

전기자동차 충전기록 데이터 처리에 관한 연구

황운영¹ 진효정² 김소연² 이정훈²¹제주대학교 회계학과²제주대학교 전산통계학과

{etace17, hgnico, jhlee}@jejunu.ac.kr, carol7378@naver.com

A study on data processing of electric vehicle charging archives

Yunweong Hwang¹ Hyojeong Jin² Soyeon Kim² Junghoon Lee²¹Dept. of Accountancy, Jeju National University²Dept. of Computer Science and Statistics, Jeju National University

요 약

본 논문에서는 환경공단이 제공하는 전기자동차 충전기의 운영기록 분석하는 과정에서 주기적으로 공개 데이터를 수집하여 지역 데이터베이스에 저장하고 데이터의 오류를 정제하는 방안에 대해 연구한다. 전력시스템 부하에 직접적으로 영향을 주는 급속충전기 운영기록만을 추출하고 날짜필드에서의 오류 혹은 역진을 포함하는 트랜잭션을 제거한 후 일차적으로 충전시간길이를 히스토그램으로 분석한다. 대부분의 충전이 20분 이내에 완료되었지만 23%는 충전완료 후에도 충전기에서 플러그를 제거하지 않은 것으로 보인다.

1. 서론

스마트시티 플랫폼은 도시생활의 다양한 분야에서 기능적인 컴퓨터 기술을 종합하여 시민들의 편의성을 증진시키고자 하는데 이를 위해서는 다양한 데이터들을 수집 분석하는 단계가 필수적이다[1]. 더욱이, 2013년부터 활발하게 운영중인 정부의 공개 데이터 사이트(data.go.kr)에서는 교통, 신재생 에너지, 관광 등 다양한 분야의 데이터들을 파일 혹은 OpenAPI 등의 형태로 제공하고 있어서 이를 이용한 데이터분석 및 관련 서비스 혹은 비즈니스 모델 등의 개발이 가능하다[2].

최근 전기자동차의 보급 확대와 함께 환경공단에는 충전인프라에 관련된 데이터도 공개하고 있다. 이 데이터는 최근 10분간 상태가 변경된 충전기의 정보들만 주기적으로 공개되므로 지속적으로 조회하여 지역 서버에 저장하고 분석하여야 한다. 또 3월 말에 공개 데이터의 포맷이 급격히 변경되었으며 각 필드내에 이상치가 많이 발견되어 이들에 대한 제거 작업이 필요하다. 본 논문에서는 전기자동차와 충전인프라 보급이 가장 활성화된 제주지역에 대해 충전 데이터를 정제하는 방안에 대해 연구하고 이를 바탕으로 전기자동차 충전이 전력시스템 부하에 미치는 영향을 추정한다.

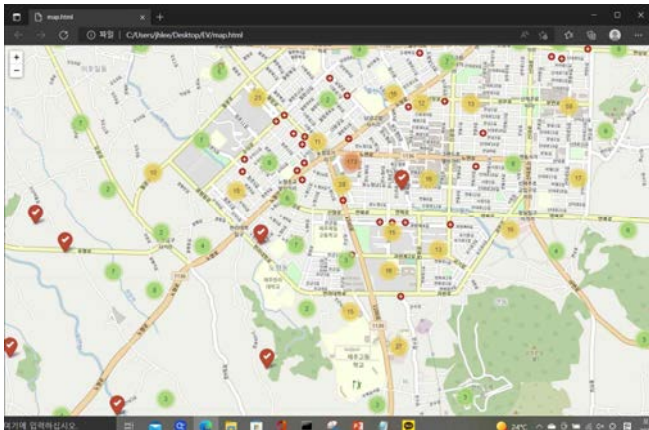
2. 데이터 셋의 특성

환경공단에서 제공하는 공개 데이터는 정적인 정보와 동적인 정보를 구분하는데 정적인 정보는 충전소명, 충전소 ID, 충전기 ID, 충전기타입, 주소, 위도, 경도, 운영기관명 등을 포함한다. 반면 주기적으로 갱신되는 동적인 정보에는 운영기관, 충전소, 충전기, 충전기상태, 상태갱신일시, 최근충전시작일시, 최근충전종료일시, 현재충전시작일시 등이 포함된다. <그림 1>은 데이터 수집과정을 보여주며 공개 데이터 사이트에서 발급받은 접근 키를 내장한 파이썬 프로그램을 통해 웹 페이지에서 상태 정보를 읽어와 로컬 데이터베이스에 저장한다. MySQL 상에 정적 정보와 동적정보 테이블이 구축되었으며 필요에 따라 join된다[3].



