VOL. 13, NO. 5, MARCH 2018 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

BIG DATA PROCESSING IN CHARGING INFRASTRUCTURES FOR SMART TRANSPORTATION SYSTEMS

Junghoon Lee and Gyung-Leen Park
Department of Computer Science and Statistics, Jeju National University, Jeju-City, Republic of Korea
E-Mail: jhlee@jejunu.ac.kr

ABSTRACT

This paper presents a data processing framework for a charging infrastructure monitoring system, embracing how to acquire open raw data, to manage in the database, and to conduct diverse analysis. Real-time status reports from 49 fast chargers in the target vicinity are consistently retrieved, parsed, and stored in our local machine equipped with a variety of open software solutions. The platform implements our own map interface as well as exploits R display packages. The analysis procedure investigates the occupancy rate according to hour-of-day, day-of-week, and charger type. Its results reveal that the city achieves about 18 % utilization, while the chargers are not mainly used by commuters. In addition, the charger-by-charger analysis finds out that 12.2 % of chargers show over 40 % utilization during the hot hours. It opens the possibility of integrating other big data sets such as geographic information, weather condition, and power consumption.

Keywords: electric vehicle, charger monitoring stream, big data analysis, occupancy rate, hot hours

1. INTRODUCTION

Recently, we are witnessing not so fast but steady penetration of EVs (Electric Vehicles) in many countries. As an important player in the modern power system called *smart grid*, EVs can achieve energy efficiency and reduce air pollution [1]. Moreover, their main weak point in short driving ranges is being overcome by the extensive installation of charging infrastructures, starting from urban areas. Charging facilities are essentially monitored and managed by a central authority, be it public or private, taking advantage of mature communication technologies. The real-time control system can respond to the external events, such as malfunction of some parts, power limit violation, or even price signal changes. Digitally controlled chargers and EVs make the transportation system as a part of power networks, even allowing us to develop a city-wide coordination mechanism [2]. Moreover, this system keeps accumulating the status reports, making it possible to conduct diverse analysis on the past behaviour of the infrastructure to make a new operation plan. Here, the seamless interaction between physical and cyber spaces becomes essential [3].

In the meantime, some cities and countries open such data to the public to encourage the development of a new service in civil domains. This service ranges from a simple display of current object status to a sophisticated prediction and provisioning. To develop a more useful service, it is necessary to integrate other data streams such as weather parameter change, power consumption dynamics, renewable energy generation, and the like. In smart grids, many efforts are under development on such big data [4]. Moreover, even non-data specialists can build an efficient prediction model by a variety of high-quality tools including R, MySQL, Hadoop, Tensorflow, and the like [5]. Those tools are made available as open software along with easy and comprehensive user interfaces. Particularly, analysts can easily benefit from up-to-date machine learning and massive data management techniques. The deep technical background is no more an indispensable prerequisite for data analysis.

Through the portal site opened by Korean national government, it is possible to develop an automatic data acquisition module following the predefined endorsement procedure and the announced standard access mechanism. As more records are acquired and stored in our local machines, we can conduct various data analysis to obtain performance statistics on the target charging infrastructure, visualize the results, build accurate prediction models, and design new business models [6]. Moreover, we can integrate other data streams on a necessary basis. It must be mentioned that our previous work has also analysed the monitoring data stream in charging stations [7]. Managed by a different charging facility cooperative, this dataset does not publish exact location of each charger, and its charging operations do not look sufficiently stable. On the contrary, this paper targets at the chargers consistently managed by administrative authorities, hence the data volume stably keeps growing.

The rest of this paper is organized as follows: After outlining the paper in Section 1, Section 2 reviews some related work on big data processing for charging infrastructure. Section 3 develops our data processing framework consisting of data acquisition, management, and analysis. Section 4 presents the analysis result and discusses what can be found. Finally, Section 5 concludes this paper with a brief introduction of future work.

