

Levenberg-Marquardt 알고리즘을 이용한 시내버스 지연요소 추정

이진우** · 이현미** · 이현수*

Estimation of City Bus Delay Element using Levenberg-Marquardt

Jin-Woo Lee** · Hyun-Mi Lee** · Hyeon-Soo Lee*

요약

최근 국내·외에서 버스운행의 효율화 및 D2D(Door to Door)서비스, 대중교통의 자율주행을 위해 교통데이터를 분석하여 다양한 분석결과를 도출해내고 있다. 하지만 대중교통, 특히 버스 지연시간의 예측을 위해 다양한 연구가 수행되고 있으나 단순분석, 데이터 취득의 한계로 현재까지의 연구는 미흡한 상태이다. 본 연구에서는 버스의 운행정보를 기반으로 요일, 날씨, 시간대 등의 데이터를 추가적으로 수집·가공하여 지연시간 추정을 수행하였다. 본 논문에서 제안하는 방법은 향후 변수를 추가하여 정확도를 향상시킴으로써 자율주행 대중교통 및 대중교통 관제시스템에 활용이 가능하다.

ABSTRACT

Recently, traffic data is analyzed for efficiency of bus operation, D2D(Door to Door) service, and self-driving of public transportation. However, various studies have been carried out to predict the delay time of public transportation, especially buses, but the research to date has been insufficient due to limitations of simple analysis and data acquisition. In this study, delay time estimation is performed by collecting and processing data such as day of the week, weather, and time of day based on bus operation information. The proposed method in this paper can be applied to autonomous public transport and public traffic control system by improving the accuracy by adding variables in the future.

키워드

Prediction of Delay Time, City Bus, Levenberg-Marquardt, Pearson Correlation Coefficient
지연 시간 예측, 시내 버스, Levenberg-Marquardt, 피어슨 상관 계수

1. 서론

최근 국내·외에서 교통 데이터 분석에 대한 관심이 집중되고 있으며 이에 따라 관련 서비스를 제공하는 전문기업이 등장하고 있는 추세이다. 대부분의 지자체

에서는 버스이용의 편의성을 증진시키기 위해 버스운행관리시스템(BMS: Bus Management System)과 버스정보안내시스템(BIS: Bus Information System)을 구축하여 운영 중이다. 일부 지자체의 경우 교통카드 시스템, 버스정보시스템, 첨단교통관리시스템 등에서

* 교신저자 : ㈜LBC소프트 ICT연구팀장

• Received : May. 02, 2017, Revised : May. 31, 2017, Accepted : Jun. 16, 2017

** (주)LBC소프트(jwlee@lbcsoft.com, hmlee0625@gmail.com)

• Corresponding Author : Hyeon-Soo Lee
Smart ICT Team, LBCSOFT,
Email : ccie25@hotmail.com

• 접수일 : 2017. 05. 02
• 수정완료일 : 2017. 05. 31
• 게재확정일 : 2017. 06. 16

스마트 시티정책의 ITS구축 일환으로 방대한 데이터를 수집하고 있다¹⁾²⁾.

외국의 경우 버스운행관리시스템과 버스정보안내시스템을 구축하여 실시간 버스 도착정보 및 위치정보는 제공하고 있다³⁾⁴⁾. 현재 Sensor Fusion[1]을 통한 실시간 위치정보, Zigbee를 통한 실시간 위치정보[2] 등 다양한 수단을 통해 수집되는 데이터를 통해 차량의 실시간 위치[1-2]를 제공하고 있다. 또한 도착정보, 여유 좌석 정보 등의 서비스[3-4]는 이미 상용화 수준에 도달한 상태이다.

