

인터벌 데이터를 이용한 수요그룹 분류 연구

유인협*, 이진기*, 고종민*, 김선익*

*한전 전력연구원

e-mail:{ihyu, jklee, kojim, sikim}@kepri.re.kr

Classification of Electricity Demand Groups using Interval Data

In Hyeob Yu*, Jin Ki Lee*, Jong Min Ko* Sun Ic Kim*

*Korea Electric Power Research Institute

요 약

최근에 들어서 전력산업에 규제완화가 도입되면서 환경이 급변하고 있는 실정이다. 여러 가지의 환경 변화가 예상되지만, 그 중에서도 공급자간에 경쟁 도입이 전력산업 참여자간에 주요 이슈로 부상하고 있다. 이와 같은 변화는 전력시스템의 기술 개발 뿐만 아니라 경영전략에도 큰 영향을 미치고 있으며, 대 수요자 서비스의 제공이 전략의 핵심이 되고 있다. 따라서 공급자는 보다 나은 서비스를 제공하기 위해서, 수요자 정보의 수집 및 분석을 해야 할 필요가 있다. 이와 같은 수요자 정보의 분석은 여러 분야가 있지만 그 중에서도 수요특성을 파악하는 것이 가장 기본이 된다. 본 논문에서는 전력 수요자의 부하 특성을 분석하고 평가하기 위하여 수요특성별로 그룹으로 분류하는 방법을 개발하고, 분류된 그룹의 특징을 검토하였다.

1. 서론

전력사 또는 전력 수요자는 부하의 특성을 분석하고, 그 결과를 이용하여 수요관리, 수요예측 및 계획, 전력 부가 서비스, 가격 관련 설계 사항 등 다양한 용도로 활용할 수 있다. 이와 같은 부하 분석을 효과적으로 하기 위해서는 전력 수요자를 그룹으로 분류하고 각 그룹의 특성을 나타낼 수 있는 표준 모델의 도출하여 분석하는 것이 필수적이다. 이와 같은 표준 부하 모델을 구하기 위해서는 고객의 전력 수요정보가 필요하다. 전력 수요정보는 크게 구분하여 요금 정보와 사용량 정보로 나누어진다. 전통적인 부하 분석은 기계식 전력량계에 의해 수집된 요금정보에 포함되는 월간 전력수요를 이용하여 수행되어 왔으나, 이는 수요의 특성을 구분하기에는 불충분한 정보였다. 최근 들어 전자식 전력량계의 설치가 확산되어 Load Profile 혹은 Interval Data라고 불리는 15분 간격으로 된 수요정보의 수집이 가능하게 되었다. 따라서 전자식 전력량계에 의해 취득된 검침정보를 분석하여 전력수요의 특성을 파악

하는 것은 부하의 분석 등에 있어서 기본 요소가 된다. 본 논문에서는 전자식 전력량계로부터 수집된 Load Profile을 이용하여 수요자 그룹을 분류하고, 분류된 그룹에 대한 구성 고객의 특성을 분석하였으며 이를 소개한다.

2. 수요그룹 분류

1. 데이터 전처리

일반 대용량고객의 전력설비에 설치되어 있는 전자식 전력량계는 15분 주기의 수요정보를 자체 내에 저장하고 있으며, 원격 검침 시스템(AMR)에서 1 시간 간격으로 그 정보를 수집하고 있다. 현재 AMR 시스템에는 약 10만호의 고객이 포함되어 있으며 Load Profile 데이터가 수집되고 있다. 이 정보를 이용하여 분석을 하려면, 그 용량이 아주 커서 대용량 데이터 처리 기술이 필요하다. 또한 이와 같은 정보들은 오류 및 누락 정보가 다수 포함되어 있다. 이들에 대한 주요 원인은 계량기로부터 데이터를 수집하는 서버까지의 경로 상에서 발생하는 통신상의

