

# 공동주택에서 5종 검침 데이터의 오류 분석 및 보정

\*강송희, \*\*홍은영, \*\*\*김동식, \*\*\*\*정범진

\*스트라드비전, \*\*몬드리안에이아이

\*\*\*한국외국어대학교 전자공학과, \*\*\*\*서울과학기술대학교 전기정보연구소

e-mail: *songhee.kang@stradvision.com, eunyeong@mondrian.ai, dskim@hufs.ac.kr, bjchung@seoultech.ac.kr*

## Error Analysis and Correction of 5 Kinds of Meter Reading Data in Apartment Houses

\*Songhee Kang, \*\*Eunyeong Hong, \*\*\*Dong Sik Kim, \*\*\*Beom Jin Chung

\*StradVision, \*\*Mondrian AI

\*\*\*Department of Electronics Engineering, Hankuk University of Foreign Studies

\*\*\*\*Research Center for Electrical and Information Technology, Seoul National University of Science & Technology

### Abstract

In apartment housing environments, typically five metering data: the electric energy, water, hot water, gas, and heat data, can be periodically collected through a data concentrator unit. We can utilize those metering data to improve each metering and prediction accuracy. The collection data, however, are usually corrupted by various errors during data collections. In this paper, we first analyze the error types and correct the errors. We then observe the collection data patterns. The electric energy usage has a similar 24-hour pattern with those of water, hot water, and gas. The heat usage pattern is quite different from other cases.

### I. 서론

에너지는 지속가능한 사회를 위한 중요한 자원이며, 에너지 공급중심의 에너지 정책에서 탈피하여 스마트시티, 스마트그리드 등 다양한 분야에서 에너지의 수요과 공급의 균형을 도모하기 위한 고도화된 에너지관리시스

템의 필요성이 대두되고 있다[1]. 고도화된 에너지관리에서는 에너지 수요를 실시간 예측하고 관리하는 기술이 요구되며 이를 위해서 대규모 수용가에서 소비되는 에너지 사용량을 원격에서 실시간으로 수집하는 원격검침 시스템이 확산되고 있다[2]. 국내 주택분야에서 원격검침시스템은 2000년대 초반 아파트에서부터 도입되었으며, 세대별 전기, 수도, 가스, 온수, 난방의 5종 검침을 통해서 아파트 세대원에게 에너지 사용량 및 요금 정보 등을 제공하고 있다. 최근 들어서 에너지데이터를 보다 적극적으로 활용하여 소비자 참여형 또는 댁내 기기연동형 수요관리 등 새로운 에너지 서비스로 확대하기 위한 연구가 진행되고 있다. 에너지데이터기반 서비스를 고도화하기 위해서는 실시간으로 수집·저장·관리되는 세대별 에너지원별 계량데이터의 품질수준을 먼저 분석해야 한다. 본 논문에서는 실거주 아파트에서의 세대별 5종 계량 데이터를 수집·저장·관리하는 과정에 발생하는 데이터 처리 오류의 유형을 분석하였다. 5종 계량 데이터의 오류들을 관찰 분석하여 총 여섯 가지의 오류형태로 분류하였다. 이를 통해서 오류 형태에 따른 오류 보정을 수행하였으며, 5종 검침 데이터의 평균 소비 형태를 시각화하여 각 시간별 소비패턴변화에 대해 관찰·분석하였다.

본 논문의 구성은 다음과 같다. 제 II장은 공동주택에서 원격검침으로 수집된 5종 데이터에서 발생하는 오류들을 정의하고 그에 따른 보정 방법을 서술한다. 제 III장은 오류 유형별 보정 작업이 완료된 5종 데이터의 평균 소비패턴을 분석한다. 마지막 장에서는 논문의 결론을 서술하여 마무리한다.

## II. 데이터 오류 형태 및 보정

본 장에서는 공동주택에서 검침한 5종 데이터에서 발생하는 오류들을 분석하고 보정 방법에 대해 서술한다. 본 실험에서는 15개동 총 1,007호수를 가진 공동주택에서 1시간 단위로 검침된 5종 데이터를 사용하였으며 수집된 기간은 2017년 4월 1일 0시부터 2019년 11월 6일 13시까지 총 31개월이다.