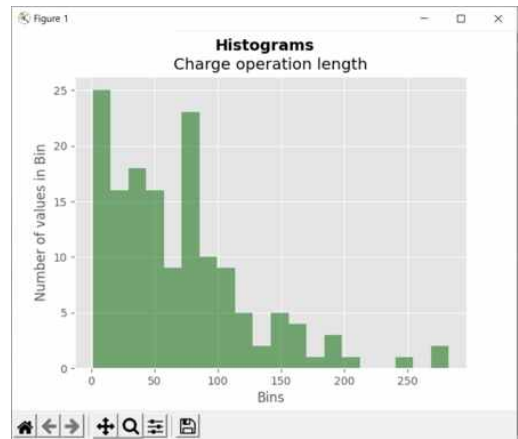
<그림 1> 충전기 상태정보 수집 구조

정적인 정보 분석에 있어서 현재 제주도에서 등록된 충전기의 수는 5,000여개 인데 이중 3,700여개는 완속충전기이고 1,555개는 급속 혹은 콤보 타입이다. <그림 2>는 각 충전기의 위치에 따라 파이썬의 folium 라이브러리를 통해 지도상에 충전기들의 밀도를 표시한 것이다. 이는 제주도에서 가장 인구 및 충전기가 밀집한 지역을 보여주고 있으며 지역에 따라 충전기의 밀도는 더큰 차이를 보인다. 또 등록된 충전기들에 있어서 운영사의 정책 혹은 고장 유무 등에 따라 실제로 상태를 보고하고 있는 충전기들은 급속의 경우 179기밖에는 되지 않는다.



<그림 2> 제주지역 전기자동차 충전기 위치 분포

종료시작일시가 역전이 된 경우는 31개가 발견되었으며 충전시간이 3시간 이상으로 파악된 경우는 347건에 달한다. 이 경우 운전자가 자동차를 플러그인해놓고 다른 용무를 본 경우라고 볼 수 있으나 전체적인 부하 분석에서는 의미가 없어 제거하는 것이 바람직하다. 또 충전 레코드의 특성상 중복된 트랜잭션을 포함하는 경우가 있어서 파이썬의 집합 기능을 이용해 제거하였다. 결국 총 10,274개의 트랜잭션이 유효한 것으로 판명되었고 이를 이용하여 충전시간의 히스토그램을 작성하면 <그림 3>과 같이 30분 이내 충전이 많은 비중을 차지한다.



<그림 3> 충전시간 길이 분포

3. 데이터 정제과정

전력시스템의 부하에는 완속보다는 급속충전기가 미치는 영향이 7-8배 이상이 된다. 또 완속충전기는 정상적으로 보고를 하지 않은 경우가 많기 때문에 급속충전기에서 발생하는 데이터에 중점을 둔다. 현재의 충전상태를 앱이나 웹으로 제공하는 서비스를 넘어 충전이력을 분석하는데 있어서는 충전트랜잭션의 시작과 종료시간 분석에 중점을 둔다. 시간의 흐름을 따르는 절차적인 과정 분석은 SQL보다는 파이썬과 같은 언어를 통해 예외분석이나 순차적인 상태의 추적 등을 수행하는 것이 효율적이다. 먼저 전체 로그에서 급속충전기들의 ID를 조회하고 이를 운영 로그에 조인하여 급속충전기에 관련된 트랜잭션을 분리한다.

이 트랜잭션들을 검토해보면 특히 날짜 부분에서 오류가 많다. 2022년 5월 3일부터 현재까지의 26,014개의 급속트랜잭션 중에서 NULL 혹은 기준일 이전의 값을 갖는 날짜 필드를 갖고 있는 트랜잭션은 1,658 개다. 이러한 경우는 파이썬의 예외처리 기능을 이용하여 탐지하고 제외한다. 또 충전시작일시와

3. 결론

본 논문에서는 공개 데이터 사이트에 공개된 제주 지역 급속전기자동차 충전기 운영 데이터를 다운로드하여 부정확하거나 손실된 필드를 포함하고 있는 트랜잭션 레코드들을 제거하였다. 이를 바탕으로 충전길이 분포를 확인하였고 추후 최대부하구간 탐지, 광고를 위한 시간구간의 특정 등을 수행할 수 있다.

Acknowledgment

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학지원사업의 연구결과로 수행되었음(2018-0-01863)

참고문헌

- [1] <https://www.smartcity.go.kr>
- [2] <https://www.data.go.kr>
- [3] J. Park, et al., Data analysis framework for electric vehicle charging networks in Jeju city," ICESI, 2022.