

머신 러닝을 이용한 Cellular Modem System 의 Validation Location 최적화에 대한 연구

신성수, 이원무*, 전호영**
삼성전자

sungsoo.shin@samsung.com, *wonmu.lee@samsung.com, **hoy.jeon@samsung.com

A Study on the Cellular Modem System's Validation Location Optimization Using Machine Learning

Shin Sung Soo, Jeon Ho Young*, Lee Won Mu**
Samsung Electronics

요약

본 논문은 머신 러닝을 활용하여 Cellular Modem System 필드 검증 및 평가를 위한 Location 선정 최적화에 대한 연구 내용을 다룬다. Cellular Modem System 의 필드 검증 및 평가는 셀 수 없이 많은 환경 변수를 고려하여야 하며, 이는 지역이나 RAT(Radio Access Technology), 기지국, 사업자 등에 따른 특성과, 신호의 세기, 간섭과, 이동성에 의해 발생하는 Fading Channel 환경과 예측 불가능한 신호 왜곡 등에서 비롯된다. 본 연구에서는 이와 같은 다양한 환경 변수를 학습하여 최적화된 Location 을 산출하는 방법에 대해 알아본다.

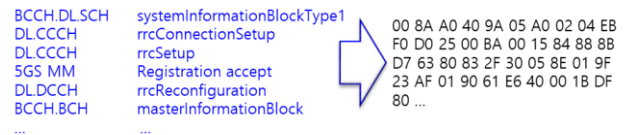
I. 서론

Cellular Modem System 은 RAT(Radio Access Technology)이나 Network 기지국, 통신 사업자에 따라 지원하는 Feature 및 기능이 달라지는 특성과, 신호의 세기, 간섭과 이동성에 의해 발생하는 Fading Channel 환경 등 시변하는 NW Channel 왜곡을 고려하여 설계되고, 이로 인한 다양한 환경 변수에 대한 검증 및 평가는 Cellular Modem 상용화에 있어서 가장 중요하고 필수적인 요소이며 큰 리소스가 소모된다. 이를 위해, 최근에는 AI 학습을 통해 검증 환경을 효율화 하기 위한 연구가 이루어지기 시작했다[1].

일반적인 필드 검증 및 시험으로 모든 경우의 수를 다루기에는 시간과 비용에 대한 제약, 리소스 문제가 뒤따른다. 현재는 다양한 환경 변수에 대한 경우의 수를 고려하기 위해 국가 및 도시, 사업자, Band 정보 등을 기준으로 Location 을 선정하지만, 한정된 조건으로 수많은 환경 변수를 cover 할 수 있는 적합한 Location 선정은 불가능 하다. 그래서, 시간 및 리소스가 가용하는 대로 많은 Location 에서 검증 및 평가를 진행하는 것이 일반적이다. 하지만, 이렇게 선정한 Location 은 모든 환경 변수를 Hit 할 수 없고, 동일한 환경 변수만 검증하는 경우가 발생하여 불필요한 중복 검증이 발생하게 된다.

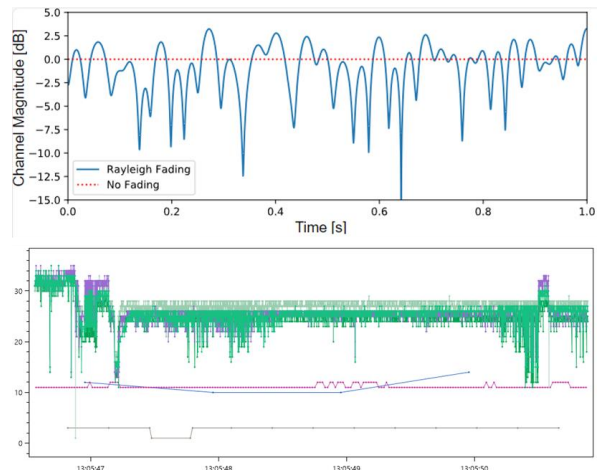
Validation Location 선정에 실제로 고려되어야 할 환경 변수는 다음과 같이 다양하다. 첫 번째로 NW 상위 계층의 RRC(Radio Resource Control) 메시지 이다. 이는 Cellular 단말기 Modem 의 UE Capability information 이나 NW 기지국의 지원 Feature 및 Capability 에 의해 결정되는 환경 변수로, RRC Setup 및 Reconfiguration 메시지 등이 해당되며, 할당된 메시지는 시간이나 Channel 에 의한 영향이 아닌 오직 NW 에

의해서 결정되기 때문에 본 논문에서는 이러한 변수를 그림 1 과 같이 시불변 환경 변수로 가정하도록 한다.



[그림 1] 시불변 환경 변수의 HEX code 화

다음으로는 NW 및 단말기 Modem 이 PDCCH(Physical Downlink Control Channel)/PUCCH(Physical Uplink Control Channel)을 통해 주고 받는 PHY Parameter 들과 NW Channel 에 의한 신호 세기, 간섭, 왜곡과 이동성에서 발생하는 Fading Channel 등 RF 감도 열화가 있다. 이러한 환경 변수 들은 그림 2 와 같이 매 Sub Frame 마다 무작위로 변하기 때문에 불확실성을 가진 시변 환경 변수로 볼 수 있다.