### A. 누적량 앞자리 수 오류

수집된 데이터에서 누적량의 앞자리 숫자 1이 9로 바뀌는 오류를 의미한다. 표 1과 같이 2019년 9월 23일의 전기 누적량의 앞자리 수는 전 데이터의 누적량과 상관 없이 9로 바뀌는 경우를 관찰할 수 있다. 이는 ISO 488 등을 통하여 공동주택에서 데이터를 수집할 때, 이진 데이터 0001에 1bit 오류가 발생하여 1001로 바뀐 경우가 된다. 일반적으로 이러한 1bit 오류는 전송 과정에서 패리티 비트 체크로 확인이 가능하다. 또한 이러한 오류는 이전 수신한 누적량 데이터를 사용하여 검출 및 복원이 가능하다.

보정 전			
년-월-일	시간	전기 누적량	전기 사용량
2019-09-23	21	19930.7	0.7
2019-09-23	22	99932.2	80001.5
2019-09-23	23	19932.5	-79999.7
2019-09-24	0	19932.8	0.3
보정 후			
년-월-일	시간	전기 누적량	전기 사용량
2019-09-23	21	19930.7	0.7
2019-09-23	22	19932.2	80001.5
2019-09-23	23	19932.5	0
2019-09-24	0	19932.8	0.3

표 1. 누적량 앞자리 수 오류 예와 1차 보정

본 논문에서 사용한 전체 데이터에서 약 0.008ppm 정도의 빈도로 매우 드물게 이러한 오류가 발생한다. 보정 방법은 표 1과 같이 9로 잘못 표기된 맨 앞자리 수를 1로 수정하고 표 1과 같이 큰 음수 값으로 잘못 표기

된 사용량의 값을 0으로 1차 보정하였다. 사용량의 값에 대한 오류는 추후에 다시 보정한다 (E 오류 참조).

### B. 신호 누적 오류

수집된 데이터에서 약 한 달 동안 사용량 데이터가 0인 경우를 의미한다. 표 2에서 보는 바와 같이 2019년 3월 24일 오전 7시 이후 데이터의 사용량은 0으로 유지되다가 2019년 4월 17일 오전 9시에 444.6이라는 큰 값이 나오는 것을 확인할 수 있다. 이러한 오류는 어느 기간 동안 누적량 데이터를 수신하지 못하는 경우 발생할 수 있다.

본 논문에서 사용한 5종 검침 데이터에서 이러한 오류는 약 0.001% 정도로 발생한다. 이에 따른 보정 방법은 해당 기간의 지난 달과 다음 달의 같은 날짜의 시간대의 사용량 평균을 비율로 고려해 누적된 사용량을 분배하였다.

보정 전				
년-월-일	시간	전기 누적량	전기 사용량	
2019-03-24	7	32060.2	0.6	
2019-03-24	8	32060.2	0	
2019-03-24	9	32060.2	0	
...				
2019-04-17	8	32060.2	0	
2019-04-17	9	32504.8	444.6	
보정 후				
년-월-일	시간	전기 누적량	전기 사용량	평균 사용량 비율
2019-03-24	7	32060.2	0.6	
2019-03-24	8	32061.4866	0.7624	0.00171
2019-03-24	9	32062.0107	0.52415	0.00117
...				
2019-04-17	8	32504.1074	0.6671	0.0015
2019-04-17	9	32504.8	0.6926	0.00155

표 2. 신호 누적 오류 예와 보정

### C. 누적량 리셋 오류

전자식 계량기는 사용량을 실시간으로 계측하고 그 값을 내부 기억장치에 누적하여 저장한다. 원격검침 시스템에서는 일정 검침 주기마다 이러한 누적량을 수집하는데, 계량기의 저장장치에서 알 수 없는 요인으로 인해서 누적량이 리셋되어 0으로 바뀔 수 있다. 표 3의 보정 전 내용과 같이 2019년 6월 25일 16시의 누적량 값을 확인해보면 -652로 음수 값을 가지는 것을 확인할 수 있다. 해당 오류는 약 0.18% 비율로 발생하여 빈도수가 상대

적으로 높다. 해당 오류의 보정 방법은 음수 값이 발생한 시간대의 사용량을 0으로 수정 한다. 누적량은 이전 시간의 누적량을 더하여 누적량이 작아지는 경우가 발생하지 않도록 보정해주었다.

