

Spatio-temporal Load Analysis Model for Power Facilities using Meter Reading Data

辛 鎮 浩[†] · 金 榮 一^{*} · 李 倖 在^{**} · 梁 日 權^{***} · 柳 根 鎬[§]
 (Jin-Ho Shin · Young-Il Kim · Bong-Jae Yi · Il-Kwon Yang · Keun-Ho Ryu)

Abstract - The load analysis for the distribution system and facilities has relied on measurement equipment. Moreover, load monitoring incurs huge costs in terms of installation and maintenance. This paper presents a new model to analyze wherein facilities load under a feeder every 15 minutes using meter reading data that can be obtained from a power consumer every 15 minute or a month even without setting up any measuring equipment. After the data warehouse is constructed by interfacing the legacy system required for the load calculation, the relationship between the distribution system and the power consumer is established. Once the load pattern is forecasted by applying clustering and classification algorithm of temporal data mining techniques for the power customer who is not involved in Automatic Meter Reading(AMR), a single-line diagram per feeder is created, and power flow calculation is executed. The calculation result is analyzed using various temporal and spatial analysis methods such as Internet Geographic Information System(GIS), single-line diagram, and Online Analytical Processing(OLAP).

Key Words : Power Load Analysis, Spatio-temporal, Data Mining, Meter Reading Data, AMR, GIS

1. 서 론

오늘날 전력산업이 규제완화와 경쟁체제로의 이행을 위한 변화에 따라 전력계통과 전력설비의 효율적인 운영의 중요성이 부각되고 있다. 전력시장에 새롭게 도입된 경쟁의 개념은 사회후생의 확대와 전력요금의 인하라는 목표를 위해 진보적이며 효과적인 계통운용을 모색하도록 요구하고 있다. 이러한 전력설비 효율화에 대한 기술요구에 원천적으로 대응하기 위해서는 종래의 계측장치를 설치하고 계측값을 보여주는 방법이 아닌 새로운 시공간 부하분석에 관한 기술개발이 필요하다.

현재 배전선로의 신설 및 설비의 운영부문에 매년 막대한 예산을 투입하고 있으나, 지역에 따라서는 과·부족 투자에 따른 정전사고, 저전압지역 발생, 유휴 전력설비의 발생 등 공급설비의 불균형 현상이 발생되고 있다[1]. 이러한 현상의 발생원인은 배전계통 및 설비의 순간순간 변화하는 부하와 최대부하 시점의 구간/변압기 단위의 부하 상황을 파악하기 어렵기 때문으로 판단된다. 또한 현재의 부하분석 방법은 변압기의 월/년 단위의 최대부하와 일률적인 보정상수를 사용하기 때문에 일부 오차가 발생되고 있다. 이렇게 배전설비

운영시 현재 배전계통에 대한 상세한 부하분석모델이 없기 때문에 투자계획 수립에도 어려움을 겪고 있으며, 최악의 상황(인출단 최대부하)을 대비한 사업계획을 수립하는 경향이 있다.

전국적으로 배전계통의 주요설비 현황을 살펴보면 회선 7천여개, 선로공장 39만c-km, 전주 760만개, 변압기 170만대, 개폐기 12만대에 달한다. 이러한 대규모의 설비를 회선 이하의 구간이나 변압기 단위로 순간순간 변화하는 부하를 감시하기는 대단히 어려운 실정이다[2]. 일부 개폐기나 변압기에 부하를 측정하는 장치를 설치하여 감시하고 있으나 모든 설비에 설치하기는 무리일 것이다.

본 논문에서는 대규모의 전력설비에 계측장치를 설치하지 않고도 모든 전력사용자(고객)로부터 15분 또는 월단위로 취득되는 검침데이터를 이용하여 순간순간 변화하는 회선이하 설비의 부하를 15분 단위로 분석할 수 있는 새로운 모델을 제시하고자 한다. 먼저 부하계산 및 검증에 필요한 여러 운영계 시스템을 연계하여 배전설비 지리정보, 원격검침, 월검침, 고객, 변압기 무선부하, 변전소 회선 계측 데이터를 추출하고 부하분석모델 데이터 웨어하우스를 구축하였다. 그 다음 배전설비와 고객과의 관계를 설정하고 AMR 미시행 고객의 부하패턴을 예측한 다음 회선별 단선도를 생성하고 조류계산을 수행하였다. AMR 미시행 고객은 고압과 저압으로 구분하고 고압고객은 AMR 데이터를, 저압고객은 변압기 무선부하감시 데이터를 기반으로 군집화(Clustering) 및 분류(Classification) 데이터마이닝 알고리즘을 적용하여 15분 단위의 부하패턴을 예측하였다. 이렇게 계산 처리된 결과를 사용자가 다양한 방법으로 효과적으로 분석할 수 있도록 인터넷 GIS 부하분석, 회선별 단선도 부하부서, 온라인분석처

† 교신저자, 正會員 : 韓國電力公社 電力研究員

E-mail : jinho@kepco.co.kr

* 正會員 : 韓國電力公社 電力研究員

** 正會員 : 韓國電力公社 電力研究員

*** 正會員 : 韓國電力公社 電力研究員

§ 非會員 : 忠北大學校

接受日字 : 2008年 8月 26日

最終完了 : 2008年 10月 8日

리 기법을 이용한 계통 및 설비 부하분석, 현황/통계, 시뮬레이션 등 시공간 부하분석모델을 개발하였다.

