

위치 이력 데이터를 활용한 LSTM 기반 위치 예측 연구

최원석, 정지환, 최성곤*

충북대학교

wschoi@cbnu.ac.kr, starwlghks7@chungbuk.ac.kr, *choisg@cbnu.ac.kr

LSTM-based Location Prediction using Historical Location Data

Choi Won Seok, Jeong Ji Hwan, Choi Seong Gon*

Chungbuk National Univ.

요약

본 논문은 사용자 선호도를 반영한 경로 예측 및 위치 기반 서비스 제공을 위해 사용자의 위치 이력 데이터를 활용한 LSTM(Long Short Term Memory) 기반 위치 예측 연구를 기술한다. 사용자 선호도를 반영하기 위한 사용자 위치 예측 학습 파이프라인 구조를 제시하고, 실험실 수준의 실증을 통해 위치 예측 가능성과 활용 가능성을 분석한다.

I. 서론

인공지능 기반의 자율주행차 기술은 자율/협력 주행, 교통 효율, 교통 안전, 인포테인먼트 서비스에 활용하기 위해 활발히 연구되고 있다. 특히, 교통 효율 서비스의 경로 생성 기술은 교통의 흐름, 환경 정보를 포함하여 인공지능 기반의 최적 경로를 찾기 위한 연구가 진행되고 있으며, 최근에는 개인의 주관 또는 선호도를 포함하여 사용자 맞춤형 최적 경로를 추천하는 서비스의 출현을 예상한다.[1][2]

인공지능을 활용하여 선호도 기반의 예측 관련 연구로 Wafa Shafqat 등은 관광지 추천을 위한 이중 구조 LSTM(Long Short Term Memory) 기반 컨텍스트 인지 위치 예측 시스템을 제안하였다.[3] 하지만, 경로 보다는 방문한 장소 위치 데이터를 기반으로 하여 시간의 연속성을 반영할 수 없다. Ankur Pandey 등은 위치기반 광고를 위한 CNN-LSTM 기반 하이브리드 추천 모델을 제안한다[4]. 광고 목적의 추천으로 서비스 응답 품질 예측에 중점을 두고 있어, 사용자 개인의 이동 특성을 반영하지 못한다. Yuxia Wu 등은 사용자의 장기 선호도와 단기 선호도를 고려하여 다음 방문지 예측 모델을 제안하였다.[5] 시계열적인 GPS 좌표를 사용하는 것이 아니라 관심 장소를 기반으로 다음 방문지를 예측한다. Ke Sun 등은 사용자의 장기 및 단기 위치 선호를 반영하는 모델을 제안하였다.[6] 장기 선호는 No-local 네트워크로 학습하고 단기 선호는 geo-dilated LSTM으로 학습한다. 학습 데이터가 부족하고 복잡한 구조로 인해 실시간성이 부족하다.

따라서, 본 논문은 사용자 선호도 기반 경로 예측 및 위치 기반 서비스 제공을 위해 위치 이력 데이터를 활용한 LSTM 기반 위치 예측 방법을 제안한다. 본 논문은 사용자 선호도를 반영하기 위해 LSTM 기반 사용자

위치 학습 파이프라인 구조를 제시하고, 실제 개인 위치 데이터 기반의 실험실 수준의 실증을 통해 위치 예측 가능성과 활용 가능성을 분석한다.

II. 본론

본 절에서는 사용자 선호도 기반 경로 예측 및 위치 기반 서비스 제공을 위한 학습 파이프라인 구조를 설명한다.

그림 1은 제안하는 LSTM 기반 사용자 위치 예측 학습 파이프라인 구조를 보여준다. 파이프라인은 크게 데이터 수집, 데이터 보간, LSTM layer, Dense layer로 구성된다. 데이터 수집은 사용자의 이동을 확인 할 수 있는 GPS 모듈이 장착된 장치에서 네트워크를 통해 수집 서버로 수집 한다. 사용자의 이동은 시간을 기준으로 변경되므로 기본 수집 데이터로 날짜, 시간, 위도, 경도를 수집한다.