보정 전			
년-월-일	시간	가스 누적량	가스 사용량
2019-06-25	15	652	0
2019-06-25	16	0	-652
2019-06-25	17	0	0
	...		
2019-06-27	18	0	0
2019-06-27	19	1	1

보정 후			
년-월-일	시간	가스 누적량	가스 사용량
2019-06-25	15	652	0
2019-06-25	16	652	0
2019-06-25	17	652	0
	...		
2019-06-27	18	652	0
2019-06-27	19	653	1

표 3. 누적량 리셋 오류 예와 보정

보정 전			
년-월-일	시간	수도 누적량	수도 사용량
2018-03-07	13	784.5	0
2019-06-25	14	784	0
	...		
2018-03-09	20	784	0
2018-03-09	21	785	1

보정 후			
년-월-일	시간	수도 누적량	수도 사용량
2018-03-07	13	784.5	0
2019-06-25	14	784.5	0
	...		
2018-03-09	20	784.5	0
2018-03-09	21	785	1

표 4. 누적량 역전 오류 예와 보정

D. 누적량 역전 오류

원격검침의 경우 이제까지 사용한 양의 누적량을 송신하므로, 최소한 검침 누적량은 같거나 이전 값보다 커야한다. 그러나 검침 결과를 보면, 현재 누적량이 이전 누적량의 값보다 작은 경우가 있다. 표 4의 보정 전 내용을 확인해보면 2018년 3월 7일 13시 사용량(784.5)이 14시 사용량(784)보다 큰 경우를 관찰할 수 있다. 해당 오류는 약 0.17%의 비율로 발생하였다. 보정 방법은

k-1시의 누적량과 k시의 누적량을 비교하여 k-1시의 누적량의 값이 큰 경우 k시의 누적량을 k-1시의 누적량으로 보정하였다. 이 때 계속 이전 시간의 누적량으로 보정해 나가다 k시의 누적량이 k-1시의 누적량보다 커질 때 더 이상 보정을 진행하지 않는다.

E. 사용량 계산 오류

사용량은 계산에 의해 채워지는 값으로 'k-1시의 누적량 + k시의 사용량 = k시의 누적량'의 규칙을 따른다. 해당 오류는 이전 오류 보정 과정에서 발생하는 변화로 인해 규칙에 맞지 않은 사용량 값을 의미한다.

발생비율은 C 오류와 동일하게 발생 되었다. 보정 방법은 모든 사용량에 대해서 'k-1시의 누적량 + k시의 사용량 = k시의 누적량' 규칙에 의해 값을 새로 구해 채워 넣었다.

보정 전			
년-월-일	시간	수도 누적량	수도 사용량
2018-03-09	20	784.5	0
2018-03-09	21	785	1

보정 후			
년-월-일	시간	수도 누적량	가스 사용량
2018-03-09	20	784.5	0
2018-03-09	21	785	0.5

표 5. 사용량 계산 오류 예와 보정

F. 시간 누락 오류

해당 오류는 특정 시간대의 데이터 값이 누락되는 경우를 의미한다. 모든 세대에서 같은 날짜 동일한 시간대에 주어진 기간 동안 3번 발생하는 오류이다.

보정 방법은 C 오류와 같이 이전 주와 다음 주의 동일한 요일의 사용량 평균을 비율로 하여 누락된 사용량을 분배해 주었다. 비율의 경우는 3가지가 있다.

