

클라우드와 데이터 마이닝을 이용한 차량 분석 시스템 설계

정이나 · 손수락 · 김경덕 · 이병관*

가톨릭관동대학교

A design of a Vehicle Analysis System using cloud and data mining

Yi-Na Jeong · Su-rak Son · Kyung-Deuk Kim · Byung-Kwan Lee*

Catholic Kwandong University

E-mail : lupinus07@nate.com, sonsur@naver.cm, chamsol93@naver.com, bklee@cku.ac.kr

요 약

본 논문에서는 차량에서 측정되는 모든 센서 데이터를 클라우드에 저장하고, 저장된 데이터를 분류 모델을 이용해 분석한 다음, 분석이 완료된 데이터를 실시간으로 운전자의 디스플레이에 제공하는 “클라우드와 데이터 마이닝을 이용한 차량 분석 시스템”을 설계한다. 제안하는 정보 분석을 위한 클라우드 서버는 차량에서 측정하는 센서 데이터를 클라우드 서버의 테이블에 저장하고 전달받은 데이터를 분석 모듈로 전달하는 센서 데이터 통신 모듈과 분류를 위해 전달받은 데이터를 학습 알고리즘을 이용해 분류한 분류 모델을 이용서 목적에 맞게 분석, 분류하고 운전자에게 실시간으로 정보를 제공하는 센서 데이터 분류 모듈로 구성된다. 제안된 정보 분석을 위한 클라우드 서버는 차량에서 수집되는 수많은 센서 데이터를 클라우드 서버에 저장하기 때문에 차량에 데이터가 과부하 되지 않고 데이터 분류를 위한 연산을 차량이 아닌 클라우드 서버에서 진행하기 때문에 데이터를 빠르고 효율적으로 관리할 수 있다. 또한, 운전자가 원하는 정보들을 디스플레이에 시각화하여 사람들의 자율주행차량에 대한 안정성을 증가시킬 수 있다.

ABSTRACT

In this paper, a "Vehicle Analysis System(VAS) using cloud and data mining" is proposed that store all the sensor data measured in the vehicle in the cloud, analyze the stored data using the classification model, and provide the analyzed data in real time to the driver 's display. The VAS consists of two modules. First, Sensor Data Communication Module(SDCM) stores the sensor data measured in the vehicle in a table of the cloud server and transfers the stored data to the analysis module. Second, Sensor Data Analysis Module(SDAM) analyzes the received data using the genetic algorithm and provides analyzed result to the driver in real time. The VAS stores sensor data collected in the vehicle in the cloud server without accumulating it in the vehicle, and stored data is analyzed in the cloud server, so that the sensor data can be quickly and efficiently managed without overloading the vehicle. In addition, the information desired by the driver can be visualized on the display, thereby increasing the stability of the autonomous vehicle.

키워드

V2X, 데이터마이닝, 클라우드 서버, 차량 센서 데이터, 빅 데이터

I. 서 론

자율주행 자동차는 운전자의 조작 없이 자동차 스스로 운행이 가능한 수준을 의미한다. 자율주행 자동차는 ICT(Information and Communications Technologies)의 영향으로 빠른 성장을

이루고 있다. 세계의 많은 자동차 회사들은 완전한 자율주행 자동차 출시를 목표로 삼고 있다. [1] 자율주행 자동차는 미국 자동차기술 학회(SAE)에 의해 6단계로 분류가 된다. 0단계는 비 자동화 단계, 1단계는 운전자 지원 단계, 2단계는 부분 자동화 단계, 3단계는 조건부 자동화 단계, 4단계는 고도 자동화 단계, 5단계는 완전 자동화 단계다. 현재 자율주행 자동차는 2단계까지 상용화가 되었다. 이