2. RELATED WORK

First, NEC has built a completely commercial EV charging infrastructure system in which real-time monitoring and control is possible [8]. A set of charging controllers coordinate the energy distribution and cooperation management across hundreds of chargers. Specifically, each controller not only implements comprehensive management of charging functions but also provides a user interface allowing customers to start or stop charging. It can work with those chargers obeying the

compatible protocol standard. The inter-controller communication mechanism enables a facility-wide coordination such as limiting the amount of simultaneous charging. Impressively, controllers are connected to the remote NEC cloud, where a series of status records are being stored. This high-end server features remote monitoring, user authentication, membership management, and multi-service gateway for billing. NEC charging infrastructure has a potential to analyse the big data created from a variety of charging operations and facility management.

Next, Hitachi builds information-control platforms for data collection in smart cities [9]. These platforms embrace data collection, data analysis, and application coordination. Data collection is also accomplished mainly by facility monitoring systems from diverse sources like power consumption, equipment operation, EV charging, and the like. Data analysis adds value to the collected raw data, creating high-level timely decisions in control applications. This procedure covers data interpolation, prediction, and knowledge acquisition. Application coordination supports not only an efficient integration of renewable energy and grid power but also inter-EV cooperation for balanced charging. The authors emphasize the importance of the identification of trends in electric power consumption on districts or building clusters as well as the support for accurate energy supply decisions. Trend prediction makes it possible to shift EV charging load, locates the malfunctioning part, and even invites V2G (Vehicle-to-Grid) applications.

After addressing the essentiality of performance measurement for charging infrastructures, [10] identifies their key performance indicators. As for its classification, result indicators reveal how well a company has performed from a specific perspective, while performance indicators provide guidance what to do to improve the result. This approach defines 5 stakeholder groups for charging infrastructures, namely, municipality, EV users, non-EV users, commercial parties in the EV chain, and grid operators. The main concerns in each group lie in air quality improved, accessibility to the infrastructure, level of infrastructure utilization; cost decreased, and reduced risk of power outage, respectively. Then, those result indicators are associated with their performance indicators one by one, namely, amount of energy consumed for EV charging, charging time ratio, number of under-utilized stations, cost/benefit ratio, and peak power level. It will help policy makers to optimize the roll-out and improvement of the business case for charging infrastructures. Their viewpoints can be desirably considered incharging facility data analysis as much as possible.

3. DATA ACQUISITION

KEC (Korea Environment Corporation) runs a charging facility management system over the nation and allows permitted individuals to retrieve the real-time monitoring status of chargers any time they want [11]. This data stream is available at the well-known open data portal site, namely, *www.data.go.kr*. Figure-1 depicts how our data acquisition module works. After getting an access key via an off-line request, the acquisition module embeds the key in the HTTP *get* message and receives the web page using *WinInet* library functions in a Windows PC. This program is activated every 5 minutes and web pages are stored in our local machine as XML files. Then, the parser module, implemented by means of the Python *ElementTree* library, extracts the fields of interest and generates a series of SQL statements to insert the records into MySQL tables. As we are mainly concerned only on those records belonging to our vicinity, namely, Jeju City, others are filtered out. The target area is an island whose perimeter is about 200 *km* long, possessing 49 fast chargers under the management of KEC.

Each retrieval contains 49 status records, each of which includes station id, charger id, location, charger type, timestamp, and current status (whether it is charging an EV ornot). After all, 49 records are added in our database table every 5 minutes. We define 2 tables of *OpenLoc* and *OpenCharger*. The first stores static information such as latitude, longitude, type, and region of each charger, while station id and charger id comprise the primary key. This table does not change any more, unless a new charger is added to, or some charger is excluded from the management domain. In addition, *OpenCharger* keeps appending status information for each charger. Here, the primary key further includes the timestamp. It must be mentioned that this monitoring system does not include the amount of energy spent for EV charging. The energy amount can be just estimated by the constant electricity flow rate and the connection interval. Additionally, the real-time status records of only fast chargers are made available.

