e-mail : cjw49@paran.com, madoyo78@paran.com, ykyang@kyungwon.ac.kr

A Study on Link Travel Speed Forecasting using GPS data

Jin-woo Choi, Nam-kwan Hong, Young-kyu Yang
Dept. of Computer Science, Kyung-Won University

요 약

텔레매틱스 서비스 중 가장 중요한 서비스는 운전자가 요청하는 교통 상황 정보를 신속하고 정확하게 전달해 주는 것이다. 본 논문에서는 GPS 단말기를 장착한 차량들을 통해 수집된 자료를 활용한다. 가까운 과거의 패턴 자료를 이용하여 필터링 범위를 산정한 후, 실시간으로 비정상적인 GPS 자료를 제거하여 링크의 대표 속도 값을 산출하는 방법을 제시한다.

1. 서론

도심 내 교통 혼잡의 원인은 많은 운전자들이 여유 용량을 지닌 도로가 있음에도 불구하고 일부분의 도로망으로 집중되면서 전체 도로망을 효율적으로 사용하지 못하는 데 있다. 하지만, TSP(Telematics Service Provider)에서 실시간 교통정보를 제공하게 되면 운전자들로 하여금 보다 나은 경로를 선택하도록 유도하여 전체 도로망의 운영 효율을 증가시킬 수 있다.

실시간으로 도로를 이용하고 있거나 향후 도로 이용 계획이 있는 운전자가 가장 원하는 정보는 특정 구간에 대한 소요 통행 시간이나 속도라 할 수 있다.

정확한 통행 속도 정보를 운전자에게 전달하기 위해서는 우선, 실시간으로 변화하는 교통 자료를 신속하고 정확하게 수집하고, 이를 토대로 이용자가 이해하기 쉬운 형태로 가공 하여 전달하는 절차가 필요하다.

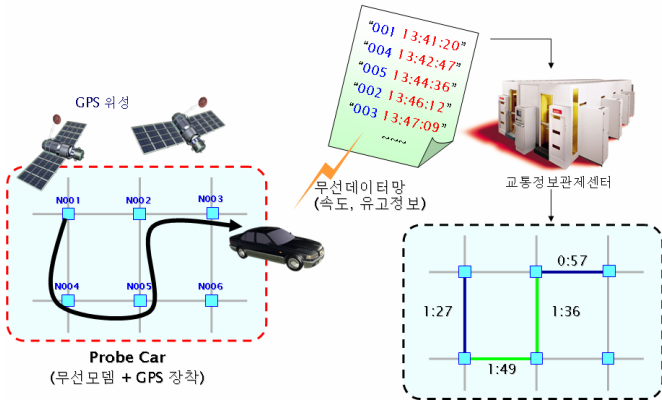
현재 운영 중이거나 시범중인 교통정보 수집체계는 크게 루프검지기나 적외선검지기, 영상검지기와 같은 고정식 검지체계와 Beacon 이나 GPS 와 같은 이동식 검지체계로 분류된다.

고정식 검지체계 중 가장 대표적이라 할 수 있는

루프검지기의 경우는 신호시스템 용도의 루프검지기를 제외하더라도 전국에 약 3000 개 이상이 운영되고 있으며, 대부분의 고속도로에 설치되어 있다. 고정식 검지기를 이용하여 교통 자료를 수집하는 방법은 자료의 신뢰성이나 운영의 효율성에서는 우수한 면을 보이나, 설치 시 도로를 굴착해야 하는 어려움이나 한 지점에서의 정보만을 수집 하기 때문에 특정 구간에 대한 대표 자료로 이용하기 어렵다는 문제점을 갖고 있다.

반면, probe 차량을 사용하는 Beacon 이나 GPS 같은 이동식 검지기를 이용해 교통 자료를 수집하는 방식은 특정 차량들의 운행 정보를 실시간으로 알 수 있어 교통의 흐름을 쉽게 파악할 수 있고, 초기 투자 이후 상대적으로 낮은 유지 비용으로 최근 각광을 받고 있다.

