

기계학습 기반 열차 자율주행 기법 개발을 위한 인프라 구축 및 데이터 분석

고경준*, 김정태*

Infrastructure Construction and Data Analysis for Machine Learning Based Automatic Train Operation Scheme

Kyeongjun Ko*, Jungtai Kim*

요 약

본 논문에서는 기존의 열차자동운전(Automatic Train Operation, ATO) 시스템에서 고려하지 못한 환경변수(온도, 습도 등) 및 차량 상태 정보 등을 반영하는 새로운 시스템 개발을 목표로 그것을 위한 인프라 구축 및 ATO 시스템 성능에 영향을 미치는 환경변수를 분석한다. ATO 시스템에서의 핵심 성능 중에 하나인 정위치 정차의 정확도를 측정하기 위한 센서 플랫폼을 제작하고 센서에 의해 수집된 데이터를 분석하여 기계학습에 이용될 환경변수를 추출한다. 실험결과는 일부 환경변수가 정차 오차에 영향을 미치는 것을 보여준다.

Key Words : Machine Learning, ATO, Stop of Train, Temperature, Humidity

ABSTRACT

In this paper, we construct an infrastructure and analyze environmental variables affecting to performance of ATO system to development new train operation system applying environmental variables, such as temperature and humidity, and car condition information which is not considered in the existing ATO system. We make a sensor platform to estimate accuracy of stop of trains and analyze data gathered by the sensor platform to extract variables for machine learning

I. 서 론

열차자동운전(ATO) 시스템은 주어진 속도 제한 내에서의 자동 운전과 역내 정위치 정차제어를 담당하는 신호장치이며, 특히 정위치 정차제어의 경우, 승객 안전 및 운영 효율과 연관된 매우 중요한 기능이다. 현재 ATO 시스템은 여러 제어 알고리즘을 탑재하여 운행하고 있으나 날씨, 탑승 인원 등 다양한 외부 환경 변수의 변화와 제동기의 노후화 등 시스템의 변화로 인하여 규정된 오차 범위를 벗어나는 경우가 빈번

하게 발생하고 있다. 그 이유는 실제 마찰계수에 직접적인 영향을 미칠 수 있는 환경 변수들이 기존의 ATO 알고리즘에 직접적으로 적용이 되지 않기 때문이다.

이런 환경 변수들을 기존의 ATO 시스템에 적용하기 위해서 기계학습을 이용할 수 있다. 기계학습이란 명시적인 프로그래밍 없이 컴퓨터가 스스로 학습하도록 하는 분야를 말한다¹⁾. 기계학습은 최근에 하드웨어 및 알고리즘의 발달에 의해 빅데이터를 효율적으로 처리할 수 있게 됨으로써 크게 주목을 받고 있고

※본 연구는 한국철도기술연구원 주요사업의 연구비 지원으로 수행되었습니다.

•* First and Corresponding Author : (ORCID:0000-0001-7575-864X)Korea Railroad Research Institute, kkj8000@krii.re.kr, 정희원

* (ORCID:0000-0002-0085-7261)Korea Railroad Research Institute, jtkim@krii.re.kr, 정희원

논문번호 : KICS2018-01-013, Received January 9, 2018; Revised March 13, 2018; Accepted March 14, 2018

많은 산업에서 활용되고 있다.^[2-4]

ATO의 기능 중에 핵심인 정위치 정차 제어는 최근에 스크린 도어 설치로 인해서 더욱 더 향상된 정확도를 요구받고 있다. 하지만 현재 열차의 정차 오차는 줄자를 이용하여 특정 시간 동안 수작업으로 측정하는 것이 대부분이고 차량마다 발생하는 정차 오차에 대한 통계값은 현재 존재하지 않는 상태이다. 기계학습을 기반으로 하여 정차 오차 정확도를 향상시키기 위해서는 정차 오차에 대한 많은 데이터를 필요로 하며 이를 위해서는 실시간으로 정차 오차를 측정할 수 있는 장치가 필요하다.

또한 기계학습 기반의 ATO 시스템 개발을 위해서는 학습해야 되는 변수를 추출해야 한다. 변수가 많은 경우에 차원의 저주(Curse of dimensionality) 현상에 빠지기 때문에 여러 후보 변수들 중에 정차 오차에 영향을 미치는 변수만 추출해서 학습해야 할 필요가 있다.

