



RNN-LSTM 기반 공휴일 정보를 고려한 단기 전력수요예측

RNN-LSTM based Short-Term Electricity Demand Forecasting using Holiday Information

저자 (Authors)	김한솔, 송형찬, 고석갑, 이병탁, 신종원 Hansol Kim, Hyungchan Song, Seok-Kap Ko, Byung-Tak Lee, Jong Won Shin
출처 (Source)	대한전자공학회 학술대회 , 2016.11, 552-555 (4 pages)
발행처 (Publisher)	대한전자공학회 THE INSTITUTE OF ELECTRONICS ENGINEERS OF KOREA
URL	http://www.dbpia.co.kr/Article/NODE07071102
APA Style	김한솔, 송형찬, 고석갑, 이병탁, 신종원 (2016). RNN-LSTM 기반 공휴일 정보를 고려한 단기 전력 수요예측. 대한전자공학회 학술대회, 552-555.
이용정보 (Accessed)	서울대학교 147.47.207.*** 2018/08/20 12:42 (KST)

저작권 안내

DBpia에서 제공되는 모든 저작물의 저작권은 원저작자에게 있으며, 누리미디어는 각 저작물의 내용을 보증하거나 책임을 지지 않습니다. 그리고 DBpia에서 제공되는 저작물은 DBpia와 구독계약을 체결한 기관소속 이용자 혹은 해당 저작물의 개별 구매자가 비영리적으로만 이용할 수 있습니다. 그러므로 이에 위반하여 DBpia에서 제공되는 저작물을 복제, 전송 등의 방법으로 무단 이용하는 경우 관련 법령에 따라 민, 형사상의 책임을 질 수 있습니다.

Copyright Information

Copyright of all literary works provided by DBpia belongs to the copyright holder(s) and Nurimedia does not guarantee contents of the literary work or assume responsibility for the same. In addition, the literary works provided by DBpia may only be used by the users affiliated to the institutions which executed a subscription agreement with DBpia or the individual purchasers of the literary work(s) for non-commercial purposes. Therefore, any person who illegally uses the literary works provided by DBpia by means of reproduction or transmission shall assume civil and criminal responsibility according to applicable laws and regulations.

RNN-LSTM 기반 공휴일 정보를 고려한 단기 전력수요예측

*김한솔, 송형찬, **고석갑, 이병탁, *신중원

광주과학기술원 전기전자컴퓨터공학부

e-mail : *hansol@gist.ac.kr, shchan420@gist.ac.kr, jwshin@gist.ac.kr,

**softgear@etri.re.kr, bytelee@etri.re.kr

RNN-LSTM based Short-Term Electricity Demand Forecasting using Holiday Information

*Hansol Kim, Hyungchan Song, and Jong Won Shin

School of Electrical Engineering and Computer Science

Gwangju Institute of Science and Technology

**Seok-Kap Ko, and Byung-Tak Lee

Electronics and Telecommunications Research Institute

Abstract

Daily electricity demand and its fluctuation have increased by abrupt climate change and excessive use of air conditioning and these has affected to forecast the short-term electricity load. Also, the electricity load pattern learning is disturbed by holidays that cause sudden the electricity demand reduction. We proposed the feature extraction algorithm for demand reduction in holidays and implemented the RNN-LSTM (Recurrent Neural Network-Long Short Term Memory) based forecasting. The results were compared with the forecasting performance of SARIMA (Seasonal Auto Regressive Integrated Moving Average). The comparative result shows that RNN-LSTM outperforms SARIMA.

I. 서론

첨단 기술의 발전과 함께 기업과 가정에서의 전력 사용량이 증가하면서 세계적으로 블랙아웃의 위험이 증가하고 있다. 하지만 급격한 기후 변화와 이에 따른 냉·난방 장치의 과도한 사용은 일일 전력 수요량과 변동 폭의 증가를 유발하였고, 이는 정확한 전력수요예측을 어렵게 하는 원인이 되었다. 따라서 기후 변화와 같이 전력 수요에 영향을 미치는 요인들에 대해 강인하고 정확한 단기 예측 모델을 개발하기 위한 연구가

국·내외로 진행되고 있다.