이와 같이 상세하게 시간단위로 회선이하의 설비를 분석하게 되면 정확하게 언제 어떤 구간이나 설비에서 얼마의 과부하가 걸리는 지, 얼마의 전압강하나 손실이 있는 지 찾아낼 수 있으며, 정전이나 고장이 발생한 경우 그 시점에 해당 선로의 부하상태는 어떠한지 분석할 수 있다. 또한 특정 시점의 부하를 증감시키거나 다른 설비로 전환하여 시간대별 부하가 어떻게 변화하는지 시뮬레이션을 수행해 봄으로써 설비운영의 효율화를 도모하고 최적화된 부하계통의 운용방안을 수립할 수 있을 것이며, 신규 설비 확충시 기존 선로의 부하를 분석하여 보다 정확한 설비 투자 계획을 수립할 수 있을 것이다.

2. 시공간 부하분석모델의 구성 및 적용기술

시공간 부하분석모델의 전체적인 구성은 그림 1과 같다. 먼저 부하계산에 필요한 데이터를 기존의 운영계 시스템을 연계하여 추출한다. 고객의 특성을 나타내는 고객기본정보와 월 단위의 검침정보, 고압고객으로부터 15분 단위로 유효전력과 무효전력을 취득한 원격검침정보를 추출한다. 그리고 공간 분석을 위한 배전계통 및 설비 정보와 일부 중요 변압기에서 제측한 상별 전압, 전류 데이터를 연계한다. 변전소 회선 인출단의 계측 정보는 부하분석모델에서 계산한 값을 비교 검증하기 위해 추출한다. 이 운영계 시스템의 데이터는 오류 및 이상치를 내포하고 있으므로 정제 작업을 거친 다음, 부하분석모델 데이터 웨어하우스의 구조와 형식에 맞게 변환하고 적재한다.

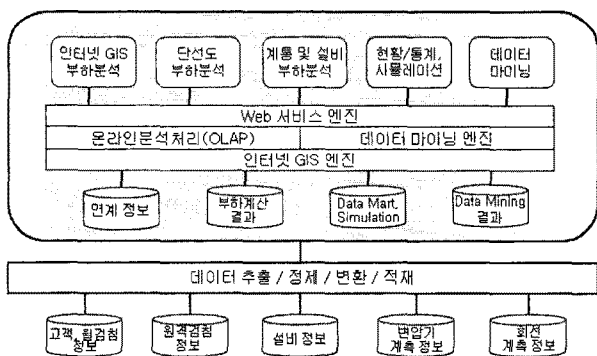


그림 1 시공간 부하분석모델 구성도
Fig. 1 Spatio-temporal load analysis model framework

이 모델의 개발을 위해 최신 IT 기술의 적용은 필수적이다. 배전계통이 공간적으로 분포되어 있는 특성이 있으므로 지리정보시스템을 활용하고, 사용자의 편리성을 고려하여 Web 환경에서 이용할 수 있는 인터넷 GIS를 적용한다. 또한 전체 배전 부하의 70%를 차지하는 계약전력 100kW 이상인 고압고객의 15분 단위의 대용량 원격검침 데이터를 연계하고, 15분 단위의 부하를 계산하고 분석하기 위해 데이터 웨어하우스와 온라인분석처리(OLAP) 기법을 적용한다. 데이터 마이닝은 이전에 알려지지 않는 새로운 지식을 발견하는 기법[3]으로 AMR 미시행 고객과 변압기의 부하패턴을 예측하는 분야에 적용한다.

3. 부하계산 과정 및 방법

검침 주기와 방법에 따라 실질적인 전력소비자인 고객을 표 2와 같이 AMR 고객, 未AMR 고압고객, 저압고객으로 구분하였다. 여기에서 고압의 구분은 공급방식이 단상2선식 3.3kV 이상이고 계약종별이 고압으로 분류된 코드를 사용하는 고객이 해당된다. AMR을 시행하지 않고 고압고객으로 분류되지 않는 고객이 저압고객이다. 모/자 관계가 있는 고객은 개별 검침하고 있기 때문에 별도로 구분하지 않았다.

