

Non-IID 환경에서 연합 학습 기반 전기 수요 예측

염성웅†, 콜레카르 시바니 산제이†, 조현준*, 김정백†
 전남대학교 인공지능 융합학과†
 전남대학교 정보보안 협동과정*

e-mail : yeomsw0421@gmail.com, shivanikolekar@gmail.com,
 hyjun0522@gmail.com, kyungbaekkim@jnu.ac.kr

Federated Learning-Based Electricity Demand Prediction in Non-IID Environment

Sungwoong Yeom†, Shivani Sanjay Kolekar†, Hyeonjun Jo*, Kyungbaek Kim†
 Department of Artificial Intelligence Convergence, Chonnam National University†
 Department of Interdisciplinary program of Information Security, Chonnam National University*

요약

최근 전기 수요 예측 분야에서 연합학습은 클라이언트의 개인 정보를 중앙 서버와 공유하지 않고 공동으로 글로벌 모델을 학습함으로써 통신 비용 감소 및 개인 정보 보호로 인해 상당한 주목을 받고 있다. 하지만, 데이터의 양과 비선형적 특성이 균등하지 않은 Non-IID 환경에서는 연합학습의 모델이 편향적으로 학습하여 성능 저하를 일으킬 수 있다. 이 논문에서는 Non-IID 문제가 발생하는 전기 소비 예측 환경에서 성능 저하 완화를 위한 결측치 보간 및 버스트 확률 고려 연합학습 기법을 제안한다.

1. 서론

재생 가능한 전력 생산 및 시장의 분산화의 증가에 따라 에너지 분야에서 빌딩 전력 수요 예측 기법은 빌딩 내 전기 수요와 공급 사이의 균형을 유지하고 전력망의 안정적인 부하를 유지하기 위한 필수적인 연구이다 [1]. 최근 빌딩 내 설치된 IoT(Internet of Things) 기기의 수가 증가함에 따라 비선형적으로 관찰되는 전기 수요를 안정적으로 예측하기 위해 시계열 딥러닝 신경망을 사용하는 연구가 활성화되고 있다.

하지만, 일부 빌딩들에서 수집된 전기 소비 이력 데이터가 적어 시계열 딥러닝 신경망을 충분히 학습하지 못할 수 있다. 빌딩 관리자는 충분하지 않은 데이터를 소유한 빌딩들을 위해 데이터 센터에서 각 빌딩의 전기 소비 이력을 수집하고 시계열 딥러닝 신경망을 일반화 후 배포하는 중앙집중식 접근 방식을 채택해왔다 [2]. 그러나 다수의 빌딩이 중앙집중식 관리에 참여한다면 중앙 빌딩 관리시스템에서는 비효율적인 데이터 처리 및 높은 통신 비용 문제가 발생할 수 있다. 또한, 소비자는 개인 정보 침해에 노출될 수 있다.

이러한 중앙집중식 접근 방식의 문제들을 해결하기 위해 중앙 서버가 아닌 로컬 빌딩 전반에 걸쳐 분산하여 학습할 수 있는 연합학습(Federated Learning) 기반 접근 방식이 빌딩 관리시스템에 채택되었다 [3]. 연합학습은 중앙 엔티티의 조정하에 모델 학습에 참여하는 빌딩들이 협력하여 글로벌 모델을 학습하는 분산 머신러닝 기법이다. 연합학습은 전기 이력 데이터 대신 로컬 빌딩에서 업데이트된 모델의 가중치를 중앙 엔티티에 송수신함으로써 개인 정보를 보호하고 연합학습 라운드 수를 조절함으로써 통신 비용을 줄일 수 있다.

하지만, 빌딩 내 IoT 기기 고장 및 데이터 결측으로 인한 데이

터 수 불균형 및 비선형적인 빌딩별 거주자의 행동으로 인한 데이터 특성 비율 불균형과 같은 통계적 이질성이 발생할 수 있다. 이 Non-IID(Not independent and identically distributed) 환경에서 연합학습은 시계열 딥러닝 신경망의 기울기가 편향되지 않게 추정할 필요가 있다. 대표적인 연합학습 기법 FedAvg는 Non-IID 문제 중 데이터 불균형 문제를 고려할 수 있지만, 비선형적 특성 비율에 대한 고려가 충분하지 않을 수 있다.