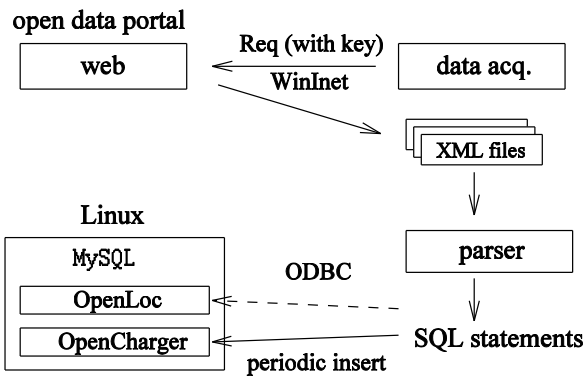

Figure-1. Data acquisition.

Since the location of each charger is obtained, it is possible to combine geographic information, also published by a governmental authority. We cannot only conduct a location based analysis but also develop charging-integrated routing services. We implement our own map viewer using Win32 API. It takes a standard road network published by Korean National Transport Information Center as a form of ESRI shape files [12]. All shapes are represented in the WGS84 coordinate system and converted to device coordinates for the user-specified boundary. The viewer supports basic map interfaces such

as zoom in, zoom out, pan, and reset of the target map objects. Figure-2 shows a subset of chargers included in the zoomed-in area. This figure also includes the utilization level of each charger from 0 to 5. Namely, five boxes are shown and the number of dark boxes is decided by the utilization level.

Figure-2. Map-based analysis.

4. STREAM ANALYSIS

Currently, the records from March 2 to 24 in 2017 are stored in our database. To begin with, Figure-3 traces the working status change of a charger during a single day, specifically, March 21. The graph is obtained by a set of SQL statements issued to the MySQL database. The curve is generated from 288 records, each of which contains 0 or 1. The working status remains 0 (non-working) until 8:30 AM, just before the start of office hours. Then, the status alternates each time an EV is connected to or leaves from the charger, until 11 PM, much beyond the end of office hours. A charging transaction seems to range from 5 to 60 minutes. For fast chargers, it takes at most 30 minutes to fully charge an EV, even in the case that its battery is almost depleted. However, we cannot exactly know whether a new EV is connected between two report periods. This situation makes a single transaction look so long. Anyway, Figure-2 shows the main characteristic of our monitoring record sets.

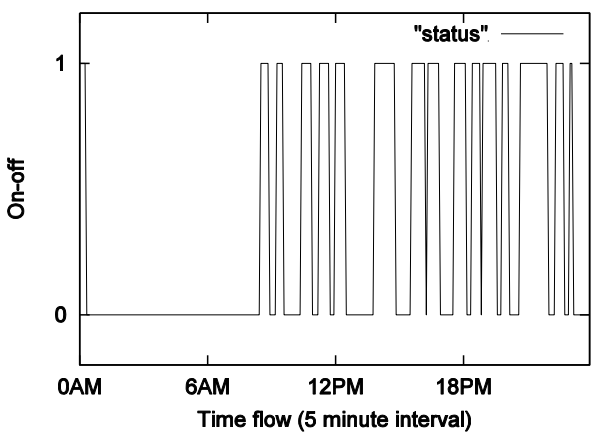

Figure-3. Working status change.

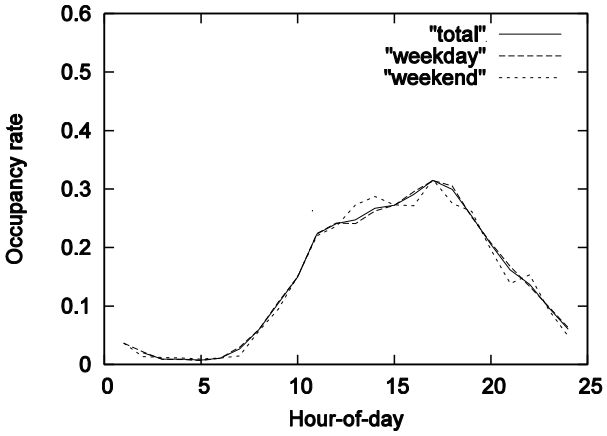

Figure-4. Hourly occupancy rate.