단기 전력수요예측 방법은 크게 통계적 기법과 인공지능형 기법으로 나눌 수 있다[1]. 통계적 기법은 현재까지의 전력 수요량과 관련 정보들(기온 등)의 수학적 조합을 이용하여 미래의 전력 수요량을 예측하는 기법으로 지수평활법, 시계열 분석법 등이 있다. 지수평활법은 현재 시점에 대해 예측된 수요량과 실제 수요량의 가중 평균으로 미래의 수요량을 예측하는 방법으로 시계열 데이터의 추세(trend)와 계절성(seasonality)을 고려한 Holt-Winters 지수평활 모델이 대표적이다[1]. 통계적 기법으로는 자동 회귀 이동 평균법(Auto Regressive Moving Averaging, ARMA)에 시계열 데이터의 비정상성과 추세, 그리고 계절성을 고려한 SARIMA (Seasonal Auto Regressive Integrated Moving Averaging) 모델이 대표적이며, 국내에서는 한국 전체 또는 각 지역별 전력 수요량을 예측하는 연구가 진행되기도 하였다[2][3]. 이러한 기법들은 모델의 구조가 간단하고 해석이 편리한 장점이 있지만 변수간의 비선형적 관계를 표현하는데 한계가 있었다.

인공지능형 기법은 신경망(Neural Network, NN)을 통해 전력 수요량과 다양한 환경적 변수들의 복잡한 관계를 쉽게 모델링해줄 수 있는 장점이 있다. 특히 NN는 최근 컴퓨팅 자원의 발전과 함께 음성 및 이미지 신호 처리 분야에서 많은 성능 향상을 이끌어 왔다. W. He는 심층 신경망(Deep Neural Network, DNN)을 기반의 시간 당 전력수요 예측 모델을 제안하

였고[4], 유승형 등은 SNN (Shallow Neural Network) 과 DNN을 이용해 시간 당 전력수요 예측 성능을 비교하였다[5]. 하지만 DNN 모델은 훈련 과정에서 데이터의 시간 정보를 충분히 고려하지 못하는 단점을 가지고 있다. 이는 모델이 고려할 수 있는 과거 데이터에 제한을 주게 되며, 특히 전력 수요 데이터와 같이 시간에 따라 그 특성이 변하는 시계열 데이터를 처리하기에 한계가 있다.

전력수요예측을 어렵게 하는 원인 중 하나는 공휴일에 발생하는 급격한 수요량 감소 현상이다. 공휴일의 수요량은 과거 같은 요일의 전력 수요량보다 적으며 설날이나 추석 등 연휴 기간에 뚜렷이 나타난다. 이는 예측 오차를 크게 할 뿐만 아니라 정상적인 수요 패턴의 학습을 방해하여 최적의 예측 모델을 얻는 것을 어렵게 한다.

본 논문에서는 데이터의 시간 정보와 모델의 학습 과정에서 데이터의 장기 의존성(long-term dependency)을 고려하기 위해 개발된 신경망인 순환형 신경망(Recurrent Neural Network, RNN) 중 장단기 메모리(Long-Short Term Memory, LSTM)를 이용하여 전력수요 예측을 진행하였다. 이후에는 위 신경망 모델을 RNN-LSTM으로 지칭한다. 또한 공휴일에 수요량이 감소하는 현상을 고려하기 위해 공휴일에 대한 특징 추출 알고리즘을 제안하였다. 제안된 모델의 성능은 통계적 기법 중 하나인 SARIMA 모델과 비교되었다.