따라서 본 논문에서는 정차 오차를 실시간으로 측정할 수 있는 센서 플랫폼 제작과 그것을 이용하여 측정된 정차 오차와 환경 변수(본 논문에서는 온도와 습도)와의 독립성 검정을 통해 관련 변수를 추출한다.

이를 위한 본 논문의 구성은 다음과 같다. 우선 본론의 1절에서 정차 오차를 실시간으로 측정할 수 있는 센서 플랫폼 제작 원리에 대해 설명한다. 2절과 3절에서는 센서 플랫폼 설치 과정과 데이터 수집 방법에 대해 각각 기술한다. 그 후에 센서 플랫폼에서 수집한 정보들을 이용하여 정차 오차와 관련이 있는 환경 변수들을 결과 부분에 도출하고 마지막에 결론으로 마무리를 짓는다.

II. 본 론

2.1 정차 오차 측정을 위한 센서 플랫폼 제작 원리

정차 오차 측정을 위한 센서 플랫폼 제작 원리는 [5]에서 설명되어 있다. 그림 1에서 볼 수 있듯이, 정차 오차 측정 센서는 센서부와 로깅부로 구성되어 있다. 센서부는 레이저 센서로 구성되어 있으며 SICK사의 DT1000 센서를 사용하였다. 이는 최대 460m까지 측정이 가능하며 속도도 같이 측정할 수 있다.

로깅부는 프로세서 모듈과 인터페이스, 외부 메모리 및 전원으로 구성되어 있다. 우선 RS-422 인터페이스는 센서부에서 측정된 위치와 속도 데이터를 로깅부로 전달하는 역할을 수행한다. 전달된 위치와 속도 데이터는 선로 구배나 곡선에 의해 레이저 센서에서 입력되는 값에 오차가 발생할 수 있기 때문에 프로

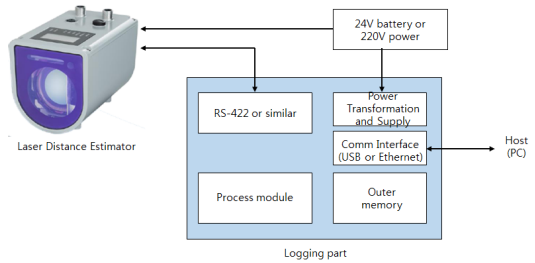


그림 1. 정차 오차 측정을 위한 센서 플랫폼 구성도
Fig 1. Block diagram of sensor platform for estimating stop deviation

세서 모듈에서 변환되어 저장된다. 외부 메모리는 저장된 위치와 속도를 다른 PC에 이동하기 위해 설치된다. 전원은 220V를 받아 어댑터를 통해 24V로 변환되며 이 전압은 RS-422 연결 시, 사용되는 케이블의 여분의 핀을 통해 센서부에도 공급된다. 그림 2는 레이저 센서 및 정차 오차 측정 센서 플랫폼 형상을 보여준다.

다음으로 센서 플랫폼에 의한 정차 오차 측정 방법을 그림 3을 이용하여 설명한다. 그림 3에서 센서 플



(a) 레이저 센서 형상



(b) 정차 오차 측정 센서 플랫폼

그림 2. 레이저 센서 및 정차 오차 측정을 위한 센서 플랫폼
Fig. 2. Figure of a laser and a sensor platform for estimating stop deviation

플랫폼이 지상과 각 θ 를 가지고 설치된다. 센서에서 나오는 레이저가 차량 전두부에 부딪히면 d' 이 바로 계산이 되고 이미 알고 있는 θ 에 의해 센서 플랫폼 지면과 열차 지면 사이의 거리 d 는 다음과 같이 $\cos \theta$ 에 의해 보정된다.

$$d = d' \cos \theta \quad (1)$$

그리고 가로 푸른색 점선으로 표시되어 있는 부분이 정차 오차 측정 기준점이라고 가정하면 미리 알고 있는 d' 에 의해 최종적인 정차 오차 d'' 는 다음과 같이 계산된다.

$$d'' = d - d''' \quad (2)$$

차량이 정차를 위해 역사에 진입하는 상황에서 센서 레이저가 차량 전두부에 부딪히기 시작할 때부터 거리가 측정되기 시작한다. 측정된 거리를 바탕으로 실시간 속도와 가속도도 같이 계산할 수 있다. 샘플링된 거리 샘플순으로 총 N 개의 샘플이 있다고 가정하면, 샘플링된 거리는 $d(1), d(2), \dots, d(N)$ 이라고 정의한다. 그러면 n 번째 샘플에서의 속도 $v(n)$ 은 다음과 같이 계산된다.

