

# IoT 디바이스 기반 공공서비스 데이터의 지능형 예측 모델

정재희\*, 지민근\*\*, 이강만\*\*

\*명지대학교 정보통신공학과

\*\*동국대학교 멀티미디어공학과

e-mail : jhjung@mju.ac.kr, mingeun@mme.dongguk.edu, gangman@mme.dongguk.edu

## An intelligent predictive model for analyzing public service data on IoT device

Jaehee Jung\*, Mingeun Ji\*\*, Gangman Yi\*\*

\*Dept. of Information and Communication Engineering, Myongji University

\*\*Dept. of Multimedia Engineering, Dongguk University

### 요약

수도, 전기, 가스와 같은 공공서비스의 과금을 위하여 전통으로 행해진 방문 검침 방법은 오류 및 과금 방식의 한계가 있다. 수동적인 검침 방법을 극복하기 위하여 계량기를 디지털화하여 원격 검침이 가능한 시대로 발전되고 있다. 원격 검침으로 쌓여진 데이터를 분석하여 누수, 누전과 같은 비정상적 상황에 대한 감지가 가능하면 공공서비스 수요 관리 및 대책 정책 수립에 도움이 될 수 있다. 따라서 제안하고자 하는 방법은 세대별 공공서비스 사용량에 대하여 각 세대별 분석을 위하여 각 세대나 업소의 사용에 영향을 미치는 요인을 파악하고 비정상적 사용 감지에 대한 예측 모델을 구축하고자 한다.

### 1. 서론

2019년 발표된 상수도 통계에 따르면 2017년도 사용기준으로 가장 큰 비중을 차지하는 상수도는 가정용수이다[1]. 하지만 여전히 수도의 사용량을 검침하기 위하여 검침원이 개인 가구나 업소를 방문하여 검침하고 기록하는 일을 담당하여 상수도 요금을 부과하고 있다. 특히 비 도심권에서 수도사용량을 검침하기 위하여 세대나 업소방문으로 과금을 계산하는 방법은 시간적 노력적 한계에 다다르고 있다.

기존의 연구는 주로 가정용수에 대한 사용량에 대한 모형 연구와 원격 검침 시스템 개발에 관한 연구가 주를 이루고 있다[2]. Myoung *et al.*[2]는 가정내에서 사용되는 물 사용량을 예측하여 물수요에 대한 관리 및 정책에 활용 될 수 있는 데이터를 마련하였다. Ju *et al.*[3]은 국외의 디지털 수도 계량기의 원격 검침 시스템에 관한 동향을 고찰하였으며, 최신의 무인 검침 동향으로 데이터를 통합서버에 저장하여 검침하고 관리하는 시스템과 데이터 전송을 위한 전송 방식에 관한 연구를 소개하고 있다.

물리적인 원격 시스템뿐만 아니라 실제 각 세대의 데이터를 실시간 분석하여 누수와 같은 이상징후가 발생가능에 대비한다면 유수율 - 정수장에서 생산되어 공급된 수도량 중 수입으로 책정될 수 있는 비율을 높일 수 있다. 따라서 제안하는 방법은 상수도 요금 부과에 대한 효율성을 증대하기 위하여 각 세대마

다의 요일 별, 시간별, 월별 사용량에 유의미한 사용량의 변화가 있는지 살펴봄으로써 각 세대마다 수도 사용량에 대한 예측 모델을 생성하여 수도 사용량의 비정상적 탐지를 하고자 한다. 즉, 세대별 무선 원격 검침 시스템을 개발하여 수도 사용량에 대한 지능형 원격 분석이 가능한 모델을 개발하고자 한다.

