

Machine Learning Approach for Pattern Analysis of Energy Consumption in Factory

Jong Hoon Sung[†] · Yeong Sik Cho^{††}

ABSTRACT

This paper describes the pattern analysis for data of the factory energy consumption by using machine learning method. While usual statistical methods or approaches require specific equations to represent the physical characteristics of the plant, machine learning based approach uses historical data and calculate the result effectively. Although rule-based approach calculates energy usage with the physical equations, it is hard to identify the exact equations that represent the factory's characteristics and hidden variables affecting the results. Whereas the machine learning approach is relatively useful to find the relations quickly between the data. The factory has several components directly affecting to the electricity consumption which are machines, light, computers and indoor systems like HVAC (heating, ventilation and air conditioning). The energy loads from those components are generated in real-time and these data can be shown in time-series. The various sensors were installed in the factory to construct the database by collecting the energy usage data from the components. After preliminary statistical analysis for data mining, time-series clustering techniques are applied to extract the energy load pattern. This research can attributes to develop Factory Energy Management System (FEMS).

Keywords : Factory Energy, Power Consumption, Machine Learning, Factory Energy Management System

머신러닝 기법을 활용한 공장 에너지 사용량 데이터 분석

성종훈[†] · 조영식^{††}

요약

본 연구에서는 머신러닝 기법을 활용하여 공장에서 발생하는 에너지 사용량에 대한 데이터 분석 및 패턴 추출에 대해 다룬다. 통계학이나 기존의 방법들은 몇 가지 물리적 특성을 반영하는 수학적 모델을 구축하는 반면, 머신러닝을 통한 접근방법은 데이터 학습을 통하여 모델의 계수들을 결정하게 된다. 기존의 방법들은 특정한 구조를 갖는 수학적 모델을 구축해야 한다는 어려움이 있으며 과연 데이터의 특징들을 잘 반영하는지에 대한 의문이 존재했다. 그러나 머신러닝을 통한 방법은 사람이 구축하기 어려운 작업들을 용이하게 구축한다는 장점을 가지고 있기 때문에 데이터 간의 관계를 파악하기에 더 효율적이라는 장점을 가지고 있다. 공장의 에너지 소비에 직접적으로 영향을 끼치는 요소들이 존재하며 이러한 전력 소비는 시간에 따른 데이터로 나타나게 된다. 각 요소들로부터 발생하는 소비 전력을 계측하고 데이터 베이스를 구축하기 위해 각 요소에 센서를 장착하였다. 취득된 데이터에 대해 전처리 과정 및 통계적인 분석을 거친 뒤, 머신러닝을 통해 패턴을 분석하는 과정을 거쳤다. 이를 통해 공장에서 발생하는 소비 전력 데이터에 대한 패턴 분석을 진행하였다.

키워드 : 공장에너지, 소비전력, 머신러닝, 공장 에너지 관리 시스템

1. 서론

에너지 문제는 오랜 시간동안 사회적으로 공통된 주제였으며, 기술이 고도화됨에 따라 전 세계적으로 에너지 소비 절감 및 에너지 효율성 향상에 대한 관심이 더욱 올라가며 관련 기술 개발에 대한 수요 또한 높아지고 있다. 이러한 에너지 문제

에 대한 관심은 국내에서도 마찬가지이며 그에 따른 에너지 문제에 관련된 연구 또한 그 동안 꾸준히 진행되어 왔다. 그 결과 에너지 의존도가 높은 우리나라는 특히 전기 에너지에 대한 사용량 및 의존도가 매우 높은 것으로 나타났다[1]. 실제로 2007년에는 기록된 에너지 사용 수요의 3분의 2가 전기 에너지 생성을 위한 것으로 기록되었다. 심지어 세계 에너지 소비의 80%가 넘는 비율이 주로 도시에서 발생한다고 한다[2, 3].

그 중에서도 전 세계적으로 총 에너지 사용의 20~40%는 건물에서 소비된다고 한다[4-6]. 그만큼 건물에서 발생하는 에너지 소비량은 무시하지 못할 수준이란 것인데, 국내에서 발생하는 전력 소비의 22% 이상은 건물에서 발생한다고 한

※ 본 연구는 한국에너지기술연구원(KETEP)의 지원을 받아 수행한 연구과제임 (No.20172010000260).