*corresponding author

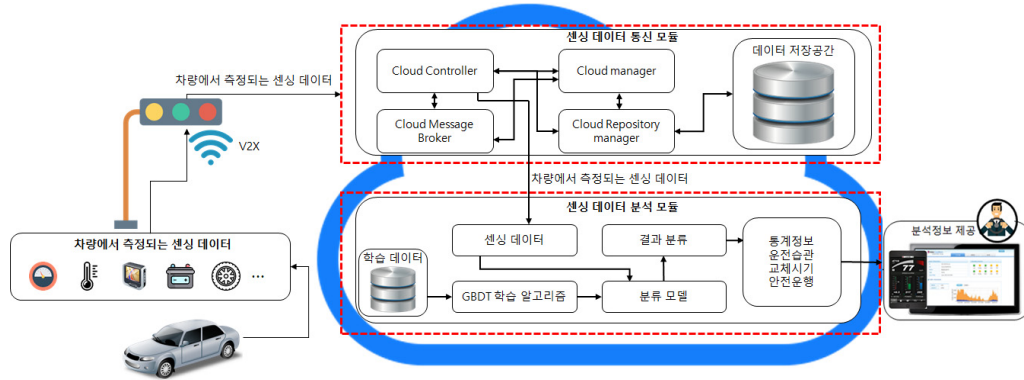


그림 1 자율주행차량의 정보 분석을 위한 데이터 순환 과정

후의 단계들이 상용화를 이루기 위해서는 차량의 안정성을 증가시켜야 한다.

자율주행차량의 안정성을 증가시키기 위해서는 차량을 주행하면서 측정되는 데이터를 효과적으로 이용할 수 있어야 한다. 자율주행차량의 2단계 과정에서는 주행에 관련된 데이터만 분석한다. 하지만 그 외에 데이터들을 수집하고 분석하면 다양한 목적에 맞게 데이터를 응용할 수 있다. 하지만 차량에서 수집되는 모든 데이터를 저장하기에는 차량의 내부 저장공간의 무리가 있다. 또한, 많은 센서 데이터를 빠르고 효과적으로 분석하기 위해서는 차량에 존재하는 연산장치로는 무리가 있다.

이러한 문제를 해결하기 위하여 본 논문은 차량에서 수집되는 차량 센서 데이터를 클라우드 서버로 저장을 하고 저장된 센서 데이터를 학습 데이터로 산출한 분류 모델을 이용하여 목적에 맞게 분류한다. 분류가 완료된 데이터를 운전자에게 실시간으로 제공할 수 있도록 클라우드와 데이터 마이닝을 이용한 차량 분석 시스템을 설계한다.

본 논문의 2장은 클라우드와 데이터 마이닝을 이용한 차량 분석 시스템 설계에 직, 간접적으로 영향을 받은 연구들을 설명하고, 3장은 클라우드와 데이터 마이닝을 이용한 차량 분석 시스템 설계한다. 4장은 클라우드와 데이터 마이닝을 이용한 차량 분석 시스템에 대한 설명과 실험의 결과를 정리하여 향후 연구 방향을 제시한다.

II. 관련 연구

아래 연구들은 기존에 차량의 진단을 위하여 제시된 방법들과 진단 내용을 시각화하는 방법들이다.

[2]에서는 운전자의 운전 행태나 성향을 자동으로 분석한 후 경제 운전을 위한 방법을 운전자에게 알려주기 위해서 차량에 대한 운행일시, 운행거리, 운행시간, 주행속도, 공회전시간, 급가속/급감속 횟수, 연료소모량 등의 운행정보를 수집하였

고, 데이터마이닝을 이용하여 운전자의 운행패턴이 경제 운전에 어떤 영향을 미칠 수 있는지 분석하였다. [2]는 주행 중 운전자에게 지속해서 공회전과 과속 정보, 급가속/급감속 횟수를 차량 단말에 표현하여 제공하고, 공회전과 과속 비율이 일정 임계치를 초과할 경우 경고 정보를 제공함으로써 경제 운전에 악영향을 미칠 수 있는 운전 습관을 미리 예방할 수 있다.

[3]에서는 현재 사용되는 패턴 교통정보가 과거의 정보를 가공하여 교통정보를 예측하기 때문에 특별한 상황(유고, 날씨 등)에서는 예측이 정확하지 않은 문제점을 해결하려는 방법으로, 빠른 길을 찾기 위해 실시간으로 운전자들이 요청하는 경로 탐색 데이터를 분석하여 가까운 미래 운전자들이 위치할 도로의 교통 혼잡도를 미리 파악하여 패턴 교통정보보다 정확한 예측 교통정보를 제시하였다. [3]결과 첫째, 연구지역의 정체 경로인 양재에서 마포 간 차량 속도 비교에서는 기존 상습 정체 도로의 속도가중치 정확도가 3km/h에서 18km/h의 오차율이 발생하였지만, [3]의 Real 예측 교통정보를 적용한 결과는 1km/h에서 5km/h의 오차율이 발생하였다. 둘째, 경로 품질에서 기존의 경로보다 최대 약 9분, 평균 약 3분 일찍 목적지에 도착하여 예측 교통정보 결과의 신뢰성을 입증할 수 있었다. 셋째, 기존의 경로 탐색 결과보다 혼잡도를 예측하여 혼잡이 발생할 도로에 대해 회피되는 경로 탐색 결과를 도출할 수 있었다. 따라서 [3]의 경로 탐색 비교를 통해 교통량에 대한 예측정보를 획득할 수 있었으며 이를 활용하여 실시간 빠른 길 탐색이 가능하고, 향후 차량 흐름을 분산시키는 데도 도움이 될 수 있다.본론은 필요에 따라 3-4 개의 장으로 편집할 수 있습니다.