오류, 수집된 정보의 변환 과정 그리고 국내의 각 지역별 서버에서 중앙서버로 전송되는 과정에서 오류 등으로 구분되고 있다. 또한 매일 빈번히 일어나는 고객 정보의 변화에 대한 데이터베이스 교체작업이 시차를 가지는 등의 문제도 포함되고 있다. 이상과 같은 누락 또는 오류 정보를 처리하기 위한 방안을 설정하였다. 이를 위한 주요사항은 다음과 같다. 월별, 계절별, 연간 등의 분석기간내의 누락 정보가 20%이상일 경우에는 해당 고객을 대상에서 제외한다. 그리고 일간 데이터인 96개의 수요 값 중에서 부분적으로 누락된 정보는 경계 값을 이용한 선형 증감으로 가정하여 보정을 한다. 그리고 데이터 중에서 나타나는 돌출 값 중 평균의 $\pm 3\sigma$ 를 초과하면 역시 누락정보에 포함시킨다...

2. 분류 알고리즘 개발

고객의 수요정보인 Load Profile은 15분 간격의 수요를 나타내는 일간 96개의 값으로 이루어져 있으며, 분류를 위해서는 알고리즘에 적용하기 위해 사전에 변수들의 정의가 필요하다. 우선 각 고객들의 일간 Load Profile 데이터를 분석 기간 내에서 평균 값을 구한다. 이렇게 하여 특정 구간에 대한 각 고객의 대표적인 수요정보를 추출하게 되며 이를 일간 Load Diagram이라 한다. 수요자들을 Class들로 구분하기 위하여 일간 Load Diagram에 관련된 특정 인덱스를 정의하고 이를 이용하여 구분한다. 다음은 분류과정을 간단히 설명한다. 먼저 Load Diagram의 집합을 $l = \{l_h \quad h=1, \dots, H\}$ 이라 정의한다. 그리고 M 고객을 가진 특정그룹의 일간 Load Diagram을 $L = \{l^{(m)} \quad m=1, \dots, M\}$ 이라 한다. M Customer를 K 그룹으로 분류하는 과정을 고려하면, L의 Subset인 $L^{(k)}$ 는 $n^{(k)} \quad k=1, \dots, K$ 개의 Load Diagram들을 포함한다. 각 Class에 대해 대표적인 Load Diagram $r^{(k)}$ 는 $L^{(k)}$ 내에 있는 Load Diagram들의 평균으로 구한다. 따라서 각 Class들의 대표적 Load Diagram의 집합인 $R = \{r^{(k)}, k=1, \dots, K\}$ 이 생성된다. 각 그룹을 구분하기 척도로 Diagram들 간의 거리를 사용하므로 아래의 거리들에 대한 정의가 필요하다.

가. 두 Load Diagram 사이의 거리

$$d(l^{(i)}, l^{(j)}) = \sqrt{\frac{1}{H} \sum_{h=1}^H (l_h^{(i)} - l_h^{(j)})^2}$$

나. 분류된 특정 그룹의 대표 Load Diagram과 $L^{(k)}$ 사이의 거리

$$d(r^{(k)}, L^{(k)}) = \sqrt{\frac{1}{n^{(k)}} \sum_{m=1}^{n^{(k)}} d^2(r^{(k)}, l^{(m)})}$$