데이터 전처리는 시간 기반 선형 보간과 Kalman filter smoothing, One-hot encoding, Normalization으로 구성된다. 먼저, 측정한 GPS 데이터는 전송 지연, 전송 실패, 장치 전원 깨짐 등의 원인으로 특정 구간의 측정 값이 없을 수 있다. 결측치로 인한 학습 및 추론 영향을 줄이기 위해 보간을 수행한다. 보간은 시계열적 연속성을 유지하기 위해 시간 기반 선형 보간법을 사용하여 비어있는 위도, 경도 데이터를 생성한다. 다음으로 GPS 측정의 오차를 고려하여 측정된 위도, 경도 값에 대해 Kalman filter smoothing을 수행하여 노이즈를 제거하여 더 안정적인 학습을 가능하게 한다. 사용자 선호도 특성은 사용자의 이동, 위치 등의 행동은 그 행동을 하는 시점 및 상황과 관련된 날짜, 요일, 시간, 날씨 등이 요인이 될 수 있다. 본 논문에서는 사용자의 날짜, 시간, 요일별 이동 패턴을 수치로 모델에 알려주기 위해 One-hot encoding을 통해 구분하도록 한다. 위도, 경도,

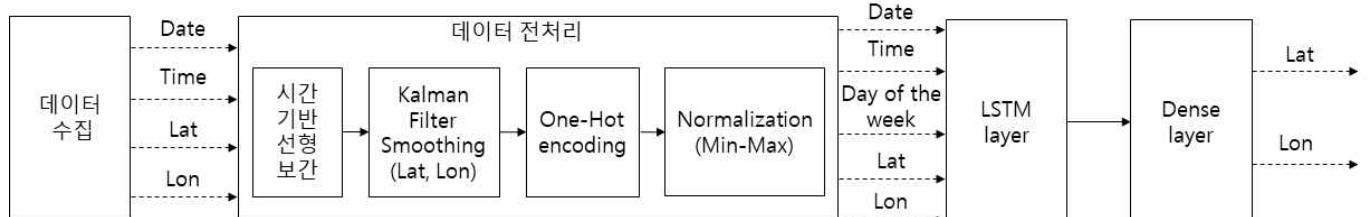


그림 1. 제안하는 LSTM 기반 사용자 위치 예측 학습 파이프라인 구조

날짜, 시간, 요일 등 각각의 입력 값에 따라 LSTM 학습이 안정적으로 수행되도록 스케일을 동일하게 Min-Max Normalization을 수행한다.

정규화된 날짜, 시간, 요일, 위도, 경도 값을 LSTM 계층에 입력 값으로 넣고 학습을 수행한다. LSTM이 마지막 시퀀스의 은닉 상태를 출력하면 Dense 계층이 1차원 벡터를 위도, 경도 값으로 변환한다.

III. 실험 및 결과

본 절에서는 제시하는 LSTM 기반 사용자 위치 학습 파이프라인 구조를 실증하고 실제 개인 위치 데이터를 기반으로 실험하였다.



그림 2. 사용자 위치 이력 수집 장치

그림 2는 사용자의 이력 데이터 수집을 위해 개인 자동차에 설치한 사용자 위치 이력 수집 장치를 보여준다. 라즈베리파이 4에 GPS 모듈을 장착하여 1초 마다 데이터를 수집하도록 구성하였다.

표 1. 실험에 사용한 GPS 모듈 정보

Parameter	Specification
Engine	66 channel GPS, L1CA, QZSS, SBAS
Performance	-162dBm Tracking sensitivity
Accuracy	Position Single Point L1, 3.0m CEP without SA
Baud Rate	9600bps
TTFX	Signal Reacquisition: Hot 1s / Cold: 35s

표 1은 실험에 사용한 Ascenkorea GPS620의 정보를 보여 준다. 이 GPS 모듈은 66채널 GPS와 L1CA, QZSS, SBAS를 지원하고. 수신 감도는 -162 dBm이며, 단일 포인트 L1 기준으로 SA 해제 시 CEP 3.0 m 수준의 정확도를 제공한다. 또한, 통신 인터페이스는 9600 bps의 전송 속도를 사용하고, Time To First Fix(TTDX)는 hot start 시 약 1초, cold start 시 약 35초가 소요된다.