† 정회원: ㈜에스더블유엔 부설연구소 이사

†† 비회원: ㈜AMEP 기술연구소장

Manuscript Received: February 7, 2019

Accepted: March 13, 2019

* Corresponding Author: Jong Hoon Sung(jhsung@swm.ai)

다[7]. 건물 소비 전력의 대부분은 일반 가정집들이 아닌 사무용 건물과 공장에 집중되어 있다. 특히 우리나라의 수출 주도형 경제 특성상 공장에서 발생하는 소비 전력은 매우 중요한 비중을 차지하고 있다고 볼 수 있다.

한편 GPU 등의 하드웨어 기술 고도화와 더불어 머신러닝에 대한 연구가 활발해졌다. 머신러닝은 사람이 찾아내기 어려운 수많은 데이터들 간의 패턴 및 관계를 찾아내는 데 용이하다는 것이 장점이며, 다양한 분야에 적용되고 있다. 최근에는 머신러닝 기법을 에너지 효율 분야에 적용하려는 연구 또한 진행되고 있으며, 실제로 이러한 건물 운용 과정에서 발생하는 에너지를 효율적으로 관리하기 위해 건물에너지관리시스템(Building Energy Management System)의 도입이 아직은 미비하지만 시범적으로 적용되는 등, 점차 늘어가고 있다[8].

그러나 아직까지 공장 에너지 관리 시스템(FEMS: Factory Energy Management)에 대한 연구 및 적용이 미비한 상황이다. 미국 등의 경우 공장의 규모 등의 이유로 비교적 많은 연구가 진행되었으나 국내에는 관련 연구가 매우 미비한 상황이다. 대부분의 전력 소모가 조명, HVAC, 사무 장비 등에서 발생한다는 어느 정도 공통점을 지닌 사무용 건물에 비해 공장은 생산 라인, 종류 등에 따라 일반 사무용 건물보다 더욱 다양한 데이터 양상을 가질 가능성이 높다. 공장에는 전력을 소비하는 많은 요소들이 존재하는데, 공장 에너지 분석 및 FEMS에 대한 연구를 위해서는 이러한 요소들에서 발생하는 정보들을 각종 기기 및 센서들을 활용하여 실시간으로 취득해야만 한다. 또한 이렇게 얻어진 에너지 관련 데이터를 효과적으로 분석하고 활용하게끔 하는 기능이 필수적이다.

따라서 본 연구에서는 대상 공장에서 얻어진 샘플 데이터를 기반으로 데이터 분석을 진행하고, 나아가 기계 학습(Machine Learning) 기법을 활용, 취득한 에너지들 간의 관계 분석 및 데이터 패턴 분석을 진행하였다.

2. 소비 전력 데이터 베이스

2.1 대상 공장




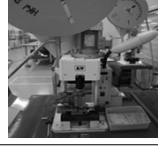


본 연구에서는 공장에서 발생하는 데이터를 대상으로 분석 및 연구를 진행하였기 때문에 적절한 대상 공장을 선정하고 그에 따른 적절한 전력 소비 요소를 파악하는 것이 중요했다. 본 연구의 대상 공장은 다섯 개의 층으로 이루어져 있으며, 조명 관련 제품 생산을 위한 생산 설비 라인을 갖추고 있다. 그 외에도 직원들의 업무를 위한 사무 공간 또한 존재하였다. 즉 본 연구의 대상 건물(공장)은 생산 설비 라인과 사무 공간 모두를 지니고 있기 때문에 추후 공장의 특성과 일반 사무용 건물의 에너지 소모 비중을 비교하는 데도 적절할 것으로 판단된다. 본 연구에서는 아직 충분한 데이터 베이스의 구축이 완료되지 않았으므로, 취득된 샘플 데이터에서만 연구를 진행하였다.

2.2 공장 전력 소비 요소 선정

공장은 일반 사무용 건물과는 달리 생산설비 라인을 갖추

고 있고 이들은 공장만의 특수한 전력 소모 요소들이다. 따라서 본 연구에서는 일반 사무용 건물과 가장 차별화되는 요소들인 생산 설비들에 대해서 우선적으로 데이터 취득을 진행하였고 각 설비들의 종류 및 수량은 다음과 같다.