$$v(n) = \frac{d(n+1) - d(n)}{t_s} \quad (3)$$

여기서 t_s 는 샘플링 주기이다. 실시간 가속도 또한 속도와 마찬가지로 방법으로 구할 수 있다. n 번째 샘플에서의 가속도 $a(n)$ 은 다음과 같이 계산된다.

$$a(n) = \frac{v(n+1) - v(n)}{t_s} \quad (4)$$

2.2 정차 오차 측정을 위한 센서 플랫폼 설치

ATO 시스템에서 환경 변수를 적용하기 위해, 환경 변수의 영향이 크게 작용할 수 있는 실외 역에서 정차 오차 측정을 위한 센서 플랫폼을 설치하였다. 센서 플랫폼 설치 방안은 그림 4와 같다. 센서는 열차와 마주 보면서 정차 지점에서 일정 거리를 두고 설치되었다. 센서 플랫폼은 지면에서 θ 의 각도를 가지고 설치되었고 θ 는 차량 진입 위치와 정차 기준 지점을 종합적으로 고려하여 적절한 값으로 수작업을 이용하여 결정하였다. (설치역에서 $\theta = 0.97^\circ$) 센서 플랫폼은 센서 부분과 본체 부분을 나누어서 설치를 진행하였다. 정차 오차를 제대로 측정하기 위해서는 선로에서 레이저가 발사되어야 하나 선로 내부에서의 설치는

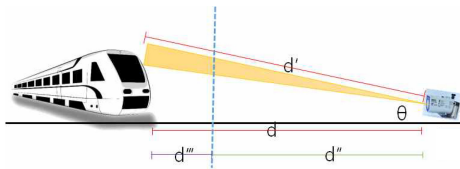


그림 3. 센서 플랫폼에 의한 정차 오차 측정 방법
Fig. 3. Description for how to estimate stop deviation using sensor platform

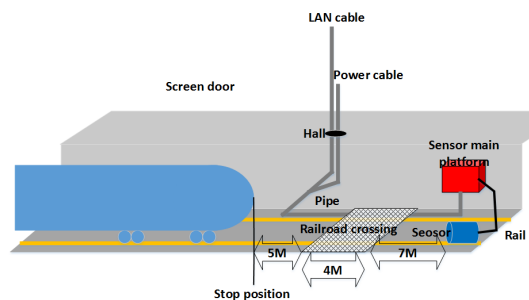
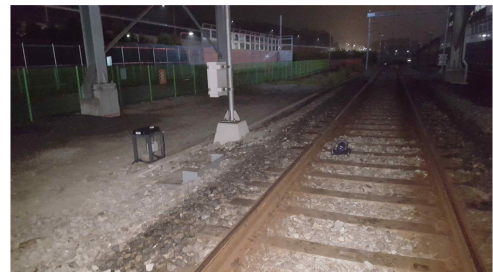


그림 4. 센서 플랫폼 설치 방안
Fig. 4. Installation diagram of a sensor platform



(a) 측정 센서



(b) 측정 센서 및 본체

그림 5. 센서 플랫폼의 실제 설치 현황
Fig. 5. Real installation figure of a sensor platform in a subway station

실제 운행 열차에 피해가 될 수 있기 때문에 최소 부분만 선로에 설치하고 나머지는 선로변에 설치하였다. 따라서 레이저가 발사되는 센서 부분만 선로 내부에, 본체는 선로변에 설치하고 케이블로 연결하였다.

또한, 설치 역의 선로변에 전원을 연결할 수 있는 부분이 부족해서 긴 파이프를 이용하여 역사 내부의 전원에 연결하였고 본체 내부의 메모리 카드에 저장되어 있는 데이터를 수집하기 위해 파이프를 통해 LAN 케이블도 역사 내부의 작업실에 연결하였다. LAN 케이블을 역사 내에 설치하는 이유는 작업자가 수시로 선로 내부에 들어갈 수 없고 데이터 수집 작업을 좀 더 용이하게 하기 위해서이다. 그리고 온도와 습도를 측정하기 위해 본체에 온습도 측정 센서를 부착하였다.

그림 5는 실제 실외 역에서 설치한 센서 플랫폼을 보여준다.