실제 수집된 GPS 데이터를 이용하여 30일간의 데이터를 이용하여 그림 1의 파이프라인 구조로 Single LSTM 기반 모델과 Stacked LSTM 기반 모델에 대해 실험을 진행하였다. 각 예측 결과는 Haversine 거리 기준 정확도와 F1 score를 측정하였다.

```
[single] 평균 Haversine 오차 : 0.0334 km
== single 샘플 5개 ==
1: True=[ 36.6253257 127.46003815 ], Pred=[ 36.62473248 127.46020342 ], Err=0.0676
2: True=[ 36.62532394 127.46004695 ], Pred=[ 36.62471761 127.4602228 ], Err=0.0692
3: True=[ 36.62532218 127.46005575 ], Pred=[ 36.62470638 127.460243 ], Err=0.0705
4: True=[ 36.62532042 127.46006455 ], Pred=[ 36.62469932 127.46026349 ], Err=0.0713
5: True=[ 36.62531866 127.46007336 ], Pred=[ 36.62469674 127.4602873 ], Err=0.0718
single F1: 0.9942
```

(1) Single LSTM 모델 기반 실험 결과

```
[stacked] 평균 Haversine 오차 : 0.1018 km
== stacked 샘플 5개 ==
1: True=[ 36.6253257 127.46003815 ], Pred=[ 36.62565908 127.4596632 ], Err=0.0499
2: True=[ 36.62532394 127.46004695 ], Pred=[ 36.62561068 127.45968931 ], Err=0.0451
3: True=[ 36.62532218 127.46005575 ], Pred=[ 36.62557659 127.45972999 ], Err=0.0406
4: True=[ 36.62532042 127.46006455 ], Pred=[ 36.62555569 127.4597812 ], Err=0.0364
5: True=[ 36.62531866 127.46007336 ], Pred=[ 36.62554582 127.45983828 ], Err=0.0328
stacked F1: 0.6915
```

(2) Stacked LSTM 모델 기반 실험 결과

그림 3. LSTM 기반 사용자 위치 예측 학습 결과

그림 3은 LSTM 기반 사용자 위치 예측 학습 결과를 보여준다. Single LSTM 기반 모델을 사용한 결과 평균 0.0334 km의 haversine 오차를 보였으며, 0.9942의 F1 score를 보였으며, Stacked LSTM 기반 모델을 사용한 결과 평균 0.1018 km 오차와 0.6915의 F1 score를 확인했다. 이는 학습 데이터가 Stacked LSTM 모델을 적용하기에 데이터가 충분하지 않아 학습이 불안정해진 것으로 보인다.

IV. 결론

본 연구에서는 사용자 선호도를 반영하기 위한 사용자 위치 예측 학습 파이프라인 구조를 제시하고, 실험 실수준의 실증을 통해 위치 예측 가능성과 활용 가능성을 분석하였다. 본 연구는 LSTM 모델 기반 개인 차량 GPS 데이터, 날짜, 시간, 요일 정보의 결합을 통해 정밀한 시간 단위의 위치 예측의 가능성을 보였다.

향후 연구에서는 온도, 습도, 날씨와 같은 외부 환경 정보 추가 및 다양한 사용자 데이터 확보를 통해 모델의 범용성과 정확도를 더욱 높일 계획이다.

ACKNOWLEDGMENT

이 논문은 2025년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(No. RS-2020-NR049604). 또한, 2025년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임(No. RS-2024-00397979, 6G 수십 테라급 초정밀 전달망 시스템 기술 개발).

*교신저자: 최성곤(choisg@cnu.ac.kr)

참 고 문 현

- Choi, Won-Seok, and Seong-Gon Choi, "Thread-Based Modeling and Analysis in Multi-Core-Based V2X Communication Device," Sustainability 14(8277), 2022.
- Bathla, G., Bhadane, K., Singh, R. K., Kumar, R., Aluvalu, R., Krishnamurthi, R., Basheer, S., "Autonomous vehicles and intelligent automation: Applications, challenges, and opportunities," Mobile Information Systems, 2022.
- Shafqat, W., & Byun, Y. C., "A context-aware location recommendation system for tourists using hierarchical LSTM model," Sustainability, vol. 12, no. 10, 2024.
- Pandey, A., Manneppalli, P. K., Gupta, M., Dangi, R., & Choudhary, G., "A Deep Learning-Based Hybrid CNN-LSTM Model for Location-Aware Web Service Recommendation," Neural Processing Letters, vol. 56, no. 5, 2024.
- Wu, Y., Li, K., Zhao, G., & Qian, X., "Long-and short-term preference learning for next POI recommendation," In Proceedings of the 28th ACM international conference on information and knowledge management, pp. 2301-2304, Nov. 2019.
- Sun, K., Qian, T., Chen, T., Liang, Y., Nguyen, Q. V. H., & Yin, H., "Where to go next: Modeling long-and short-term user preferences for point-of-interest recommendation," In Proceedings of the AAAI conference on artificial intelligence, vol. 34, no. 01, pp. 214-221, Apr. 2020.