1. 지난주와 다음 주 사용량의 평균이 특정 비율을 가지는 경우, 빈 시간의 사용량을 특정 비율 기반으로 누적 사용량을 배분한 후 누적량을 계산해 채워준다. 표 6의 수정전 12시의 전기 사용량이 비율에 맞게 계산 될 경우 0.162가 되는데 이때 누적량이 20547.236으로 20547.2보다 크기 때문에 넘치는 값을 맨 마지막 사용량에서 조절해주었다.
2. 표 6의 보정 전 가스 사용량을 확인해보면 누락된 시간의 누적 사용량이 0인 경우, 빈 시간의 누적량은 누락 이전 시간의 누적량으로 대신하고 사용량을 0으로 한다.
3. 표 6의 보정 전 수도 사용량을 확인해보면 지난주와

보정 전										
년-월-일	시간	전기 누적량	전기 사용량	가스 누적량	가스 사용량	수도 누적량	수도 사용량			
2018-03-07	9	20546.3	0.3	376.6	0	1021.5	0			
2018-03-07	12	20547.2	0.9	376.6	0	1021.6	0.1			
보정 후										
년-월-일	시간	전기 누적량	전기 사용량	평균 사용량 비율	가스 누적량	가스 사용량	평균 사용량 비율	수도 누적량	수도 사용량	평균 사용량 비율
2018-03-07	9	20546.3	0.3		376.6	0		1021.5	0	
2018-03-07	10	20546.9	0.612	0.64	376.6	0		1021.53	0.03	0.3
2018-03-07	11	20547.0	0.162	0.18	376.6	0		1021.56	0.03	0.3
2018-03-07	12	20547.2	0.126	0.18	376.6	0		1021.6	0.04	0.4

표 6. 시간 누락 오류 예와 보정

다음 주 사용량이 모든 시간에서 0인 경우, 1/n으로 누적 사용량을 배분한 후 누적량을 계산해 채워준다.

G. 그 밖의 오류

A부터 F까지의 6가지 오류는 기본적인 검침 데이터의 형식에 어긋나므로 검출이 용이하다. 그러나 이러한 기본 형식을 준수하고 있지만 데이터의 내용에 이상이 있는 경우가 있다. 예를 들면, 날짜마다 1,007개 호수에 대한 시간 당 평균을 구하였을 때, 그림 1과 2의 예에서 보는 바와 같이 중간에 갑자기 올라갔다 0으로 내려가는 이상 패턴이 관찰된다. 이는 전체 세대의 계량데이터를 수집하는 장치인 DCU(data concentrator unit)에서 데이터를 검침하는 과정에서 문제가 생기는 경우로, 이 시간대에서는 모든 세대의 5종 데이터에서 이상이 발생하고, 이러한 오류를 세대별 에너지원별 요일별 특성을 고려하지 않고 임의로 처리함으로써 이전 데이터 패턴과 상이한 형태의 사용 패턴을 가지게 된다. 이러한 오류는 보다 심도있게 관찰하여 보정해줄 필요가 있다.

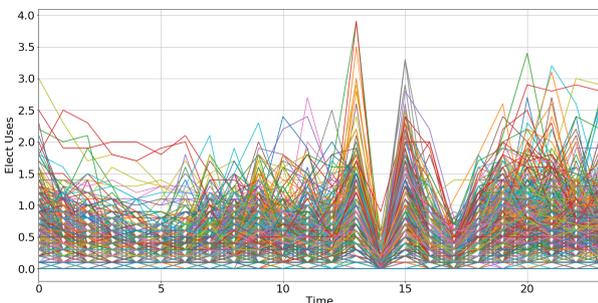


그림 1. 2018년 12월 12일 유의미한 전기 사용량을 가지는 900 가구의 사용량 예

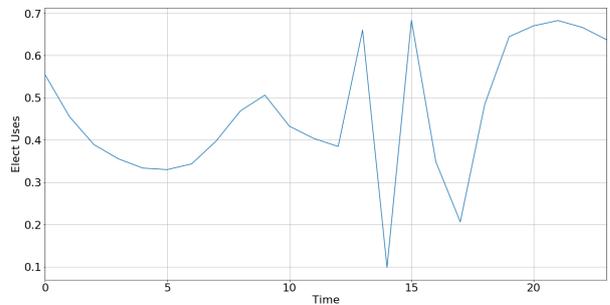


그림 2. 그림 1의 900 가구의 시간대별 평균 전기사용량.

