과거이력기반 일일전력사용량 예측 알고리즘

김민경[†], 주용재[*], 노재구[*], 최준균[†]
[†]한국과학기술원, [*]한전 전력연구원,
e-mail : mkkim1778@kaist.ac.kr, yongjae.joo@kepco.co.kr,
jknoh@kepco.co.kr, jkchoi59@kaist.edu

A Day-ahead Electricity Forecasting Algorithm based on Historical Profiles

Minkyung Kim[†], Yongjae Joo[*], Jaegu Noh[*], Jun-Kyun Choi[†]
[†]KAIST, [*]KEPCO Research Institute

요 약

AMI 시스템의 도입으로 스마트미터로부터 얻은 계량데이터를 활용한 다양한 고객 서비스가 시행되고 있으며, 고객별 전력 사용량 예측 서비스는 그 중 하나이다. 고객에게 예측 서비스를 제공하기 위해서는 정확하면서도 신속하게 방대한 데이터에서 예측 값을 도출해야 할 것이다. 이러한 필요성에 따라 본 논문에서는 과거이력기반 일일전력사용량 예측 알고리즘을 제안하였고, 가법적 계절지수평활법(Additive Holt-Winters Method)과 비교하여 알고리즘의 정확도 및 예측 소요시간을 검증하였다.

1. 서론

AMI(Advanced Metering Infrastructure) 시스템의 도입으로 스마트미터 계량데이터를 활용한 다양한 고객 서비스가 제공되고 있으며, 그 일환으로 고객별 전력 사용량 예측 서비스를 들 수 있다. 고객이 원하는 시점 언제든지 신속하고 정확한 예측 서비스를 제공해야 할 것이므로, 전력 사용량 예측 서비스에서는 예측의 정확성과 예측 소요시간이 중요한 관점이 될 것이다.

본 논문에서는 주어진 시계열 데이터에서 데이터의 특성으로 모델을 만들고 그 모델을 통해 예측을 해 나가는 기존 기법에서 탈피하여, 가지고 있는 과거 데이터를 적절히 조합하는 방법으로 위 두 요건을 만족시켜 보려 한다.

본 논문은 AMI 기반 서비스 환경에서의 기존 예측 기법에 대한 고찰, 제시 알고리즘의 개요, 실 데이터를 사용한 예측의 정확성 및 소요시간 검증으로 구성된다.

2. 기존 시계열 예측기법 고찰

스마트미터로부터 수집된 계량데이터는 시계열을 이룬다. 기존 시계열 예측 기법에는 평활법과 같은 전통적 시계열분석방법(표 1)과 ARIMA와 같은 확률적 시계열분석방법이 있다. 이러한 기존 기법들은 과거 데이터를 이용하여 시계열의 변동 요인을 찾고, 변동 요인을 바탕으로 모델을 만들어 이를 예측에 이용한다. 이 때 과거의 패턴이 미래에도 계속 될 것이라는 가정은 필수적이다. 모델 생성에 사용되는 시계열의 변동 요인에는 추세 변동, 계절 변동, 순환 변동, 불규칙 변동이 있으며, 변동의 특성에 따라 적절한 모델을 생성해야 할 것이다.

(표 1) 기존 시계열 분석 기법 예시[1]

전통적 시계열 분석 기법	
단순이동평균법	불규칙 변동만을 포함하는 시계열
$F_{n+1}=\frac{1}{m}\sum_{t=n-m+1}^{n} Z_t$	
선형이동평균법	선형 추세를 갖는 시계열
$F_{n+l}=a_n+b_n l$ $=(2MA_n-MA_n')+\frac{2}{m-1}(MA_n-MA_n')l$ $MA_n=\frac{1}{m}\sum_{t=n-m+1}^{n} Z_t$ $MA_n'=\frac{1}{m}\sum_{t=n-m+1}^{n} MA_t$	
이차지수평활법	이차 곡선 형태의 추세를 갖는 시계열
$F_{n+1}=a_n+b_n l+\frac{1}{2}c_n l^2$ $a_n=3SM_n-3SM_n'+SM_{n-1}''$ $b_n=\frac{\alpha}{2(1-\alpha)^2}[(6-5\alpha)SM_n-(10-8\alpha)SM_n'+(4-3\alpha)SM_n'']$ $c_n=\frac{\alpha^2}{(1-\alpha)^2}(SM_n-2SM_n'+SM_n'')$ $SM_n=\alpha Z_n+(1-\alpha)SM_{n-1}$ $SM_n'=\alpha SM_n+(1-\alpha)SM_{n-1}'$	
가법적 계절지수평활법	선형 추세성, 계절 변동을 갖는 시계열
$F_{n+l}=a_n+b_n l+S_{n+l-L},\ 1\le l\le L$ $a_n=\alpha(Z_n-S_{n-L})+(1-\alpha)(a_{n-1}+b_{n-1})$: 수평성분 $b_n=\beta(a_n-a_{n-1})+(1-\beta)b_{n-1}$: 추세성분 $S_n=\gamma(Z_n-a_n)+(1-\gamma)S_{n-L}$: 계절성분	
Z_n : 시점 n의 실제값/ F_n : 시점n의 예측값/ m : 반영 데이터 수 MA_n : 이동 평균/ MA_n' : 이동평균의 평균 SM_n : 지수 평활값/ SM_n' : 이중 지수 평활값/ α : 평활 상수 L : 계절성의 길이/ S_n : 계절 인자/ β,γ : 평활 상수	

하지만 스마트미터로부터 고객 당 15분 간격으로 모이는 방대한 양의 데이터를 가지고, 수많은 고객이 원하는 시점에서 신속하게 예측 사용량을 제공해야 한다는 점에 주목해보자. 또 과거의 패턴이 미래에 계속 될 것이라는 가정 하에 예측을 하게 되지만, 과거 한 시점에서 만들어진 모델로부터 모든 예측을 할 수는 없다. 최근의 패턴을 반영하기 위해 모델