From now on, the occupancy rate will be the main observational perspective. It is the ratio between the number of records indicating a charger is working and the number of total records for a specific set. Figure-4 plots the citywide occupancy rate according to hour-of-day. The whole records are calculated without differentiating chargers. As expected, during the night time, almost every charger is idle, while the chargers hardly serve an EV until 7 AM, that is, the occupancy rate is less than 3%. The rate increases subsequently, reaching the peak of 30.0 % at 5 PM over the entire city. It is just before the end of office hours. However, even after the office hours, the occupancy rate does not drop sharply but stays above 10 % until 10 PM. Hence, we discriminate the rates for weekdays and weekends to check if commuters' charging demand has impacts on this rate distribution. However, as can be seen in the figure, three curves have no significant difference, showing that many non-commuters are charging EVs in the evening.

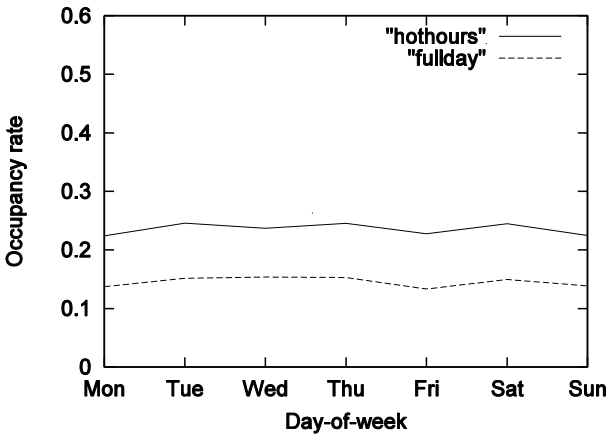

Figure-5. Day-of-week investigation.

Additionally, Figure 5 shows the occupancy rate according to day-of-week. In the figure, curves are largely at regardless of whether it is a weekend or weekday. We further define *hot hours*, during which the rate exceeds 10 %, and this interval falls in the range from 9 AM to 9 PM. The experiment investigates the occupancy rate in

this interval as most EV charging is concentrated. The difference between the two curves reaches 9.4 % on Tuesday, but daily difference is not so significant. If chargers are used mainly by commuters, the rate would be different between weekdays and weekends. This result also indicates that current EV users are tourists, local residents, and the like.

Next, Figure-6 exhibits how many chargers are occupied according to the charger type. Currently, two types are available in our region, one for (DC Chademo, AC 3-phase) labeled by *Chademo* and the other for (DC Chademo, AC 3-phase, DC Combo) labeled by *Chademo combo*. The combo type is more widely used these days. Out of 49 chargers, 20 belong to the first type, while the others to the second. The charging type field is obtained from the table joined with *OpenLoc*.

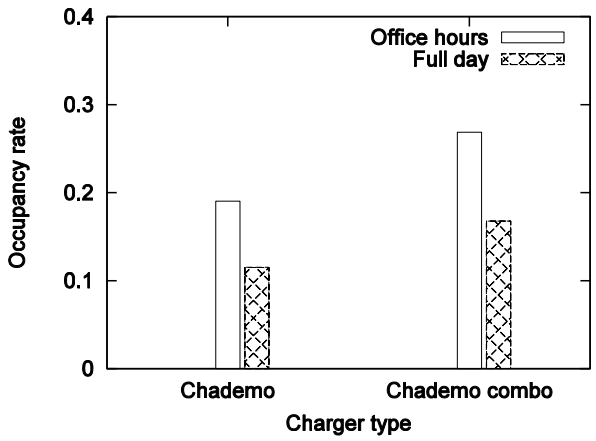

Figure-6. Charger type effect.