표 1 검침 고객 구분
Table 1 Customer by meter reading

구분코드	고객 구분	설명
1	AMR 고객	AMR 시행하는 고압 고객
2	未AMR 고압고객	AMR 시행하지 않고 있고 월 단위 검침하는 고압 고객
3	저압고객	월 단위 검침하는 저압고객

AMR 고객은 15분 단위로 전력사용량이 검침이 되지만 미AMR 고압고객과 저압고객은 월검침이 시행되므로 원격검침 데이터의 시점과 일치하도록 15분 단위로 부하패턴을 계산하는 작업이 필요하다. 군집화와 분류 데이터 마이닝 기법을 적용하여 월검침 데이터를 15분 단위의 부하 프로파일을 생성하는 알고리즘을 개발하였다. 미AMR 고압고객은 AMR 데이터를 기반으로, 저압고객은 변압기 무선부하감시 시스템 데이터 기반으로 부하패턴을 생성한다. 각 고객의 15분 단위 부하패턴이 생성되면 배전계통 구간 단위의 단선도를 생성하고 각 구간에 전력을 공급하는 고객의 부하를 집계한다. 그 다음 조류계산을 수행하여 회선과 구간의 피상전력, 전류, 전압, 손실을 구하고 통계정보를 산출한다. 이러한 부하계산 과정을 도시하면 그림 2과 같고, 다음 절부터 각 과정의 상세한 계산 방법을 기술한다.

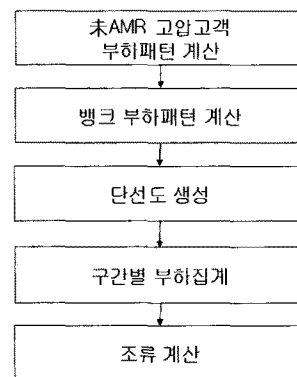


그림 2 부하계산 과정
Fig. 2 Process of the load calculation

3.1 未AMR 고압고객 및 뱅크 부하패턴의 계산

AMR 부하 데이터의 군집분석을 통해 고압 고객별 대표 부하 프로파일을 생성하고, 미AMR 고압고객의 월별 부하패턴을 예측하는 과정을 그림 3에 나타내었다.

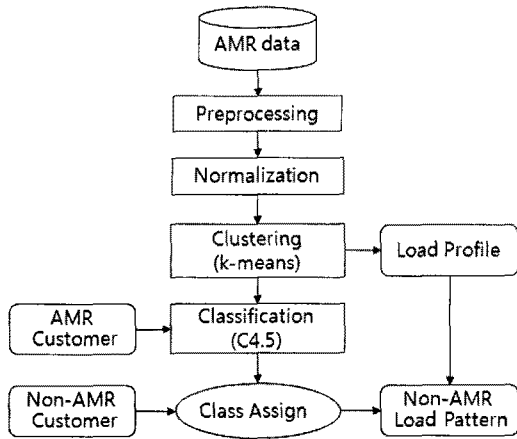


그림 3 미AMR 고압고객 부하패턴 계산 과정
 Fig. 3 Process of the load pattern calculation for non-AMR high voltage customers

구축된 부하분석 데이터 웨어하우스로부터 AMR 고압고객의 기본정보와 15분 단위의 부하 데이터를 추출한다. 수집된 데이터에 포함된 오류 및 Outlier 데이터는 군집분석의 성능을 크게 저하시킬 수 있기 때문에 데이터 정제를 위한 전처리 작업이 필요하다. AMR 원시 데이터에서 하루 96개 데이터가 존재하지 않거나, 하루 유효전력의 합이 1보다 작은 당일 데이터는 제외한다. 15분 단위의 최소 검침량은 0.08kWh로 1kWh 이하는 정상적으로 검침되지 않은 것으로 간주하였다. 또한 Outlier 처리를 위하여 데이터 정제 기법 중 SOM(Self-Organizing feature Map) 군집화 알고리즘 [3]을 적용하였다. 구성 매트릭스는 10 by 10 (100 군집)으로 한 클러스터에 포함된 데이터 객체가 1개 이하인 군집 결과는 outlier로 간주하여 제외하였다.