이 논문에서는 Non-IID 환경에서 결측값 보간 및 비선형성이 관찰되는 성능을 개선하기 위해 STL 분해 기반 버스트 확률 고려 연합학습 기법을 제안한다.

2. 전기 수요 예측을 위한 STL 분해 기반 결측치 보간 및 버스트 확률 고려 연합학습

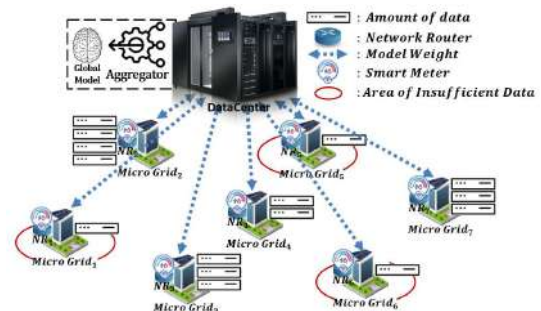


그림 1. Non-IID 환경에서 빌딩별 전기 수요 예측을 위한 연합학습 아키텍처

그림 1은 제안하는 빌딩 전기 수요 예측을 위한 연합학습 아키텍처를 보여준다. 각 빌딩은 스마트 미터와 상대 습도 및 온도를 캡처하는 실내 센서들과 연결되어 전기 소비 이력 및 빌딩 실내 상태 정보를 포함하는 데이터 x 를 1시간 단위로 기록된다. 데이터 x 에서는 빌딩 거주자들의 행동에 따라 계절성 및 버스트와 같은 비선형성이 관찰될 수 있다. 하지만, 빌딩 내 설치된 스마트 미터 및 실내 센서들은 작동 오류로 인해 결측이 발생할 수 있다. 우리는 데이터 x 에서 결측값 보간 및 비선형성 특징 추출을 위해 STL 분해 기법을 사용한다 [4]. STL 분해 기법은 계절성 S , 추세 T , 그리고 잔차 R 를 가산분해 한다. 우리는 데이터로부터 분해된 잔차 R 로부터 결측값 보간 및 버스트 B 추출을 위해 사용한다. 만약 t 번째 데이터 x_t 에서 결측값이 발생하였다면 t 번째 잔차 R_t 는 $-(S_t + T_t)$ 값을 가진다. 이 경우, 우리는 선형적 보간하기 위해 잔차 R_t 를 0으로 대체한다. t 번째 버스트에 대한 수식은 아래와 같다.

$$B_t = \begin{cases} 1, R_t \geq 2 * \sum_{i=1}^N \frac{|R_i|}{N} \\ 0, R_t < 2 * \sum_{i=1}^N \frac{|R_i|}{N} \end{cases} \quad (1)$$

t 번째 데이터 x_t 의 버스트 정도에 시계열 딥러닝 모델의 예측 성능이 저하될 수 있다.

제안하는 전기 수요 예측 기법은 시계열 신경망 모델인 LSTM을 기반으로 전기 수요를 추정한다. 제안하는 모델은 3개의 LSTM 레이어와 3개의 Dense 레이어로 구성된다. 이 세 개의 Dense 계층 중 처음 두 계층은 ReLU(Recified Linear Unit) 활성화 함수를 사용하고 마지막 계층은 선형 활성화 함수를 사용한다. 각 레이어는 10개의 노드를 포함한다. 제안하는 LSTM 모델은 특정 시간부터 9시간 전까지 10개의 연속된 전기 수요 정보를 사용하여 다음 1시간 후에 대한 전기 수요를 예측한다. 관찰된 전기 수요에 대한 예측의 안정성을 보장하기 위해 제안된 LSTM 모델에 대한 입력 벡터를 전기 수요, 실내 온도, 실내 습도, 월별 인덱스, 시별 인덱스, 요일 유형 6가지 도메인으로 구성한다. 제안하는 LSTM 모델은 전기 수요를 예측하기 위해 입력 벡터로 학습한다.