특히, 인공위성을 이용한 GPS 방식의 경우 수신기만 있으면, 쉽게 현재 위치를 파악할 수 있어 BIS(Bus Intelligent System)이나 CNS(Car Navigation System)과 같은 분야에서 활발하게 사용 되어지고 있고, 최근 들어서 실시간 교통 정보 제공을 위한 매체로도 활발하게 발전 되어가는 추세이다.



(그림 1) GPS 를 이용한 링크 통행 시간 수집 방법

하지만, 현재 수집되는 정보만으로는 수집되는 모든 GPS 데이터가 해당 링크에서 정상적으로 통행한 데이터인지의 여부는 알 수가 없다.

본 연구의 목적은 특정 링크에서 수집된 GPS 데이터 중 의미가 있는 정보만을 취해서 링크의 정확한 통행 속도를 산출해 내는 방법을 알아보는 데 있다.

2. 수집 데이터

1) 수집 일시와 구간

본 연구에서는 2005년 6월 8 ~ 30일 사이에 강남대로(한남 IC-신사역, 길이 200m)를 대상으로 GPS 수신기를 장착한 수집 차량(영업용 택시)들에 의해 수집된 리얼텔레콤(주)의 속도 자료를 바탕으로 진행한다.

수집 시간은 05 ~ 24시이며, 수집 자료의 형태는 5분 단위 간격으로 해당 링크를 통과한 차량들의 개별 속도(Km)로 구성되어 있고, 최대 10개의 자료가 수집될 수 있다. 수집된 속도가 없는 경우는 0으로 표시되며, 수집 상의 오류로 인한 데이터는 배제하였다. 예제로 <표 1>에 6월 8 ~ 30일 사이의 자료 몇 개를 나타내었다.

<표 1> 6월 8 ~ 30일, GPS 수집 자료(단위:Km)

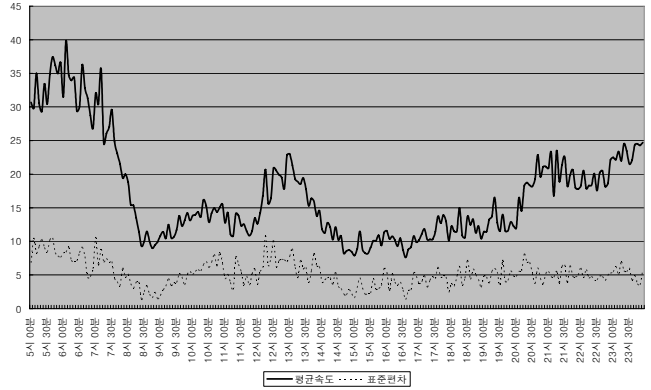
8일	05:00~05:05	13	30	0	0	0	0	0	0	0	0
	05:05~05:10	27	20	0	0	0	0	0	0	0	0
	05:10~05:15	19	0	0	0	0	0	0	0	0	0
...
30일	23:45~23:50	44	13	19	37	37	18	13	0	0	0
	23:50~23:55	21	60	35	35	12	27	24	13	22	7
	23:55~24:00	36	19	56	56	20	13	42	12	46	0

2) 수집 데이터 패턴 분석

6월 8 ~ 30일 사이의 주말을 제외한 월 ~ 금요일(총 17일) 사이의 하루 동안의 시간대별 속도 패턴을 살펴보면 거의 비슷한 모양으로 나타나는 것을 볼 수 있다.

(그림 2)에 새벽 5시 ~ 자정까지의 속도 패턴이 실선으로 표시되어 있으며, 점선은 시간대별 표준편차를 나타내었다. 전체 표준편차는 5.24 Km 정도로 나타났으며, 이는 평일을 기준으로 같은 시간대에 차량의 흐름은 유사하다는 것을 증명해 줄 수 있다.

름은 유사하다는 것을 증명해 줄 수 있다.



(그림 2) 실험 기간 중 하루 동안의 차량 흐름

3. 링크 대표 속도 산출

해당 링크를 통과한 차량 중, 정차나 돌발상황 등으로 인한 방해 요인 없이 정상적으로 통과한 차량들의 속도는 거의 일정하다. 그러나, GPS 데이터는 모든 차량들의 속도 데이터를 수집 하기 때문에, 정상 흐름으로 링크를 통과하지 못한 데이터들을 필터링 해 주는 작업이 필요하다.