II. 본론

2.1 SARIMA 모델

시계열 데이터가 정상성(stationary)을 가질 때, 시간 t 에서의 전력 수요량에 대한 ARMA 모델은
$$X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \dots + \theta_q \epsilon_{t-q}$$
 이고, 이 때, X_t 는 시간 t 에서의 전력 수요량, ϵ_t 는 시간 t 에서의 오차항(교란항), ϕ, θ 는 각각 AR 계수와 MA 계수, p 와 q 는 각각 AR 차수와 MA 차수를 말한다. 위 식을 시간 지연 항(B)을 추가하여 정리하면

$$\phi_p(B)X_t = \theta_q(B)\epsilon_t,$$

$$\phi_p(B) = 1 - \phi_1 B - \dots - \phi_p B^p$$

$$\theta_q(B) = 1 + \theta_1 B + \dots + \theta_q B^q$$

이 된다. 하지만 실제 환경에서 시계열 데이터는 비정상성을 가지게 되며, 모델에 적용시키기 위해 차분(differencing)을 통해 제거된다. 이 모델을 ARIMA(p, d, q)라고 부르며, 다음과 같이 표현된다.

$$\phi_p(B)(1-B)^d X_t = \theta_q(B)\epsilon_t,$$

d 는 차분 차수이다. 또한 시계열 데이터는 계절성이라 불리는 주기적인 패턴을 보이며, 이를 고려한 모델은

$$\phi_p(B)\Phi_{p'}(B^s)(1-B)^d(1-B^s)^{d'}X_t = \theta_q(B)\Theta_{q'}(B^s)\epsilon_t$$

$$\Phi_{p'}(B^s) = 1 - \Phi_1 B^s - \dots - \Phi_{p'} B^{p's}$$

$$\Theta_{q'}(B^s) = 1 - \Theta_1 B^s - \dots - \Theta_{q'} B^{q's}$$

으로 표현되고, Seasonal ARIMA(p, d, q)(p', d', q')[m]라고 불린다. 위 식에서 s 는 seasonal 계수, p', d', q' 은 각각 seasonal AR, 차분, MA 차수이다. 본 실험에서 SARIMA 모델의 계수와 차수는 경험적으로 결정되었으며, 모델의 파라미터($\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q, \Phi_1, \dots, \Phi_{p'}, \Theta_1, \dots, \Theta_{q'}$)는 최대 우도 추정(Maximum Likelihood Estimator, MLE)을 통해 얻어졌다.

2.2 RNN-LSTM based 모델

DNN는 데이터 간의 비선형적 관계를 손쉽게 표현해 줄 수 있는 장점을 가지고 있지만 입력 데이터 간의 관계를 독립적으로 가정하기 때문에 시간에 따른 정보를 학습시키는데 한계가 있다. RNN은 각 은닉층에 저장된 특정 시점의 데이터 정보를 다음 시점으로 전달하도록 개발된 모델로 데이터 간의 비선형적 관계 뿐만 아니라 시간 정보까지도 고려할 수 있는 모델이다. 하지만 현실적으로 RNN의 학습 과정에서 기울기 소실 문제(vanishing gradient problem)가 발생하여 데이터의 장기 의존성(long-term dependency)을 고려하는데 한계가 있다. 본 연구에서는 데이터의 장기 의존성을 고려한 RNN의 변형 모델인 RNN-LSTM을 사용하였다.

전력수요예측을 위해 제안된 모델은 RNN-LSTM과 선형 회귀(linear regression) 모델이 직렬 연결된 구조이다. 입력 데이터는 RNN-LSTM의 각 시간 단계(time step)에 입력되어 처리되고, 출력값들은 선형 회귀(linear regression) 모델에 의해 향후 1일의 전력수요를 예측하게 된다. 예측 모델에 사용된 비용 함수(cost function)로는 예측값과 실제 수요량과의 평균 제곱 오차(mean square error)를 사용하였다. 예측 모델의 학습은 크게 선형 회귀 모델의 학습과 RNN-LSTM의 학습으로 나뉜다. 선형 회귀 모델의 학습은 역전파(back-propagation) 알고리즘으로, RNN-LSTM 모델은 시간에 따른 역전파(back-propagation through time) 알고리즘으로 각각 학습되었다.