### 2. 방법

#### 2.1 데이터

본 논문에서 사용된 공공서비스 데이터는 2018. 08. 21부터 2019. 05. 31 까지 수집된 수도 사용량 데이터로 총 869 개의 세대로부터 취득 되었다. 전자식 유량계에서 각 세대마다 하루에 3 번 (0~8 시, 9~16 시, 17~24 시) 서버로 전송된 데이터를 사용하였으며, 전송된 데이터는 다양한 속성을 가지고 있으나, 필드 값 미사용 경고나 계량기 무응답, 계량기 데이터 오류 코드와 같은 코드가 전송되는 레코드 값은 학습에 사용되는 데이터에서 제외하였다.

#### 2.2 전 처리를 통한 학습 데이터 생성

서버로 전송된 데이터는 세대에서 순서 없이 무작위 방식으로 저장되기 때문에, 각 세대별 데이터의 재정렬이 필요하다. 분석의 목적이 각 세대나 업소별 실시간으로 검침되어 전송된 데이터가 비정상적인 범

위인지 확인하기 위함이기 때문에, 예측을 위한 학습 데이터가 필요하다. 학습 데이터는 세대별로 정렬되어 전 처리된 데이터이다. 전 처리된 데이터가 충분히 축적되어 있을 경우 실시간으로 입력되는 데이터는 테스트로 사용되고, 한번도 전 처리된 데이터가 없는 상태라면 초기 학습데이터로 사용 된다. 이때 요일 별 분석을 위하여 최소 2 주의 데이터를 필요로 하기 때문에 2 주 미만으로 세대에서 전송되지 않는 경우는 필터링하여 학습 및 테스트 데이터로 사용하지 않는다.

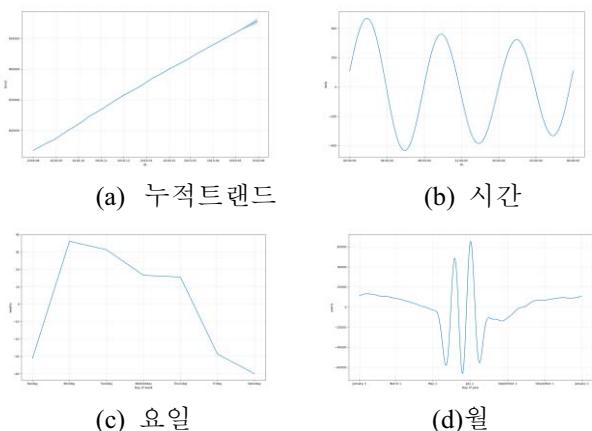


그림 1. “140C5BFFFF1300D7+17100585”

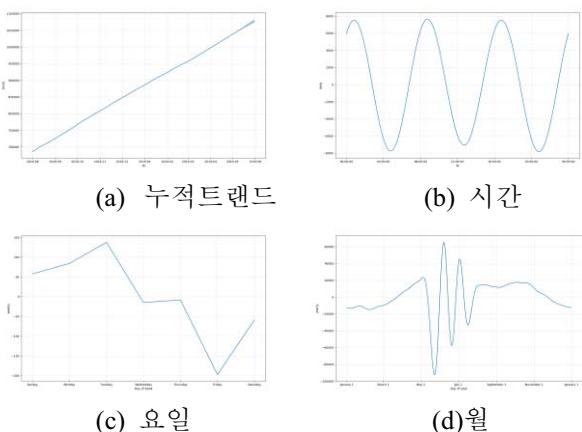


그림 2. 140C5BFFFF130097+17100634

위 그림 1 은 계량기에서 전송되어 각 세대의 고유 값을 나타내는 ResourceID: 140C5BFFFF1300D7, WM\_ID: 17100585 의 누적 트렌드 및 예측 값의 요일, 시간, 월에 대한 각 요인 분석을 의미한다. 그래프의 데이터 흐름을 통해서 해당 ID 의 요일, 시간, 월별의 사용량을 대략적으로 파악이 가능하다. 그림 1(a)는 누적 트렌드를 의미하여, 해당 ID 의 2018년 8월부터 2019년 5월까지의 전체적인 누적 수도 사용량을 나타낸다. 2019년 5월 이후의 하늘색 영역으로 표시된 부분은 예측 값을 의미하여, 예측 값의 최대 최소 범위와 함께 표시되고 있다. 그림 1(b)는 시간별 그래프