기존 Load Profile 연구[4-6]에서는 평일과 휴일을 구분하여 하루 단위의 전력 사용량 벡터를 구성하였으나, 이럴 경우 월초나 월말의 평일의 부하가 동일하게 생성되기 때문에 매일 매일 변화하는 부하 패턴을 분석하기 어렵다. 따라서 본 논문에서는 월 전체의 15분 단위의 부하를 고객별로 하나의 벡터로 식 (1)과 같이 재구성하였다.

$$V^{(m)} = \sum_{i=1}^{30} \{V_0^{(m)}, \dots, V_h^{(m)}, \dots, V_H^{(m)}\} \quad (1)$$

여기서, m =고객, $h=0, \dots, 2345$, $i(day)=1, \dots, 30$

이때 원시 데이터를 그대로 사용하게 되면 군집화가 전력 사용량의 분포에 따라 형성된다. 따라서 위 식의 고객별 벡터의 최대값이 1이 되도록 정규화한 다음 군집분석을 수행하였다. 많은 군집화 알고리즘 중에서 본 논문에서는 대용량 데이터에서 빠른 군집 구성, 사용자 기반의 군집 수 결정방식, 군집의 적합성을 판단하는 제한을 적용 용이성 등을 고려하여 k-means 알고리즘[3]을 적용하였다.

군집분석의 결과로 다수의 군집이 생성되면 클래스의 분류 모델을 구축한다. 분류는 서로 다른 클래스의 객체들을 구별하기 위한 설명 도구로서의 역할과 알려지지 않은 레코드들의 클래스 레이블을 예측하기 위해 사용된다. 즉, 각 군집을 AMR 고객특성에 따라 분류하고 미AMR 고압고객이

입력되었을 때 클래스 레이블을 예측한다. 여기에서 고객 특성정보는 계약종별, 계약전력, 전기사용용도, 산업분류, 공급방식, 월검침량이며 분류기의 입력변수로 사용된다. 분류기의 종류로는 의사결정트리, 베이저안 분류기, 신경망, SVM(Support Vector Machine) 그리고 규칙기반 분류기[3] 등이 있는데, 본 논문에서는 성능을 평가하여 의사결정트리 분류기를 사용하였다. 의사결정나무(Decision Tree)는 의사결정규칙을 나무구조로 도표화하여 분류와 예측을 수행하는 분석방법이다. 이 방법은 분류 또는 예측의 과정이 나무구조에 의한 추론규칙(Induction Rule)에 의해서 표현되기 때문에, 다른 방법들에 비해서 분석자가 그 과정을 쉽게 이해하고 설명할 수 있다는 장점을 가지고 있다. 미AMR 고압고객의 부하 예측은 C4.5 결정트리 모델에 의해 월별 패턴 클래스를 할당 받아 예측한다. 이 예측된 패턴은 0과 1 사이의 정규화된 부하패턴이므로 원래의 부하량으로 복원 되어야 하며 부하 Ratio를 식 (2)으로 계산한다.

$$C_k' = T \times \frac{C_k}{\sum_{i=1}^n C_i} \quad (2)$$

여기서, C : 유효 전력량,
 k : 특정 시점, n : 전체 시점의 수,
 T : 미AMR 고압고객의 월 검침량

저압고객의 원격검침은 대부분 주택용 고객을 대상으로 소규모로 시범운영되고 있어 현재는 활용하기 어렵다. 따라서 본 논문에서는 30분 단위로 계측되고 있는 변압기 무선 부하감시시스템의 데이터를 이용하고 해당 변압기에서 공급하고 있는 저압고객의 특성정보를 사용하여 Банк 단위로 부하패턴을 생성하였다. Банк 부하패턴의 생성은 미AMR 고압고객의 부하패턴 생성 과정과 유사하게 수행하는데, 변압기 무선부하감시시스템의 전류 데이터를 사용하여 Банк의 대표 부하패턴을 생성하고 계측되지 않은 Банк에 대해 부하를 예측한다. 분류 모델 구축을 위한 Банк와 고객 특성정보는 전등수용호수, 동력수용호수, 부하지역특성, 변압기 용량, 고객의 계약전력, 전기사용용도, 고압/고압구분, 월검침량을 사용한다.

3.2 단선도 생성 및 조류계산

추출된 설비도 데이터에서 구간을 구성하고 구간 단위의 부하집계와 조류계산을 위해 회선별 단선도를 생성하였다. 배전계통이 Loop되는 구간과 Isolation되는 구간을 찾아내고 기존 설비도와 고압계통경과도, 회로도라 비교하여 해당 구간을 연결하고 보정한 다음 구간의 전원측과 부하측을 결정하였다. 지중 다회로 개폐기에서 고압고객에게 직접 수전되는 단자는 그 부하를 앞단의 구간에 포함시켜 구간부하를 집계하도록 하였다. 단선도를 생성하는 알고리즘은 3진 Tree 회귀적(Recursive) 방법으로, 변전소 CB정보로부터 Root 노드를 생성하고 다회로 개폐기와 분기점에서 구간을 분리하고, 지중뱅크와 개폐기가 개방상태이고 ALTS나 하단의 구간정보가 없으면 상위 전원측 설비로 Pointer를 이동하면서 3진 트리를 생성하게 된다.