버스 통행시간 예측/예보 시스템에 대한 연구개발은 아직 이루어지지 않았고 그 요인에 대한 분석도 미흡한 상태이다. 대중교통수단을 통한 경로 검색 시 예상 소요시간 산출에 있어 버스 평균속도와 정류장간 거리를 이용한 단순계산에 불과하여 다양한 변수가 고려되지 않았으며, 신뢰성이 확보되지 못한 실정이다[3-4]. 따라서 버스의 운행시간이 지연되는 중요 요인 및 정밀한 지연시간 예측에 대한 연구가 요구되고 있으며 이를 위한 데이터를 수집하기 할 수 있는 인프라의 구축 또한 요구되고 있다.

이에 본 논문에서는 피어슨 상관분석을 통해 대중교통의 지연요인을 분석하였고, ANN(Artificial Neural Network)[6]과 함수 최적화 기법인 Levenberg-Marquardt 알고리즘[7-9]을 결합한 학습 모델을 활용하여 버스의 도착시간에 대한 모델을 제안한다.

II. 시내버스 지연시간에 따른 특징추출

2.1 시내버스 데이터 수집

본 논문에서 대중교통의 지연요인분석 및 지연시간 예측을 위해 서울에서 운행 중인 106번 버스를 대상으로 운행 정보를 수집하였다. 그림 1과 같이 106번 버스는 의정부시에서 출발하여 서울 혜화동 로터리를 통해 이후 역순으로 의정부시로 돌아가는 경로이며 106번 버스의 왕복운행거리는 45km 이며 편도로 43

개의 정류장을 경유하여 순환한다. 운행시간은 시간대별로 판이하지만 대체로 120~170분정도로 순환한다.



그림 1. 관측된 시내버스의 경로
Fig. 1 Path of observed city bus

그림 1은 구글맵에 106번 버스의 경로를 매핑시킨 결과이다. 그림 2는 106번 버스가 운행하며 생성한 데이터 셋이다. 수집된 데이터 셋은 그림 2와 같이 각 정류장에서 출발한 시간을 기준으로 수집되었다.

201503.01 버스번호	서울70시2	서울70시2	서울70시2	서울70시2	서울70시2	서울70시2	서울70시2
출발 대원역	4:47:12	7:41:43	10:59:06	14:44:05	18:01:43	21:24:13	5:06:13
출발 가재울교차로	4:49:00	7:42:10	10:59:33	14:44:35	18:04:03	21:26:33	5:06:57
출발 가남역.삼베드로병원	4:49:52	7:44:07	11:01:36	14:46:43	18:06:31	21:29:00	5:08:37
출발 의정부보훈지청	4:50:29	7:45:13	11:02:47	14:47:39	18:07:33	21:29:57	5:09:51
출발 의정부경찰서앞	4:51:13	7:46:31	11:04:01	14:48:27	18:10:18	21:31:00	5:10:58
출발 의정부역.홍선지하도입구	4:52:17	7:47:17	11:04:56	14:49:13	18:11:15	21:32:22	5:11:29
출발 의정부3동우체국	4:54:19	7:48:55	11:07:05	14:51:38	18:13:51	21:34:46	5:13:29
출발 의정부3동주민센터입구	4:55:33	7:52:02	11:09:44	14:54:47	18:15:35	21:35:38	5:14:38
출발 호원2동주민센터입구	4:56:39	7:53:32	11:11:05	14:56:26	18:16:56	21:37:13	5:15:57
출발 회동역	4:58:29	7:56:13	11:13:36	14:58:42	18:18:32	21:38:27	5:18:25
출발 호원동.한주34차아파트	4:59:03	7:56:43	11:13:57	14:59:15	18:19:12	21:39:07	5:18:53
출발 자동차검사소.한주아파트	4:59:35	7:57:19	11:14:26	14:59:49	18:19:35	21:39:43	5:19:15
출발 회동중학교	5:00:20	7:59:19	11:15:05	15:01:57	18:20:23	21:41:51	5:20:05
출발 안달.병무청.호원예비군	5:02:45	8:00:21	11:15:54	15:03:10	18:21:15	21:42:55	5:20:59
출발 경수원.신한대학교	5:04:44	8:02:22	11:17:43	15:04:47	18:22:46	21:44:56	5:22:25
출발 호원고교.호암초교	5:05:26	8:04:02	11:19:11	15:06:38	18:24:05	21:46:33	5:24:05
출발 롯데.신도아파트	5:06:23	8:04:46	11:19:41	15:07:23	18:24:48	21:47:15	5:24:51
출발 다락원.서울인강학교앞	5:08:59	8:08:00	11:22:28	15:10:06	18:27:42	21:49:14	5:26:35
출발 도봉산역(중)	5:11:49	8:10:37	11:24:33	15:12:05	18:28:59	21:50:32	5:28:10
출발 도봉산역(중)	5:11:49	8:12:27	11:25:30	15:14:17	18:31:05	21:51:36	5:29:21
출발 서울북부지방병원.검찰청	5:14:14	8:13:55	11:27:26	15:15:41	18:32:21	21:54:10	5:31:43
출발 신도봉사거리(중)	5:14:43	8:15:13	11:27:53	15:16:48	18:33:32	21:54:36	5:33:14
출발 신도봉시장.도봉구청입구	5:16:47	8:18:36	11:29:49	15:19:40	18:36:39	21:56:43	5:34:11