를 의미한다. 현재 데이터는 하루에 3 번 (02:00, 10:00, 18:00 근처) 수집된다. 실제 데이터의 시간별 트렌드와 다소 차이가 있을 수 있지만, 어느 시간에 가장 많은 수도를 사용하는지를 분석이 가능한 상대 그래프이다. 그림 1(c)은 해당 세대의 요일 별 수도 사용을 의미한다. 월일이 약 35 정도로 표시되고 있고 금요일이 -30 정도로 표시되고 있는데, 이 값은 절대 값이 아닌 상대 값을 의미하여 월요일이 금요일 보다 상대적으로 더 많은 수도 사용함을 의미한다. 그림 2(d)는 월별 그래프이지만 다년간 데이터가 수집되지 않아 5 월부터 8 월까지는 푸리에 급수에 의해 균사치로 추정된 값이기 때문에 실제 데이터의 트렌드와는 다소 차이가 있을 수 있다.

그림 2 는 ResourceID: . 140C5BFFFF130097, WM\_ID: 17100634 의 다른 세대의 시간, 요일, 월별 사용량의 경향과 누적 경향 및 예측 값을 보여주고 있다. 그림 1 과 2 에서 보는 바와 같이 각 세대별 수도를 사용하는 패턴의 경향이 서로 다른 것을 볼 수 있다. 따라서 제안하는 방법은 가법 모델을 이용하여 시간 요일 월이 모두 사용량에 영향을 미치는 요인으로 변수화 가능 하여 모두 합하여 각 세대의 사용량 분석에 사용하였다.

### 2.3 각 세대별 테스트

각 세대별 테스트를 위하여 세대별 가법 모델을 생성하였으며, 이미 누적된 학습된 데이터를 이용하여 사용예측이 가능한 최소 값과 최대 값의 범위영역을 예측한다. 예측된 범위에서 실제 검침되어 전송되는 누적량을 벗어 나는지에 확인 가능한 예측 시스템을 구축 할 수 있다.

### 3. 결론

각 세대별로 누적되어지는 수도량에는 시간별 월별 요일별 서로 다른 사용패턴이 존재 함을 확인하여 이를 차후 세대당의 사용량 예측 모델의 변수로 사용할 수 있다. 모델을 바탕으로 예측의 범위를 벗어난 경우 특히 예측 범위의 최대 값을 초과할 겨우 누수와 같은 현상을 의심할 수 있다. 특히 단발성이 아닌 지속적인 현상이 발생할 경우에 경험치적 학습을 추가하여 세대별 예측 범위를 벗어난 예외상황에 대한 정교한 학습모델을 생성 할 수 있을 것으로 기대한다.

### 참고문헌

- [1] 환경부 17년도 상수도 통계  
[http://me.go.kr/home/web/policy\\_data/read.do?pagerOffset=0&maxPageItems=10&maxIndexPages=10&searchKey=title&searchValue=%EC%83%81%EC%88%98%EB%8F%84&menuId=10264&orgCd=&condition.deleteYn=N&seq=7316](http://me.go.kr/home/web/policy_data/read.do?pagerOffset=0&maxPageItems=10&maxIndexPages=10&searchKey=title&searchValue=%EC%83%81%EC%88%98%EB%8F%84&menuId=10264&orgCd=&condition.deleteYn=N&seq=7316)
- [2] S. Myoung, D. Kim , J. Jo, Estimating Bathroom Wateruses based on Time Series Regression, 19(8), Journal of the Korea society of computer and information, 2014
- [3] J.C.Joo, H.Ahn, C.H.Ahn, K-R.Ko, H. Oh, Recent Developments and Field Application of Foreign Waterworks Automatic Meter Reading Journal of Korean Society of Environmental Engineers 2012