III. 5종 검침 데이터의 패턴 분석

본 장에서는 제 II장에서 서술한 오류들을 보정하여 각 5종 데이터의 시간별 전체 평균 사용량을 구하여 에너지 소비패턴을 간단히 비교하였다.

그림 3의 전기 사용 곡선을 보면, 대체적으로 출근시간인 오전 5-7시 경에 전기 사용량이 증가하는 패턴을 보이고 있고 이후에는 점차 전기 사용량이 감소하다가 퇴근 시간인 20시 이후 급격하게 전기 사용량이 증가하는 경향을 보이고 있다. 이러한 패턴은 전기 에너지뿐만 아니라 수도, 온수, 가스 에너지의 경우에도 비슷한 사용 패턴을 보여주고 있다.

반면에 난방 계량데이터의 경우 계절에 의한 영향이 뚜렷하기 때문에 전기, 수도, 온수, 가스의 4종 계량 데이터와 다른 패턴을 보여주고 있다. 난방 계량데이터는 10시-15시에 급격히 감소하는 경향을 보이다가 20시 이후에 급격하게 증가하는 형태를 보여주고 있다. 난방 검침데이터 경우 다른 검침데이터에서 보여준 소비 패턴과

다른 패턴을 보여주고 있는 것을 확인할 수 있다.

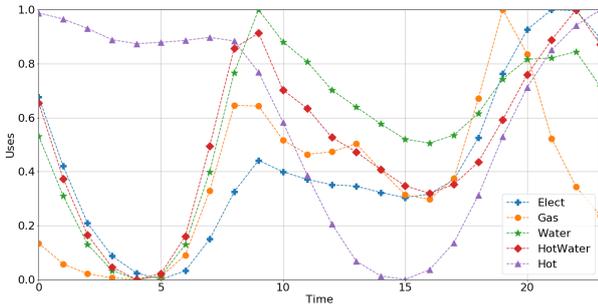


그림 3. 시간별 평균 5종 데이터 사용량

#### IV. 결론

본 논문에서는 공동주택에서 수집된 5종 원격검침 데이터를 관찰하여 다양한 오류 형태를 분류하고 오류 이를 보정하였다. 보정된 5종 데이터를 분석한 결과, 시간별 평균 전기, 수도, 온수, 가스 데이터 사용량은 세대의 출퇴근 생활패턴에 맞추어서 에너지 사용량이 변화하는 패턴을 보여주었고, 계절의 영향을 받는 난방 에너지 사용패턴은 그 외 4종 에너지 데이터와는 다른 형태를 보여주었다. 향후 연구로는 이렇게 보정된 5종 에너지 검침 데이터를 사용하여 에너지원별 상관도를 관찰하고, 에너지원별 세분화 및 딥러닝 기반 에너지 사용량 예측 기술 등을 개발하여 에너지 데이터의 활용도를 확대하고자 한다.

#### ACKNOWLEDGMENT

본 연구는 산업통상자원부(MOTIE)와 한국에너지기술평가원(KETEP)의 지원을 받아 수행한 연구 과제입니다.(No. 20191210301580)

#### 참고문헌

[1] 산업통상자원부, “제 3차 에너지 기본계획”, 2019.  
 [2] 양현재, 김동식, 정범진, “공동주택에서 전기에너지 사용량과 다양한 검침 데이터간의 상관도 분석,” 대한전자공학회 하계학술대회 논문집, 2020.  
 [3] E. A. Feinberg and D. Genethliou, “Load Forecasting”, Applied Mathematics for Restructured Electric Power Systems, pp. 269-285, 2005.