본 연구에서는 Quartile(사분위 값)을 이용해 Fence 값을 구하고, 이를 넘어서는 데이터를 비정상적인 데이터로 간주하여 필터링 하는 방법을 사용하였다.

6월 8 ~ 28일 사이의 평일에 해당하는 자료를 통해 Fence 값을 구하고, 6월 29일 자료를 통해 테스트 한다.

1) 전처리

05시부터 24시까지 5분 단위 별로 속도를 높은 속도 순으로 정렬하고, 5분 단위 별 0이 아닌(관측된) 자료의 개수를 구한다.

$$N(i) = \text{해당 시간대의 데이터 수}$$

이 때, $i=1,2,\dots,228$
 $(1=05:00\sim05:05, 2=05:05\sim05:10, \dots, 228=23:55\sim24:00)$

2) Quartile 구하기

Quartile 은 자료의 산포도를 잘 나타내주는 값으로서 그 의미는 다음과 같다.

- $Q1$ = 자료의 25%가 그 값보다 작고 자료의 75%가 그 값보다 크게 되는 값
- $Q2$ = 자료의 50%가 그 값보다 작고 자료의 50%가 그 값보다 크게 되는 값, 즉 중앙값(=Median)
- $Q3$ = 자료의 75%가 그 값보다 작고 자료의 25%가 그 값보다 크게 되는 값
- $IQR(Inter-Quartile Range) = Q3 - Q1$

다음과 같이 5분 단위 시간대 별로 Quartile 을 구한다.

$Q1(i)$ = 해당 시간대의 Q1 값
 $Q2(i)$ = 해당 시간대의 Q2 값
 $Q3(i)$ = 해당 시간대의 Q3 값
 $IQR(i)$ = 해당 시간대의 IQR 값

3) Fence 값 계산

앞에서 구한 Quartile 값을 이용해 5 분 단위 시간대 별로 High Fence, Low Fence 를 구하고, 이를 활용해 비정상적인 차량의 데이터를 제거할 수 있다. 이는 특정 링크의 단위 시간에 대한 해당 Fence 값만 구해놓으면 향후 링크를 지나가는 차량의 속도 측정 시에도 다른 복잡한 계산이 필요 없이 비정상적인 차량의 속도를 실시간으로 제거할 수 있으므로 매우 효율적이라 할 수 있다. Fence 값을 구하는 방법은 다음과 같다.

$$F_{high}(i) = Q3 + M_{high} \times IQR(i)$$

(단, $F_{high}(i) > 80$ 이면, $F_{high}(i) = 80$)

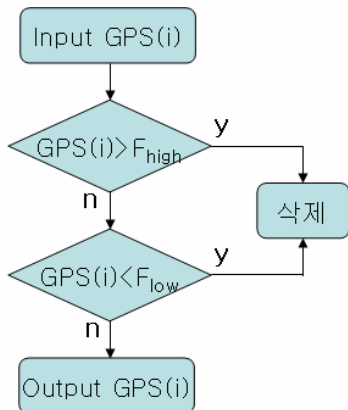
$$F_{low}(i) = Q1 - M_{low} \times IQR(i)$$

(단, $F_{low}(i) < 0$ 이면, $F_{low}(i) = 0$)

여기서 M_{high} 와 M_{low} 는 필터링 강도(Measure)로써, 이를 조정해 필터링 정도를 조절할 수 있고, $F_{high}(i)$ 는 해당 링크 제한 속도를 감안해 80 이상이 될 수 없도록 하였는데, 이는 지나치게 과속하는 차량들이 필터링 될 수 있도록 하기 위함이다.

4) 링크 대표 속도 산출

과거의 패턴 자료를 통해 $F_{high}(i)$ 과 $F_{low}(i)$ 를 구하고, 이를 이용해 해당 링크의 필터링 범위를 선정할 수 있다. 해당 링크를 통과하는 GPS차량의 자료 중 필터링 범위에 해당하지 못하는 자료는 실시간으로 제거 되도록 구현하였다.