구간 단위의 단선도가 생성되면 각 구간의 부하를 집계한다. 구간에서 공급하는 AMR 고객, 미AMR 고압고객, Банк

에 대해 3.1절에서 생성한 부하를 합산하게 되면 시간대별 구간 부하가 생성되고 조류계산을 수행할 수 있게 된다.

전력조류의 계산은 특정 모선에서의 모선전압, 계통손실, 전력조류 데이터를 얻는 것이다. 즉, 부하전력, 무효전력, 배전계통 중의 전력과 무효전력의 흐름으로 선로와 변압기의 손실과 각 모선의 전압과 위상각을 구하는 것이다[7]. 배전계통은 운영 특성상 방사상(Radial)구조로 되어 있기 때문에 트리 구조의 계산방법이 적절하다. 이 논문에서는 Forward Sweeping 기법을 적용하였으며, 입력변수는 각 구간에 합산된 유효전력과 무효전력이고, 출력변수는 회선과 각 구간의 피상전력, 전류, 전압, 전압강하, 손실이다.

4. 부하분석모델 구현 결과

인터넷 GIS 부하분석은 공간상에서 부하를 분석하기 위해 설비도면 위에 회선, 구간, 변압기에 대한 시간대별 부하 현황과 통계를 분석할 수 있도록 구현한 모듈이다. 그림 4는 3개의 회선에 대해 일 단위 최대부하를 구간 단위로 부하밀도와 함께 구간 피상전력을 색상으로 구분하여 표현한 것이다. 공간 분석에서 부하시간의 선택은 특정시점만 표현이 가능하며, 여기에서 시간은 정시, 일, 주, 월은 각각의 기간 중 최대부하 시점을 의미한다. 시공간 부하분석은 설비, 부하, 표현방식의 선택이 조합되어 설비도 상에 표현되는데, 회선 분석은 여러 인접 회선에 대해 부하의 현황 파악이 용이하며, 구간 분석은 각 구간에 대한 부하의 분포 특성을 분석할 수 있고, 변압기 분석은 특정 지역내 변압기의 부하를 색상구분이나 객체크기에 따라 표현하여 공간상에서 설비 위치정보와 함께 부하를 분석할 수 있도록 구현하였다.

부하밀도는 3단계 명암으로 부하가 많은 설비 주변을 짙게 표현함으로써, 축적을 크게 조정하게 되면 부하밀집 지역을 파악하는 데 유용할 수 있다. 또한 설비 부하의 고압/저압 비율, 계약종별 비율을 파이(Pie) 차트와 바(Bar) 차트로 설비도 상에 표현함으로써 공간상에서 통계분석이 가능하도록 개발하였다.

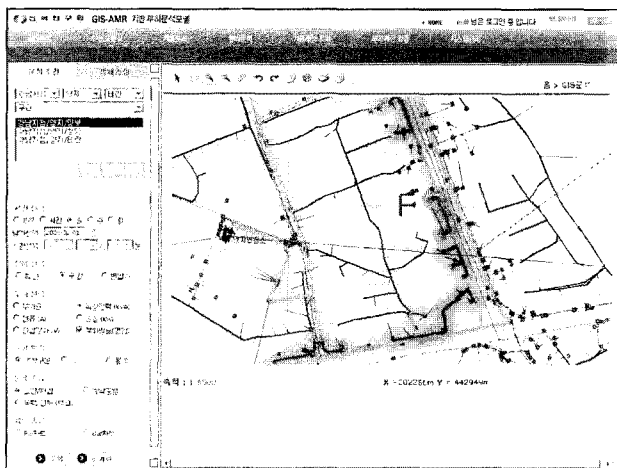


그림 4 인터넷 GIS 부하분석
Fig. 4 Load analysis using internet GIS

인터넷 GIS 부하분석은 실제 지형공간에 분포된 형태로

계통의 부하분석은 용이하나, 하나의 회선에 대해 전체 구간의 부하를 한눈에 파악하기는 불편한 점이 있다. 따라서 회선별 단선도를 통해 회선과 구간에 대한 시간대별 부하분석이 가능하도록 개발하였다. 그림 5는 하나의 회선을 3진 트리로 구성된 단선도에서 회선과 구간에 대한 부하(전류, 전압강하, 손실, 전력)를 선택적으로 표현하고, 시간이동 버튼을 이용하여 15분전, 15분후, 어제 같은 시분, 내일 같은 시분으로 쉽게 시간을 이동하면서 부하를 분석할 수 있는 특징이 있다. 그리고 회선과 특정 구간에 대해 당일의 15분 단위 부하패턴, 특정시분에 대한 월간 부하패턴, 일 최대부하를 이용한 월간 부하패턴, 월 최대부하를 이용한 연간 부하패턴을 분석할 수 있다.