다. L 내의 $l^{(i)}$ Diagram과 고정 Diagram과의 거리

$$d(l^{(i)}, c) = \sqrt{\frac{1}{H} \sum_{h=1}^H (l_h^{(i)} - c)^2}$$

이 값이 최소일 때에 Squared Mean Error 면에서 최적의 Interpolation이 됨

라. L과 M 구성원사이의 기하 평균거리

$$\hat{d}(L) = \sqrt{\frac{1}{2M} \sum_{m=1}^M d^2(l^{(m)} - L)}$$

그리고 다수의 데이터로 이루어진 Load Diagram의 특성을 추출하기 위해서는 인덱스가 필요하며, 이들을 보면 Diagram의 최대, 최소, 평균값들로 이루어진 Non-Uniformity Coefficient, Fill-up Coefficient와 피크 및 비 피크 시의 기여도를 나타내는 Modulation Coefficient 들이 있다. 본 알고리즘에서는 분류그룹의 수에 대한 초기화가 필요 없으며, 또한 각 그룹의 센터도 자동으로 찾아낸다. 각 Diagram에 적용하는 인덱스는 Weighting Factor가 부가되었다. 사용된 Weighting Factor는 $\sigma_j^2 / \bar{\sigma}^2$ 로 정의되며, 여기서 σ_j^2 는 jth 인덱스의 표준편차이고, $\bar{\sigma}^2$ 는 각 인덱스 표준편차의 평균값이다. 표준편차가 큰 값을 가진 인덱스는 Weighting Factor에 의해 영향이 확대가 된다. 그림 1은 그룹분류의 과정을 flow chart로 나타내고 있다. 우선 각 고객의 diagram들에 대해서 순차적으로 위에 정의된 인덱스를 산출하여 특성 벡터를 형성한다. 이들을 이용하여 각 그룹과 거리를 산출한다. 이중 최소 거리를 선정하여 거리의 threshold값과 비교한다. Threshold 값보다 클 경우에는 새로운 그룹을 형성한다. 그리고 threshold 값보다 작은 경우에는 그 최소거리에 해당하는 그룹으로 포함되게 된다. 따라서 해당 그룹은 그룹 내의 요소가 증가되었으므로 이에 따른 그룹 중심을 변경하게 된다. 그리하여 그룹수와 각 그룹에 속하는 패턴을 선정하게 된다. 다음에 각 그룹의 중심이 변경됨으로써 각 그룹간의 거리도 변하게 됨으로 처음 과정에서 정해진 그룹들에 대하여 다시 거리인덱스의 threshold 값을 비교함으로써 하여 그룹요소들의 그룹변경 여부를 확인하게 된다. 그리하여 분류과정을 종료하게 된다.

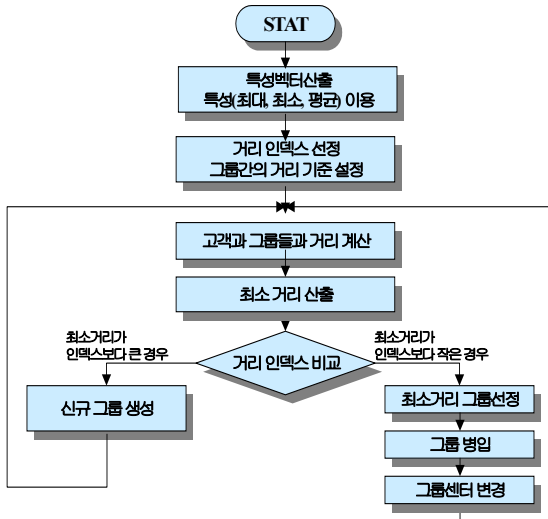


그림 1. 분류 알고리즘

3. 분류그룹 분석

본 연구에 사용된 데이터는 국내에서 운영되고 있는 원격 검침 시스템의 사용량 및 고객정보이다. 그리고 현재 약 10만호 정도의 고객이 원격 검침 시스템에 포함되어 있다. 본 분류시험에는 샘플링이 없이 그 시스템에 포함된 전 고객을 대상으로 하였다. 이상의 데이터를 이용하여 2003.07에 대해 그룹분류시험을 하였으며 평일에 대한 수요패턴의 분류 결과가 그림 2에서 나타나고 있다. 그림에서는 분류 그룹이 8개인 경우를 표시하고 있다. 그림에서 가로축은 일간 패턴을 나타내므로 시각을 나타내고 있으며 세로축은 패턴 비교를 위해서 15분 주기의 수요 값이 0과 1 사이의 값으로 변환되어 있음을 나타낸다. 그리고 각 그룹에 속하는 수요자들의 수가 그룹명과 함께 범례를 표시하는 상자에 나타나 있다. 패턴의 분류 결과를 보면 각 그룹간의 특징이 잘 나타나고 있음을 알 수 있다.