Figure-7 plots the day-by-day behaviour for charger facility utilization. First of all, the occupancy rate consistently remains between 10 % and 20 % for the whole observation interval. This result can encourage charging infrastructure builders to continue their business. In addition, the figure shows explicit daily fluctuation, which does not come from day-of-week, according to the previous experiment. It is necessary to investigate other effects such as weather conditions [13].

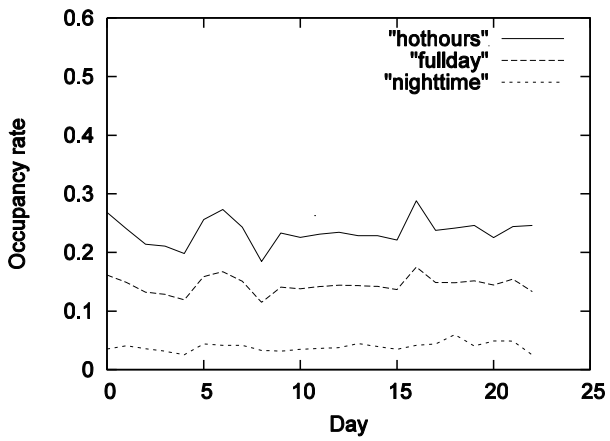

Figure-7. Daily occupation rate.

Now, Figure-8 shows per-charger statistics. Chargers are ordered just by their identifiers (station, charger), so there is no interdependency in their sequence. Here again, two occupancy rate curves, one for full day and the other for hot hours, are plotted. The curve for hot hours more fluctuates. 6 chargers exceed 40 %, while some staying around 10 %. The full day curve undergoes less severe fluctuations, as the night time usage is equally small for all chargers. Only two chargers show occupancy rate more than 10 % during the night time.

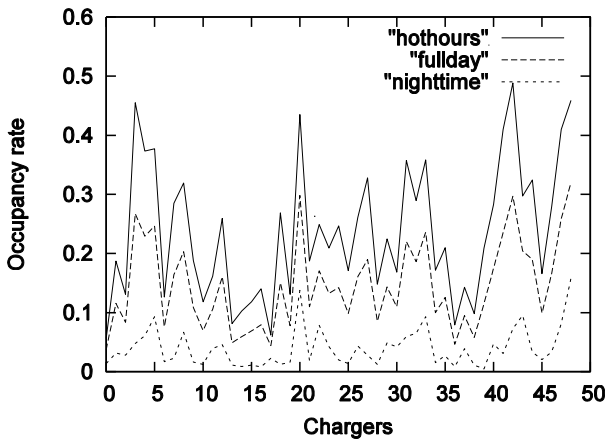

Figure-8. Charger-by-charger occupancy rate**.**

Finally, Figure-9 illustrates the rate for each charger in a single graph. The shapes are generated by the R package, after importing the SQL query result into the R working space, where abundant set of visualization functions are provided. We can see that the daily occupancy rate is largely flat for each charger. The difference between chargers is maintained during the whole interval.

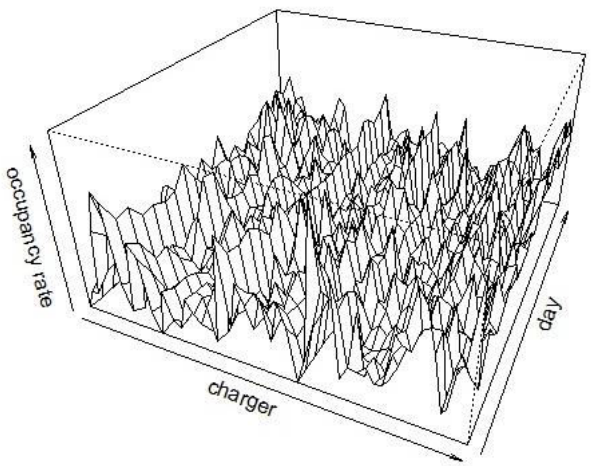

Figure-9. Perspective view.