계통 및 설비의 부하분석은 온라인분석처리 기법을 이용하여 다양한 각도에서 다차원 부하분석을 할 수 있는 모듈이다. 회선, 구간, 변압기에 대해 시간대별(15분, 시간, 일, 주, 월, 같은 시간대, 최신), 부하별(부하율, 피상전력, 유효전력, 무효전력, 전류, 손실, 전압강하)로 표와 그래프를 이용하여 부하패턴을 분석할 수 있도록 구현하였다. 부하 표와 그래프를 가장 요약된 레벨로부터 가장 상세한 레벨까지 차원의 계층에 따라 분석에 필요한 요약 수준을 바꿀 수 있는 Drill Up/Down, 시간대와 부하의 축을 서로 바꾸면서 분석하는 Pivoting, 다차원 배열의 부분집합으로 만들어 분석하는 Slice/Dice, 리포트의 형태와 조건을 바꾸는 Surfing 등의 온라인분석처리 기법을 이용하여 다차원 분석할 수 있다. 또한 정전이나 고장정보가 입력되면 그 시점의 부하를 분석할 수 있도록 인터넷 GIS부하분석, 단선도 부하분석, 회선분석, 구간분석과 연동된다.

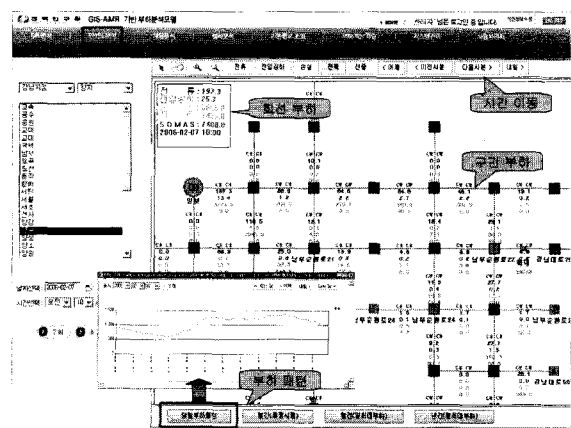


그림 5 회선별 단선도 부하분석
Fig. 5 Load analysis using single-line diagram per feeder

회선과 구간에 대한 현황/통계는 고압과 저압, 계약종별(주력용, 일반용, 교육용, 산업용, 기타)로 부하비율과 피상전력을 분석하고, 변압기에 대해서는 동력과 전동, 계약종별로 분석한다. 그리고 부하율과 피상전력에 대한 최대/평균/최소값과 시간대별로 현황/통계에 대한 표와 그래프도 다차원 분석이 가능하다. 또한, 계통과 변압기에 대해 부하를 증가 또는 감소 시켜보는 부하변동 시물레이션과 회선에 대한 부하평준화 시물레이션, 변압기에 대해 다른 변압기로 부하를 넘겨보는 부하전환 시물레이션이 가능하도록 개발하였다.

6. 사례 연구

이 절에서는 미AMR 고압고객의 부하패턴 생성 과정에서 군집 수 결정과 분류 모델의 평가, 회선에 대한 조류계산 데이터와 변전소에서 측정한 회선 부하 데이터의 비교 실험 결과를 기술한다. 서울시의 강남구와 서초구에 전력을 공급하는 한국전력공사 강남지점의 2007년 1월부터 10월까지의 3,349 AMR 고객, 792 미AMR 고압고객, 27만 저압고객을 대상으로 시험하였다.

군집분석에서는 최적의 k값(군집 수) 결정이 가장 중요한 요소로서, AMR 대표 부하패턴의 군집 수 결정은 재현성 평가 방법을 사용하였다. 재현성 평가 모델은 신경망이나 결정 트리 분류 및 회귀와 같은 지도학습 모델링에서 사용되는 데이터 분할 기법을 활용한 것이다. 자료 분할은 동일한 군집화 방법의 반복을 가능하게 해주므로 이를 활용하여 재현성 평가를 할 수 있다. 재현성 평가 절차는 다음과 같다.

- ① 주어진 자료를 임의로 2개로 분할한다. 하나는 훈련 데이터이고, 다른 하나는 테스트 데이터이다.
- ② 훈련 데이터를 군집화하여 군집화 모델을 산출한다. 그리고 테스트 데이터의 각 객체를 훈련 데이터로부터 생성된 군집화 모델에 적용하여 분리한다. 즉 테스트 데이터를 가장 가까운 중심의 군집에 할당한다.
- ③ 테스트 데이터를 동일한 방식으로 군집화하여 자체 모델을 산출한다. 그리고 테스트 데이터의 각 객체를 몇 개의 군집 중 하나에 할당한다.
- ④ 훈련 데이터와 테스트 데이터의 두 군집화 결과를 이용하여 교차분류표를 생성한다. 교차분류표에서 각 군집별 최대값이 주 경향을 나타내는 값이며, 나머지가 주 경향에서 벗어난 데이터이다.