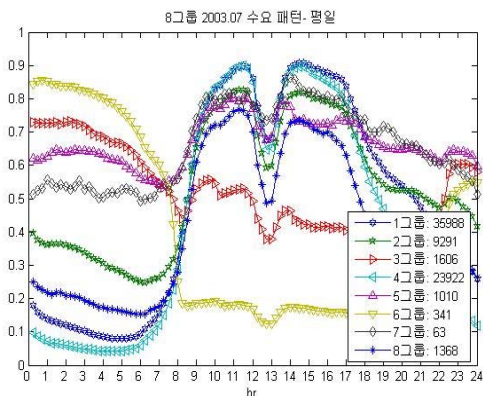


그림 2. 2003.07 평일의 수요 패턴

8개의 그룹을 보면 크게 3 종류로 볼 수 있다. 1, 2, 4, 8 그룹들은 주간에는 수요가 높은 분류에 속하고 3, 5, 7 그룹들은 주야간에 별 차이가 없는 분류에 속하고 6그룹은 심야 전력을 사용하는 분류를 나타낸다.

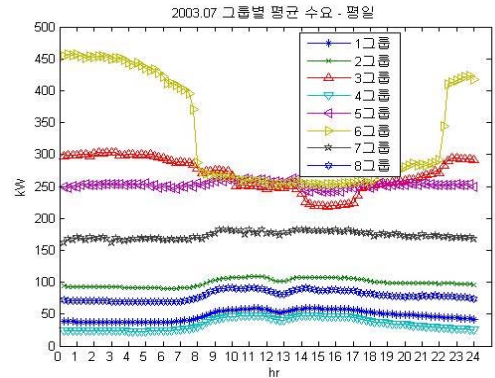


그림 3. 2003.07 평일의 그룹별 평균수요

그림 3은 분류된 8개 그룹에 대하여 각 그룹에 속하는 고객들의 평균 수요를 나타내고 있다. 여기서는 평균 수요의 비교를 나타내기 위해 8개의 그룹을 동시에 나타내고 있으며, 각 그룹의 실제적인 시간대별 수요를 나타내고 있다. 그러나 이 그림에서는 심야 전력그룹을 제외하고는 각 그룹의 수요크기의 형태 구분이 어렵다. 그래서 그림 4에서 그룹1의 평균 수요만을 나타내고 있다.

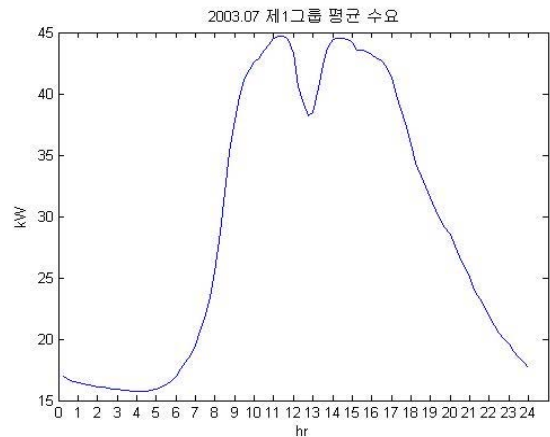


그림 4. 그룹1의 2003.07의 평균 수요

그림 4는 35988 고객으로 이루어진 그룹1의 계약전력의 구성비를 나타내고 있다. 100-500 kW의 고객이 70%를 차지하고 있는 것으로 나타나 있으며 전형적인 고압 고객의 수요패턴으로 평가할 수 있다.

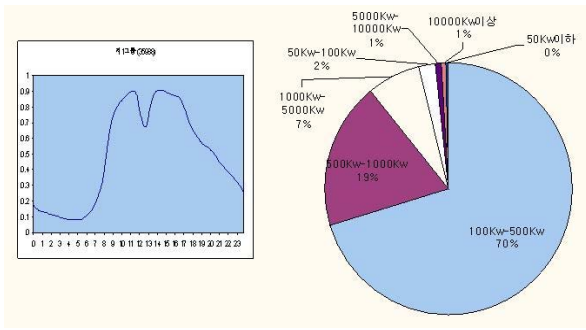


그림 4. 그룹1의 계약전력별 구성비

그림 5는 그룹2에 속한 수요자들의 산업분류별 구성비를 나타내고 있다. 패턴에서의 구분을 그룹1과 비교하면 이 그룹에 속하는 업종은 일과 후의 수요가 그룹 1에 비해 높은 것을 알 수가 있다. 그리고 포함된 산업분류가 다양하여 산업분류별에 의한 수요패턴의 구분은 어렵다는 것을 나타내고 있다.