III. 클라우드와 데이터 마이닝을 이용한 차량 분석 시스템 설계

본 논문에서 제안하는 클라우드와 데이터 마이닝

을 이용한 차량 분석 시스템 설계는 차량에서 측정되는 모든 센서 데이터를 V2X 기반으로 클라우드에 전송한다. 클라우드는 차량에서 전달받은 센서 데이터를 데이터 저장공간에 저장하기 위해서 센서 데이터 통신 모듈로 데이터를 수신한다. 센서 데이터 통신 모듈은 내부에 저장된 센서 데이터를 분석을 위해 센서 데이터 분석 모듈로 전달한다. 센서 데이터 분석 모듈은 전달받은 데이터를 분류 모델에 대입하여, 분류 결과를 산출한다. 산출된 분류 결과는 운전자가 모니터링하고 있는 스마트기기에 실시간으로 제공한다. 그림 1은 자율주행차량의 정보 분석을 위한 데이터 순환 과정이다.

3.1 센서 데이터 통신 모듈

센서 데이터 통신 모듈은 차량에서 측정되는 모든 센서 데이터를 수신하고 데이터 저장공간에 저장한다. 또한, 전달받은 센서 데이터를 센서 데이터 분석 모듈로 전달하는 역할을 한다. 센서 데이터 통신 모듈은 크게 5개로 분류된다. 첫째, Cloud Controller는 차량으로부터 전달받은 데이터를 실제로 처리하는 역할을 한다. 또한, 독립적이고 고유의 업무만의 처리에 이용한다. 둘째, Cloud Message Broker는 차량으로부터 전달받은 데이터를 분석하여, 클라우드의 기능들을 적절하게 사용한다. 그리고 기능을 사용한 결과를 Cloud Controller로 전달한다. 셋째, Cloud Manager은 클라우드 시스템을 제어하고 관리하는 역할을 한다. 넷째, Cloud Repository Manager은 센서 데이터를 관리하는 역할을 한다. 다섯째, 데이터 저장공간은 차량에서 전송된 센서 데이터가 저장되는 장소로 테이블 형식으로 데이터가 저장된다. 테이블의 각 속성은 그림 2와 같이 센서의 종류에 따라 분류된다. RPM, 엔진 온도, GPS, 배터리 전압, 타이어 공기압 등의 차량에 존재하는 센서에 따른 데이터가 저장되는 공간을 할당한다. 센서 데이터는 1000ms 기준으로 전달이 되고, 데이터 저장공간에 저장된다.

센서 데이터 통신 모듈은 전반적인 클라우드의 통신을 관리한다. 이때, 통신은 외부, 내부를 모두 포함하고 있다. 차량에서 측정되는 센서 데이터는 센서 데이터 통신 모듈 내부의 동작이 완료된 후 데이터 분석을 진행하기 위해서 센서 데이터 분석 모듈로 데이터를 전달한다.

	RPM	엔진온도	GPS	배터리 전압	타이어 공기압	...
1	4320	88	37.1234 126.9687	14.4	300	...
2	3850	85	37.1234 126.9687	14.6	298	...
...
n