입력 벡터의 첫 번째, 두 번째 및 세 번째 도메인은 전기 수요, 실내 온도 및 실내 상대 습도이다. 각 도메인 특성은 STL 분해를 통해 계절성, 트렌드 및 잔차로 분해하고 이를 입력 벡터로 구성된다. 입력 벡터의 네 번째 및 다섯 번째 도메인은 시별 및 월별 인덱스이다. 월별 인덱스는 1에서 12 사이의 값을 갖고 시별 인덱스는 1에서 24 사이의 값을 갖는다. 마지막으로, 입력 벡터의 여섯 번째 도메인은 요일 유형이다. 요일 유형은 공휴일을 포함한 8가지의 범주를 고려하여 원-핫 인코딩을 통해 요일 특성과 전기 수요 간의 관계를 임베딩한다. 제안한 모델은 연합학습을 통해 학습된다.

연합학습은 각 빌딩의 데이터 샘플 수와 관계없이 모델을 공동으로 학습할 수 있도록 제안되었다. 중앙 서버는 글로벌 모델을 초기화하고 연합학습에 참여하는 빌딩 집합 S 에 포함된 빌딩들에 글로벌 모델의 가중치 w_0^g , 학습률 η 및 배치 크기 b 파라미터를 초기화한다. 각 로컬 빌딩들은 라운드 별로 모델 가중치를 업데이트한다. 이때, 빌딩별로 불균형한 버스트의 수를 고려하기 위해 제안하는 모델의 예측 성능이 달라질 수 있다. i 번째 빌딩의 모델 가중치 업데이트에 대한 수식은 아래와 같다.

$$w_t^{s_i} = w_t^g - \eta \times \sum_{x \in S_i} P_x(B_t = 1) \times \nabla \text{loss}(w; b) \quad (2)$$

i 번째 빌딩에서 학습된 모델 가중치 $w_t^{s_i}$ 는 중앙 서버로 전송된다. 중앙 서버에서는 n 개의 빌딩들로부터 업데이트된 가중치들 $w_t^{s_1}, \dots, w_t^{s_n}$ 을 집계하고 이를 FedAvg[7]를 통해 학습한다.

$$w_t^g = \sum_{s_i \in S} \frac{|N_{s_i}|}{\sum_{s_j \in S} |N_{s_j}|} w_t^{s_i} \quad (3)$$

이때, N_{s_i} 는 i 번째 빌딩에서 수집된 데이터 크기를 의미하며 글로벌 모델의 가중치 w_t^g 를 FedAvg를 통해 평균화한다.

ACKNOWLEDGMENT

“이 논문은 2022년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임”(IITP-2022-0-01203)

참고 문헌

- [1] A. R. Khan, A. Mahmood, A. Safdar, Z. A. Khan and N. A. Khan, “Load forecasting dynamic pricing and DSM in smart grid: A review”, *Renew. Sustain. Energy Rev.*, vol. 54, pp. 1311-1322, Feb. 2016.
- [2] P. Zhang, X. Wu, X. Wang and S. Bi, “Short-term load forecasting based on big data technologies”, *CSEE J. Power Energy Syst.*, vol. 1, no. 3, pp. 59-67, Sep. 2015.
- [3] 염성웅, 이미리내, 쉬바니, 김경백, “전기 수요 예측을 위한 연합학습 통신 비용 트레이드 오프(Communication Cost Trade-off in Federated Learning for Electrical Demand Prediction)”. In *Proceedings of 2022 통신망융합관리(KNOM) 학술대회*, May 12, 2022, 강원대학교, 강원도.
- [4] 염성웅, 김형태, 콜레카르 산자이 시바니, 김경백 “LSTM 기반 멀티스텝 트래픽 예측 기법 평가 (Assessing LSTM-based multi-step traffic prediction methods)”. *KNOM REVIEW*, 24권 2호, pp.13-23, December, 2021.