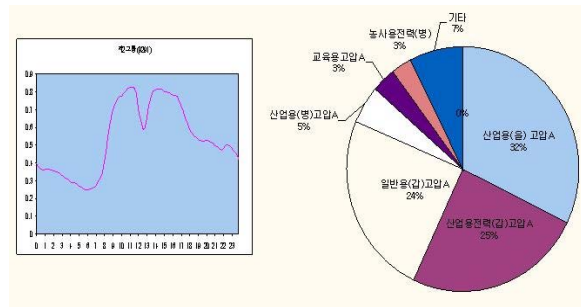


그림 5. 그룹2의 산업분류별 구성비

그리고 그림 6에서는 각 그룹별 2003년 7월에 대한 실제 수요의 일 평균값을 누적하여 나뉜 것이다. 따라서 각 그룹의 면적은 일간 사용량이 되며 시간대별 총 사용에 대한 기여 분을 나타낸다. 따라서 이 자료는 시간대별로 피크 기여도 분석 등의 자료로 유용하게 사용될 수 있다.

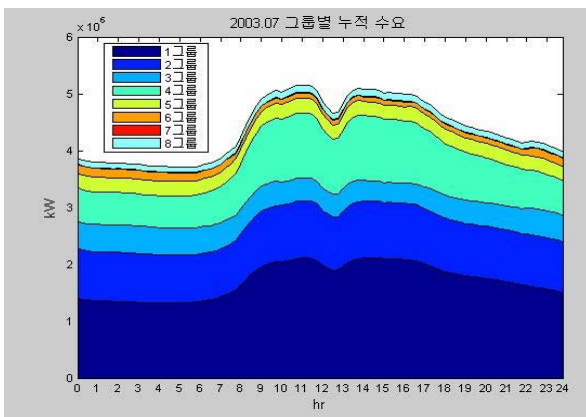


그림 6. 2003.07 평일의 그룹별 누적 수요

4. 결 론

전통적으로 수요자 정보의 분석은 월간 전력사용량을 사용하여 통계적인 분석 수준에 그치고 있었다. 본 논문에서는 AMR 시스템의 정보를 이용하여 보다 실제적인 평가를 할 수 있는 기반을 구축하기 위하여 원격검침을 시행하는 전 고압 수요자의 정보를 이용하여 시간대별 수요에 대한 패턴을 구분하였다. 여기서 개발한 분류 알고리즘의 시험 결과를 이용하여 다양한 관점에서 분석이 가능할 것으로 예상된다. 그리고 개발된 분류 알고리즘을 적용하면 각 수요 그룹에 대한 표준 부하 모델을 선정 할 수 있다. 또한 그 모델들을 이용한 부하의 분석 시스템의 개발, 수요관리, 시스템 계획, 부하 관리, 판매전략 수립 등 전력 부가서비스의 개발에 대한 기반을 구축할 수 있다.

참고 문헌

- [1] Y.-H. Pao and D. J. Sobajic, "Combined use of unsupervised and supervised learning for dynamic security assessment," *IEEE Trans. Power Syst.*, vol. 7, pp. 878-884, May 1992.
- [2] C. S. Chen, J. C. Hwang, Y.M. Tzeng, C.W.Huang and M.Y.Cho, " Determination of Customer Load Characteristics by Load Survey System at Taipower" *IEEE Trans. Power Delivery*, vol. 7, pp. 1430-1436, July 1996.
- [3] Gerbec, D., Gasperic, S., Smon, I. and Gubina, F. "An approach to customers daily load profile determination" *Power Engineering Society Summer Meeting, 2002 IEEE*, Volume: 1 , pp. 587-591 July 2002