그림 2. 수집된 시내버스 데이터 셋
Fig. 2 Collected data set of city bus

수집된 데이터 셋을 기반으로 지연요인을 분석하기 위하여 해당 날짜의 날씨정보 및 요일정보를 추가로 수집하였으며, 시간정보를 표 1과 같이 구체화하여 저장하였다. 실제 대중교통은 시간단위의 변화보다는 출근, 퇴근 등의 시간의 집합으로 표현된 구간이 가지는 의미가 크기 때문에 지연요인의 분석을 위해 매시간

1) <http://topis.seoul.go.kr/>
 2) <http://www.gbis.go.kr>
 3) <http://socialtransport.org/>
 4) <https://tfl.gov.uk/>

단위로 분석하는 것보다 7개의 구간으로 분석하는 것이 의미 있을 것으로 판단하였다.

표 1. 시간대 설정
Table 1. Configuration of Time Slot

Type	Time Slot
Dawn	6-7
Morning Rush Hour	7-9
Morning	9-12
Noon-hour	12-13
Afternoon	13-17
Evening rush hour	17-20
Night	20-24

2.2 피어슨 상관계수를 통한 시내버스 지연요인

피어슨 상관계수는 변수간의 상관관계를 나타내기 위해 널리 사용되는 상관계수이다. 피어슨 상관계수를 구하기 위해서 식(1)을 활용한다. x와 y는 각 변수에 속하는 변량을 의미하며 n은 총 변량의 개수, z는 표준점수를 의미한다.

$$r = \frac{\sum z_x z_y}{n - 1} \quad (1)$$

본 논문에서는 대중교통의 운행시간에 영향을 미치는 요소를 분석하기 위해서 날씨, 요일, 날씨, 시간대를 변수로 설정하였다.

표 2. 피어슨 상관계수 결과
Table 2. Result of Pearson Correlation Coefficient

	Date	Day of Week	Weather	Time Slot	Operation Time
Date	1	-	-	-	-
Day of Week	0.057	1	-	-	-
Weather	0.052	0.02	1	-	-
Time Slot	0.01	0.02	0.001	1	-
Operation Time	0.01	0.19	0.132	0.313	1

분석된 결과는 표2와 같다. 대중교통의 지연시간에는 운행시간(Operation Time)과 시간대(Time Slot)의 relation값이 0.313으로 가장 뚜렷한 양의 선형관계를 나타냈다. 다음으로는 운행시간(Operation Time)과 날씨(Weather)가 0.132로 양의 선형관계를 나타냈으며 운행시간과 날씨(Date)와의 상관관계는 0.01로 매우 약한 상관관계를 나타냈다.