그림 2. 데이터 저장공간의 테이블 구성도

3.2 센서 데이터 분석 모듈

센서 데이터 분석 모듈은 센서 데이터 통신 모듈에서 전달하는 차량의 데이터를 분석하여 목적에 맞게 분류한다. 센서 데이터 분석 모듈은 센서 데이터를 분석하기 위해서 학습 데이터를 알고리즘을 이용한다. 그리고 학습결과로 분류 모델을 생성한다. 센서 데이터 분석 모듈은 데이터를 학습하기 위해서 지도 학습 알고리즘 중의 하나인 그래디언트 부스팅 결정 트리(GBDT) 알고리즘을 사용한다. GBDT는 깊이가 깊지 않은 단일 결정 트리를 여러 개 묶어 하나의 모델로 만든 것이다. 알고리즘 1은 GBDT의 과정을 나타낸다.

```

알고리즘 1. GBDT
input : training set  $\{(x_i, y_i)\}_{i=1}^n$ , a differentiable
loss function  $L(y, F(x))$ , number of iterations
M.

 $F_m(x) = F_{m-1}(x) + \arg \min_{h_m \in \mathcal{H}} \left[ \sum_{i=1}^n L(y_i, F_{m-1}(x_i) + h_m(x_i)) \right]$ 
for( int m=0; m<M; m++){
//computes pseudo-residuals
for (int i=0; i<n; i++){
double  $r_{im} = -\left[ \frac{\partial L(y_i, F(x_i))}{\partial F(x_i)} \right]_{F(x)=F_{m-1}(x)}$ 
}
training tree generate function  $h_m(x)$ 
using  $(x_i, r_{im})_{i=1}^n$ 
double  $\gamma_m = \arg \min_r \sum_{i=1}^n L(y_i, F_{m-1}(x_i) + \gamma h_m(x_i))$ 
 $F_m(x) = F_{m-1}(x) + \gamma_m h_m(x)$ 
}

Output :  $F_m(x)$ 
    
```

센서 데이터 분석 모듈은 GBDT알고리즘을 이용하여 데이터를 학습한 결과로 분류 모델을 생성한다. 센서 데이터 통신 모듈에서 전송한 차량의 센서 데이터를 분류 모델에 대입하면, 일정한 결과

를 분류할 수 있다. 데이터가 분류된 결과로 다양한 정보를 분석할 수 있다. 예를 들어, 정보의 통계, 운전 습관, 부품의 교체 시기, 안전운행 등을 알 수 있다. 분석한 정보들은 운전자가 모니터링하고 있는 디스플레이에 실시간으로 제공된다.

IV. 결 론

본 논문은 자율주행차량의 안전성을 증가시키기 위해서 기존에 상용화가 되어 있는 주행에 관련된 센서만 자율주행에 이용하고 있는 한계를 극복하려 한다. 그러므로, 차량에서 측정되는 대용량의 센서 데이터를 효과적으로 저장하고 다양한 목적에 맞게 저장된 데이터를 분석, 분류하기 위해서 클라우드와 데이터 마이닝을 이용한 차량 분석 시스템을 제안하였다. 클라우드와 데이터 마이닝을 이용한 차량 분석 시스템은 센서 데이터 통신 모듈을 사용하여 차량에서 측정되는 데이터를 저장하고 관리한다. 그리고 센서 데이터 분석 모듈은 클라우드 서버의 저장공간에 저장된 데이터를 분류 모델을 이용해서 목적에 맞게 분류한다. 또한, 분류가 완료된 데이터를 운전자가 실시간으로 모니터링할 수 있게 디스플레이에 제공한다. 제안한 시스템의 기대효과는 다음과 같다.

첫째, 차량에서 수집되는 수많은 센서 데이터를 효과적으로 저장할 수 있으며, 저장된 데이터를 클라우드 서버로 연산하므로 빠른 속도로 분류할 수 있다. 그러므로, 차량의 부하를 방지할 수 있다.

둘째, 운전자가 원하는 모든 데이터를 디스플레이에 시각화하면서 자율주행차량에 대한 안정성을 증가시킬 수 있다.

본 논문에서 제안한 클라우드 서버는 차량의 센서 데이터를 V2X 기반으로 통신하고, 차량의 내부 디스플레이를 이용해 분석한 정보를 제공한다. 하지만, V2X 통신이 불가능한 차량과 디스플레이가 내부에 존재하지 않는 차량은 고려하지 않았다. 향후, V2X가 없는 차량과 내부에 디스플레이가 없는 차량에 대해서 해당 클라우드 서버를 적용할 방안이 연구되어야 한다.

Acknowledgement

이 논문은 2018년도 정부(과학기술정보통신부)의 재원으로 수행된 NRF-2018R1A2B6007710의 연구 결과로 수행되었음.

References

- [1] [Internet]. Available :<http://www.segye.com/newsView/20161003001346>.
- [2] Hyun-Suk Kim, Jong-Woo Choi, Dae-Woo Kim, Ho-Sung Park, Sung-Kee Noh, Cheong-Hee Park, "An Application of Data Mining Techniques in the Driving Pattern Analysis", *The journal of the Korea institute of intelligent transport systems*, Vol. 8, No. 6, pp.1-12, 2009
- [3] Jun Hyun Kim, Kee Wook Kwon, "A Study on Predictive Traffic Information Using Cloud Route Search", *Journal of the Korean Society of Surveying, Geodesy, Photogrammetry and Cartography*, Vol. 33, No. 4, pp.287 - 296, 2015