Table 1. Power Consuming Components

	Name	EA
	Circuit Tester	1
	Welding Machine	3
	Flexible Cutter	1
	Hand-operated Compressor	4
	Button Inserter	1
	Automatic Compressor	1

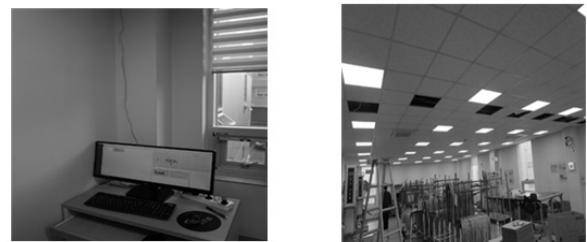


Fig. 1. Other Electricity Consumption Components

이 외에도 앞서 언급했듯이 본 연구의 대상 건물에는 사무용 공간이 마련되어 있다. 따라서 업무용 PC, 공장 내 조명, 공기조화설비 등 생산설비 이외에도 추가적인 전력 소비 요소들이 존재한다. 그러나 아직 충분한 데이터 취득 환경이 갖추어지지 않았으므로 본 논문에서는 표에 제시된 생산설비에 대해서 데이터 취득을 진행하였다. 생산 설비들은 디지털화적인 비

교적 최신 장비 외에도 오래된 장비들이 존재한다. 최신 장비들로부터는 데이터 취득이 비교적 용이했던 반면, 연식이 오래된 장비들은 추가적인 센서 부착을 통하여 아날로그 데이터를 디지털 데이터로 변환하는 작업을 통해 데이터 취득하였다.

2.3 IoT 데이터 수집 및 데이터 형식

대상 공장의 대표적인 전력 소모 요소들을 선정하였고, 각 시설 장비들로부터 유효한 데이터들을 얻기 위해 설비 기계 센서 등을 부착하여 데이터 취득 환경을 구축하였다. 그 결과 다음과 같이 작업에 관련된 정보(압력, 에러 Count 등)를 얻어낼 수 있었다.

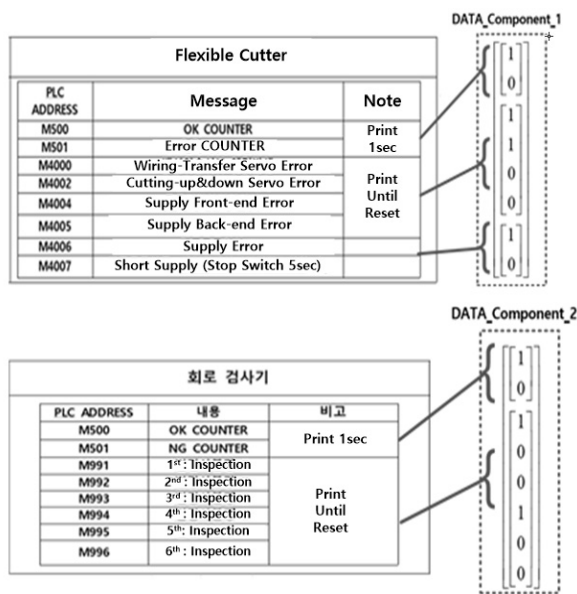


Fig. 2. Energy Consumption Data Example

위 그림은 전력소모 요소들 중 플렉시블 절단기와 회로 검사기에서 얻어진 데이터의 포맷을 보여준다. 그림에서 볼 수 있듯이 각 설비 요소들의 데이터는 작업에 관련된 항목들을 boolean 형태로 갖고 있으며 이들은 one-hot encoding 형태의 vector 데이터로 얻어질 수 있다. 이러한 데이터는 매 sampling time마다 얻어지게 되며 설비 요소별 총 작업량 등에 대한 추가적인 정보를 얻는 데 도움이 될 수 있다.

3. 머신 러닝 기법을 활용한 데이터 분석

3.1 데이터 전처리

얻어진 데이터들은 시간에 따라 각 요소별로 column에 나타내게 된다. 이러한 데이터는 time-series 형태 혹은 longitudinal data 라 불리운다. 데이터를 바로 사용하기 보다는 이에 앞서 전처리(pre-processing) 과정이 필요하게 된다. 데이터가 기록되는 과정에서 누락이 되거나(NaN값) 하는 경우가 발생하기 때문에, 데이터 누락이 있는지 확인한다. 또한 각 데이터 별로 물리적 단위 및 기록되는 단위에 따라 값의 level이 달

라지게 되는데, 이러한 값의 레벨 차이가 심한 경우 상대적으로 데이터의 영향이 무시되는 경우가 발생하기도 하며, 올바른 분석이 이루어지지 않게 된다. 따라서 값의 level을 적절하게 맞춰주는 작업 또한 필요하다.