2.3 데이터 수집 방법

센서 플랫폼은 50ms 주기를 ($t_s = 50$) 가지고 csv 파일 형태로 데이터를 로깅한다. 그리고 센서 레이저에 의해 차량이 감지되었을 때만 실시간 거리, 속도, 가속도 등이 로깅되기 때문에 모든 시간에 대해서 데이터가 존재하는 것은 아니다. 센서에 로깅되는 데이터는 unix time, 실시간 거리, 실시간 속도, 실시간 가속도, 온도, 습도 등이 있는데 이런 정보를 가지고 정차 시점 및 정차 지점을 찾아내는 것이 매우 중요하다. 본 논문에서는 실시간 측정 거리 기준으로 50cm 이내의 거리가 100번 이상(5초 이상) 로깅되는 범위를 지정하고 그 범위 내에서, 실시간 속도가 최소값을 가지는 시점을 정차 시점으로 파악한다. 그리고 그 때의 실시간 거리를 정차 오차라고 간주하고 그 때의 온도 및 습도를 필요 데이터로 사용한다.

분석은 2017년 9월 6일부터 10월 25일까지의 수집 데이터를 이용하여 진행하였다. 이 기간 동안 설치역에 지나간 차량은 총 7396편성이다.

III. 실험

환경 변수가 현재의 ATO 시스템에 영향을 미치는 지를 분석하기 위해 온도, 습도와 정차 오차와의 관련성을 분석하였다. 이것을 위해 카이제곱검정을 이용하여 유의 확률(P-value) 계산하였다. 우선 데이터 수집 기간 동안의 온도와 습도의 통계값은 표 1과 같다.

표 1을 바탕으로 카이제곱검정을 위한 표를 작성한다. 우선, 정위치 정차 성공 여부에 대한 기준을 정하

고 온도와 습도에 대한 기준도 함께 결정한다. 표 2와 표 3은 정위치 정차 성공 여부에 대한 기준을 $\pm 10\text{cm}$ 로 가정하고 온도와 습도의 기준을 각각의 평균값을 기준으로 나눈다고 결정한다. 표 2와 표 3은 이런 기준에서 온도/습도와 정차 오차와의 연관성 분석을 위한 카이제곱검정 결과를 각각 보여준다.

표 2와 표 3을 보면 온도와 습도가 정차 오차 $\pm 10\text{cm}$ 기준으로 유의 수준 5% 이내에 들어온다는

표 1. 데이터 수집 기간 동안 설치역의 온도 및 습도의 통계값

Table 1. Statistical values of temperature and humidity during period collecting data in the station

	Temperature(°)	Humidity(%)
Max	30.7	86.4
Min	6.5	17.2
Mean	20.93	52.54
Std	4.16	13.02

표 2. 정차 오차 $\pm 10\text{cm}$ 를 기준으로 온도와 정차 오차와의 카이제곱검정 결과

Table 2. Result of chi-square test between temperature and stop deviation with $\pm 10\text{cm}$ of stop deviation

	stop deviation \leq 10cm (Success)	stop deviation $>$ 10cm (Fail)	Sum
Temperature \leq 20.93	2359	1151	3510
Temperature $>$ 20.93	2519	1367	3886
Sum	4878	2518	7396
P-value	0.033		

표 3. 정차 오차 $\pm 10\text{cm}$ 를 기준으로 습도와 정차 오차와의 카이제곱검정 결과

Table 3. Result of chi-square test between humidity and stop deviation with $\pm 10\text{cm}$ of stop deviation

	stop deviation \leq 10cm (Success)	stop deviation $>$ 10cm (Fail)	Sum
Humidity \leq 52.54	2517	1210	3727
Humidity $>$ 52.54	2361	1308	3669
Sum	4878	2518	7396
P-value	0.004		

것을 알 수 있다. 즉, 온도와 습도가 정차 오차 $\pm 10\text{cm}$ 기준으로 정차 오차에 유의미하게 영향을 미친다고 결론을 내릴 수 있다. 특히, 습도는 온도보다도 정차 오차에 더 큰 영향을 미친다는 것을 알 수 있다. 습도는 선로 마찰계수에 직접적으로 영향을 미치는 수분과 관련이 있기 때문에 습도가 정차 오차에 큰 영향을 미친다고 생각할 수 있다.

표 4와 표 5는 정위치 정차 성공 기준을 $\pm 5\text{cm}$ 로 가정할 때의 온도와 습도에 대한 카이제곱검정 결과를 각각 보여준다.