2.3 공휴일에 대한 특징 추출 알고리즘

설날, 추석, 선거일 등, 공휴일에는 일반일보다 전력수요량이 급격히 감소하는 현상을 보인다. 이는 데이터의 특이성처럼 작용하여 정확한 시계열 데이터의 패

		전체 학습 수					
		100,000	200,000	300,000	400,000	500,000	600,000
조합	입력차원(7(3일)) 고려기간(3주)	2.391	2.261	2.229	2.189	2.227	2.238
		153.131	144.327	142.312	139.266	141.925	142.302
	입력차원(9(4일)) 고려기간(1주)	2.221	2.161	2.133	2.135	2.163	2.143
		142.790	138.599	136.006	135.956	138.153	137.059

표 2. 표 1에서 가장 좋은 성능을 보인 두 가지 조합("RNN-LSTM 각 시간 단계 당 입력차원 : 7(일), 특징 추출을 위해 고려된 기간 : 4주", "RNN-LSTM 각 시간 단계 당 입력차원 : 9(일), 특징 추출을 위해 고려된 기간 : 1주")에 대한 전체 학습 수에 따른 전력수요예측 성능(위 : MAPE, 아래 : MAE)

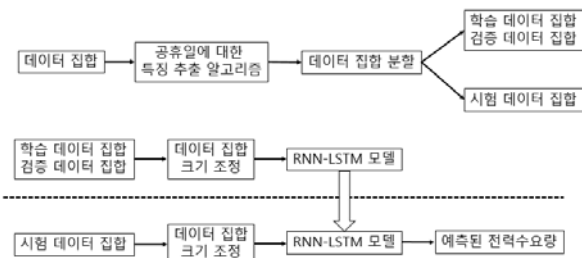


그림 1. 전력수요 데이터의 공휴일에 대한 특징 추출 과정과 RNN-LSTM 모델의 학습과 시험을 위한 데이터 집합 분할 과정(위), 학습 및 검증 데이터 집합을 이용한 RNN-LSTM 모델의 학습 과정과 학습된 모델과 시험 데이터 집합을 이용한 향후 전력수요 예측 과정(아래)

터이 학습되는 것을 어렵게 한다. 본 연구에서는 공휴일에 발생한 전력 수요 감소의 특징을 이전 연도의 동일한 공휴일의 전력 수요량과 그 공휴일의 과거 일정 기간 동안 같은 요일의 평균 전력 수요량의 차이로 정의하였다. 그림 1은 RNN-LSTM 모델을 이용한 전력수요 예측 과정의 전체 구조도이다.

III. 구현

실험에서는 SARIMA 모델과 RNN-LSTM 모델을 각각 사용하여 향후 1일씩, 약 2년 동안 전력수요 예측을 진행하였다. 사용한 전력수요 데이터는 2005년에서 2014년까지의 한국의 일일 전력 수요량으로 전체 데이터의 80%(05년~12년)와 20%(13년~14년)가 각각 학습 데이터 집합(training set)과 시험 데이터 집합(testing set)을 만들기 위해 사용되었다. 학습 및 시험 데이터 집합의 한 원소에는 RNN-LSTM의 시간 단계 개수만큼의 데이터 소집합이 존재한다. 각 소집합은 RNN-LSTM의 시간 단계에 들어가는 데이터를 말하며, 시간에 따라 1일씩 옮겨진 전력수요 데이터가 들어간다. 시험 데이터 집합의 평균은 6441.652, 표준편차는 660.043이다. 또한 학습 데이터 집합 중 20%는 검증 데이터 집합(validation set)으로 사용되었다. 모든 데이터는 크기 조정(scaling)을 위해 0.001이 곱해졌다. 공휴일의 특징은 이전 연도의 동일한 공휴일에 대해 과거 몇 주를

		공휴일 특징 추출을 위해 고려된 과거 기간			
		1주	2주	3주	4주
RNN-LSTM 각 시간 단계 당 입력차원	3 (1일)	2.581	2.215	2.318	2.342
		163.712	141.572	147.322	148.769
	5 (2일)	2.269	2.330	2.279	2.306
		143.696	147.681	144.272	146.580
	7 (3일)	2.490	2.327	2.227	2.344
		158.946	148.849	141.92	148.962
	9 (4일)	2.163	2.253	2.251	2.281
		138.153	143.509	143.994	145.415