3.2 데이터 군집화 알고리즘

군집화(Clustering) 혹은 군집분석(Cluster analysis)는 비지도학습의 한 방법으로서, 데이터 분포들 중에서 유사하다고 판단되는 값들끼리 묶어 몇 가지 군집으로 나누는 작업을 의미한다. 군집분석을 통해서 에너지 소비 패턴과 공장 내의 장비 및 시설들의 사용 특성을 알 수 있다. 이러한 패턴 정보는 플랜트에 대한 다른 정보가 없을 때 유용하게 활용될 수 있다. 또한 나아가 추후 데이터의 값을 예측할 때 패턴 정보를 활용할 수도 있다. 본 연구에서는 데이터 군집화를 통해 소비 패턴을 몇 가지 군집으로 추출하였다.

1) K-means Clustering

K-means clustering 알고리즘은 대표적인 군집화 알고리즘의 하나이다[9]. 우선 군집의 개수를 설정하면, 각 군집은 하나의 중심(centroid)을 갖게 되는데, 이 때 데이터들은 각 데이터들로부터 가장 가까운 중심에 해당하는 군집에 속하게 된다. 즉 관측된 데이터들을 (x_1, x_2, \dots, x_n) 이라 하고 각 data들은 m-dimensional real vector라 할 때, k개의 군집으로 나누는 k-means clustering은 n개의 관측치를 집합 $S = \{S_1, S_2, \dots, S_k\}$ 로 분배하는 작업이 된다. 이를 수식으로 나타내면 다음과 같다.

$$\arg \min_s \sum_{i=1}^k \sum_{x \in S_i} \|x - \mu_j\|^2 \quad (1)$$

본 연구에서는 각 벡터가 time-series 데이터가 되는 것이고 k-means clustering을 통해 이러한 시간에 따른 데이터들이 어떠한 패턴 k개로 나눌 수 있는지 확인한다.

2) Time-series Distance Measures

앞서 살펴본 것과 같이 K-means 알고리즘은 분포의 중심으로부터 데이터가 얼마나 떨어져 있는지를 비교하여 군집을 결정한다. 기존의 2차원 데이터 분포도에서의 K-means 알고리즘은 일반적으로 Euclidean distance를 사용하게 되며 이는 time-series 데이터에 대해서도 마찬가지로 적용할 수 있다. 이 경우는 특정 시간에 대한 데이터 값의 차이를 distance로 둘 수 있다. 이 때의 Euclidean distance는 다음과 같이 정의된다.

$$ED(\vec{x}_i, \vec{y}_j) = \sqrt{\sum_{n=1}^m (x_i(t_n) - y_j(t_n))^2} \quad (2)$$

Euclidean distance는 일반적으로 널리 쓰이며 가장 단순하면서도 효율적으로 거리를 계산하는 데 유용하다. 그러면서도 적절히 높은 정확도를 보여준다[10].

그러나 시계열 데이터는 거의 동일한 양상을 갖는 데이터

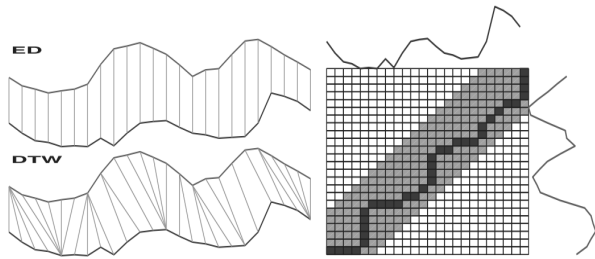


Fig. 3. Similarity Computation Using DTW[11]

가 시간만 옆으로 shift 되는 경우도 많기 때문에 단순 y 축 값들의 차이로 계산하는 것은 적절하지 않을 수 있다. 실제로 시계열 데이터 간의 유사도를 측정할 때에도 단순히 동일한 시간에서의 데이터 간 차이로 계산하게 되는 경우 적절한 결과를 얻지 못하는 경우가 많다.