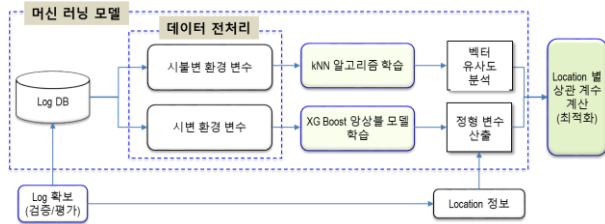
[그림 2] Fading Channel(상)과 Reference signal 및 Shared Channel 신호의 왜곡 모습(하)

본 연구에서는 머신 러닝 기법을 이용하여 Cellular Modem 환경에서 발생하는 다양한 변수를 학습하여 최적화된 Validation Location 을 산출하는 방법에 대해 알아본다.

II. 본론

II-1 학습 데이터 및 학습 방안

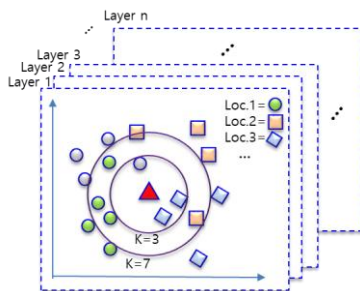
기존의 검증 및 시험 Log 는 머신 러닝 기법의 학습 데이터로 사용된다. 먼저 Log DB 에서 변수의 변화가 적은 시불변 환경 변수와 매 순간 값이 변하는 시변 환경 변수로 데이터를 전처리 한다. 그리고 그림 3 과 같이 각 타입의 환경 변수의 특성에 적합한 머신 러닝 알고리즘을 사용하여 학습을 진행한다.



[그림 3] 머신 러닝을 이용한 Location 최적화 방법

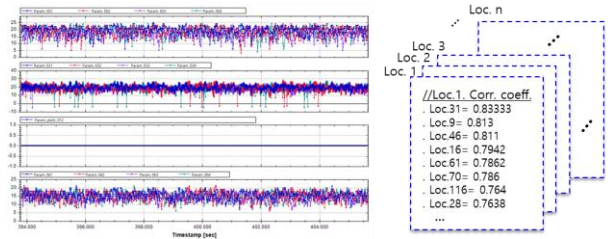
II-2 머신 러닝 모델 선정

시불변 환경 변수는 각 변수에 대해 모든 Location 벡터와의 거리를 비교하여 타겟 벡터와의 유사성을 비교할 수 있는 kNN(k-Nearest Neighbors) 알고리즘을 사용한다[2]. 이 때 유사성의 범위를 지정하는 k 값은 가용한 리소스 등 정황에 맞는 설정이 요구된다.



[그림 4] 다중 Layer kNN 알고리즘 적용

시변 환경 변수는 매 Sub Frame 이나 Block 단위로 각 변수의 값이 무작위로 바뀌기 때문에 데이터의 변화가 극적이고 사이즈가 기하급수 적으로 늘어난다. 회귀 알고리즘 중 XG Boost(Extreme Gradient Boosting) 앙상블 모델은 대용량 데이터 학습에 빠른 성능을 보여주며 과 적합 방지 및 결측 처리가 간단하여 시계열 환경 변수의 학습 및 분석에 적합하다[3]. 학습을 통해 정형화된 변수를 산출하고, 각 Location 변수와의 상관계수를 구하여 유사성을 파악한다.



[그림 5] 정형화 변수 산출 및 Location 별 상관 계수

II-3 최적화 Location 선정

kNN 알고리즘 및 XG Boost 모델을 통해 산출한 Location 간의 유사성을 바탕으로 최적화 Location 을 추출한다. 상관계수가 낮을수록 Location 간 유사성이 낮아져, 최적화 Location 선정에 적합하다. 본 연구에서는 190 개의 Location 에 대한 학습을 통해 환경 변수 간 상관계수를 확인하였다. 이 중 60 개의 Location 은 상관계수가 0.8 수준으로 높게 나타나, 검증 및 평가에서 제외를 고려해 볼 수 있을 것으로 판단된다. 물론, 학습 데이터와 Location 정보에 따라 상관계수와 추출 결과가 달라지기 때문에 본 수치는 하나의 예시임을 고려해야 한다.

III. 결론

기존의 일반적인 Validation Location 선정 방식은 경험 및 리소스에 의존하고, 고려하는 변수가 한정되어 효과적이지 못했다. 하지만, 본 연구의 방법으로 Location 을 최적화 할 경우, 기존 Location 수 대비 더 적은 수로 기존과 동등 수준의 검증 및 평가가 가능할 것으로 보여, 리소스 비용 및 시간 절감에 기여할 것으로 기대된다. 단, 연구한 방법의 Location 최적화는 학습 데이터의 Location 범위 내에서만 이루어진다. 그래서 더 최적화 된 Location 을 산출하기 위해서는 다양한 환경을 포함하는 양질의 Location 변수 학습이 필요하다. 또한 연구 결과를 바탕으로, 실제 필드 적용을 통한 개선 연구도 계속되어야 한다.

ACKNOWLEDGMENT

본 논문은 삼성전자 재원으로 수행된 연구임.

참고 문헌

- [1] Dimitar Minovski. "Throughput Prediction Using Machine Learning in LTE and 5G Networks" Manuscript received 12 October 2020; revised 7 June 2021; accepted 8 July 2021. Date of publication 26 July 2021; date of current version 3 February 2023. (Corresponding author: Dimitar Minovski.) Digital Object Identifier no. 10.1109/TMC.2021.3099397
- [2] Thomas M.; Hart, Peter E. (1967). "Nearest neighbor pattern classification" (PDF). IEEE Transactions on Information Theory. 13 (1): 21– 27
- [3] Tianqi Chen. "XGBoost: A Scalable Tree Boosting System", KDD '16, August 13–17, 2016,