표 1. RNN-LSTM의 각 시간 단계 당 입력 차원(예측을 위해 고려되는 과거 일수)과 공휴일 특징 추출을 위해 사용된 과거 기간에 따른 전력수요예측 성능(위 : MAPE, 아래 : MAE)

MAPE	2.296
MAE	1706.631

표 3. ARIMA(8,1,1)(0,1,1)[7] 모델을 이용하여 2014년에서 2015년까지 향후 1일씩 전력수요 예측

고려하느냐에 따라 성능 차이가 있었다. 실험에서는 과거 1주, 2주, 3주, 4주에 따라 성능이 비교되었다. RNN-LSTM의 각 시간 단계에 입력되는 데이터의 일의 특수일 특징이 순차적으로 조합된 형태이다. 실험에서 RNN-LSTM의 시간 단계는 7을 사용했으며, 각 시간 단계에서 입력의 차원은 고려되는 과거 일수에 따라 각각 3(1일), 5(2일), 7(3일), 9(4일)로 설정되었다. 모델의 over-fitting 문제를 고려하기 위해 전체 학습 횟수는 100,000부터 600,000까지 10만씩 증가시키며 시행되었다. 1회 학습은 mini-batch 크기가 1일 때, RNN-LSTM 모델이 한번 업데이트 되는 것을 의미한다. 본 실험에서는 별도의 최적화 과정을 통해 학습률(learning rate)은 0.1, mini-batch 크기는 200을 사용하였다. 성능 평가는 평균 절대 비율 오차(Mean Absolute Percentage Error, MAPE)와 최대 절대 오차(Maximum Absolute Error, MAE)를 사용하였다. 표 1은 RNN-LSTM 모델의 각 시간 단계에서의 입력 차원과 공휴일의 특징 추출에서 고려되는 과거 데이터 개수에 따른 예측 성능을 보여준다. 표 2는 표 1의 실험 결과 중 가장 좋은 성능을 보인 두 가지 조합에 대해 전체 학습 횟수에 따른 예측 성능이다.

SARIMA 모델은 예측일 이전 1095일(3년)의 데이터를 이용하여 매 시점마다 새로 학습되며, ARIMA(8,1,1)(0,1,1)[7] 모델을 사용하였다. 모델의 최적화

과정은 본 논문에서는 생략하였다. 표 3은 ARIMA (8,1,1)(0,1,1)[7]를 이용한 전력수요 예측 성능이다.

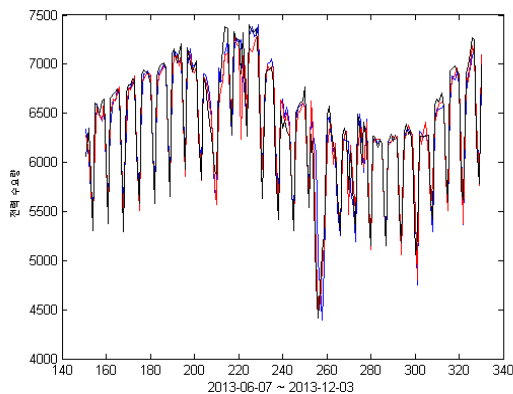


그림 2. 2013년 06월 07일부터 12월 03일까지 180일 동안 RNN-LSTM과 SARIMA를 이용해 전력 수요량을 예측한 그래프(검정 : 실제 수요량, 빨강 : RNN-LSTM, 파랑 : SARIMA)