표 2 AMR 부하패턴 재현성 평가 결과

Table 2 Reproducibility evaluation result for the AMR load pattern

k값	테스트 데이터 수	주 경향에서 벗어난 데이터 수	주 경향에서 벗어난 데이터의 Percentage
k=30	1553	41	2.64%
k=31	1553	23	1.48%
k=32	1553	23	1.48%
k=33	1553	114	7.34%

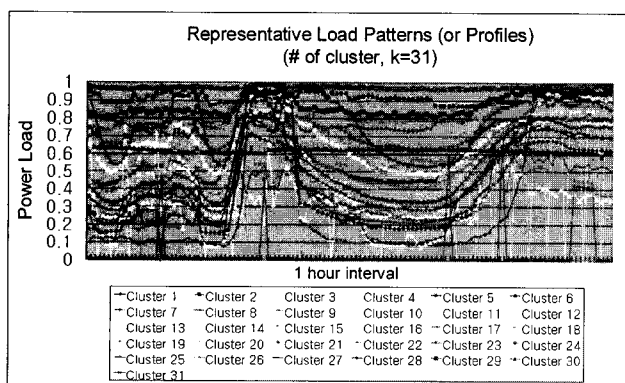


그림 6 AMR 대표 부하패턴

Fig. 6 AMR representative load pattern

2007년 1월 AMR 데이터의 재현성 평가 결과는 표 2와 같이 나타났으며, 주 경향에서 벗어난 데이터의 percentage가 가장 작은 k값으로 31을 선택하였다. 그림 6은 동월 AMR 고객에 대한 31개 군집의 대표 부하패턴이다. 이 군집화 결과를 미AMR 고압고객에 적용하기 위해 생성된 Cluster Label과 고객 특성정보를 이용하여 분류 모델을 생성한다. 모델 생성을 위한 알고리즘은 Decision tree(C4.5), Bayesian network, naive Bayesian 그리고 SVM를 사용하였으며, 분류 모델의 평가는 10-fold cross validation 기법을 적용하였다. 분류 모델의 성능평가는 TP(True Positive), TN(True Negative), FP(False Positive), FN(False Negative)로 표현되는 혼동 행렬(Confusion matrix)[8]을 이용하여 아래 식으로 계산하였다.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

그림 7과 같이 C4.5 알고리즘의 정확성이 92.528로서 다른 분류 모델에 비해 가장 높게 나타났다. 따라서 미AMR 고압고객에 대한 부하패턴 예측 모델로 C4.5 의사결정 트리를 선택하였다.

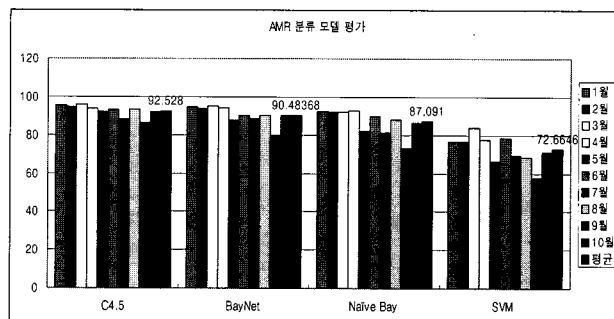


그림 7 미AMR 고압고객 분류 모델 평가

Fig. 7 Classification model evaluation for non-AMR high voltage customers

이 논문에서 제안한 부하분석모델의 정확성을 검증하기 위해 회선에 대한 조류계산 결과와 변전소에서 측정한 회선 부하 데이터를 그림 8과 같이 비교 실험하였다.

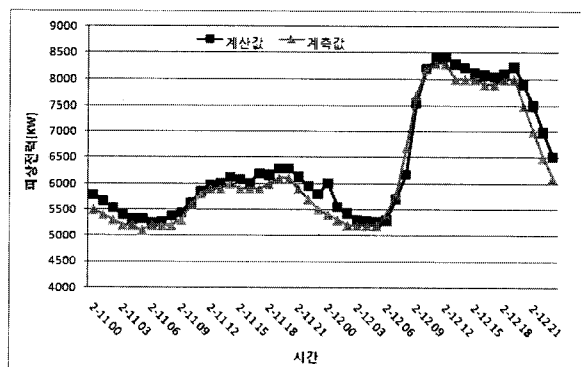


그림 8 회선 계산값과 측정값 비교

Fig. 8 Comparison of calculation and measurement value for feeder

두 값의 유사도를 측정하기 위해 코사인 유사도(cosine similarity)를 아래 식과 같이 계산하였다.

$$\cos(x, y) = \frac{x \cdot y}{\|x\| \|y\|} \quad (4)$$

여기서 \cdot 는 벡터의 도트곱(dot product)으로 $x \cdot y = \sum_k x_k y_k$ 를 나타내고, $\|x\|$ 는 벡터 x 의 길이로 $\|x\| = \sqrt{\sum_k x_k^2}$ 을 나타낸다. 위 그래프의 코사인 유사도를 계산하면 0.9996이다. 코사인 유사도는 1에 가까울수록 유사도가 높은 것이므로 회선부하의 계산값과 예측값은 유사도가 높다고 할 수 있다. 또한 두 그래프의 평균 차이값이 202kW로 변전소 예측값이 100kW 단위로 표현되는 것을 고려하면 조류계산에 의한 회선 부하값은 정확성이 높다고 할 수 있다.