이러한 문제를 해결하기 위해 동적 시간 굽힘(DTW: Dynamic Time Warping)은 시간 index를 가진 요소 뿐 아니라 그 주변의 다른 index의 값과의 거리도 비교하게 된다. 이러한 DTW를 거리 정보로서 K-means에 적용할 수 있다. DTW는 다음과 같이 정의된다.

$$DTW(\vec{x}_i, \vec{y}_j) = \min \sqrt{\sum_{n=1}^r \omega_n} \quad (3)$$

이 때 w는 warping path로 $W=\{w_1, w_2, \dots, w_r\}$ ($r \geq m$)로 정의된다. path는 다이나믹 프로그래밍(dynamic programming)을 통해 다음과 같이 계산된다.

$$\gamma(i, j) = ED(i, j) + \min \gamma(i-1, j-1), \gamma(i-1, j), \gamma(i, j-1) \quad (4)$$

수식에서 살펴볼 수 있듯이 DTW의 path는 Euclidean distance와 dynamic programming의 결합으로 나타내어진다. 즉, DTW는 DE에서 확장된 것으로 보아도 무방하다.

3) 군집 수 결정

k-means 알고리즘의 군집 수 k는 hyper parameter로 유도되거나 도출되는 값이 아닌, 사용자가 임의로 정하는 값이다. try and error를 통해서 적절한 값을 선정할 수도 있지만 본 연구에서는 Bayesian information criterion(이하 BIC)을 k값을 결정하

는 기준으로 활용한다. BIC는 통계학에서 모델을 결정하는 기준으로 많이 사용되며 수학적으로 다음과 같이 정의된다.

$$BIC = \ln(n)k - 2\ln(\hat{L}) \quad (5)$$

이 때 n은 데이터 개수, k는 k-means에서의 k값을, \hat{L} 은 likelihood function의 최댓값을 의미한다. 본 연구에서는 likelihood function으로 Gaussian mixture model을 사용하였다. 따라서 BIC가 가장 작은 값을 k값으로 결정하여 사용할 수 있다.

이 외에도 Hierarchical tree를 구성하여 각 요소들이 어떠한 계층으로 분류될 수 있는지 살펴보고 그 안에서 군집 수를 결정하는 방법도 가능하다.

4. 연구 결과

4.1 데이터 군집화

군집화(Clustering) 혹은 군집분석(Cluster analysis)는 비지도학습의 한 방법으로서, 데이터 분포들 중에서 유사하다고 판단되는 값들끼리 묶어 몇 가지 군집으로 나누는 작업을 의미한다. 군집분석을 통해서 에너지 소비 패턴과 공장 내의 장비 및 시설들의 사용 패턴이 어떻게 나타나는지 살펴보았다.

1) 전력 사용량 및 전처리

각 전력 소모 요소들에서 소비되는 전력 소모량 데이터에 대해 우선 전처리를 진행하였다. 누락값(NaN)의 경우, 전반적인 값이 0 근처인 경우에는 합리적일 수 있으나 그렇지 않은 경우 0으로 처리한다면 데이터 왜곡이 되어 올바른 처리를 할 수 없다. 따라서 앞 뒤 데이터 값을 통한 보간값으로 채워 넣었다. 또한 각 요소에서 발생하는 소비 전력량의 레벨이 너무 다른 경우 개형을 직접적으로 비교할 수 없기 때문에, 각 값을 0과 1사이의 level로 normalize 하였다.

2) 군집수 결정

얻어진 데이터들을 군집분류하기 위해 BIC를 사용하였고, 그 결과는 다음과 같다.

BIC의 값이 k가 5~7일 때 작게 나타나고, 이 때 k=6으로 설정하여 clustering을 진행하였을 때 각 cluster의 대표 궤적은 다음과 같다.