2.3 성능지수 예측 모형

본 논문에서는 교통량 예측을 위한 모델로서 LM(Levenberg-Marquardt) 알고리즘[5-7]을 활용하였다. LM알고리즘은 오차 함수를 테일러 급수로 전개하고 2차 미분을 통하여 근사 최소점을 찾는 방법[7]으로 curve fitting, neural network등에서 비선형 함수 최적화 문제에 널리 사용되는 방법이다. 그림3은 LM알고리즘을 ANN에 결합한 학습모델의 구조도이다. 입력노드의 구성은 2.1절에서 분석을 위한 변수들인 날씨, 요일 등으로 구성하였다.

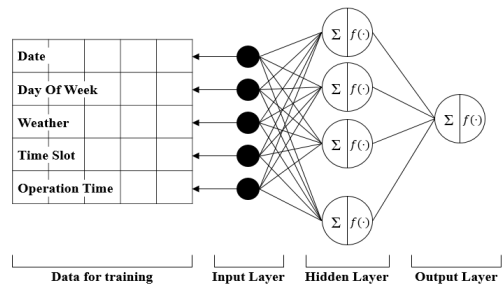


그림 3. 지연시간 예측을 위한 다층 신경망 구조
Fig. 3 Multi-Layer Perceptron for Prediction of Delay Time

LM 알고리즘의 오차함수는 측정값과 함수와의 오차에 대한 제곱 합으로 식(2)와 같이 정의된다[8].

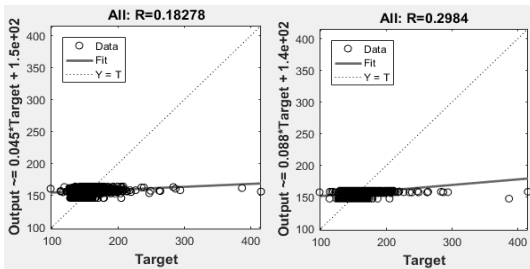
$$F(a) = \sum_{i=1}^N [y_i - y(x_i; a)]^2 \quad (2)$$

y_i 는 측정값을 의미하고 $y(x_i; a)$ 는 매개변수 a에 대한 함수 값을 나타낸다.

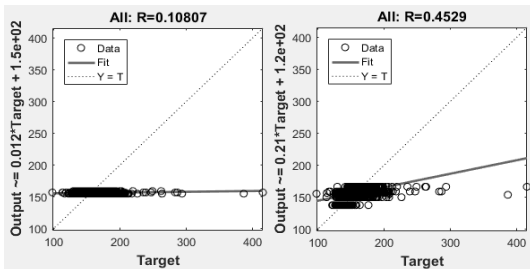
III. 실험결과 및 토의

3.1 LM알고리즘 결과

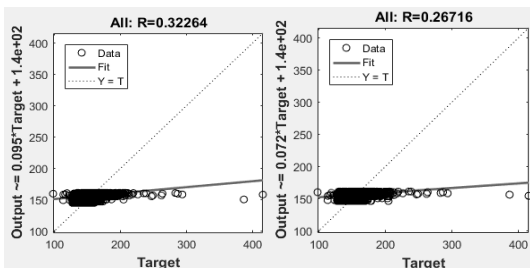
본 논문에서는 효율적인 교통량 예측을 위하여 변수들의 조합을 통하여 회귀분석을 수행하였다. 그림 4는 변수들의 조합들을 LM 알고리즘으로 분석하였고 그 결과를 Regression 그래프로 표현하였다. 그림 4의 캡션에서 DW는 요일, WE는 날씨, TS는 시간대를 의미한다. 그림 4의 (a)(b)(c)(d)는 단일 변수의 분석 결과이고 그림 4의 (e)(f)(g)(h)(i)는 변수들의 조합의 분석결과이며 그림 4의 (m)은 관측된 전체 데이터의 분석결과이다.