(그림 3) Fence 값을 이용한 GPS 자료의 필터링

4. 결과

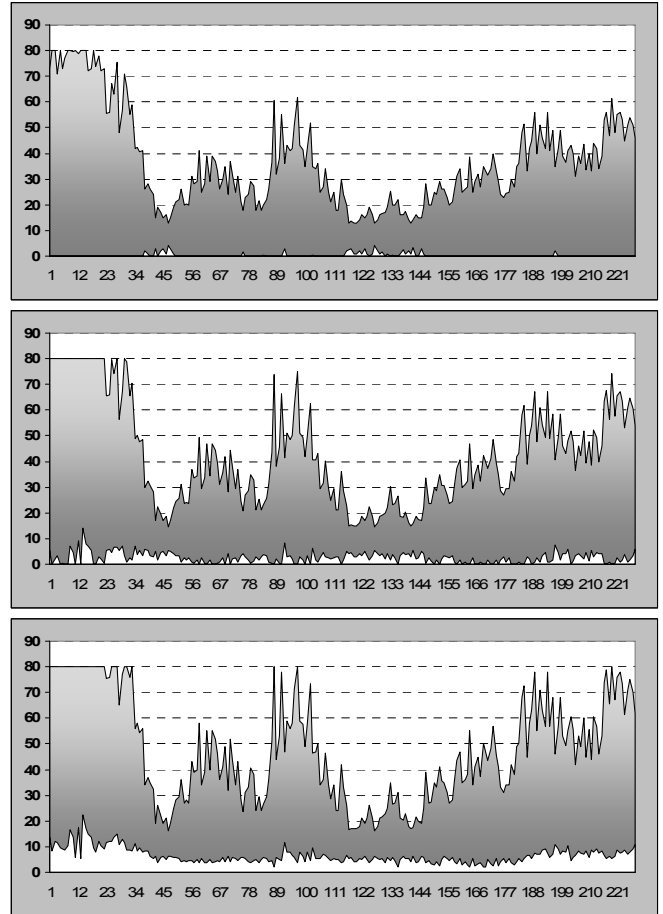
수집 기간 동안 수집된 전체 데이터의 수는 14,206 개이며, 2005년 6월 29일의 988개 자료를 대상으로 테스트를 수행하였다.

M_{high} , M_{low} 의 변화에 따라 결과가 어떻게 달라지는지를 보기 위해 M_{high} , M_{low} 를 각각(1.0, 1.0), (1.5, 0.5),

(2.0, 0.2)의 세 단계로 변화를 주면서 실험을 해 보았다.

1) 필터링 범위

M_{high} , M_{low} 를 달리하면서 정상적인 흐름으로 지나간 차량인지의 여부를 판단하는 필터링 범위를 조절할 수 있다. 다음 그림은 M_{high} , M_{low} 를 위와 같이 변화시켰을 때, 필터링 되는 범위가 달라지는 것을 보여주고 있다. 위의 선이 $F_{high}(i)$ 이고, 아래 선이 $F_{low}(i)$ 이 되겠다.



(그림 4) 단계별 필터링 범위

2) 필터링 결과

범위에 들어오지 못하는 GPS 데이터는 비정상적인 데이터로 간주되어, 링크 대표 속도 값을 산출하는 데 영향을 미치지 못하도록 삭제된다.

각 단계별로 제거되는 데이터의 비율은 차례대로 7.19, 5.45, 12.04% 이다.

3) 링크 대표 속도 값 산출

필터링 범위와 결과를 분석해보면 첫 번째 단계에서는 높은 속도들의 자료가 많이 제거되고, 세 번째 단계에서는 낮은 속도들의 자료가 많이 제거되었을 것이라고 예상해 볼 수 있는데, 이는 단순 평균 값과 필터링 후 산출한 평균값을 비교해보면 잘 알 수 있

다. 단계별로 5 분 단위 링크 대표 속도 값(실선)을 산출해 단순 평균 값(점선)과 함께 나타낸 다음 그림에 잘 나타나있다.