k=6으로 나뉘었다는 것은 대상 공장의 설비 요소들의 중

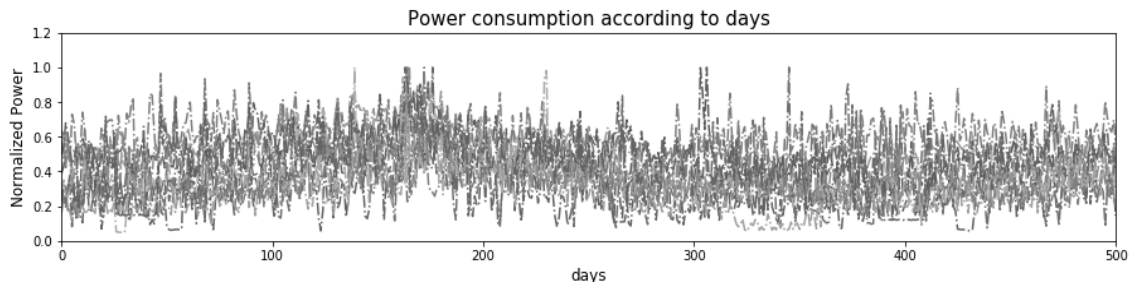


Fig. 4. Power Consumption According to Time

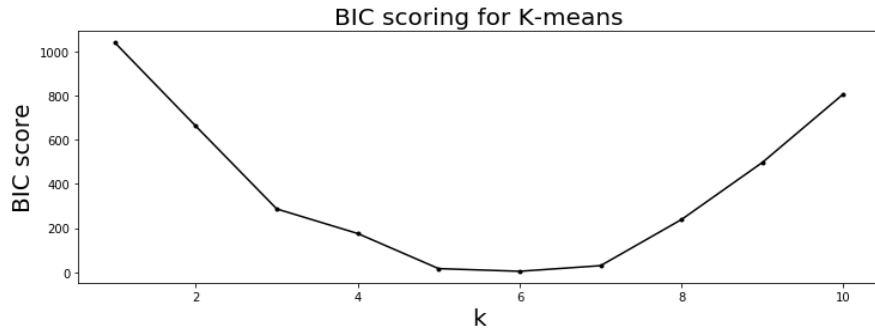


Fig. 5. BIC Score for K-means Clustering

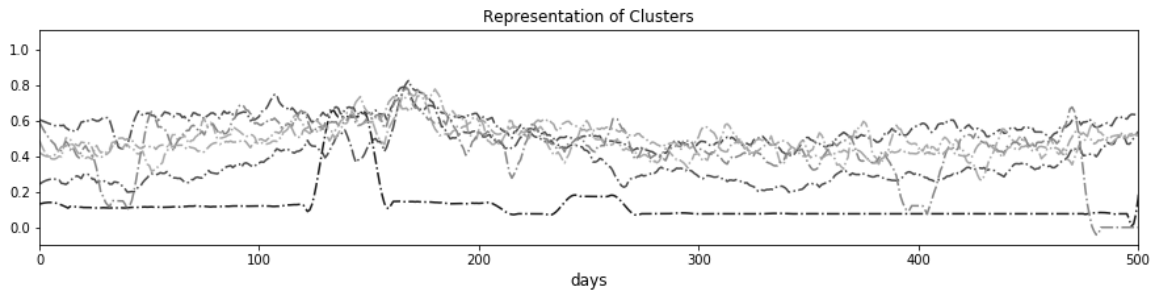


Fig. 6. Clustered Result When k=6

류와 맞아 떨어지며, 이는 동일한 설비 요소인 경우 유사한 전력 소모 패턴을 가질 것이라는 예측과 맞아 떨어진다. 그러나 빨간 선으로 나타난 군집을 제외한 나머지 요소들 간의 군집 구분이 명확하지 않다.

3) 요일에 따른 패턴 분석

앞서 대상 공장의 대표 소비 요소들에 대한 전력 소모 패턴을 알아보았다. 이 장에서는 공장에서 소모되는 총 전력소모량이 요일에 따라서 어떠한 패턴을 갖는지 확인해보도록 한다. 다음 Fig. 7은 대상 공장에서 소비되는 모든 전력 요소들의 normalized된 데이터들을 일주일 간격으로 그린 결과를 보여준다.

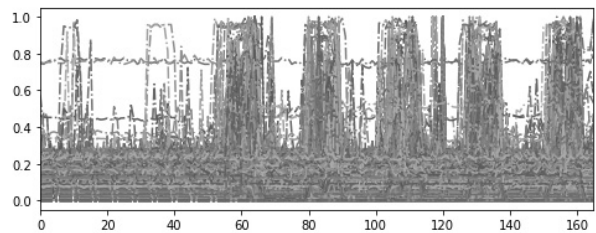


Fig. 7. Normalized Data of All Elements

또한 Fig. 8은 공장에서 소비되는 총 소비 전력량을 시간에 따라 그린 것이고 이를 일주일 간격으로 다시 그리면 Fig. 9와 같다. 결과에 의하면 주말인 토요일, 일요일에는 공장을 거의 가동하지 않는 것으로 보이며 주중동안에 규칙적으로 공장을 가동하는 것으로 보인다.