7. 결 론

본 논문에서는 대규모의 모든 전력설비에 계측장치를 설치하지 않고도 전력사용자로부터 주기적으로 취득되는 검침 데이터를 이용하여 순간순간 변화하는 배전계통의 회선이하 설비의 부하를 분석할 수 있는 모델을 개발하였다. 이 모델의 개발을 위해 운영계 시스템에서 부하계산에 필요한 데이터를 연계하고 데이터마이닝 기법을 적용하여 미AMR고객의 부하패턴을 계산하고 조류계산을 수행한 다음 인터넷 GIS, 회선별 단선도, 부하 표와 그래프, 시뮬레이션 등 다양한 방식으로 회선, 구간, 변압기의 부하를 분석할 수 있도록 개발하였다. 시공간 분석 기법을 적용하여 이해도 높은 현황과 문제점 파악, 시각적 수요분석 및 부하관리, 공간적 부하 분포와 밀도 특성 분석이 가능하고, 고객, 설비, 부하, 고장간의 상관관계와 영향성 분석이 가능하다. 이를 통해 정확하고 신뢰할 수 있는 배전선로의 부하분석과 설비운영의 효율화 및 배전계통의 품질향상에 기여할 수 있을 것으로 기대된다.

감사의 글

본 연구는 지식경제부 전력산업기반기금의 지원으로 수행되었습니다.

참 고 문 헌

- [1] 한국전력공사 전력연구원, "최적 배전계획 수립을 위한 전산시스템 개발 및 적용연구" 최종보고서, 2000
- [2] 한국전력공사 전력연구원, "배전용 변압기 최적 부하관리 방안 연구" 최종보고서, 2003
- [3] Pang-ning Tan et al., "Introduction to Data Mining 2nd," publisher Pearson addison wesley, 2006
- [4] Vera Figuciredo, "An Electric Energy Consumer Characterization Framework based on Data Mining Techniques," IEEE Trans. on power systems, Vol. 20, No. 2, pp. 596-602, May 2005
- [5] Alejandro Ugedo, "Decision Trees Applied to the Management of Voltage Constraints in the Spanish

- Market", IEEE Trans. on power systems, Vol. 20, No. 2, pp. 963-972, May 2005
- [6] Pitt B. and Kirchen D., "Applications of Data Mining Techniques to Load Profiling," In Proc. IEEE PICA, pp. 131-136, 1999
- [7] Jian Liu, "Power Flow Analysis on Simplified Feeder Modeling", IEEE Transactions on Power Delivery, Vol. 19, pp. 279-287, 2004
- [8] David Taniar, "Research and Trends in Data Mining Technologies and Applications", IDEA Group Publishing, 2007

저 자 소 개



신 진 호 (辛 鎭 浩)

1970년 7월 26일생
1996년 한밭대 전자계산학과 졸업
2004년 충북대 전자계산학과 졸업(석사)
현재 한전 전력연구원 근무
E-mail : jinho@kepco.co.kr



김 영 일 (金 榮 一)

1972년 11월 27일생
1998년 충남대 컴퓨터공학과 졸업
2000년 충남대 컴퓨터공학과 졸업(석사)
현재 한전 전력연구원 근무
E-mail : yikim@kepco.co.kr



이 봉 재 (李 倬 在)

1959년 1월 13일생
1982년 성균관대 수학과 졸업
1985년 한국전력공사 입사
현재 한전 전력연구원 근무
E-mail : leeboja@kepco.co.kr



양 일 권 (梁 日 權)

1954년 4월 24일생
1976년 조선대 전기공학과 졸업
1992년 미국 Indiana대 컴퓨터과학과 졸업
현재 전력연구원 정보통신그룹장
E-mail : yangilk@kepco.co.kr



류 근 호 (柳 根 鎬)

1953년 11월 11일생
1976년 숭실대 전산학과 졸업
1980년 연세대 전산전공(공학석사)
1988년 연세대 전산전공(공학박사)
현재 충북대학교 전기전자 컴퓨터공학부 교수
E-mail : khryu@dblab.chungbuk.ac.kr