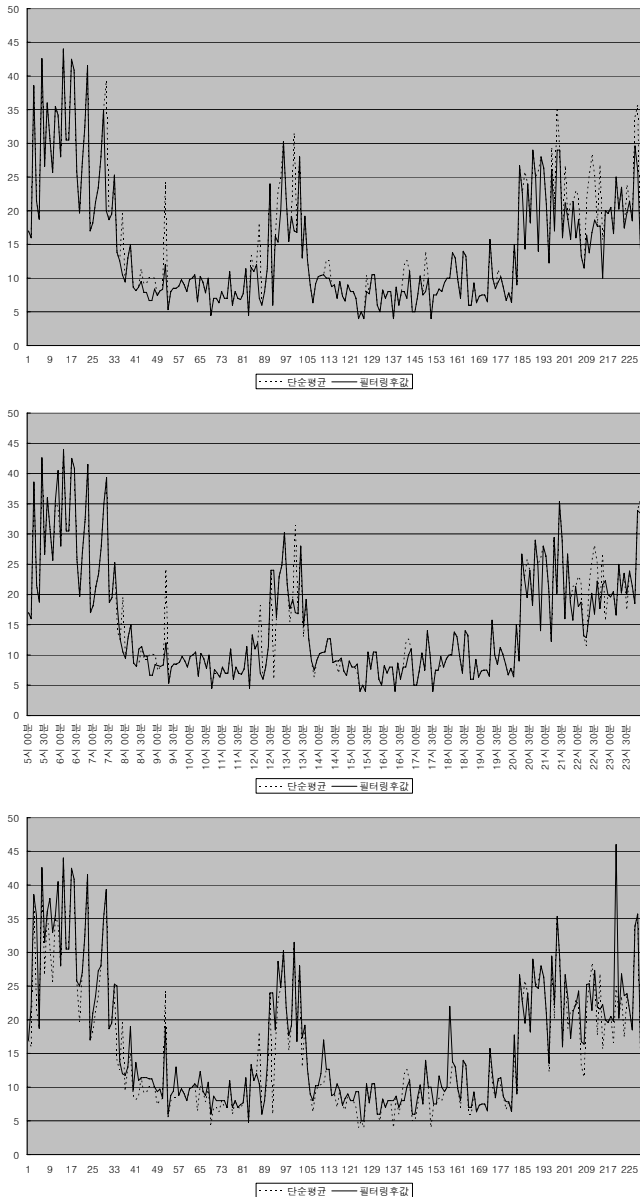
앞으로의 연구 과제는 실제 차량 주행 자료, 또는 다른 수집 장비 데이터와의 비교를 통한 본 알고리즘의 확실한 검증 및 보완이다.

사 사

"본 연구는 정보통신부 및 정보통신연구진흥원의 대학 IT 연구센터 지원사업의 연구결과로 수행되었음" (IITA-2005-C1090-0502-0009)

참고문헌

- [1] 최기주, 신치현, "GPS 와 GIS 를 이용한 링크통행 시간 예측기법" 대한교통학회지, 제 16 권 제 2 호, 1998.
- [2] 김도경, "루프검지기를 이용한 도시간선도로의 실시간 통행속도 추정방법론", 서울시립대 석사학위 논문, 1998.
- [3] 정연식, 최기주, "GPS probe 및 루프검지기 자료의 융합을 통한 통행시간 추정 알고리즘 개발", 대한교통학회지, 제 17 권 제 3 호, 1999.
- [4] 김영찬, 최기주, 김도경, 오기도, "단일루프검지기를 이용한 간선도로 실시간 통행속도 추정 방법론", 대한교통학회지 제 15 권 제 4 호, 1997.
- [5] 정우진, 이종수, 고진웅, 박평수, "퍼지논리 및 GPS 정보를 이용한 링크통행속도의 예측", 한국퍼지및지능시스템학회논문지, 제 13 권 제 3 호, 2003.



(그림 5) 단계별 5 분 단위 링크 대표 속도

5. 결론

바쁜 현대인들에게 자동차에서 낭비하는 시간을 조금이라도 줄여줄 수 있는 텔레매틱스 서비스의 중요성이 날로 증가하고 있다. 이를 위해서, 정확하고 신속한 교통 정보의 제공이 필수적인 요소임은 분명하다.

본 연구에서는 Quartile 값을 이용해 GPS 단말기에서 수집되는 차량의 속도 자료를 실시간으로 필터링하여 링크의 대표 속도 값을 산출해내는 방법을 구현하였다. 이는 향후, GPS 단말기를 장착한 차량이 증가할수록 더욱 널리 쓰일 수 있는 방법이라 기대한다.