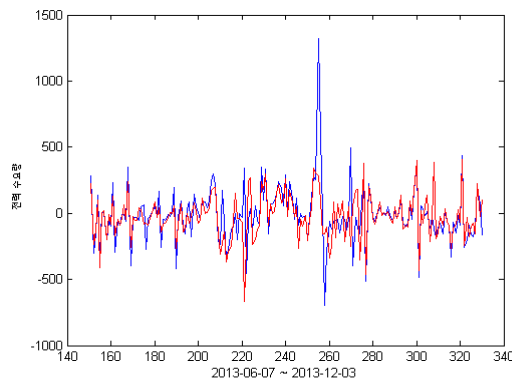


그림 3. 2013년 06월 07일부터 12월 03일까지 180일 동안의 실제 수요량과 두 모델의 예측치의 오차(빨강 : RNN-LSTM, 파랑 : SARIMA)

그림 2는 2013년 06월 07일부터 12월 03일까지 실제 수요량과 RNN-LSTM 모델과 SARIMA 모델로 수요량을 예측한 값의 그림이며, 각 모델의 예측오차를 그림 3에 나타내었다. 실험 결과, RNN-LSTM 모델의 각 시간 단계 당 입력 차원은 9(4일), 공휴일 특징 추출을 위해 고려되어야 하는 과거 기간은 4주, 전체 총 학습 횟수를 300,000회 진행했을 때의 성능이 가장 좋았다. 그림 3에서 양의 값은 예측 모델이 실제 수요량보다 더 크게 예측한 상황을 말하며 시설들에 대한 과잉 투자를 야기하게 된다. 음의 값은 실제 수요량보다 더 적게 예측한 상황을 말하며 블랙아웃의 원인이 된다. 그림 3은 RNN-LSTM 모델이 SARIMA 모델보다 더 정확한 예측을 하고 있음을 보여준다. 또한

RNN-LSTM 모델이 SARIMA 모델보다 공휴일(위 그림에서는 광복절(220번째)과 추석(254~256번째))에서 비교적 안정적으로 예측을 하고 있음을 알 수 있다.

IV. 결론 및 향후 연구 방향

최근 전력 사용량이 급증하면서 여름철과 겨울철에 전력 부족으로 인한 블랙아웃 현상이 사회적 문제로 부각되고 있으며, 이에 따라 정확한 단기 전력수요 예측의 중요도가 높아지고 있다. 기존에 주로 사용되던 기법인 SARIMA 모델은 공휴일 등 갑작스러운 전력 수요의 변동을 잘 추적하지 못하는 문제점을 안고 있었다. 본 연구에서는 RNN-LSTM 모델을 이용하여 시계열 데이터 사이의 비선형적 관계를 모델링하였고, 전력 수요 패턴에 크게 영향을 끼치는 공휴일의 특징을 추출하여 갑작스러운 전력 수요량의 변화에도 강인하게 추적할 수 있는 모델을 제안하였다. 제안된 모델은 특히 설날, 추석 등과 같은 연휴를 가지는 공휴일들에 대해 강인한 예측 성능을 보였다. 하지만 광복절, 어린이날 등의 공휴일에서는 아직 큰 오차가 발생하고 있으며 이에 대한 개선이 필요하다, 또한 기온, 강수량 등의 요소들을 고려한 연구도 진행되어야 할 것으로 보인다.

Acknowledgment

This work was supported by ETRI R&D Program (15YI1200), funded by the government of Korea.

참고문헌

- [1] 김철현, “다중 계절성 지수평활법을 활용한 국내 단기 전력수요 예측,” 에너지경제연구원, 2013.
- [2] 김시연, 정현우, 박정도, 백승묵, 김우선, 전경희, 송경빈, “계절 ARIMA 모형을 이용한 104주 주간 최대 전력수요예측,” 조명·전기설비학회논문지, vol. 28, No. 1, pp. 50-56, 2014.
- [3] 안병훈, 최희련, 이흥철, “계절 ARIMA 모형을 이용한 국내 지역별 전력사용량 증장기수요예측,” 한국산학기술학회논문지, vol. 16, No. 12, 2015.
- [4] W. He, “Deep neural network based load forecast,” Computer modeling and New technologies, vol. 18, No. 3, pp.258 - 262, 2014.
- [5] 유승형, 노재구, 김홍석, “심층신경망 기반 전력수요예측 모델에 대한 연구,” 한국통신학회 동계종합 학술발표회, 2016.