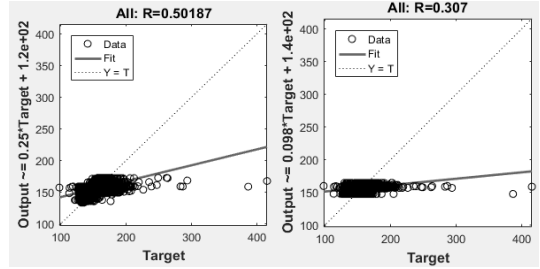
(a) Date(날짜) (b) Day of Week(요일)



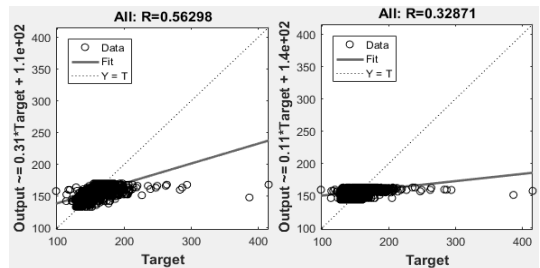
(c) Weather(날씨) (d) Time Slot(시간대)



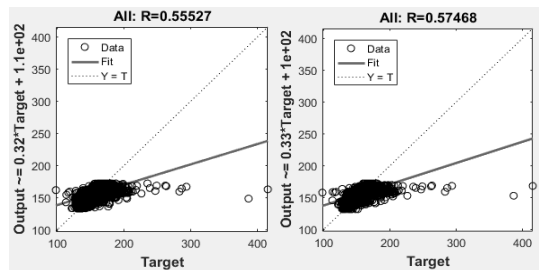
(e) Date+DW(날짜+요일) (f) Date+WE(날짜+날씨)



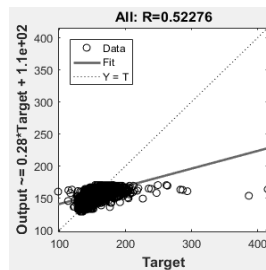
(g) Date+TS(날짜+시간대) (h) DW+WE(요일+날씨)



(i) DW+TS(요일+시간대) (j) Date+DW+WE(날짜+요일+날씨)



(k) Date+DW+TS(날짜+요일+시간대) (l) DW+WE+TS(요일+날씨+시간대)



(m) Whole Data(전체 변수)

그림 4. 분석 결과
Fig 4. Result of analysis

조합된 변수들은 표 3에서 보이는 것과 같이 단일 변수들과 각 변수들의 개수별 조합으로 실험을 구성하였다. R값은 출력과 목표 값과의 관계를 나타낸다. 표 3에서 보이는 것같이 단일 변수들 중에서는 시간대(Time slot)이 지연시간과의 상관관계가 가장 높게 나타났으며 변수들의 조합 중 요일(Day of Week)+날씨(Weather)+시간대(Time Slot)의 조합이 지연시간과 가장 높은 상관관계를 나타냈다. 수집된 전체 변수의 조합(Date+Day of Week+weather+Time Slot)은 기대한 바와는 다르게 지연시간에 대해 낮은 상관관계를 나타냈다. 이는 지연시간을 예측하기 위해서 현재 고려된 변수들 중에서 단일변수로는 시간대가 가장 중요한 변수이며 조합된 변수들의 집합으로는 요일과 날씨, 시간대의 변수가 가장 주요한 변수들의 집합임을 의미한다.

표 3. LM 알고리즘 결과
Table 3. Result of LM Algorithm

Features	R
Date	0.18278
Day of Week	0.29844
Weather	0.10807
Time Slot	0.4529
Date+Day of Week	0.32264
Date+Weather	0.26716
Date+TimeSlot	0.50187
Day of Week+Weather	0.307
Day of Week+TimeSlot	0.56298
Date+Day of Week+Weather	0.32871
Date+Day of Week+TimeSlot	0.55527
Day of Week+Weather+TimeSlot	0.57468
Whole	0.52276