표 4와 표 5를 표 2, 표 3과 비교해 보면 온도와 습도가 정차 오차 $\pm 10\text{cm}$ 기준보다 정차 오차 $\pm 5\text{cm}$ 기준에서 보다 정차 오차에 영향을 미칠 확률이 더 높다는 것을 알 수 있다. 이를 통해 정위치 정차 성공 기준이 낮을수록 온도와 습도가 정차 오차에 더 많은 영향을 미친다고 결론내릴 수 있다.

표 4. 정차 오차 $\pm 5\text{cm}$ 를 기준으로 온도와 정차 오차와의 카이제곱검정 결과
Table 4. Result of chi-square test between temperature and stop deviation with $\pm 5\text{cm}$ of stop deviation

	stop deviation $\leq 5\text{cm}$ (Success)	stop deviation $> 5\text{cm}$ (Fail)	계
Temperature ≤ 20.93	1381	2129	3510
Temperature > 20.93	1426	2460	3886
계	2807	4589	7396
P-value	0.02		

표 5. 정차 오차 $\pm 5\text{cm}$ 를 기준으로 습도와 정차 오차와의 카이제곱검정 결과
Table 5. Result of chi-square test between humidity and stop deviation with $\pm 5\text{cm}$ of stop deviation

	stop deviation $\leq 5\text{cm}$ (Success)	stop deviation $> 5\text{cm}$ (Fail)	Sum
Humidity ≤ 52.54	1514	2213	3727
Humidity > 52.54	1293	2376	3669
Sum	2807	4589	7396
P-value	2.09×10^{-6}		

본 논문에서 정위치 정차 기준을 $\pm 5\text{cm}$ 나 $\pm 10\text{cm}$ 로 결정한 이유는 그 이상의 경우에 해당하는 역을 지나가는 차량들이 대부분 기준안에 들어오기 때문에 통계적으로 큰 의미를 찾기 힘들고, 광역·도시철도 급행화와 차상 중심 신호시스템 개발을 위해서는 정차 오차를 최대한 줄이는 것이 필요하기 때문이다.

IV. 결 론

본 논문에서는 기계학습 기반 열차자율주행 시스템 개발을 위해 필요 데이터를 수집하고 정차 오차에 영향을 미치는 변수를 도출하였다. ATO 시스템의 핵심적인 기능 중의 하나인 정위치 정차에 영향을 미치는 환경변수를 도출하기 위해 자체적으로 센서 플랫폼을 제작하여 실외 역에 설치하였고 그것을 이용하여 현재 시간, 실시간 속도, 실시간 거리, 온도, 습도 등을 50ms 주기로 수집하였다. 카이제곱검정을 통해 정차 오차와 수집된 데이터들의 상관관계를 분석하여 정차 오차에 영향을 미치는 환경변수를 도출하였다. 실험 결과는 온도와 습도가 정차 오차에 영향을 미친다는 것을 보여주고 특히 습도가 온도보다 더 많은 영향을 미친다는 것을 설명한다. 또한 현재 ATO 시스템에서 정위치 정차의 성공 기준을 좀 더 강화할수록 온도 및 습도가 정차 오차에 더 많은 영향을 미친다는 것을 알 수 있다.

References

- [1] S. Arthur, "Some studies in machine learning using the game of checkers," *IBM J. Res. and Development*, vol. 3, no. 3, pp. 210-229, Jul. 1959.
- [2] IBM Watson, www.ibm.com/watson/
- [3] O. M. Parkhi, A. Vedaldi, and A. Zisserman, "Deep face recognition," in *Proc. British Machine Vision*, pp. 41.1-41.12, 2015.
- [4] aws.amazon.com/ko/big-data/use-cases/
- [5] J. T. Kim, Y. G. Kim, Y. G. Kim, and S. J. Kim, "A study on the development of realtime measurement device for the position and speed of train," in *Proc. KICS Fall Conf.*, pp. 82-83, Daegu, Korea, Nov. 2017.

고 경 준 (Kyeongjun Ko)



2006년 2월 : 서울대학교 전기
공학부 졸업
2012년 8월 : 서울대학교 전기
컴퓨터공학부 박사
2013년 9월~현재 : 한국철도기
술연구원 선임연구원

<관심분야> 통신공학, 기계학습, 빅데이터

김 정 태 (Jungtai Kim)



1997년 2월 : 서울대학교 전기
공학부 졸업
1999년 2월 : 서울대학교 전기
공학부 석사
2011년 8월 : KAIST 전기및전
자공학과 박사
2012년 8월~현재 : 한국철도기
술연구원 선임연구원

<관심분야> 열차제어, 신호처리, 임베디드 시스템