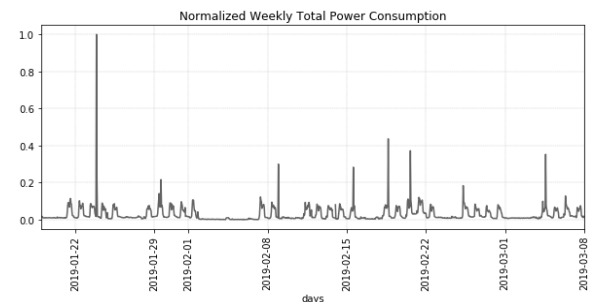


Fig. 8. Normalized Data of Total Power Consumption

공장에서 발생하는 전력소모량이 주중, 주말 등에 따라 어떻게 운영되는지 보기 위해 데이터를 특정 요일, 주중, 주말에 따라 구분하고 일주일에서 각 요일의 사용 퍼센티지로 데이터를 전처리 하였다. 이 때 요일, 주중, 주말의 전력 소비율에 대한 상관관계를 분석하면 다음과 같다.

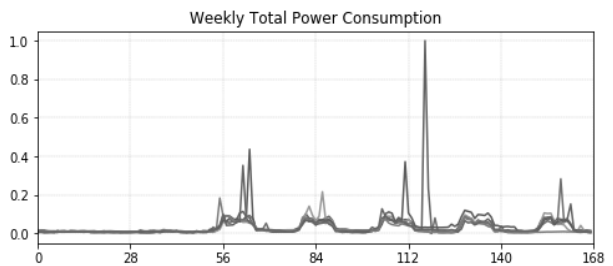


Fig. 9. Weekly Normalized Data of Total Power Consumption

상관관계 결과를 살펴보면, 전체 총 전력 소비량과 가장 밀접한 관계가 있는 것은 화요일과 주중으로 나타났다. 또한 주말과 수요일부터 일요일까지는 총 소비전력량과 음의 상관관계가 있는 것으로 나타났는데, 이는 공장의 총 소비전력량이 늘었을 경우 대부분의 업무는 주간에 처리하기 때문에, 그에 따라 상대적으로 주중에 비해 주말 가동은 줄어든 것으로 해석할 수 있다.

total KW	1	1	0.29	0.67	-0.36	-0.35	-0.47	-0.24	-0.43	0.42	-0.42
average per day	1	1	0.27	0.69	-0.39	-0.38	-0.49	-0.24	-0.39	0.39	-0.39
% Monday	0.29	0.27	1	0.2	-0.49	-0.6	-0.58	-0.55	0.055	0.29	-0.29
% Tuesday	0.67	0.69	0.2	1	-0.7	-0.087	-0.45	-0.64	-0.39	0.63	-0.63
% Wednesday	-0.36	-0.39	-0.49	-0.7	1	0.26	0.63	0.47	-0.11	-0.21	0.21
% Thursday	-0.35	-0.38	-0.6	-0.087	0.26	1	0.48	-0.031	-0.23	0.17	-0.17
% Friday	-0.47	-0.49	-0.58	-0.45	0.63	0.48	1	0.33	-0.44	0.088	-0.088
% Saturday	-0.24	-0.24	-0.55	-0.64	0.47	-0.031	0.33	1	0.31	-0.79	0.79
% Sunday	-0.43	-0.39	0.055	-0.39	-0.11	-0.23	-0.44	0.31	1	-0.83	0.83
% weekday	0.42	0.39	0.29	0.63	-0.21	0.17	0.088	-0.79	-0.83	1	-1
% weekend	-0.42	-0.39	-0.29	-0.63	0.21	-0.17	-0.088	0.79	0.83	-1	1
total KW											
average per day											
% Monday											
% Tuesday											
% Wednesday											
% Thursday											
% Friday											
% Saturday											
% Sunday											
% weekday											
% weekend											

Fig. 10. Correlation Matrix Between Components

5. 결 론

본 연구에서는 대상 공장의 대표 설비요소들에 대해 데이터를 취득하고, 기계 학습 기법을 기반으로 각 설비요소의 소비 전력 데이터에 대해서 분석해보았다. 머신러닝 방법을 적용하여 대표 설비요소들의 대표적인 전력 소비 패턴들을 도출하였으며 이러한 패턴을 몇 개로 정의하는 것이 적절한지 검토하였다. 기존의 통계학적 분석방법론과는 달리 머신러닝 적용방법을 적용함으로써 추후 더욱 누적될 데이터양 및 확장될 데이터 취득 요소들에 대해서 효율적인 분석이 이루어질 것으로 보인다.