V. 결론 및 향후과제

본 연구에서는 취득이 용이한 날씨, 요일, 날씨, 시간대 정보를 통해 시내버스의 지연시간을 추정하는 방법을 제안하였다. 제안 방법은 시내버스의 각 정류장에 도착시간을 수집하고, 수집 대상에 대한 날씨, 요일, 날씨, 시간대 정보를 추출하여 추정을 수행하였다. 제안방법에서 추정된 지연시간 및 지연요소는 실

제 시내버스 운행 중 다음 정거장까지의 시간을 예측하기 위한 모델에 활용할 수 있다. 또한 지연요소에 대한 추가연구를 통하여 버스의 운영계획에 활용할 수 있다. 향후 구간 별 도로상황 및 버스의 상태정보, 승하차 승객의 수 등 비교적 취득이 어려운 변수들을 보강이 필요하며 시계열적인 분석을 보강하여 이를 통하여 시내버스의 지연시간 추정의 정확도를 높임으로써 자율주행 대중교통 및 대중교통 관제시스템에 활용이 가능할 것으로 기대한다.

감사의 글

본 논문은 2016년도 국토교통부에서 국토교통과학기술진흥원의 교통물류사업의 지원으로 수행되었음.(과제번호 16CTAP-C097506-02).

References

- [1] J. Park and K. Choi, "Overview of sensor fusion techniques for vehicle positioning," *J. of the Korea Institute of Electronic Communication Sciences*, vol. 11, no. 2, 2016, pp. 139-144.
- [2] Y. Ko, "An Experimental Study about the BIS(Bus Information System) using ZigBee Communication," *J. of the Korea Institute of Electronic Communication Sciences*, vol. 7, no. 1, Feb. 2017, pp. 47-52.
- [3] K. Nam, "Study on Intelligent bus management system using beacon-based BIS," *J. of the Korea Institute of Electronic Communication Sciences*, vol. 12, no. 1, Feb. 2012, pp. 157-162.
- [4] C. Steven, Y. Ding, and C. Wei, "Dynamic bus arrival time prediction with artificial neural networks," *J. of Transportation Engineering*, vol. 128, no. 5, 2002, pp. 429-438.
- [5] B. Yu, Y. Zhongzhen, and Y. Baozhen, "Bus arrival time prediction using support vector machines," *J. of Intelligent Transportation Systems*, vol. 10, no. 4, 2006, pp. 151-158.
- [6] B. Kwon and N. Cho, "Development of an Early Warning System Based on Artificial Intelligence," *Internet Explorer Interfaces*, vol. 25, no. 3, 2012, pp. 319-326.
- [7] M. Hagan and M. Menhaj, "Training

Feedforward Networks with the Marquardt Algorithm," *IEEE Trans. Neural Networks*, vol. 5, no. 6, 1994, pp. 989-993.

- [8] Y. Kwak, "Accelerating Levenberg-Marquardt Algorithm using Variable Damping Parameter," *J. of the Korea Society of Computer and Information*, vol. 15, no. 4, 2010, pp. 57-63.
- [9] J. Park and S. Kim, "Application of Levenberg Marquardt Method for Calibration of Unsteady Friction Model for a Pipeline System," *J. of the Korean Water Resources Association*, vol. 46, no. 4, 2013, pp. 389-400.

저자 소개



이진우(Jin-Woo Lee)

2009년 명지대학교 교통공학과 졸업(석사)

2015년 ~ 현재 (주)엘비씨소프트 ICT연구팀

※ 관심분야 : ITS통신시스템, ICT, C-ITS



이현미(Hyun-Mi Lee)

2014년 London School of Economics and Political Science 대학원(석사)

2017년02월 ~ 현재 (주)엘비씨소프트 ICT연구팀

※ 관심분야 : ITS통신시스템, ICT, C-ITS



이현수(Hyeon-Soo Lee)

2015년 아주대학교 교통공학과 수료(공학박사)

2016년 ~ 현재 (주)엘비씨소프트 ICT연구팀 팀장

※ 관심분야 : ITS통신시스템, ICT, C-ITS