5. CONCLUDING REMARKS

In this paper, we have presented our big data processing scheme for the dataset obtained from a charging infrastructure monitoring system, Real-time

status reports from 49 fast chargers in the target vicinity are consistently retrieved by a periodic batch program, parsed by a Python program, and stored in our local database tables. The platform implements our own map interface and exploits R for location based visualization. All these analysis steps are performed taking advantage of open data and software modules. The analysis procedure investigates the occupancy rate for the parameters including hour-of-day, day-of-week, and charger type. The analysis reveals that the city achieves 18 % utilization, while the chargers are not mainly used by commuters. In addition, charger-by-charger analysis finds out that 6 out of 49 chargers show over 40 % utilization during hot hours.

As future work, we are planning to develop business models related to charging infrastructures, particularly focusing on the integration of renewable energy and V2G (Vehicle-to-Grid) services [14]. For example, those chargers not used frequently will be the gateway of vehicle-oriented energy and renewable energy.

ACKNOWLEDGEMENTS

Prof. Gyung-Leen Park is the corresponding author. This research was supported by the 2018 scientific promotion program funded by Jeju National University.

REFERENCES

[1] A. Ipakchi, F. Albuyeh. 2009. Grid of the future. IEEE Power & Energy Magazine.pp. 52-62.

[2] J. Yin, P. Sharma, I. Gorton, B. Akyol. 2013. Large-scale data challenges in future power grids. IEEE Int'l Symposium on Service-Oriented System Engineering. pp. 324-328.

[3] A. Amato, R. Aversa, B. Martino, S. Venticinque. 2016. A cyber physical system of smart micro-grids. International Conference on Network-Based Information Systems. pp. 165-172.

[4] Z. Asad, M. Chaundhry. 2016. A two-way street: Green big data processing for a greener smart grid. IEEE System Journal.

[5] R. Bonnin. 2016. Building Machine Learning Projects with TensorFlow. Packt Publishing Ltd.

[6] R. Pinto, J. Pombo, M. Calado, J. Mariano. 2015. An electric vehicle charging station: Monitoring and analysis of power quality. 9th International Conference on Compatibility and Power Electronics.

[7] J. Lee and G. Park. 2016. Electric vehicle charger management system for interoperable charging facilities. JurnalTeknologi. 78(5): 117-121.

[8] A. Tsuyoshi. 2015. EV charging infrastructure system that facilitates commercialization of EV charging. NEC Technical Journal. 10(1).

[9] K. Iwamura, H. Tonooka, Y. Miznuo, Y. Mashita. 2014. Big data collection and utilization for operational support of smarter social infrastructure. Hitachi Review. 63(1).

[10] J. Helmus, R. Hoed. 2016. Key performance indicators of charging infrastructure. International Battery, Hybrid and Fuel Cell Electric Vehicle Symposium.

[11] J. Lee, G. Park, Y. Han, S. Yoo. 2017. Big data analysis for an electric vehicle charging infrastructure using open data and software. ACM eEnergy. pp. 252-253.

[12] J. Lee, G. Park. 2016. Analysis of stream data from electric vehicles for energy consumption statistics Advanced Science Letters. 22(11): 3454-3458.

[13] J. Yamazaki, D. Yoshiro, H. Fukuhara, T. Hayashi. 2015. A comprehensive data processing approach to the future smart grid. 4th Int'l Conf. on Renewable Energy Research and Applications. pp. 1033-1036.

[14] A. Ovalle, A. Hably, S. Bacha, G. Ramos, J. Hossain. 2017. Escort evolutionary game dynamics approach for integral load management of electric vehicle fleets. IEEE Transactions on Industrial Electronics. 64(2): 1358-1369.