전력 소비 패턴 뿐 아니라 본 연구에서는 대상 공장에서 발생하는 소비 전력이 주중, 주말, 요일에 따라서 어떠한 상관관계를 지니는지 알아보았다. 이러한 상관계수를 통하여 추후 요일에 따른 대상 공장의 전력 소모양상을 추정할 수 있을 것으로 보인다. 또한 추후 데이터 취득 요소가 증가하고 다량의 데이터가 누적되었을 경우 더욱 효과적인 분석이 가능할 것으로 보이며, 나아가 각 전력 데이터 간의 상관관계 및 경향성 분석을 통해 소비 전력 예측도 가능할 것으로 보인다.

References

[1] Il-Su Seol, Sun-Woong Kim, and Dong-You Choi, "A Study of the Current Status of Domestic Building Energy Management System and the Correct way for Improvement," *Proceedings of KIIT Summer Conference*, 2015.

[2] Ji-Young Eum, Soo-Hwan Choi, Si-Sam Park, and Yong-Ki Kim, "Development of Mathematical Model for the Energy Demand Pattern Analysis of City Buildings," The Korean Institute of Electrical Engineers, 2015.

[3] Ki-Ho Kim and Bae Kim, "Sustainable Development for the City : City Design Initiatives through Greenways," *Asia Design Journal*, Vol.5, pp.136-165, 2010.

[4] L.P. Lombard, J. Ortiz, and C. Pout, "A review on buildings energy consumption information," *Energy and Buildings*,

Vol.40, pp.394-398, 2008.

[5] K. Amarasinghe, D. Wijayasekara, H. Carey, M. Manic, D. He, and W. Chen, "Artificial Neural Networks based Thermal Energy Storage Control for Buildings," *Proc. 41st Annual Conference of the IEEE Industrial Electronics Society, IEEE IECON 2015*, Yokohama, Japan, Nov.09-12, 2015.

[6] D. Wijayasekara and M. Manic, "Data-Fusion for Increasing Temporal Resoluition of Building Energy Management System Data," *Proc. 41st Annual Conference of the IEEE Industrial Electronics Society, IEEE IECON 2015*, Yokohama, Japan, Nov. 09-12, 2015.

[7] Tae-Won Lee, Yong-Ki Kim, and Jae-Sang Seo, "The Major Functions and Using Method of the Building & Energy Management and Information System(BEMIS)," The Society of Air-Conditioning and Refrigerating Engineers of Korea, 2012.

[8] Hyeun Jun Moon, Sung Kwon Jung, and Seung Ho Ryu, "Building Cooling and Heating Energy Consumption Pattern Analysis Based on Building Energy Management System (BEMS) Data Using Machine Learning Techniques," The Society of Air-Conditioning And Refrigerating Engineers of Korea, 2015.

[9] M. Steinbach, G. Karypis, and V. Kumar, "A comparison of document clustering techniques," *KDD workshop on text mining*, 2000.

[10] Xiaoyue Wang, Abdullah Mueen, Hui Ding, Goce Trajcevski, Peter Scheuermann, and Eamonn Keogh, "Experimental comparison of representation methods and distance measures for time series data," *Data in, Knowl*, 2013.

[11] J Paparrizos and L Gravano, "Fast and accurate time-series clustering," *ACMTransactions on Database System*, vol. 42, 2017.



성 중 훈

<https://orcid.org/0000-0003-0586-7116>
 e-mail : jhsung@swm.ai
 1998년 아주대학교 전자공학부(학사)
 2000년 아주대학교 전자공학과(석사)
 2015년~현 재 ㈜에스더블유엠
 부설연구소 이사

관심분야 : Machine Learning and AI



조 영 식

<https://orcid.org/0000-0002-0622-9493>
 e-mail : pobicho@amep.co.kr
 1995년 명지대학교 전기공학과(학사)
 2016년 숭실대학교 전기공학과(석사)
 2016~현 재 (주)AMEP 기술연구소장
 관심분야 : Big Data, Energy Saving, AI